Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Gratification Delay and Later Outcomes

Article *in* Psychological Science · May 2018

DOI: 10.1177/0956797618761661

CITATIONS
20

READS
2,431

3 authors, including:

Tyler Watts
New York University
17 PUBLICATIONS **283** CITATIONS

SEE PROFILE

Greg Duncan
University of California, Irvine
414 PUBLICATIONS **31,714** CITATIONS

SEE PROFILE

Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Gratification Delay and Later Outcomes

Tyler W. Watts[1]
New York University

Greg J. Duncan
University of California, Irvine

Haonan Quan
University of California, Irvine

Manuscript published at *Psychological Science*
May 25, 2018

1. *Corresponding Author*
Steinhardt School of Culture, Education and Human Development
New York University
627 Broadway, 8th Floor
New York, NY, 10003
tyler.watts@nyu.edu
212-998-5156

Abstract

We replicated and extended Shoda, Mischel, and Peake's (1990) famous "marshmallow" study, which showed strong bivariate correlations between a child's ability to delay gratification just before entering school and both adolescent achievement and socioemotional behaviors. Concentrating on children whose mothers had not completed college, we found that an additional minute waited at age 4 predicted a gain of approximately $1/10^{th}$ of a SD in age-15 achievement. But this bivariate correlation was only half the size of those reported in the original studies, and was reduced by two-thirds in the presence of controls for family background, early cognitive ability, and the home environment. Most of the variation in adolescent achievement came from being able to wait at least 20 seconds. Associations between delay time and age-15 measures of behavioral outcomes were much smaller and rarely statistically significant.

Keywords: gratification delay, marshmallow test, achievement, behavioral problems, longitudinal analysis, early childhood

In a series of studies based on children who attended a preschool on the Stanford University campus, Mischel, Shoda, and colleagues showed that under certain conditions, a child's success in delaying the gratification of eating marshmallows or a similar treat was related to later cognitive and social development, health, and even brain structure (Mischel et al., 2010; Shoda, Mischel, Peake, 1990; Tsukayama et al., 2009). Although only part of a larger research program investigating how children develop self-control, Mischel and Shoda's delay time/later outcome correlations and the preschooler videos accompanying them have become some of the most memorable findings from developmental research. Gratification delay is now viewed by many to be a fundamental "noncognitive" skill which, if developed early, can provide a lifetime of benefits (see Mischel et al., 2010 for review).

Since the publication of Mischel and Shoda's seminal studies (e.g., Mischel, Shoda, Peake, 1988; Mischel, Shoda, Rodriguez, 1989; Shoda et al., 1990), other researchers have examined the processes underlying the ability to delay gratification. Some have modified the Marshmallow Test to illuminate the factors that affect a child's ability to delay gratification (e.g., Imuta, Hayne & Scarf, 2014; Kidd, Palmeri, & Aslin, 2013; Michaelson & Munakata, 2016; Rodriguez, Mischel, & Shoda, 1989; Shimoni, Asbe, Eyal, & Berger, 2016); others have investigated the cognitive and socio-emotional correlates of gratification delay (e.g., Bembenutty & Karabenick, 2004; Duckworth, Tsukayama, & Kirby, 2013; Romer, Duckworth, Sznitman, & Park, 2010). These studies have added to a growing body of literature on self-control suggesting that gratification delay may constitute a critical early capacity. For example, Moffitt and Caspi demonstrated that self-control – typically understood to be an umbrella construct that includes gratification delay, but also impulsivity, conscientiousness, self-regulation and executive function – averaged across early and middle childhood predicted outcomes across a host of adult

domains (Moffitt et al., 2011). Duckworth and colleagues (2013) showed that the relation between early gratification delay and later outcomes was partially mediated by a composite measure of self-control, which has further fueled interventions designed to promote skills that fall under the "self-control" umbrella (e.g., Diamond & Lee, 2011). However, despite the proliferation of work on gratification delay, and the related construct of self-control, Mischel and Shoda's longitudinal studies still stand as the foundational examinations of the long-run correlates of the ability to delay gratification in early childhood.

Revisiting these studies reveals several limiting factors that warrant further investigation. First, Mischel and Shoda's reported longitudinal associations were based on very small and highly selective samples of children from the Stanford University community (n's= 35-89; Mischel et al., 1988; Mischel et al., 1989; Shoda et al., 1990). Although Mischel's original work included over 600 preschool-aged children (Shoda et al., 1990), follow-up investigations focused on much smaller samples (e.g., for their investigation of SAT and behavioral outcomes, Shoda and colleagues were able to contact only 185 of the original 653 children). Moreover, these children originally underwent variations of the gratification-delay assessment; Mischel experimented with trials in which the treat was obscured from a child's vision, and some of the children were supplied with coping strategies to help them delay longer. They found positive associations between gratification delay and later outcomes only for children participating in trials in which no strategy was coached and the treat was clearly visible – a circumstance they called the "diagnostic condition."

For the 35 to 48 children who were tested in the "diagnostic condition" and for whom adolescent follow-up data were available, Shoda and colleagues (1990) observed large correlations between delay time and SAT scores ($r(35) = .57$ for math; $r(35) = .42$ for verbal) and

between delay time and parent-reported behaviors (e.g., "[my child] is attentive and able to concentrate," r(48) = .39). These bivariate correlations were not adjusted for potential confounding factors that could affect both early delay ability and later outcomes. Because these findings have been cited as motivation both for interventions designed to boost gratification delay specifically (e.g., Kumst & Scarf, 2015; Murray, Theakston, & Wells, 2015; Rybanska et al., 2017) and for interventions seeking to promote self-control more generally (e.g., Diamond & Lee, 2011; Flook, Goldberg, Pinger, & Davidson, 2014; Rueda et al., 2012), it is important to consider possible confounding factors that might lead bivariate correlations to be a poor projection of likely intervention effects.

In the current study, we pursued a conceptual replication of Mischel and Shoda's original longitudinal work. Specifically, we examined associations between performance on a modified version of the Marshmallow Test and later outcomes in a larger and more diverse sample of children, and we employed empirical methods that adjusted for confounding factors inherent in Mischel and Shoda's bivariate correlations. Several considerations motivated our effort. First, replication is a staple of sound science (Campbell, 1986; Duncan, Engel, Claessens, & Dowsett, 2014). Second, Mischel and Shoda's highly selective sample of children limits the generalizability of their results. Finally, if researchers are to extend Mischel and Shoda's work to develop interventions, a more sophisticated examination of the long-run correlates of early gratification delay is needed. Interventions that successfully boost early delay ability might have no effect on later life outcomes if associations between gratification delay and later outcomes are driven by factors unlikely to be altered by child-focused programs (e.g., SES, home parenting environment).

Current Study

We used data from the NICHD Study of Early Child Care and Youth Development to explore associations between preschoolers' ability to delay gratification and age-15 academic and behavioral outcomes. We focused most of our analysis on a sample of children born to mothers who had not completed college for two reasons. First, it allowed us to investigate whether Mischel and Shoda's longitudinal findings extend to populations of greater interest to researchers and policymakers concerned with developing interventions (e.g., Mischel, 2014a). Second, empirical concerns over the extent of truncation in our key gratification delay measure in the college-educated sample limited our ability to assess reliably the correlation between gratification delay and later abilities. Because of these differences, we consider our study to be a conceptual, rather than traditional, replication of Mischel and Shoda's seminal work (Robbins, 1978).

Method

More complete information regarding study data and measures can be found in the online supplementary material. Here, we provide a brief overview of key study components.

Data

Data for the current study were drawn from the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (SECCYD), a widely used dataset in developmental psychology (NICHD Early Child Care Research Network, 2002). Participants were recruited at birth from ten U.S. sites across the country, providing a geographically diverse although not nationally representative, sample of children and mothers. Participants have been followed across childhood and adolescence, with the last full round of data collection occurring when children were 15 years old.

The current study relies on data collected when children were 54 months of age, and our

outcome variables were measured during the grade-1 and age-15 assessments. Our analysis sample was limited to children who had a valid measure of age-54-months delay of gratification, as well as non-missing achievement and behavioral data at age 15 (n=918). For conceptual and analytic reasons (detailed below), we then split our sample based on mother's education, and we focused much of our analyses on children whose mothers did not report having completed college when the child was one month old (n= 552 – a sample that is ten times larger than the sample size in the Shoda et al., (1990) study).

In Table 1, we present selected demographic characteristics for children included in our analytic sample, split by whether the child's mother received a bachelor's degree. For purposes of comparison, we also present the same set of characteristics for a nationally-representative sample of kindergarteners collected 2 to 3 years after our sample's 54-month wave of data collection (nationally representative data were drawn from the publically-available Early Childhood Longitudinal Survey- Kindergarten Cohort, 1998-1999; more information regarding this dataset can be found in the online supplementary information).

Perhaps not surprisingly, the children of college completing mothers were largely White (91%), with 55% of them reporting family income that was at least 4 times above the poverty line (i.e., "income to needs" ratio over 4.0), and none of them reporting income at or below the poverty line (i.e., "income to needs" ratio at or below 1.0). The subsample of children with mothers without a college degree were more comparable to the nationally representative sample. In both samples, about 16% of children were Black, mother's age at birth was approximately 27 years, 14% of mothers did not complete high school, and between 17% and 18% of families were living at or below the poverty line. However, Hispanic children were still underrepresented in this sample, underscoring the fact that although diverse, our data were not nationally

representative.

Table 1

Demographic Comparisons Between the Analytic Samples and a Nationally-Representative Sample of Kindergarten Children (ECLS-K: 1998)

	NICHD SECCYD		ECLSK: 1998
	Children of Non-Degreed Mothers	Children of Degreed Mothers	Nationally Representative Sample
Male	0.49	0.46	0.51
Black	0.16	0.02	0.16
Hispanic	0.07	0.03	0.19
White	0.73	0.91	0.57
Mother's Age at Child Birth (years)	26.84	31.67	27.28
	(5.61)	(4.01)	(6.61)
Mother's Education			
Did Not Complete High School	0.14	0	0.14
Graduated from High School	0.32	0	0.29
Some College	0.54	0	0.33
Bachelor's Degree or Higher	0.00	1.00	0.23
Income to Needs Ratio			
Income/ Needs < = 1	0.18	0	0.17
Income/ Needs > 1 & <= 2	0.27	0.05	0.26
Income/ Needs > 2 & <= 3	0.25	0.19	0.16
Income/ Needs > 3 & <= 4	0.15	0.21	0.16
Income/ Needs > 4	0.15	0.55	0.24
Mother Unemployed	0.29	0.23	0.32
Number of Children in Home	2.32	2.16	2.49
	(1.03)	(0.83)	(1.16)
Mother Married	0.67	0.93	0.70
Observations	552	366	21,242

Note. Mean values are presented with standard deviations in parentheses. The ECLS-K estimates were derived from data made publically available by NCES (see online supplementary information file and nces.gov/ecls/). All ECLS-K measures shown were collected during the fall of kindergarten (i.e., 1998), and SECCYD measures were collected during the 54-month interview (i.e., preschool; 1995-1996), except for "mother's education" and "mother's age at child's birth," which were both collected at the 1-month interview. The ECLS-K variables were weighted using the C1CW0 weight to generate nationally representative estimates.

Measures

Delay of gratification. A variant of Mischel's (1974) self-imposed waiting task (i.e., the "Marshmallow Test") was administered to children when they were 54 months of age. An interviewer would present children with an appealing edible treat based on the child's own stated

preferences (e.g., marshmallows, M&Ms, animal crackers, etc.). Children were then told that they would engage in a game in which the interviewer would leave the child alone in a room with the treat. If the child waited for 7 minutes, the interviewer would return and the child could eat the treat and receive an additional portion as a reward for waiting. Children who chose not to wait could ring a bell to signal the experimenter to return early, and they would then only receive the amount of candy originally presented. The measure of delay of gratification is then recorded as the number of seconds the child waited, with 7 minutes being the ceiling.

The measure of gratification delay used here differed from the one employed by Mischel (1974) in several noteworthy ways. First, the 7-minute cap was much shorter than Mischel's maximum assessment length; the children in Mischel's sample were asked to wait between 15 and 20 minutes, depending on the study, before the assessment ended. In our sample, approximately 55% of children hit the 7-minute ceiling on the measure, presenting a potential analytic challenge to our models. However, we found that the ceiling was much more problematic for higher- than lower-SES children. Children whose mothers obtained college degrees hit the ceiling at a rate of 68%, compared with 45% for children whose mothers did not complete college ($p < .001$; see Table 2).

We adopt several approaches to dealing with this truncation problem, principally exploring possible non-linearities in the "time waited"/outcome associations by dividing the distribution of waiting times into discrete intervals. We also focused much of our analyses on the children of mothers who did not complete college, as far fewer of the children in this sample hit the ceiling on the minutes waited measure and, as explained above, this group of children complements the sample of children included in the Mischel and Shoda studies. But because the subsample of children with college-educated mothers allows for a more direct replication of

Mischel and Shoda's famous work (e.g., Shoda et al., 1990), we also present results for them, bearing in mind the limitations imposed by the substantial delay truncation.

Finally, it should also be noted that children in the NICHD study were given only the version of the task that Shoda and colleagues (1990) called the "diagnostic condition" (i.e., the children were not offered strategies and were able to see the treat as they waited).

Academic achievement. Academic achievement was measured using the Woodcock-Johnson Psycho-Educational Battery Revised (WJ-R) test (Woodcock, McGrew, & Mather, 2001), a commonly used measure of cognitive ability and achievement (e.g., Watts, Duncan, Davis-Kean, & Siegler, 2014). For math achievement at grade 1 and age 15, we used the *Applied Problems* subtest, which measured children's mathematical problem solving. At grade 1, reading achievement was measured using the *Letter Word Identification* task, a measure of word recognition and vocabulary, and at age 15, reading ability was measured using the *Passage Comprehension* test. The *Passage Comprehension* test asked students to read various pieces of text silently and then answer questions about their content.

For all the WJ-R tests, we used the standard scores, which were normed to have a mean of 100 and SD of 15 in each respective wave. We took the average of the grade-1 math and reading measures and the age-15 math and reading measures, respectively, to create composite measures of academic achievement.

Behavioral problems. Following Shoda et al. (1990), we relied primarily on mothers' reports of child behavior. Mother-reported internalizing and externalizing behavioral problems were assessed using the Child Behavioral Checklist (CBCL; Achenbach, 1991) at age 54 months, first grade, and age 15. The CBCL is a widely used measure of behavioral problems, and it included approximately 100 items rated on 3-point scales that captured aspects of internalizing

(i.e., depressive) and externalizing (i.e., anti-social) behavior. As with academic achievement, at grade 1 and age 15 we averaged together the externalizing and internalizing measures to create a behavioral composite score that, before standardization, ranged from 32 to 83, with higher scores indicating higher levels of behavioral problems. We also tested models that used a host of alternative behavioral measures taken from youth reports and direct assessments at age 15; these measures and models are described in the online supplementary material.

Additional covariates. All covariates included in our models are listed in Table 3, and we grouped the covariates into two distinct sets of control variables: "Child Demographic and Home Controls" and "Concurrent 54-Month Controls."

Child demographic and home controls. Child demographic characteristics (i.e., gender and race), birth weight, mother's age at the child's birth, and mother's level of education were collected at the one month interview via interview with study mothers. Family income was collected from study mothers at the one-, six-, 15-, 24-, 36- and 54-month interviews. We took the average of all non-missing income data over this span, and then log-transformed average family income to restrict the influence of outliers. Mother's Peabody Picture Vocabulary Test (PPVT) score was assessed in a lab visit when the focal child was 36 months old. The PPVT is a commonly used measure of intelligence.

We also included early indicators of child cognitive functioning, as measured at age 24 months by the Bayley Mental Development Index (MDI; Bayley, 1991) and at age 36 months by the Bracken Basic Concept Scale (BBCS; Bracken, 1984). The MDI measured children's sensory-perceptual abilities, as well as their memory, problem solving, and verbal communication skills. The BBCS was an early measure of school readiness skills, and it required students to identify basic letters and numbers.

Child temperament was measured at age 6 months using the Early Infant Temperament Questionnaire (Medoff-Cooper, Carey, & McDevitt, 1993), a 38-item survey to which mothers responded. This questionnaire asked mothers to rate their child on a six point Likert-scale with items focused on the child's mood, adaptability, and intensity. We took the average score across these items as our measurement of temperament, with higher scores indicating more agreeable dispositions.

Finally, the set of controls measured prior to age 54-months also included indicators of the quality of the home environment, as measured by an observational assessment called the Home Observation for Measurement of the Environment (HOME) inventory (Caldwell & Bradley 1984). The HOME was assessed when the focal child was approximately 36 months old, and it was designed to capture aspects of the home environment known to support positive cognitive, emotional, and behavioral functioning. We used 9 subscales of the HOME in our models: the first eight subscales are commonly used with the HOME measure (Learning Materials, Language Stimulation, Physical Environment, Responsivity, Academic Stimulation, Modeling, Variety, and Acceptance), and the 9th subscale, called "Responsivity- Empirical Scale," was derived by NICHD SECCYD study from factor analyses of the HOME items. This final scale was distinct from the traditional "Responsivity" scale, as it included items from the "Language Stimulation" scale that also measured mother responsivity and sensitivity to the child.

Concurrent 54-month controls. For models that included controls for concurrent cognitive and behavioral skills, we also included subscales taken from the age 54-month *WJ-R* test. As our measure of early reading, we included the *Letter-Word Identification* task, which tested children's ability to sound out simple words, and the *Applied Problems* test at age 54-months was our measure of early math skills. For preschool children, the *Applied Problems* test

requires children to count and solve simple addition problems. We also used the *Memory for Sentences* and *Incomplete Words* subtests as measures of cognitive ability. The *Incomplete Words* test measured auditory closure and processing, and children listened to an audio recording where words missing a phenome were listed off. They were then asked to name the complete word. Finally, the *Picture Vocabulary* test was a measure of verbal comprehension and crystallized intelligence. In this task, children were asked to name pictured objects. All of these tasks have been widely used as measures of children's early cognitive skills and their measurement properties have been widely reported (e.g., Watts et al., 2014).

Finally, we also included the mother's report of children's externalizing and internalizing problems from the Child Behavioral Checklist at age 54 months. Much like the measure used for age-15 behavioral problems, the 54-month survey included a battery of items designed to assess children's anti-social and disruptive behavior (i.e., externalizing) and depressive symptoms (i.e., internalizing).

Analysis

Our primary goal was to estimate the association between early gratification delay and long-run measures of academic achievement and behavioral functioning. Like the work of Shoda and colleagues (1990), our data did not include a measure of gratification delay in which between-child differences were generated from some exogenous intervention, so we do not claim that the associations we estimate reflect causal impacts. Instead, our goal was to assess how much bias might be contained in longitudinal bivariate correlations between gratification delay and later outcomes as a result of failure to control for characteristics of children and their environments. Regression-adjusted correlations should provide better guidance regarding whether interventions boosting gratification delay might also improve later achievement and

behavior.

To accomplish our analytic goals, we modeled later academic achievement and behavior (measured at both grade 1 and age 15) as a function of an age-54-months measure of gratification delay. We then tested models that added controls for background characteristics and measures of the home environment (see Panel 1 of Table 3) before moving to models that also included measures of cognitive and behavioral skills assessed at age 54 months (see Panel 2 of Table 3).

These two approaches reflect different assumptions regarding how variation in gratification delay ability might arise. Models with controls measured between birth and age 36 months still allow for variation in age-54-months gratification delay caused by the differential development of general cognitive or behavioral skills (e.g., executive function, self-control, etc.) between 36 and 54 months. Put another way, these models contain controls only for factors that even ambitious preschool child-focused interventions are unlikely to alter (e.g., birthweight, temperament at 6 months of age, early home environment).

In contrast, the models with "concurrent 54 months" covariates control for variation in a range of cognitive capacities and behavioral problems developed by age 54 months. They help to isolate the possible effects of an intervention that targets only the narrow set of skills involved with gratification delay (e.g., a program that merely provided children with strategies to help them delay longer; see Mischel, 2014b, p.40), but not concurrent general cognitive ability or socioemotional behaviors.

Although it is impossible to know exactly how individual differences in gratification delay emerge (e.g., changes in parenting, development of cognitive skills), by controlling for factors unlikely to be altered by interventions (e.g., ethnicity, parental background), we can purge our estimates of bias due to observable characteristics that are correlated with gratification

delay and later outcomes. If remaining unobserved factors also contribute to gratification delay and later outcomes (e.g., changes in parenting), and if these unobserved factors are unlikely to be altered by a particular intervention, then bias in our estimates may still remain. Yet, our estimates should serve as an improvement over the unadjusted correlations reported previously (e.g., Shoda et al., 1990).

In all models shown, continuous variables were standardized so that coefficients can be read as effect sizes, and all models with control variables included a set of dummy variables for each site to adjust for any between-site differences. In order to account for missing data on control variables, we used SEM with Full Information Maximum Likelihood in Stata 15.0 to estimate all analytic models. Finally, we report all estimated p-values to the thousandth decimal place (with p-valued below .001 displayed as "< .001"), and we describe any estimate corresponding to a p-value less than 0.05 as "statistically significant." Though we recognize the arbitrariness of focusing only on results with a p-value less than .05, we selected this alpha level because it was the minimum threshold for statistical significance used in the studies we attempted to replicate and extend (i.e., Mischel et al., 1988; Mischel et al., 1989; Shoda et al., 1990). Consequently, any differences in conclusions reached between our studies and the previous literature should be attributed to design and sample differences rather than alpha level choices.

Results

Descriptive Findings

Table 2 provides descriptive results for key analysis variables, including the 54-months delay of gratification measure, split by mother's education level. In the sample of children with non-degreed mothers, children waited an average of 3.99 minutes (SD= 3.08) before ending the

task. We also present the proportion of children falling within certain ranges on the measure, with the "7 minute" category representing children who successfully completed the trial. In the lower-SES sample, 45% of children waited the maximum of 7 minutes, and 23% waited less than 20 seconds (i.e., 0.33 minutes). In the higher-SES sample, only 10% of children waited less than 20 seconds, and the average time waited was 5.38 minutes [statistically significantly longer than the lower-SES group ($p < .001$)].

Table 2
Descriptive Characteristics of Key Analysis Variables

	Children of Non-Degreed Mothers	Children of Degreed Mothers		
	M (SD)	M (SD)	β	P-Value of Difference
Delay of Gratification (minutes waited)	3.99 (3.08)	5.38 (2.62)	0.45	.001*
Delay of Gratification (categories)				
7 minutes	0.45	0.68	0.21	.001*
2 to 7 minutes	0.16	0.12	-0.02	.324
0.333 to 2 minutes	0.16	0.10	-0.06	.012*
< 0.333 minutes	0.23	0.10	-0.13	.001*
Outcome Measures- Grade 1				
Achievement Composite	108.42 (13.71)	117.29 (13.47)	0.63	.001*
Behavior Composite	49.15 (8.43)	47.40 (7.87)	-0.18	.008*
Outcome Measures- Age 15				
Achievement Composite	101.23 (11.63)	112.72 (13.19)	0.82	.001*
Behavior Composite	47.12 (9.37)	44.50 (8.66)	-0.27	.001*
Observations	552	366		

Note. Mean values are presented in each cell, and standard deviations are in parentheses. For the delay of gratification categories (e.g., "< 0.333 minutes") the proportion of students falling within each category is presented. The sample is split based on mother's education, and p-values were derived from a series of regressions in which each characteristic was regressed on a dummy for "whether mother graduated from college" and a series of site fixed effects. Values shown in the "β" column represent effect sizes measuring the standardized differences between the two groups.
* p< .05

Because the 7-minute ceiling presented a potential analytic challenge for both samples, we estimated models that substituted the four dummy categories shown in Table 2 for the continuous "minutes waited" variable as a way to assess nonlinearities in the relationship between delay time and academic and socioemotional outcomes. Importantly, these models also provide information on how much our analysis might be compromised by the seven-minute truncation.

Table 3 presents descriptive information for the various control measures used in the analysis, and means are presented separately for children that did and did not complete the delay task. In both the higher- and lower-SES samples, performance on the gratification delay task was highly correlated with differences on most observable characteristics considered. For example, for children from non-degreed mothers, those who completed the gratification delay task were from higher income families ($p < .001$) than non-completers, had mothers with higher PPVT scores ($p < .001$), and had higher scores on dimensions of the H.O.M.E. observational assessment (p ranged from .04 to < .001). Null or smaller differences were generally observed for the children of degreed mothers, perhaps owing to the lack of heterogeneity in this subsample.

Regression Results

Results for children of non-degreed mothers. Table 4 presents coefficients and standard errors from models that estimate the association between 54-months gratification delay and our first-grade and age-15 achievement and behavioral composites for the sample of children from non-degreed mothers. Panel 1 of Table 4 displays results for a standardized continuous measure of gratification delay (i.e., the number of minutes waited during the Marshmallow Test). As Column 1 reflects, the bivariate association between minutes waited and academic

Table 3
Control Variable Descriptive Characteristics

	Children of Non-Degreed Mothers				Children of Degreed Mothers			
	Waited 7 Minutes		β	P-Value of Difference	Waited 7 Minutes		β	P-Value of Difference
	Yes	No			Yes	No		
Panel 1: Child Demographic and Home Controls								
Child Background								
Male	0.47	0.51	-0.04	.338	0.45	0.50	-0.05	.409
White	0.82	0.64	0.18	.001*	0.94	0.85	0.10	.007*
Black	0.07	0.24	-0.15	.001*	0.00	0.05	-0.05	.024*
Hispanic	0.06	0.07	-0.01	.545	0.03	0.03	-0.00	.962
Other	0.04	0.05	-0.01	.530	0.03	0.07	-0.05	.058
Child's Age at Delay Measure (mos.)	56.11	56.01	0.13	.105	55.99	55.99	0.07	.519
	(1.11)	(1.14)			(1.13)	(1.15)		
Birth Weight (g)	3490.23	3449.02	0.09	.320	3516.63	3572.53	-0.13	.268
	(478.56)	(540.26)			(520.52)	(527.17)		
Bracken Standard Score (36 mos.)	9.06	7.67	0.47	.001*	10.67	10.14	0.19	.043*
	(2.56)	(2.86)			(2.20)	(2.35)		
Bayley (24 mos.)	93.89	85.91	0.53	.001*	100.88	95.21	0.41	.001*
	(12.40)	(14.40)			(11.78)	(14.10)		
Child Temperament (6 mos.)	3.18	3.25	-0.17	.053	3.13	3.09	0.07	.531
	(0.42)	(0.38)			(0.37)	(0.43)		
Log of Family Income (1 mo - 54 mos.)	0.89	0.57	0.38	.001*	1.54	1.42	0.14	.057
	(0.61)	(0.73)			(0.51)	(0.56)		
Mother's Age at Birth (years)	27.75	26.07	0.29	.001*	31.58	31.87	-0.06	.438
	(5.66)	(5.46)			(4.05)	(3.91)		
Mother's Education (years)	13.00	12.68	0.12	.017*	17.02	16.82	0.07	.234
	(1.41)	(1.50)			(1.31)	(1.26)		
Mother's PPVT	96.43	90.47	0.30	.001*	114.10	105.63	0.44	.001*
	(13.38)	(17.03)			(15.62)	(16.51)		

H.O.M.E. Score (36 mos.)								
Learning Materials	7.20	5.86	0.53	.001*	8.64	8.41	0.12	.168
	(2.36)	(2.51)			(1.59)	(2.20)		
Language Stimulation	6.13	5.67	0.46	.001*	6.38	6.17	0.21	.046*
	(1.04)	(1.24)			(0.84)	(1.13)		
Physical Environment	6.16	5.64	0.40	.001*	6.35	6.33	0.07	.372
	(1.04)	(1.54)			(0.83)	(0.91)		
Responsivity	5.67	5.17	0.31	.001*	6.09	5.81	0.21	.033*
	(1.28)	(1.52)			(0.99)	(1.30)		
Academic Stimulation	3.43	2.97	0.38	.001*	3.74	3.57	0.17	.112
	(1.21)	(1.29)			(0.97)	(1.29)		
Modeling	3.13	2.82	0.29	.001*	3.64	3.51	0.11	.285
	(1.10)	(1.14)			(0.93)	(1.04)		
Variety	6.80	6.14	0.45	.001*	7.54	7.29	0.17	.088
	(1.34)	(1.50)			(1.17)	(1.36)		
Acceptance	3.39	3.22	0.18	.038*	3.70	3.57	0.13	.162
	(0.85)	(1.04)			(0.59)	(0.82)		
Responsivity- Empirical Scale	5.54	5.14	0.37	.001*	5.77	5.55	0.21	.026*
	(0.91)	(1.29)			(0.52)	(0.91)		
Panel 2: Concurrent 54-Month Controls								
54 mos. WJ-R Scores								
Letter-Word Id.	99.03	93.22	0.42	.001*	105.93	102.31	0.26	.011*
	(11.98)	(12.63)			(12.19)	(11.94)		
Applied Problems	104.80	95.67	0.57	.001*	112.36	106.06	0.40	.001*
	(12.88)	(15.72)			(12.13)	(12.31)		
Picture Vocabulary	100.54	93.74	0.43	.001*	109.11	103.47	0.36	.001*
	(13.07)	(13.80)			(13.45)	(13.58)		
Memory for Sentences	93.21	85.43	0.43	.001*	100.99	92.34	0.49	.001*
	(15.59)	(17.67)			(18.73)	(17.45)		
Incomplete Words	98.08	92.72	0.41	.001*	102.18	98.05	0.35	.001*
	(12.91)	(13.52)			(11.69)	(11.98)		

54 mos. Child Behavioral Checklist								
Internalizing	47.36	47.94	-0.06	.477	46.55	46.81	-0.01	.988
	(9.11)	(8.51)			(8.84)	(8.17)		
Externalizing	51.14	53.09	-0.21	.020*	50.44	50.99	-0.06	.604
	(9.34)	(9.84)			(9.11)	(8.53)		
Observations	251	301			250	116		

Note. Mean values are presented in each cell, and standard deviations are in parentheses. The p-value column compares children who successfully completed the task and waited 7 minutes to students who did not, and the "β" column presents effect sizes measuring the standardized differences between the two groups. P-values were generated from a series of bivariate regressions in which each variable was regressed on a dummy indicating whether the child completed the marshmallow test, and series of site dummy variables was also included to adjust for site differences. P-values below .001 have been rounded to .001.
* p< .05

achievement was 0.28 (*SE* = 0.04, $p < .001$), considerably less than the .57 correlation Shoda and colleagues found for SAT math scores and the .42 correlation they found for verbal scores. These linear results suggest that children's grade-1 achievement would improve by approximately 1/10th of a SD for every additional minute waited at age 4. When the controls measured prior to age 54 months (i.e., second column of Table 3) were added to the model, the standardized association fell to 0.10 (*SE* = 0.03, $p = .002$), and when concurrent 54-months controls were added (i.e., third column of Table 1), the association fell to a statistically non-significant 0.05 (*SE* = 0.03, $p = .114$).

Columns 4 through 6 show analogous models for the age-15 measure of achievement. The magnitudes of the age-15 correlations were remarkably similar to the first-grade correlations. The age-15 achievement correlation in the absence of other controls was of moderate size and statistically significant ($\beta = 0.24$, *SE* = 0.04, $p < .001$), but fell substantially when controls for earlier characteristics were added ($\beta = 0.08$, *SE* = 0.03, $p = .016$) and became non-significant when 54-months controls were added ($\beta = 0.05$, *SE* = 0.03, $p = .140$). Given that Shoda and colleagues found almost as strong correlations with later behavior as with later achievement, we were surprised to find virtually no relationship – even in the absence of controls – between gratification delay and the composite score of mother-reported internalizing and externalizing at either grade 1 or age 15 (right half of Table 4).

Children who waited less than 20 seconds (i.e., the lowest category) served as the comparison group for our models that represented delay times in a set of dummy variables (see Table 2 for the proportion of students in each category). As shown in Panel 2 of Table 4, models of outcomes at both grade 1 and age 15 that lack control variables show a strong gradient between gratification delay and later achievement. Relative to children who waited less than 20

Table 4
Associations Between Age 54-month Gratification Delay and Later Measures of Academic Achievement and Behavior for Children of Mothers Without College Degrees

	Achievement Composite						Behavior Composite					
	First Grade			Age 15			First Grade			Age 15		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL 1												
Delay minutes (continuous)	0.279*	0.102*	0.047	0.236*	0.081*	0.050	-0.060	-0.015	0.023	-0.062	-0.026	0.003
	(0.038)	(0.033)	(0.030)	(0.037)	(0.034)	(0.032)	(0.043)	(0.044)	(0.044)	(0.046)	(0.047)	(0.042)
PANEL 2												
Delay minutes (categorical)												
<0.333 minutes	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
0.333- 2 minutes	0.298*	0.189	0.127	0.353*	0.230*	0.178	0.055	0.090	0.079	-0.140	-0.071	-0.106
	(0.126)	(0.105)	(0.093)	(0.122)	(0.103)	(0.098)	(0.144)	(0.138)	(0.105)	(0.152)	(0.148)	(0.132)
2 to 7 minutes	0.424*	0.206	0.041	0.457*	0.300*	0.235*	-0.088	-0.020	0.039	-0.182	-0.109	-0.053
	(0.126)	(0.104)	(0.093)	(0.123)	(0.103)	(0.099)	(0.144)	(0.137)	(0.106)	(0.151)	(0.145)	(0.131)
7 minutes	0.720*	0.284*	0.141	0.646*	0.234*	0.150	-0.121	-0.007	0.072	-0.193	-0.095	-0.048
	(0.098)	(0.086)	(0.078)	(0.098)	(0.088)	(0.084)	(0.112)	(0.114)	(0.087)	(0.120)	(0.123)	(0.111)
p-value of test of equality of all categories	.001*	.012*	.247	.001*	.015*	.093	.477	.866	.837	.428	.861	.885
p-value of test of equality of 2^{nd}, 3^{rd} and 4^{th} categories	.001*	.563	.475	.015*	.752	.630	.382	.700	.923	.927	.969	.882
Child demographic and H.O.M.E. controls	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.
Concurrent 54-month controls	-	-	Inc.	-	-	Inc.	-	-	Inc.	-	-	Inc.

Note. n= 552. Standard errors are in parentheses. Continuous variables were standardized, so coefficients can be interpreted as effect sizes. Estimates shown in the first column of each set (i.e., columns 1, 4, 7, and 10) only contained the measure of delay of gratification and a given outcome measure. Estimates shown in the second column of each set (i.e., columns 2, 5, 8, and 11) added child demographic characteristics, H.O.M.E. scores, and site dummy variables. Estimates shown in the third column of each set (i.e., columns 3, 6, 9, and 12) added other behavioral and cognitive measures also measured at age 54 months. P-values were generated from post-hoc Chi-square tests in order to assess whether respective sets of variables were different from one another. P-values below 0.001 have been rounded to 0.001.

* p< .05

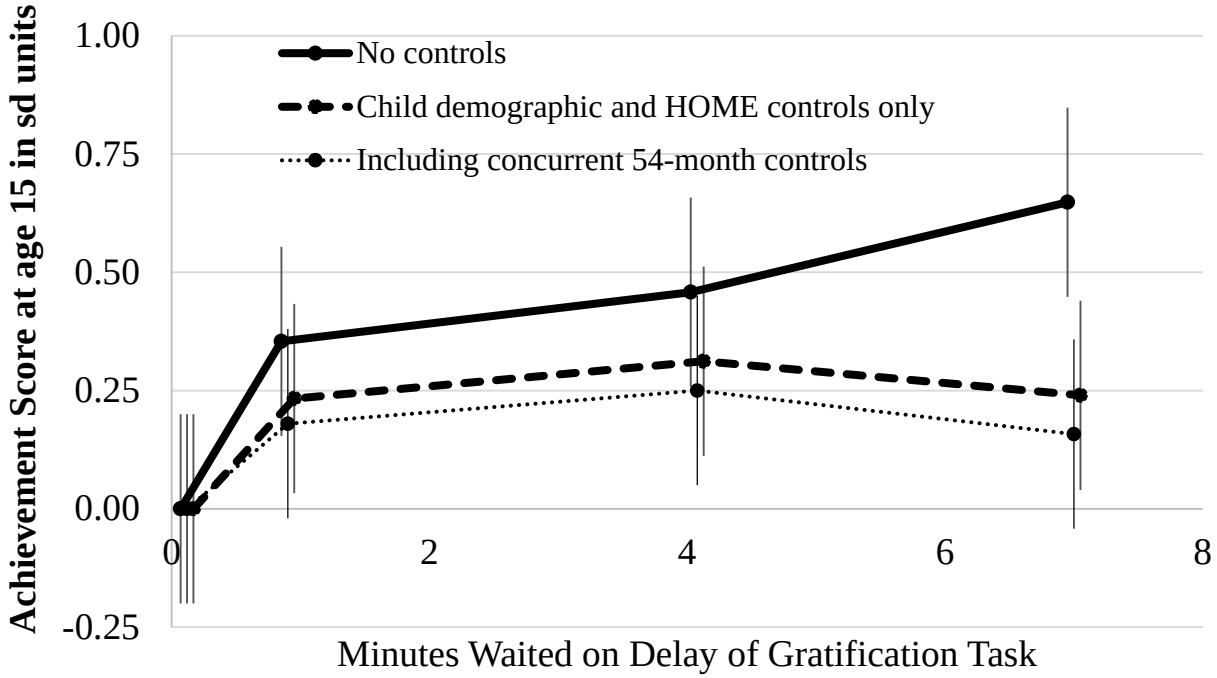
seconds, children who waited between 20 seconds and 2 minutes scored about 1/3 of a SD higher on the achievement measure at grade 1 and age 15, and this difference grew to nearly 3/4th of a SD for the group that waited the entire 7 minutes. The "p-value of test of equality of 2nd, 3rd, and 4th categories" entry in the first column shows that the coefficients produced by the three groups of children who waited longer than 20 seconds differed significantly from one another ($p < .001$), as did coefficient differences across all four categorical variables (the p-value for which is shown in the "p-value of test of equality of all categories" row).

At both grade 1 and age 15, when controls for early child and family characteristics were added to the model (Column 2 for grade 1; Column 5 for age 15), the coefficients estimated for all three delay-time groups fell by roughly 50%. Surprisingly, the addition of the background controls also flattened out the gradient of the prediction across the gratification delay distribution. Relative to the <20 second reference group, achievement differences for children who waited more than 20 seconds, but not the full 7 minutes, were strikingly similar to the difference for children who waited the full 7 minutes. At age 15, the threshold nature of the relationship was most apparent; the coefficients produced by the three groups that waited longer than 20 seconds all fell between 0.23 and 0.30, and were not close to being statistically significantly different from one another ($p = .752$).

When concurrent 54-months controls were added, coefficients fell even further. At age 15, only the coefficient produced by the group describing children who waited 2 to 7 minutes retained statistical significance ($\beta = 0.24$, $SE = 0.10$, $p = .018$), though once again the set of coefficients on the included categories of delay time did not differ from one another ($p = .630$). As with the models shown in the right half of Panel 1, we found no statistically significant relationships between gratification delay and the first-grade and age-15 behavioral composites.

In our focal case of age-15 achievement, the return for delaying gratification appeared to be driven by differences between children who managed to wait at least 20 seconds and those who did not. Figure 1 illustrates this threshold effect with three lines showing the coefficients produced by our gratification-delay categories in the age-15 achievement models. The solid line shows coefficients drawn from the no-control model (i.e., Column 4 of Panel 2), the dashed line shows coefficients from the model with early controls (i.e., Column 5 of Panel 2), and the dotted line shows coefficients produced by models with the 54-months controls (i.e., Column 6 of Panel 2).

Figure 1: Predicted Achievement Scores by Minutes of Delay for Children of Mothers With No College Degree



Note. Error bars represent 95% confidence intervals. For each of the 4 discrete groups described in the note of Table 2, we graphed the deviation in achievement composite scores from the reference group (delay <20 seconds) against the within-group average amount of time waited on the x-axis. The average wait times for the "No Controls" and "Child Demographic and HOME Controls Only" groups are displaced by ±.025 to distinguish the sets of error bars. The "High Delay" group's coefficients are plotted at 7 minutes, although the 7-minute truncation prevents us from

knowing what the mean value of minutes waited would have been for this group in the absence of the 7-minute limit.

The uncontrolled line has a steep initial jump, followed by a more gradual increase for wait times longer than 20 seconds. Both "with controls" lines decrease after four minutes. Using 7 minutes to anchor the "more than 7 minutes" group is probably an underestimate, but it is clear from the downward trajectory that no assumptions about the distribution of wait times above 7 minutes would produce a strong positive slope for the last segment of the line. Thus, in the case of children with mothers who lack college degrees, the truncation of delay time at seven minutes does not affect the conclusion that children with the highest delay times show similar achievement levels at age 15 as other children who are able to delay for at least 20 seconds.

Results for children from mothers with college degrees. In Table 5, we present key results for children of mothers with college degrees. As with the results in Table 4, we again present results for the continuous measure of gratification delay (Panel 1) and the categorical measures split along parts of the gratification delay distribution (Panel 2). For the continuous measure, we again found evidence of positive unadjusted associations between gratification delay and later achievement at both first grade (β = 0.18, *SE* = 0.06, p = .001) and age 15 (β = 0.17, *SE* = 0.06, p = .007), and the categorical results suggested that much of this association was somewhat linear through the distribution. For the age 15 models, these relations became statistically indistinguishable from 0 once controls were added, and the point estimate for the "> 7 minute" was surprisingly small and negative (β = -0.04, *SE* = 0.15, p = .816). As with the models shown in Table 4, we again found no evidence of associations between gratification delay and the behavioral measures at first grade or age 15 in the high-SES sample.

Despite statistically non-significant results, point estimates were sometimes positive and substantial (e.g., the "2 to 7 minutes" group coefficient shown in Column 1; β = 0.40, *SE* = 0.21,

Table 5
Associations Between Age 54 Gratification Delay and Later Measures of Academic Achievement and Behavior for Children of Mothers with College Degrees

	Achievement Composite						Behavior Composite					
	First Grade			Age 15			First Grade			Age 15		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL 1												
Delay minutes (continuous)	0.178*	0.120*	0.048	0.167*	0.062	0.007	-0.049	-0.059	-0.050	0.031	0.038	0.043
	(0.056)	(0.053)	(0.045)	(0.062)	(0.059)	(0.054)	(0.057)	(0.061)	(0.046)	(0.059)	(0.063)	(0.055)
PANEL 2												
Delay minutes (categorical)												
<0.333 minutes	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
0.333- 2 minutes	0.327	0.039	0.148	0.079	-0.131	-0.085	-0.069	-0.088	-0.184	-0.065	0.027	-0.083
	(0.220)	(0.198)	(0.168)	(0.245)	(0.216)	(0.197)	(0.227)	(0.228)	(0.173)	(0.231)	(0.232)	(0.200)
2 to 7 minutes	0.397	0.147	0.134	0.216	0.028	-0.032	-0.277	-0.240	-0.265	-0.318	-0.217	-0.227
	(0.206)	(0.184)	(0.155)	(0.227)	(0.199)	(0.182)	(0.210)	(0.209)	(0.157)	(0.218)	(0.216)	(0.185)
7 minutes	0.562*	0.301	0.193	0.404*	0.077	-0.036	-0.194	-0.208	-0.214	-0.007	0.068	0.052
	(0.166)	(0.154)	(0.131)	(0.183)	(0.166)	(0.152)	(0.168)	(0.174)	(0.131)	(0.174)	(0.180)	(0.155)
p-value of test of equality of all categories	.005*	.100	.521	.059	.674	.979	.515	.584	.350	.267	.367	.227
p-value of test of equality of 2^{nd}, 3^{rd} and 4^{th} categories	.238	.153	.843	.149	.477	.948	.629	.753	.867	.147	.206	.115
Child demographic and H.O.M.E. controls	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.
Concurrent 54-month controls	-	-	Inc.	-	-	Inc.	-	-	Inc.	-	-	Inc.

Note. n= 366. Standard errors are in parentheses. See Table 4 note. Estimates in this table can be directly compared with estimates from Table 4. The sample was limited to children whose mothers had at least 16 years of completed education (i.e., completed college).
* p< .05

p = .054) but the standard errors were nearly double those estimated for children of non-degreed mothers (Table 4). This is due in part to the somewhat smaller sample size for the high-SES sample but also to the lack of variation in the gratification delay measure for this sample. Thus, although we found even less evidence of associations between gratification delay and measures of later achievement when considering only the children of mothers with college degrees, it is difficult to draw strong conclusions from these models given the imprecise nature of their coefficient estimates.

Additional results and sensitivity checks

Heterogeneity. Because we found little evidence supporting associations between early delay ability and later outcomes for the higher-SES sample, we next tested whether the different pattern of results observed between the higher- and lower-SES samples constituted a statistically significant difference. In Table 6, we present models that included interaction terms between the various measures of gratification delay (i.e., the continuous and categorical measures) and the indicator for whether the subject's mother completed college. None of the interactions tested were statistically significant, and our series of "joint F-tests" indicated that the set of interactions for the categorical measures of gratification delay did not statistically significantly contribute to any of the models (p-values ranged from .342 to .968). However, as with the models that were run solely on the sample of children with college educated mothers, standard errors were quite large for the interaction terms, indicating a substantial level of statistical imprecision. Unfortunately, the wide confidence intervals on many of the interaction terms render it impossible to provide a definitive answer to whether the relation between early delay ability and later achievement differs by SES.

Table 6

Associations Between Age 54 month Gratification Delay and Later Measures of Academic Achievement with Interactions Between Gratification Delay and SES

	Achievement Composite						Behavior Composite					
	First Grade			Age 15			First Grade			Age 15		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL 1												
Delay minutes (continuous)	0.279*	0.115*	0.050	0.236*	0.083*	0.040	-0.059	-0.019	0.012	-0.062	-0.023	0.009
	(0.038)	(0.035)	(0.030)	(0.040)	(0.037)	(0.034)	(0.042)	(0.043)	(0.033)	(0.044)	(0.046)	(0.040)
High-SES Indicator	0.509*	0.050	0.032	0.747*	0.270*	0.266*	-0.187*	0.026	0.031	-0.286*	-0.119	-0.127
	(0.064)	(0.068)	(0.059)	(0.067)	(0.071)	(0.066)	(0.070)	(0.084)	(0.064)	(0.074)	(0.088)	(0.077)
Interaction	-0.101	-0.043	-0.035	-0.069	-0.007	-0.018	0.010	-0.038	-0.058	0.094	0.040	0.017
	(0.067)	(0.058)	(0.050)	(0.069)	(0.061)	(0.057)	(0.073)	(0.071)	(0.054)	(0.076)	(0.075)	(0.066)
PANEL 2												
Delay minutes (categorical)												
<0.333 minutes	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
0.333- 2 minutes	0.298*	0.182	0.109	0.353*	0.202	0.151	0.055	0.060	0.050	-0.140	-0.082	-0.097
	(0.127)	(0.110)	(0.096)	(0.131)	(0.115)	(0.107)	(0.140)	(0.137)	(0.104)	(0.148)	(0.145)	(0.127)
2 to 7 minutes	0.424*	0.215	0.053	0.457*	0.288*	0.199	-0.088	-0.046	0.006	-0.182	-0.103	-0.024
	(0.127)	(0.110)	(0.097)	(0.132)	(0.115)	(0.108)	(0.140)	(0.137)	(0.105)	(0.146)	(0.143)	(0.126)
7 minutes	0.721*	0.308*	0.147	0.646*	0.222*	0.121	-0.121	-0.025	0.034	-0.193	-0.087	-0.028
	(0.099)	(0.090)	(0.079)	(0.105)	(0.097)	(0.091)	(0.109)	(0.112)	(0.086)	(0.116)	(0.120)	(0.106)
High-SES Indicator	0.585*	0.154	0.041	0.951*	0.428*	0.417*	-0.097	0.163	0.191	-0.375	-0.185	-0.138
	(0.174)	(0.156)	(0.136)	(0.178)	(0.163)	(0.151)	(0.187)	(0.190)	(0.144)	(0.195)	(0.199)	(0.174)
Interactions												
High SES * < 0.333 min.	0.029	-0.164	0.032	-0.274	-0.337	-0.266	-0.124	-0.127	-0.160	0.075	0.119	0.035
	(0.252)	(0.218)	(0.190)	(0.259)	(0.226)	(0.210)	(0.275)	(0.269)	(0.205)	(0.284)	(0.276)	(0.243)
High SES * 2 to 7 min	-0.027	-0.138	0.010	-0.241	-0.293	-0.258	-0.188	-0.185	-0.199	-0.136	-0.090	-0.156
	(0.240)	(0.206)	(0.179)	(0.246)	(0.213)	(0.198)	(0.260)	(0.252)	(0.192)	(0.272)	(0.261)	(0.229)
High SES * 7 min	-0.159	-0.119	-0.033	-0.242	-0.119	-0.134	-0.073	-0.167	-0.203	0.186	0.115	0.049
	(0.192)	(0.165)	(0.144)	(0.197)	(0.173)	(0.161)	(0.207)	(0.201)	(0.153)	(0.217)	(0.212)	(0.186)

P-value from interaction term joint F-test	.668	.870	.968	.640	.342	.507	.899	.859	.610	.450	.753	.720
Child demographic and H.O.M.E. controls		Inc.	Inc.		Inc.	Inc.		Inc.	Inc.		Inc.	Inc.
Concurrent 54-month controls			Inc.			Inc.			Inc.			Inc.

Note. *n*= 918. Standard errors are in parentheses. The "joint F-test" evaluated whether the set of interaction terms jointly contribute to the model. In other words, it tested whether the set of interactions were statistically significantly different from "0."

* p< .05

Measurement considerations. In Table 7, we present correlations between the Marshmallow Test and all analysis variables for the full sample of children considered in our analyses (n = 918; see the supplementary file for correlation matrices for both the low-SES and high-SES samples, respectively). In Table 7, we also included the 54-month measure of the Continuous Performance Task (CPT), which is a commonly used indicator of attention and impulsivity, and we included the Duckworth et al. (2013) parent- and teacher-report index of 54-month self-control (see the supplementary file for measurement details). We included these additional measures to further investigate how the Marshmallow Test might relate to theoretically relevant constructs (see Diamond & Lee, 2011). Surprisingly, the Marshmallow Test had the strongest correlation with the *Applied Problems* subtest of the *WJ-R* (r(916) = 0.37, $p < .001$), and correlations with measures of attention, impulsivity and self-control were lower in magnitude (ranging from 0.22 to 0.30, $p < .001$). Although these correlational results were far from conclusive, they suggest that the Marshmallow Test should not be thought of as a mere behavioral proxy for self-control, as the measure clearly relates strongly to basic measures of cognitive capacity.

Table 7

Correlations Between All Analysis Variables

PANEL 1		1	2	3	4	5	6	7	8	9	10
Gratification Delay (54)											
1	Continuous	1.00									
2	<0.333 min.	-0.69	1.00								
3	0.333- 2 min.	-0.47	-0.18	1.00							
4	2 to 7 min.	-0.07	-0.19	-0.16	1.00						
5	7 min.	0.90	-0.51	-0.43	-0.45	1.00					
Related Measures											
6	Self-control (54)	0.24	-0.15	-0.15	-0.03	0.24	1.00				
7	Attention (54)	0.22	-0.18	-0.07	-0.08	0.24	0.15	1.00			
8	Impulsivity (54)	-0.30	0.26	0.06	0.05	-0.28	-0.28	-0.26	1.00		
Outcome Measures											
9	Achievement (G1)	0.31	-0.26	-0.08	-0.03	0.28	0.33	0.30	-0.27	1.00	
10	Achievement (15)	0.30	-0.25	-0.09	-0.02	0.27	0.32	0.20	-0.23	0.64	1.00
11	Behavior (G1)	-0.08	0.06	0.05	-0.02	-0.07	-0.30	-0.08	0.05	-0.09	-0.11

12	Behavior (15)	-0.06	0.08	0.01	-0.04	-0.04	-0.23	-0.06	0.06	-0.11	-0.13
Demographic Controls											
13	Male	-0.05	0.06	-0.02	0.02	-0.05	-0.20	-0.01	0.23	-0.01	0.05
14	Black	-0.25	0.21	0.07	0.05	-0.24	-0.16	-0.12	0.20	-0.29	-0.33
15	Hispanic	-0.03	-0.00	0.06	-0.02	-0.02	-0.04	-0.02	0.03	-0.05	-0.03
16	Other	-0.04	0.00	0.03	0.04	-0.05	-0.00	0.02	-0.02	0.02	0.01
17	Age	0.03	-0.04	0.03	-0.02	0.02	0.03	0.06	-0.02	0.04	-0.05
18	Log of Income	0.30	-0.26	-0.08	-0.03	0.27	0.26	0.19	-0.19	0.37	0.40
19	Mother's Age	0.20	-0.18	-0.05	-0.00	0.18	0.18	0.12	-0.14	0.22	0.32
20	Mother's Ed (yrs)	0.25	-0.19	-0.09	-0.04	0.24	0.27	0.16	-0.20	0.35	0.42
21	Mother PPVT	0.28	-0.22	-0.09	-0.08	0.28	0.29	0.12	-0.18	0.35	0.48
22	Site 1	-0.04	0.02	0.00	0.06	-0.06	-0.06	0.06	-0.02	0.03	-0.14
23	Site 2	0.00	-0.06	0.05	0.01	0.00	0.04	0.03	-0.03	0.06	0.10
24	Site 3	0.07	-0.05	-0.03	-0.02	0.07	-0.04	0.02	-0.09	-0.04	-0.08
25	Site 4	-0.00	0.02	-0.01	-0.01	-0.00	0.02	0.04	0.09	-0.02	0.05
26	Site 5	-0.06	0.02	0.06	-0.00	-0.06	0.02	0.03	0.01	-0.02	-0.05
27	Site 6	0.03	-0.01	-0.04	-0.01	0.04	0.06	0.04	-0.03	0.04	0.09
28	Site 7	-0.05	0.04	0.00	0.01	-0.04	-0.02	-0.10	0.12	-0.05	0.02
29	Site 8	0.06	0.00	-0.05	-0.09	0.09	0.10	-0.00	-0.08	-0.01	0.05
30	Site 9	-0.04	-0.00	0.04	0.04	-0.06	-0.07	-0.01	0.00	0.05	0.02
31	Birthweight (g's)	-0.01	0.02	0.01	-0.06	0.02	-0.02	0.05	-0.01	0.11	0.10
32	Bracken	0.28	-0.22	-0.10	-0.04	0.26	0.32	0.26	-0.29	0.54	0.50
33	Bayley	0.34	-0.27	-0.08	-0.06	0.31	0.29	0.24	-0.24	0.42	0.39
34	Temperament	-0.08	0.11	0.00	-0.02	-0.06	-0.14	-0.04	0.08	-0.11	-0.12
H.O.M.E. Controls											
35	Learn. Mater.	0.29	-0.23	-0.11	-0.02	0.27	0.31	0.15	-0.23	0.38	0.40
36	Lang. Stim.	0.21	-0.18	-0.05	-0.04	0.20	0.17	0.08	-0.14	0.25	0.21
37	Phys. Env.	0.20	-0.13	-0.13	0.02	0.17	0.15	0.13	-0.12	0.23	0.21
38	Responsivity	0.19	-0.13	-0.08	-0.05	0.20	0.18	0.14	-0.12	0.19	0.17
39	Academ. Stim.	0.21	-0.17	-0.06	-0.01	0.18	0.15	0.05	-0.15	0.23	0.20
40	Modeling	0.17	-0.11	-0.06	-0.05	0.16	0.17	0.10	-0.07	0.23	0.25
41	Variety	0.25	-0.15	-0.14	-0.04	0.24	0.22	0.12	-0.21	0.28	0.29
42	Acceptance	0.12	-0.07	-0.07	-0.04	0.13	0.21	0.13	-0.17	0.16	0.19
43	Respons. Emp.	0.20	-0.14	-0.08	-0.05	0.20	0.16	0.12	-0.10	0.20	0.16
54-month Controls											
44	Letter Word (54)	0.28	-0.22	-0.09	-0.03	0.25	0.29	0.25	-0.24	0.60	0.49
45	App. Prob. (54)	0.37	-0.28	-0.16	-0.01	0.33	0.35	0.32	-0.32	0.62	0.56
46	Pic. Vocab. (54)	0.28	-0.21	-0.08	-0.09	0.28	0.25	0.22	-0.18	0.42	0.50
47	Mem. Sent. (54)	0.29	-0.25	-0.09	-0.02	0.26	0.28	0.22	-0.21	0.42	0.43
48	Inc. Words (54)	0.23	-0.17	-0.08	-0.06	0.22	0.15	0.19	-0.17	0.39	0.34
49	Internalizing (54)	-0.04	0.04	0.02	-0.00	-0.04	-0.17	-0.05	0.08	-0.07	-0.08
50	Externalizing (54)	-0.10	0.07	0.07	-0.02	-0.09	-0.39	-0.07	0.09	-0.10	-0.12

PANEL 2		11	12	13	14	15	16	17	18	19	20
11	Behavior (G1)	1.00									
12	Behavior (15)	0.55	1.00								
Demographic Controls											
13	Male	-0.00	-0.04	1.00							
14	Black	0.06	0.00	-0.00	1.00						
15	Hispanic	0.01	0.04	0.03	-0.08	1.00					

16	Other	-0.01	-0.01	-0.03	-0.07	-0.05	1.00				
17	Age	0.03	0.04	-0.00	0.03	0.01	-0.04	1.00			
18	Log of Income	-0.16	-0.17	-0.02	-0.36	-0.08	-0.01	-0.01	1.00		
19	Mother's Age	-0.18	-0.21	-0.04	-0.28	-0.10	-0.04	-0.04	0.54	1.00	
20	Mother's Ed (yrs)	-0.13	-0.17	-0.04	-0.22	-0.11	-0.03	-0.00	0.61	0.52	1.00
21	Mother PPVT	-0.10	-0.07	-0.01	-0.37	-0.11	-0.09	-0.03	0.49	0.46	0.57
22	Site 1	0.09	0.07	-0.00	0.11	-0.07	-0.07	-0.03	-0.09	-0.09	-0.06
23	Site 2	-0.06	-0.07	0.00	-0.11	0.23	-0.01	-0.18	0.16	0.06	0.02
24	Site 3	0.04	0.02	-0.02	-0.02	0.08	-0.03	0.10	-0.07	-0.05	0.02
25	Site 4	-0.07	-0.04	0.02	-0.05	-0.04	0.04	0.00	0.05	0.07	-0.04
26	Site 5	-0.02	-0.02	0.02	0.13	-0.06	-0.02	0.22	-0.05	0.03	0.02
27	Site 6	-0.07	-0.08	-0.02	0.11	-0.06	-0.01	-0.10	0.12	0.09	0.13
28	Site 7	-0.02	-0.03	0.02	0.01	0.00	0.02	0.14	-0.09	-0.07	-0.08
29	Site 8	-0.02	0.05	-0.00	-0.05	0.00	0.11	-0.19	0.08	0.10	0.10
30	Site 9	0.03	0.03	0.01	-0.05	-0.06	-0.05	0.04	-0.06	-0.08	-0.12
31	Birthweight (g's)	0.02	0.09	0.12	-0.14	0.04	-0.07	0.01	0.04	0.05	0.07
32	Bracken	-0.09	-0.10	-0.15	-0.32	-0.07	-0.02	0.01	0.45	0.32	0.42
33	Bayley	-0.08	-0.13	-0.17	-0.32	-0.08	-0.01	-0.02	0.40	0.23	0.36
34	Temperament	0.12	0.15	-0.04	0.17	-0.01	0.05	-0.01	-0.19	-0.19	-0.13
H.O.M.E. Controls											
35	Learn. Mater.	-0.10	-0.11	-0.05	-0.39	-0.12	-0.08	0.03	0.49	0.35	0.48
36	Lang. Stim.	-0.01	-0.06	-0.03	-0.11	-0.12	-0.11	0.01	0.27	0.12	0.24
37	Phys. Env.	-0.09	-0.08	0.01	-0.24	0.00	0.01	-0.03	0.28	0.18	0.23
38	Responsivity	-0.09	-0.07	-0.02	-0.22	-0.06	-0.07	-0.09	0.32	0.26	0.28
39	Academ. Stim.	0.00	-0.01	-0.03	-0.17	-0.09	-0.04	0.01	0.24	0.12	0.25
40	Modeling	-0.07	-0.04	-0.05	-0.15	-0.06	-0.06	-0.05	0.31	0.23	0.33
41	Variety	-0.10	-0.07	-0.03	-0.27	-0.09	-0.05	0.01	0.41	0.27	0.39
42	Acceptance	-0.16	-0.14	-0.05	-0.10	-0.01	0.00	-0.04	0.23	0.21	0.24
43	Respons. Emp.	-0.06	-0.05	-0.02	-0.18	-0.06	-0.04	-0.06	0.31	0.20	0.26
54-month Controls											
44	Letter Word (54)	-0.07	-0.07	-0.10	-0.20	-0.08	0.05	-0.01	0.38	0.19	0.38
45	App. Prob. (54)	-0.04	-0.12	-0.10	-0.32	-0.07	0.01	-0.02	0.42	0.28	0.40
46	Pic. Vocab. (54)	-0.09	-0.04	0.10	-0.33	-0.10	-0.01	0.02	0.42	0.32	0.40
47	Mem. Sent. (54)	-0.11	-0.09	-0.04	-0.18	-0.11	0.04	0.02	0.29	0.21	0.28
48	Inc. Words (54)	-0.05	-0.12	-0.00	-0.18	-0.10	0.01	0.03	0.24	0.18	0.24
49	Internalizing (54)	0.53	0.38	0.03	0.04	-0.01	0.03	0.09	-0.06	-0.09	-0.10
50	Externalizing (54)	0.63	0.47	-0.08	0.05	0.01	-0.00	0.01	-0.12	-0.13	-0.13

PANEL 3		21	22	23	24	25	26	27	28	29	30
21	Mother PPVT	1.00									
22	Site 1	-0.10	1.00								
23	Site 2	0.04	-0.10	1.00							
24	Site 3	-0.02	-0.10	-0.11	1.00						
25	Site 4	-0.02	-0.10	-0.11	-0.11	1.00					
26	Site 5	-0.01	-0.11	-0.11	-0.11	-0.11	1.00				
27	Site 6	0.10	-0.10	-0.10	-0.11	-0.10	-0.11	1.00			
28	Site 7	-0.01	-0.11	-0.11	-0.11	-0.11	-0.12	-0.11	1.00		
29	Site 8	0.14	-0.11	-0.11	-0.11	-0.11	-0.12	-0.11	-0.12	1.00	
30	Site 9	-0.11	-0.11	-0.11	-0.11	-0.11	-0.12	-0.11	-0.12	-0.12	1.00
31	Birthweight (g's)	0.13	-0.02	0.03	-0.06	0.04	-0.02	-0.03	-0.01	0.03	0.02

32	Bracken	0.40	-0.05	0.09	0.00	0.04	0.00	0.05	-0.14	0.05	-0.03
33	Bayley	0.34	-0.05	0.10	-0.02	0.11	0.02	0.02	-0.18	-0.06	-0.01
34	Temperament	-0.19	0.06	-0.08	-0.01	-0.01	-0.04	0.02	0.02	0.02	0.03
H.O.M.E. Controls											
35	Learn. Mater.	0.47	-0.06	-0.09	-0.01	-0.04	-0.01	0.08	-0.04	0.00	0.05
36	Lang. Stim.	0.24	0.06	-0.16	-0.15	-0.20	0.02	0.18	0.01	0.05	0.09
37	Phys. Env.	0.19	-0.07	-0.06	-0.05	0.03	0.09	0.02	-0.25	-0.05	0.19
38	Responsivity	0.27	-0.11	-0.01	-0.06	-0.02	-0.11	0.30	-0.30	0.12	0.06
39	Academ. Stim.	0.24	-0.01	-0.18	-0.06	-0.07	0.04	0.12	-0.11	0.05	0.09
40	Modeling	0.29	-0.00	0.01	-0.09	-0.08	-0.00	0.15	-0.08	0.03	-0.06
41	Variety	0.37	0.02	-0.14	-0.06	-0.04	-0.03	0.16	-0.09	0.07	0.08
42	Acceptance	0.20	-0.10	-0.00	-0.06	0.01	-0.04	0.14	0.04	0.04	-0.14
43	Respons. Emp.	0.25	0.04	-0.03	-0.07	-0.03	-0.15	0.17	-0.12	0.08	0.02
54-month Controls											
44	Letter Word (54)	0.34	-0.02	0.01	-0.01	0.03	-0.01	0.10	-0.08	0.03	-0.05
45	App. Prob. (54)	0.43	-0.08	0.03	0.01	0.07	-0.02	0.09	-0.08	0.04	-0.03
46	Pic. Vocab. (54)	0.48	-0.12	0.01	0.01	0.07	-0.01	0.05	-0.04	0.10	-0.03
47	Mem. Sent. (54)	0.30	-0.08	-0.06	-0.06	0.07	0.08	0.06	-0.05	0.04	0.03
48	Inc. Words (54)	0.27	-0.07	-0.06	-0.04	0.02	0.05	0.04	0.00	-0.05	0.13
49	Internalizing (54)	-0.10	0.01	-0.04	-0.02	0.03	0.05	-0.05	0.01	-0.02	-0.03
50	Externalizing (54)	-0.11	0.04	0.01	0.03	-0.03	0.01	-0.03	-0.01	-0.06	0.01

PANEL 4		31	32	33	34	35	36	37	38	39	40
31	Birthweight (g's)	1.00									
32	Bracken	0.08	1.00								
33	Bayley	0.06	0.52	1.00							
34	Temperament	-0.04	-0.15	-0.12	1.00						
H.O.M.E. Controls											
35	Learn. Mater.	0.07	0.47	0.43	-0.12	1.00					
36	Lang. Stim.	0.10	0.28	0.23	-0.04	0.51	1.00				
37	Phys. Env.	0.01	0.25	0.24	-0.08	0.41	0.28	1.00			
38	Responsivity	0.08	0.31	0.25	-0.11	0.38	0.38	0.26	1.00		
39	Academ. Stim.	0.08	0.33	0.26	-0.02	0.55	0.55	0.31	0.33	1.00	
40	Modeling	0.07	0.24	0.24	-0.10	0.37	0.31	0.26	0.28	0.28	1.00
41	Variety	0.04	0.36	0.37	-0.09	0.56	0.41	0.33	0.33	0.43	0.35
42	Acceptance	0.05	0.22	0.19	-0.05	0.28	0.20	0.17	0.19	0.14	0.32
43	Respons. Emp.	0.04	0.24	0.19	-0.12	0.35	0.48	0.26	0.77	0.29	0.27
54-month Controls											
44	Letter Word (54)	0.07	0.61	0.40	-0.08	0.40	0.29	0.23	0.26	0.34	0.23
45	App. Prob. (54)	0.09	0.57	0.56	-0.13	0.43	0.24	0.27	0.25	0.25	0.18
46	Pic. Vocab. (54)	0.12	0.46	0.44	-0.12	0.43	0.25	0.23	0.27	0.28	0.21
47	Mem. Sent. (54)	0.08	0.39	0.43	-0.08	0.31	0.20	0.19	0.17	0.22	0.14
48	Inc. Words (54)	0.09	0.30	0.36	-0.10	0.30	0.24	0.23	0.16	0.22	0.14
49	Internalizing (54)	0.04	-0.03	-0.03	0.14	-0.07	-0.04	-0.02	-0.04	0.02	-0.04
50	Externalizing (54)	0.04	-0.11	-0.05	0.12	-0.12	-0.06	-0.09	-0.10	-0.07	-0.11

PANEL 5		41	42	43	44	45	46	47	48	49	50
41	Variety	1.00									
42	Acceptance	0.23	1.00								
43	Respons. Emp.	0.28	0.23	1.00							

54-month Controls											
44	Letter Word (54)	0.32	0.19	0.21	1.00						
45	App. Prob. (54)	0.31	0.22	0.21	0.58	1.00					
46	Pic. Vocab. (54)	0.38	0.16	0.22	0.46	0.52	1.00				
47	Mem. Sent. (54)	0.30	0.15	0.13	0.42	0.47	0.46	1.00			
48	Inc. Words (54)	0.27	0.11	0.16	0.36	0.45	0.37	0.49	1.00		
49	Internalizing (54)	-0.07	-0.07	-0.06	-0.01	-0.04	-0.08	-0.06	-0.04	1.00	
50	Externalizing (54)	-0.11	-0.14	-0.11	-0.04	-0.05	-0.09	-0.10	-0.05	0.58	1.00

Note. n=918. All non-missing cases for each pairwise correlation were included. Tables 2 and 3 include the full variable names and labels. The supplementary material presents correlations for all variables shown separately by mother's education. The "G1" abbreviation stands for "grade 1," "15" stands for "age 15," and "54" stands for 54 months.

In the supplementary file, we further assessed to what extent self-control and attention could account for the associations between gratification delay and later achievement. In Table S3, we included the 54-months measures of attention and impulse control taken from the CPT in the Table 4 models and found that inclusion of the CPT measures accounted for only 21-27% of the effect for the "> 7 minute group". In Table S4, we ran a parallel analysis using the Duckworth et al. (2013) index of self-control, and again we found that coefficients were hardly reduced when the self-control index was included. The small change in the coefficient for the gratification delay measure between models that did and did not include indicators of attention, impulsivity and self-control raises further questions regarding what constructs are measured by the Marshmallow Test.

Alternative outcome measures. Returning to our focal sample of children with mothers who had not completed college, we were surprised to see the lack of significant associations between our gratification-delay measure and the behavioral measures at first grade and age 15. We also tested models that used alternative indicators of behavior assessed at age 15, including measures of risky behavior from youth self-reports and assessments of impulse control. Surprisingly, we still found virtually no associations between gratification delay and behavior across any of these alternative measures (Tables S5 through S7 in the supplementary material). Furthermore, because we relied on aggregated measures of achievement and behavior, we also

tested separate models for math, reading, externalizing behaviors, and internalizing behaviors (Table S8). Results indicated that the achievement associations were similar for both the math and reading measures, and we still found no statistically significant effects on either measure of problem behaviors.

Discussion

We attempted to extend the famous findings of Mischel and Shoda (Mischel et al., 1988; Mischel et al., 1989; Shoda et al., 1990) by examining associations between early gratification delay and adolescent outcomes in a more diverse sample of children and with more sophisticated statistical models. As with the earlier studies, we found statistically significant, although smaller, bivariate associations between early delay ability and later achievement. But we also found that these associations were highly sensitive to the inclusion of controls. Moreover, we failed to find even bivariate associations between age-54-months delay and a host of age-15 behavioral outcomes, which was remarkable given the stability in self-control measures found in other studies (e.g., Moffitt et al., 2011).

It surprised us that for the children of non-degreed mothers, most of the achievement boost for early delay ability was gained by waiting a mere 20 seconds. Shoda et al. (1990) argued that the relationship between gratification delay and academic achievement might be driven by the ability to generate useful metacognitive strategies that will influence self-regulation throughout one's life. Such strategies are unlikely to have played much of a role in a child's ability to wait for only 20 seconds. Instead, our findings suggest that impulse control may be a key mechanism, although post-hoc inclusion of an explicit measure of impulse control explained some but certainly not most of the gratification delay effect.

These results provide further questions regarding what the Marshmallow Test might

measure, and how it relates to the umbrella construct of self-control. We observed that gratification delay was strongly correlated with concurrent measures of cognitive ability, and controlling for a composite measure of self-control explained only about 25% of our reported effects on achievement. These results suggest that the Marshmallow Test may capture something rather distinct from self-control. Indeed, Duckworth and colleagues (2013) also investigated the relations between gratification delay, self-control and intelligence using the data employed here, and found that both self-control and intelligence mediated the relation between early delay ability and later outcomes. Our results further suggest that simply viewing gratification delay as a component of self-control may oversimplify how gratification delay operates in young children.

When considering how our results might inform intervention development, recall that models with controls for concurrent measures of cognitive skills and behavior reduced the association between gratification delay and age-15 achievement to nearly 0. This implies that an intervention that altered a child's ability to delay, but failed to change more general cognitive and behavioral capacities, would likely have limited effects on later outcomes. If intervention developers hope to generate program impacts that replicate the long-term Marshmallow Test findings, targeting the broader cognitive and behavioral abilities related to gratification delay might prove more fruitful.

Indeed, Mischel and Shoda's original results (Shoda et al., 1990) supported similar conclusions. Recall that they reported long-run correlations between gratification delay and later outcomes only for children who were *not* provided with strategies for delaying longer. That the prediction was strong only in trials that relied on natural variation in children's ability to delay suggests that unobserved factors underlying children's delay ability may have driven the long-

run correlations. Our results support this interpretation.

Our study is not without weaknesses. The 7-minute ceiling was limiting, although our non-linear models indicated that it was unlikely to affect conclusions drawn for the lower-SES sample. For the higher-SES sample, the 7-minute ceiling prevented a direct replication of Mischel and Shoda's original work (e.g., Shoda et al., 1990), as a substantial majority of higher-SES children hit the ceiling. The lack of precision in our higher-SES results was unfortunate, though it should be noted that point estimates in fully-controlled models were often very small. At the very least, these results further suggest that bivariate associations between gratification delay and later outcomes probably contain substantial bias, even for more privileged children.

It should also be noted that variation in our age-54-months gratification-delay measure was not exogenous, so our models could not truly capture the effects that would be produced by exogenously spurred gains in early gratification-delay ability. However, our models included an extensive set of control variables that go well beyond the bivariate specifications employed in previous studies (e.g., Shoda et al., 1990). Finally, data not drawn to be nationally representative provide a shaky foundation for generalization.

In sum, our findings suggest that although early gratification delay did indeed correlate with later achievement for children whose mothers had not completed college, the magnitude of this association was highly sensitive to the inclusion of control variables and did not appear to be linear across the gratification-delay distribution. Future work on gratification delay should continue to examine the processes captured by the Marshmallow Test, and whether early gratification-delay interventions would be worthwhile investments for promoting children's long-run success.

Author Contributions

T.W. Watts and G.J. Duncan developed the study concept and design, and authored the manuscript. T.W. Watts and H. Quan performed the data analysis. All authors approved the final manuscript.

Acknowledgments

We are grateful to the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under award number P01-HD065704 and to Ana Auger, Drew Bailey, Daniel Belsky, Jay Belsky**,** Clancy Blair, Peg Burchinal, Angela Duckworth, Dorothy Duncan, Jade Jenkins, Terrie Moffitt, Cybele Raver, and Deborah Vandell for helpful comments on previous drafts.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 Profile.* Burlington, VT: Department of Psychiatry, University of Vermont.

Bayley, N. (1993). *Bayley scales of infant development* (2nd ed.). New York: Psychological Corporation.

Bracken, B. A. (1984). *Bracken basic concept scale.* Chicago: Psychological Corporation.

Bembenutty, H., & Karabenick, S. A. (2004). Inherent association between academic delay of gratification, future time perspective, and self-regulated learning. *Educational Psychology Review*, *16*(1), 35-57.

Caldwell, B. M., & Bradley, R. H. (1984). *Home Observation for Measurement of the Environment.* Little Rock, AR: University of Arkansas at Little Rock.

Campbell, D. (1986). Science's social system of validity-enhancing collective belief change and the problems of the social sciences. In D. Fiske & R. Shweder (Eds.), *Metatheory in social science* (pp. 108–135). Chicago, IL: University of Chicago Press.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959-964.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087.

Duckworth, A. L., Tsukayama, E., & Kirby, T. A. (2013). Is it really self-control? Examining the predictive power of the delay of gratification task. *Personality and Social Psychology Bulletin*, *39*(7), 843-855.

Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2014). Replication and robustness in developmental research. *Developmental Psychology*, *50*(11), 2417.

Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Shepard, S. A., … & Thompson, M. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development*, *75*(1), 25-46.

Flook, L., Goldberg, S. B., Pinger, L., & Davidson, R. J. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Developmental Psychology*, *51*(1), 44.

Imuta, K., Hayne, H., & Scarf, D. (2014). I want it all and I want it now: delay of gratification in preschool children. *Developmental Psychobiology*, *56*(7), 1541-1552.

Kumst, S., & Scarf, D. (2015). Your wish is my command! The influence of symbolic modelling on preschool children's delay of gratification. *Peerj*, *3*, e774–e774.

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109-114.

Medoff-Cooper, B., Carey, W.B., & McDevitt, S.C. (1993). The early infancy temperament questionnaire. *Journal of Developmental & Behavioral Pediatrics, 14*, 230-235.

Michaelson, L. E., & Munakata, Y. (2016). Trust matters: Seeing how an adult treats another person influences preschoolers' willingness to delay gratification. *Developmental Science*, *19*(6), 1011-1019.

Mischel, W. (1974). Processes in delay of gratification. *Advances in experimental social psychology*, *7*, 249-292.

Mischel, W. (2014a). Marshmallows and public policy. In *The Marshmallow Test* (pp. 233-253). New York, NY: Little, Brown and Company.

Mischel, W. (2014b). How they do it. In *The Marshmallow Test* (pp. 28-42). New York, NY:

Little, Brown and Company.

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., ... & Shoda, Y. (2010). 'Willpower'over the life span: decomposing self-regulation. *Social Cognitive and Affective Neuroscience*, *6*(2), 252-256.

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, *54*(4), 687.

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

Murray, J., Theakston, A., & Wells, A. (2016). Can the attention training technique turn one marshmallow into two? Improving children's ability to delay gratification. *Behaviour Research and Therapy*, *77*, 34-39.

Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., Houts, R., Poulton, R., Roberts, B.W., Ross, S. and Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693-2698.

NICHD Early Child Care Research Network. (2002). Early childcare and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal, 39*(1), 133-164.

Robins, L. N. (1978). Sturdy childhood predictors of adult antisocial behaviour: Replications from longitudinal studies. *Psychological Medicine, 8*, 611–622.

Rodriguez, M. L., Mischel, W., & Shoda, Y. (1989). Cognitive person variables in the delay of gratification of older children at risk. *Journal of Personality and Social*

Psychology, *57*(2), 358.

Romer, D., Duckworth, A. L., Sznitman, S., & Park, S. (2010). Can adolescents learn self-control? Delay of gratification in the development of control over risk taking. *Prevention Science*, *11*(3), 319-330.

Rybanska, V., McKay, R., Jong, J., & Whitehouse, H. (2017). Rituals Improve Children's Ability to Delay Gratification. *Child Development*. Advanced online publication. doi: 10.1111/cdev.12762

Shimoni, E., Asbe, M., Eyal, T., & Berger, A. (2016). Too proud to regulate: The differential effect of pride versus joy on children's ability to delay gratification. *Journal of Experimental Child Psychology*, *141*, 275-282.

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

Tsukayama, E., Toomey, S. L., Faith, M. S., & Duckworth, A. L. (2010). Self-control as a protective factor against overweight status in the transition from childhood to adolescence. *Archives of Pediatrics & Adolescent Medicine*, *164*(7), 631-635.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*(7), 352-360.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.

Online supplementary materials for

"Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Gratification Delay and Later Outcomes"

Method

Data

Data for the current study were drawn from the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (SECCYD). Participants were recruited from ten U.S. sites in both urban and rural settings. Mothers who had recently given birth were recruited from nearby hospitals, and mothers were excluded if they could not speak English or if they planned to move within the next 18 months. The first wave of major child-level data collection occurred when the focus child was 1 month of age, at which point 1,364 children remained in the study (roughly 50% of the originally recruited sample). Most of the attrition between initial recruitment and the 1-month interview was concentrated among low-SES mothers and children (see Duncan & Gibson, 2000; NICHD Early Child Care Research Network, 2002). Data collection continued at various times throughout children's lives, with the last full round of measurement occurring when children were 15 years old. The current study relies primarily on data collected when children were 54 month old, in grade 1, and 15 years old. However, for models with covariates, we also rely on data collected between the 1 month child interview and the age 36 month interview.

ECLS-K. For purposes of comparison, we also show demographic characteristics for children and families recruited for participation in the Early Childhood Longitudinal Study-Kindergarten Cohort of 1998-1999 (ECLS-K; Tourangeau, Nord, Le, Sorongon, & Najarian, 2009). The data were collected by the National Center for Educational Statistics (NCES), and the sample was designed to be nationally representative of children in kindergarten during the fall of 1998. This publically-available dataset has been broadly used to study child development and education, and information regarding data collection procedures and study measures has been widely reported.

In Table 1 of the main text, we present information taken from study families during the fall 1998 parent survey. Mothers reported their age at the time of the survey, and we subtracted the focus child's age from the reported mother's age to calculate "Mother's Age at Child Birth." NCES presented family annual income as a categorical variable, with respondents falling into income ranges (e.g., $20,000-$25,000). We then gave each participant the middle income value for each category, and used the number of children and adults in the home to calculate the "income to needs ratio" following the guidelines given by the Census Bureau (https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html). Finally, for all descriptive statistics presented, we weighted estimates using the C1C1W0 weight to recover nationally representative estimates.

Measures

Academic achievement. The Woodcock Johnson- Revised (WJ-R; Woodcock, McGrew & Mather, 2001) is a commonly used assessment of cognitive and academic ability, and it contains a number of subtests each designed to measure mathematics and verbal achievement, as well as more general cognitive abilities such as executive function. Each subtest was age-normed and administered by a trained examiner in a one-on-one interview with the child. We focus on WJ-R subtests designed to measure mathematics and reading achievement, which were administered at grade 1 and age 15 years. In some models, we also use WJ-R subtests given at age 54-months as control variables. For all WJ-R subtests, we used the WJ-R standard scores, which were standardized to have a mean of 100 and standard deviation of 15 at each age.

Mathematics achievement. At age 54-months, grade 1, and age 15-years, mathematics

achievement was measured using the WJ-R *Applied Problems* subtest, a commonly used measure of math ability (e.g., Siegler et al., 2012; Watts et al., 2014) that tested mathematical reasoning and problem solving. The examiner administering the test established a basal and ceiling for each child, and items were ordered hierarchically. At age 54 months, the *Applied Problems* items focused on counting and simple arithmetic; by age 15, items focused on algebraic concepts and fractions.

Reading achievement. At age 54 months and grade 1, reading achievement was measured using the *Letter-Word Identification* task, which tested children's ability to sound out simple words. Like the *Applied Problems* test, items on the *Letter-Word Identification* test were also ordered hierarchically, and the examiner established a basal and a ceiling for each child. Early items on the test asked children to match a pictographic representation of a word with an actual picture of the object, and the remaining items asked subjects to read aloud isolated letters and words. At age 15, reading achievement was measured using the *Passage Comprehension* subtest. Early items on this test involved reading a phrase then identifying a picture that depicted the phrase. Later items asked children to read a passage and fill in a missing key word.

Additional 54-month cognitive measures. In models with 54-month covariates, we also used the *Memory for Sentences* and *Incomplete Words* subtests as measures of cognitive ability. The *Incomplete Words* test measured auditory closure and processing, and children listened to an audio recording where words missing a phenome were listed off. They were then asked to name the complete word. Finally, the *Picture Vocabulary* test was a measure of verbal comprehension and crystallized intelligence. In this task, children were asked to name pictured objects.

Additional problem behaviors. In addition to using the mother-reported Child Behavior Checklist, we investigated the relationship between delay of gratification and other measures of behavioral functioning assessed at age 15. These measures included the Stoplight Task (Gardner & Steinberg, 2005), a measure of risk-taking in which children played a computer game that asked them to control a car attempting to reach a destination in a limited amount of time. While driving the car, children encountered stoplights that forced them to face the choice of either slowing down and losing time or running the light and possibly crashing the vehicle. From this task, we used two measures of risk taking: 1) the amount of time (in milliseconds) between the appearance of a yellow light and the application of the brake; 2) number of brake applications during the entire task. The stoplight task has been used in other developmental studies as a measure of adolescent risk taking (see Chein et al., 2011; Kim-Spoon et al., 2016)

As an alternative measure of internalizing and externalizing, we included scales taken from the age-15 Youth Self Report (YSR). The YSR was adapted from the CBCL, and it included 119 items that allowed youth to assess their own behavior. We also included a self-reported measure of impulse control, taken from the Winberger Adjustment Inventory (Weinberger & Schwarts, 1990), in which youth responded to 8 items designed to help them assess their own ability to control counterproductive impulses. Finally, we used a youth-reported measure of risky behavior called the Risky Behavior Questionnaire, which was adapted from several different measures used in large studies of child development (Conger & Elder, 1994; Halpern-Felsher, 2005). With this measure, youth responded to 61 items asking them how many times in the past year they had engaged in 55 different risk behaviors (e.g., alcohol use and sexual risk taking).

Continuous Performance Task. In supplemental models described below, we controlled for a 54-month measure of attention and impulsivity called the Continuous Performance Task (CPT; Barkley, 1994). With the CPT, children interacted with a computer

game in which they were asked to click a key every time a certain object appeared on the screen. Attention was measured by the proportion of trials in which the child correctly clicked the key in response to a target object. Impulsivity was measured by the proportion of trials in which a child incorrectly clicked a key in response to a non-target object. These controls were introduced to gauge the extent to which the age-54 month gratification delay "effect" is reduced when differences in concurrent impulsivity are taken into account.

Self-control. We also tested whether controlling for a measure of self-control would reduce the effect of gratification delay on later achievement and behavior. Following the example of Duckworth and colleagues (2013), we created a composite score of self-control from mother and teacher reports at age 54 months. Both mothers and caregivers completed the Child Behavior Questionnaire (CBQ; Rothbart, Ahadi, & Hershey, 1994), and both surveys contained sub-scales that measured "attention focusing" and "inhibitory control." As with the Duckworth et al. (2013) study, we averaged together the teacher and mother "attention focusing" and "inhibitory" control scales to create a self-control composite (see Duckworth et al. (2013) for a description of the measure's psychometric properties).

Analysis

Our primary goal is to estimate the association between early gratification delay and long-run measures of academic achievement and behavioral functioning. Like the work of Mischel and Shoda (e.g., Mischel, Shoda, Peake, 1988; Mischel, Shoda, Rodriguez, 1989; Shoda, Mischel, Peake, 1990) our data did not include a measure of gratification delay in which cross-child differences were generated from some exogenous intervention, so we do not pretend that the associations we estimate reflect causal impacts. Instead, the goal of our investigation is to estimate how much bias might be contained in longitudinal correlations between measures of delay of gratification and measures of child cognitive and behavioral functioning as a result of failure to control for characteristics of children and their environments, all measured before age 54 months, that may be causing both differences in 54-month gratification delay times and adolescent outcomes of interest.

To accomplish our analytic goals, we modeled later academic achievement and behavior as a function of an age-54-month measure of gratification delay and pre-54-month controls:

$$(1)\quad Y_{iLATE} = \alpha_0 + \beta_1 DOG_{i54} + \emptyset Back\&Fam_{iEARLY} + \delta Child_{iEARLY} + \lambda Cog_{iEARLY} + \gamma Home_{iEARLY} + \varepsilon_i$$

where Y_{iLATE} is an outcome measure of either academic achievement or behavioral functioning taken at a later time point (both grade 1 and age 15), DOG_{i54} is the 54-month measure of delay of gratification ("DOG") measured in minutes waited, $Back\&Fam_{iEARLY}$ is a vector of family and child demographic characteristics (e.g., family income, gender, race) assessed prior to the 54-month survey, $Child_{iEARLY}$ is a vector of early measures of the child's personal characteristics (e.g., temperament, birth weight), Cog_{iEARLY} is a vector of early measures of cognitive functioning, and $Home_{iEARLY}$ is a set of measures of the home environment captured by the 36-month HOME battery. Finally, the error term, ε_i, will be uncorrelated with our key estimate, β_1, only if our control variables perfectly capture all of the possible sources of confounding variation.

In addition, we tested models that added controls assessed at age 54 months in order to project how changes in delay of gratification ability, holding constant other concurrent cognitive and behavioral abilities, might affect later development:

(2) $Y_{iLATE} = \alpha_0 + \beta_1 DOG_{i54} + \emptyset Back\&Fam_{iEARLY} + \delta Child_{iEARLY} + \lambda Cog_{iEARLY} + \gamma Home_{iEARLY} + \chi Cog\&Beh_{i54} + \varepsilon_i$

where all model parameters are defined as with Equation 1, but a vector of cognitive and behavioral measures assessed at age 54 months, $Cog\&Beh_{i54}$, is added. Although it might be argued that the estimates from Equation 2 provide the best projections for how a delay of gratification intervention might affect later functioning, we recognize that the concurrent timing of their measurement with the gratification delay task may risk over-controlling for capacities that are themselves a product of past emotional self-regulation.

All continuous variables were standardized so that coefficients can be likened to effect sizes, and all models with control variables included a set of dummy variables for each site to adjust for any between-site differences. Finally, in order to account for missing data on control variables, we used SEM with Full Information Maximum Likelihood in Stata 13.0 to estimate all analytic models.

Additional Results

Additional correlational results. In Table S1, we a correlation matrix among all analysis variables for the children of mothers who did not graduate college (n = 551), and in Table S2, we present the same matrix for the children of college completing mothers (n= 365).

Alternative Models

Because models that relied on the children of college completing mothers yielded unreliable results, we focus only on the children of mothers without college degrees (i.e., the lower-SES sample) for supplemental models shown in Tables S3 through S8.

Continuous performance task. Table S3 presents results from models that added controls for age 54-month attention and impulsivity, as measured by the continuous performance task (CPT). These models also included controls measures prior to age 54 months (compare with models 2, 5, 8 and 11 from Table 3). We found that when CPT measures were added, coefficients for gratification delay in achievement models dropped slightly (approximately 21-27%). Thus, we did not find that a direct measure of impulse control completely explained the effect for gratification delay. Further, we only found the CPT measures to be predictive of later outcomes in the grade 1 achievement model.

Self-control. Table S4 also presents results from models with controls measured prior to age 54-months with our composite score of mother and teacher reported self-control also added. Similar to the CPT results, we found only partial mediation. However, our self-control measure significantly predicted achievement and behavior at both grades 1 and age 15. These results again indicate that the gratification delay measure may tap into processes distinct from self-control. For a more complete investigation of this issue, see Duckworth and colleagues (2013), as they had a more robust examination of this hypothesis using the same dataset and the same measure of self-control.

Additional behavioral outcomes. In Table S5, we present descriptive results for additional behavioral outcomes assessed at age 15.

Table S6 presents associations between gratification delay and measures of risk taking; two measures came from the stoplight task and one from the youth-reported Risky Behavior Questionnaire. Aside from the bivariate association between waiting for 7 minutes and youth-reported risk taking (β = -0.38, SE = 0.12, p = .002), we found no significant associations between gratification delay and measures of risk taking. In Table S7, we display associations

between youth-reported measures of behavior problems and delay of gratification and we again found no significant results.

Disaggregated measures of achievement and behavior problems. In the main text, we averaged together measures of math and reading achievement, and measures of externalizing and internalizing, to create composite measures of achievement and behavior, respectively. In Table S8, we present disaggregated results in which we regressed individual measures of age-15 math, reading, externalizing, and internalizing on gratification delay. Results closely mirrored the results shown in Table 4 of the main text. In models with controls measured prior to age 54-months, we found associations between the gratification delay dummy variables and measures of math and reading achievement that ranged between 0.19 and 0.29.

Compared with the composite score of behavioral problems, we found slightly larger associations between gratification delay and externalizing behavior, but no associations were statistically significant in models that contained controls. We found no associations between gratification delay and internalizing behavior.

References

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science, 14*(2).

Conger, R.D. & Elder, G.H. (1994). *Families in troubled times: Adapting to change in rural America*. New York: Aldine de Gruyter.

Duncan, G.J., & Gibson, C. (2000). Selection and attrition in the NICHD childcare study's analyses of the impacts of childcare quality on child outcomes. *Unpublished paper. Evanston, IL: Northwestern University*. Retrieved from: http://www.gse.uci.edu/person/duncan_g/docs/Duncan_Gibson_NICHD_childcare_quality_review_2000.pdf

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology, 41*(4), 625.

Halpern-Felsher, B. L., Cornell, J. L., Kropp, R. Y., & Tschann, J. M. (2005). Oral versus vaginal sex among adolescents: Perceptions, attitudes, and behavior. *Pediatrics, 115*(4), 845-851.

Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology Review, 2*, 109-145.

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology, 54*(4), 687.

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

NICHD Early Child Care Research Network. (2002). Early childcare and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal, 39*(1), 133-164.

Rothbart, M. K., Ahadi, S. A., & Hershey, K. L. (1994). Temperament and social behavior in childhood. Merrill-Palmer Quarterly, 40, 21-39.

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

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., … Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691–697. doi:10.1177/0956797612440101

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., Najarian, M., & Hausken, E. G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) combined user's manual for the ECLS-K Eighth-Grade and K-8 full sample data files and electronic codebook*. Washington, DC: U.S Department of Education.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43*(7), 352-360.3.

Weinberger, D.A., Schwartz, G.E., (1990). Distress and Restraint as Superordinate Dimensions of Self-Reported Adjustment: A Typological Perspective, Journal of Personality, 58, 381-417.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.

Table S1

Correlations Between All Analysis Variables for Children of Mothers Who Did Not Complete College

PANEL 1		1	2	3	4	5	6	7	8	9	10
Gratification Delay (54)											
1	Continuous	1.00									
2	<0.333 min.	-0.69	1.00								
3	0.333- 2 min.	-0.43	-0.24	1.00							
4	2 to 7 min.	0.01	-0.24	-0.19	1.00						
5	7 min.	0.89	-0.50	-0.39	-0.39	1.00					
Related Measures											
6	Self-control (54)	0.21	-0.12	-0.10	-0.03	0.20	1.00				
7	Attention (54)	0.22	-0.16	-0.08	-0.09	0.26	0.10	1.00			
8	Impulsivity (54)	-0.28	0.26	0.02	0.04	-0.26	-0.21	-0.22	1.00		
Outcome Measures											
9	Achievement (G1)	0.30	-0.26	-0.07	-0.01	0.27	0.30	0.29	-0.26	1.00	
10	Achievement (15)	0.29	-0.27	-0.04	0.02	0.24	0.24	0.18	-0.20	0.60	1.00
11	Behavior (G1)	-0.06	0.03	0.05	-0.01	-0.05	-0.33	-0.06	0.06	-0.09	-0.10
12	Behavior (15)	-0.06	0.07	-0.00	-0.02	-0.05	-0.22	-0.03	0.02	-0.08	-0.09
Demographic Controls											
13	Male	-0.06	0.07	-0.01	-0.01	-0.05	-0.22	-0.01	0.24	0.01	0.12
14	Black	-0.23	0.21	0.05	0.02	-0.23	-0.11	-0.10	0.19	-0.29	-0.33
15	Hispanic	-0.01	-0.02	0.05	-0.01	-0.01	-0.02	-0.01	0.01	-0.04	-0.00
16	Other	-0.01	0.01	-0.02	0.03	-0.02	-0.00	-0.01	-0.05	-0.01	-0.05
17	Age	0.04	-0.04	0.06	-0.07	0.05	0.03	0.02	-0.01	0.02	-0.00
18	Log of Income	0.25	-0.23	-0.04	-0.00	0.22	0.15	0.20	-0.14	0.35	0.32
19	Mother's Age	0.18	-0.17	-0.02	0.01	0.15	0.15	0.12	-0.13	0.21	0.26
20	Mother's Ed (yrs)	0.12	-0.11	-0.01	-0.01	0.11	0.13	0.11	-0.11	0.27	0.21
21	Mother PPVT	0.19	-0.20	0.01	-0.04	0.19	0.21	0.09	-0.14	0.34	0.38
22	Site 1	-0.06	0.03	0.01	0.08	-0.09	-0.06	0.08	-0.06	0.05	-0.16
23	Site 2	0.01	-0.09	0.05	0.03	0.01	0.06	0.05	-0.05	0.02	0.06
24	Site 3	0.09	-0.06	-0.04	-0.02	0.09	-0.05	-0.01	-0.10	-0.03	-0.05
25	Site 4	0.01	0.02	-0.00	-0.04	0.01	0.05	0.02	0.13	-0.01	0.08
26	Site 5	-0.05	0.04	0.05	-0.03	-0.05	0.05	0.00	0.04	-0.03	-0.03
27	Site 6	-0.04	0.06	-0.03	-0.05	-0.00	-0.04	0.01	-0.03	-0.03	0.00
28	Site 7	-0.05	0.02	0.02	0.02	-0.05	-0.01	-0.08	0.13	-0.11	-0.04
29	Site 8	0.02	-0.00	0.00	-0.07	0.05	0.10	-0.02	-0.07	0.01	0.00
30	Site 9	0.01	-0.03	0.01	0.04	-0.01	-0.05	0.02	-0.04	0.15	0.16
31	Birthweight (g's)	0.01	-0.00	0.00	-0.06	0.04	-0.04	0.07	-0.01	0.10	0.15
32	Bracken	0.25	-0.18	-0.10	-0.03	0.25	0.25	0.23	-0.26	0.54	0.47
33	Bayley	0.30	-0.23	-0.08	-0.05	0.28	0.24	0.24	-0.16	0.43	0.37
34	Temperament	-0.11	0.14	-0.04	-0.00	-0.09	-0.16	-0.05	0.11	-0.13	-0.16
H.O.M.E. Controls											
35	Learn. Mater.	0.28	-0.21	-0.10	-0.02	0.26	0.25	0.14	-0.18	0.38	0.36
36	Lang. Stim.	0.20	-0.14	-0.06	-0.05	0.20	0.12	0.09	-0.11	0.27	0.19
37	Phys. Env.	0.21	-0.12	-0.14	0.02	0.19	0.16	0.15	-0.12	0.29	0.24
38	Responsivity	0.16	-0.10	-0.08	-0.05	0.18	0.16	0.15	-0.10	0.25	0.17
39	Academ. Stim.	0.19	-0.15	-0.05	-0.03	0.18	0.12	0.08	-0.13	0.27	0.22
40	Modeling	0.14	-0.07	-0.08	-0.03	0.14	0.12	0.11	-0.05	0.23	0.23
41	Variety	0.22	-0.10	-0.14	-0.05	0.22	0.14	0.14	-0.19	0.29	0.27
42	Acceptance	0.09	-0.02	-0.09	-0.02	0.09	0.16	0.12	-0.14	0.13	0.18

43	Respons. Emp.	0.17	-0.10	-0.09	-0.04	0.17	0.13	0.11	-0.06	0.22	0.13
54-month Controls											
44	Letter Word (54)	0.25	-0.19	-0.08	-0.02	0.23	0.24	0.25	-0.22	0.56	0.44
45	App. Prob. (54)	0.34	-0.27	-0.11	0.02	0.30	0.29	0.31	-0.28	0.63	0.50
46	Pic. Vocab. (54)	0.24	-0.18	-0.03	-0.09	0.24	0.17	0.17	-0.13	0.37	0.46
47	Mem. Sent. (54)	0.26	-0.24	-0.06	0.03	0.23	0.24	0.21	-0.18	0.42	0.41
48	Inc. Words (54)	0.20	-0.16	-0.04	-0.04	0.20	0.10	0.22	-0.14	0.40	0.37
49	Internalizing (54)	-0.04	0.03	0.02	-0.00	-0.03	-0.17	-0.02	0.09	-0.08	-0.08
50	Externalizing (54)	-0.12	0.08	0.07	-0.02	-0.10	-0.39	-0.06	0.07	-0.10	-0.13

PANEL 2		11	12	13	14	15	16	17	18	19	20
11	Behavior (G1)	1.00									
12	Behavior (15)	0.54	1.00								
Demographic Controls											
13	Male	0.01	-0.04	1.00							
14	Black	0.06	-0.04	-0.01	1.00						
15	Hispanic	0.00	0.04	0.05	-0.12	1.00					
16	Other	-0.06	-0.01	-0.07	-0.09	-0.06	1.00				
17	Age	0.04	0.04	0.01	0.02	0.02	-0.02	1.00			
18	Log of Income	-0.15	-0.16	-0.01	-0.35	-0.04	-0.05	0.01	1.00		
19	Mother's Age	-0.24	-0.23	-0.05	-0.25	-0.06	-0.05	-0.03	0.48	1.00	
20	Mother's Ed (yrs)	-0.08	-0.10	-0.05	-0.07	-0.08	-0.08	0.03	0.45	0.36	1.00
21	Mother PPVT	-0.09	-0.02	-0.04	-0.39	-0.05	-0.05	-0.02	0.39	0.34	0.39
22	Site 1	0.09	0.08	-0.02	0.13	-0.09	-0.07	-0.07	-0.05	-0.07	0.02
23	Site 2	-0.06	-0.07	-0.01	-0.14	0.32	0.02	-0.18	0.20	0.08	0.07
24	Site 3	0.01	0.08	-0.07	-0.01	0.04	-0.01	0.09	-0.03	-0.06	0.01
25	Site 4	-0.10	-0.06	0.07	-0.09	-0.07	0.04	-0.02	0.06	0.11	-0.01
26	Site 5	-0.03	-0.06	0.00	0.19	-0.07	-0.01	0.22	-0.10	0.05	0.05
27	Site 6	0.02	-0.05	0.03	0.20	-0.04	-0.05	-0.11	0.02	0.04	0.05
28	Site 7	-0.02	-0.03	0.03	0.01	-0.02	0.04	0.20	-0.12	-0.10	-0.16
29	Site 8	-0.09	0.02	0.01	-0.04	0.03	0.07	-0.18	0.02	0.07	0.03
30	Site 9	0.01	0.00	-0.00	-0.10	-0.08	-0.06	0.04	0.04	-0.03	-0.07
31	Birthweight (g's)	0.05	0.08	0.11	-0.17	0.08	-0.10	0.06	0.04	0.06	0.08
32	Bracken	-0.09	-0.06	-0.14	-0.30	-0.00	-0.05	0.03	0.37	0.27	0.28
33	Bayley	-0.05	-0.11	-0.16	-0.31	-0.06	0.02	-0.01	0.35	0.19	0.24
34	Temperament	0.12	0.14	-0.02	0.19	-0.01	0.06	0.01	-0.22	-0.21	-0.07
H.O.M.E. Controls											
35	Learn. Mater.	-0.08	-0.08	-0.05	-0.37	-0.06	-0.09	0.05	0.45	0.28	0.38
36	Lang. Stim.	0.05	-0.05	-0.02	-0.09	-0.09	-0.13	0.04	0.27	0.07	0.20
37	Phys. Env.	-0.09	-0.08	-0.01	-0.24	0.01	-0.01	-0.04	0.31	0.19	0.19
38	Responsivity	-0.04	-0.02	-0.01	-0.22	-0.00	-0.07	-0.10	0.32	0.22	0.24
39	Academ. Stim.	0.03	-0.00	-0.01	-0.17	-0.07	-0.03	0.03	0.23	0.07	0.23
40	Modeling	-0.02	0.02	-0.04	-0.13	-0.03	-0.07	-0.03	0.26	0.19	0.29
41	Variety	-0.07	-0.03	-0.04	-0.26	-0.05	-0.05	0.01	0.38	0.21	0.27
42	Acceptance	-0.17	-0.15	-0.04	-0.07	0.02	-0.03	-0.07	0.16	0.18	0.15
43	Respons. Emp.	-0.03	-0.02	-0.03	-0.16	-0.03	-0.04	-0.06	0.31	0.15	0.24
54-month Controls											
44	Letter Word (54)	-0.03	-0.03	-0.07	-0.18	-0.06	0.04	-0.03	0.31	0.14	0.26
45	App. Prob. (54)	-0.02	-0.10	-0.10	-0.30	-0.06	0.01	-0.02	0.38	0.22	0.26
46	Pic. Vocab. (54)	-0.05	0.01	0.11	-0.34	-0.06	-0.02	0.03	0.34	0.28	0.26

47	Mem. Sent. (54)	-0.08	-0.05	-0.05	-0.16	-0.13	0.06	0.02	0.26	0.18	0.17
48	Inc. Words (54)	0.01	-0.10	0.00	-0.16	-0.09	0.05	0.04	0.23	0.15	0.14
49	Internalizing (54)	0.52	0.36	0.04	0.06	0.00	0.01	0.14	-0.07	-0.13	-0.11
50	Externalizing (54)	0.64	0.45	-0.06	0.05	0.01	-0.01	0.03	-0.11	-0.17	-0.13

PANEL 3		21	22	23	24	25	26	27	28	29	30
21	Mother PPVT	1.00									
22	Site 1	-0.11	1.00								
23	Site 2	0.12	-0.11	1.00							
24	Site 3	-0.01	-0.11	-0.10	1.00						
25	Site 4	0.00	-0.12	-0.11	-0.11	1.00					
26	Site 5	-0.04	-0.12	-0.11	-0.11	-0.12	1.00				
27	Site 6	-0.02	-0.09	-0.08	-0.08	-0.09	-0.09	1.00			
28	Site 7	-0.05	-0.12	-0.11	-0.11	-0.12	-0.12	-0.09	1.00		
29	Site 8	0.10	-0.10	-0.09	-0.09	-0.10	-0.10	-0.07	-0.10	1.00	
30	Site 9	-0.06	-0.13	-0.12	-0.12	-0.13	-0.13	-0.10	-0.14	-0.11	1.00
31	Birthweight (g's)	0.11	-0.05	0.01	-0.07	0.07	0.01	-0.00	0.00	-0.03	0.01
32	Bracken	0.36	-0.02	0.10	0.01	0.09	0.04	-0.06	-0.19	0.04	0.03
33	Bayley	0.30	-0.04	0.11	0.02	0.16	0.04	-0.10	-0.24	-0.09	0.06
34	Temperament	-0.19	0.07	-0.12	-0.02	-0.02	-0.03	0.06	0.04	0.00	0.03
H.O.M.E. Controls											
35	Learn. Mater.	0.42	-0.03	-0.06	-0.03	0.02	-0.05	0.00	-0.10	-0.02	0.12
36	Lang. Stim.	0.22	0.10	-0.18	-0.19	-0.18	-0.00	0.17	-0.02	0.07	0.13
37	Phys. Env.	0.23	-0.05	-0.06	-0.07	0.05	0.07	-0.03	-0.27	-0.08	0.22
38	Responsivity	0.25	-0.08	-0.01	-0.09	-0.01	-0.11	0.25	-0.32	0.18	0.11
39	Academ. Stim.	0.23	-0.01	-0.16	-0.07	-0.04	0.02	0.10	-0.14	0.05	0.12
40	Modeling	0.23	0.06	0.01	-0.09	-0.08	-0.02	0.09	-0.10	0.00	-0.02
41	Variety	0.29	0.03	-0.12	-0.06	0.01	-0.05	0.08	-0.11	0.05	0.15
42	Acceptance	0.15	-0.07	0.05	-0.07	0.03	-0.03	0.09	0.03	0.01	-0.13
43	Respons. Emp.	0.22	0.08	-0.04	-0.11	-0.01	-0.17	0.15	-0.14	0.14	0.05
54-month Controls											
44	Letter Word (54)	0.31	-0.00	0.00	0.03	0.08	0.01	0.01	-0.15	0.01	0.00
45	App. Prob. (54)	0.37	-0.05	0.03	0.02	0.12	-0.03	0.02	-0.11	-0.01	0.05
46	Pic. Vocab. (54)	0.40	-0.12	0.01	0.03	0.14	-0.02	-0.05	-0.07	0.04	0.02
47	Mem. Sent. (54)	0.23	-0.06	-0.07	-0.05	0.11	0.11	-0.02	-0.13	0.01	0.11
48	Inc. Words (54)	0.24	-0.08	-0.07	-0.06	0.07	0.06	-0.07	-0.04	-0.05	0.20
49	Internalizing (54)	-0.12	0.01	-0.03	0.00	0.06	0.00	0.03	0.02	-0.07	-0.07
50	Externalizing (54)	-0.10	0.04	0.01	0.04	-0.04	-0.01	0.01	0.00	-0.08	-0.00

PANEL 4		31	32	33	34	35	36	37	38	39	40
31	Birthweight (g's)	1.00									
32	Bracken	0.11	1.00								
33	Bayley	0.07	0.46	1.00							
34	Temperament	-0.05	-0.15	-0.10	1.00						
H.O.M.E. Controls											
35	Learn. Mater.	0.08	0.39	0.41	-0.13	1.00					
36	Lang. Stim.	0.07	0.24	0.19	-0.04	0.46	1.00				
37	Phys. Env.	0.04	0.26	0.27	-0.09	0.44	0.31	1.00			
38	Responsivity	0.09	0.27	0.21	-0.11	0.34	0.38	0.25	1.00		

39	Academ. Stim.	0.11	0.30	0.25	-0.02	0.55	0.55	0.32	0.33	1.00	
40	Modeling	0.05	0.18	0.22	-0.14	0.34	0.27	0.24	0.23	0.25	1.00
41	Variety	0.05	0.30	0.30	-0.08	0.56	0.41	0.37	0.28	0.46	0.31
42	Acceptance	0.05	0.17	0.12	-0.06	0.23	0.15	0.17	0.16	0.12	0.30
43	Respons. Emp.	0.04	0.20	0.17	-0.11	0.31	0.48	0.24	0.79	0.28	0.23
54-month Controls											
44	Letter Word (54)	0.09	0.55	0.37	-0.09	0.35	0.27	0.28	0.24	0.32	0.21
45	App. Prob. (54)	0.12	0.54	0.55	-0.15	0.40	0.24	0.30	0.23	0.23	0.16
46	Pic. Vocab. (54)	0.11	0.43	0.38	-0.13	0.39	0.20	0.25	0.25	0.28	0.16
47	Mem. Sent. (54)	0.10	0.38	0.42	-0.06	0.30	0.19	0.26	0.15	0.21	0.16
48	Inc. Words (54)	0.11	0.30	0.39	-0.09	0.28	0.25	0.27	0.13	0.23	0.15
49	Internalizing (54)	0.05	-0.04	0.02	0.14	-0.07	-0.00	-0.05	-0.01	0.03	-0.02
50	Externalizing (54)	0.04	-0.09	-0.01	0.13	-0.10	-0.05	-0.11	-0.10	-0.07	-0.10

PANEL 5		41	42	43	44	45	46	47	48	49	50
41	Variety	1.00									
42	Acceptance	0.16	1.00								
43	Respons. Emp.	0.25	0.18	1.00							
54-month Controls											
44	Letter Word (54)	0.31	0.17	0.19	1.00						
45	App. Prob. (54)	0.28	0.21	0.19	0.57	1.00					
46	Pic. Vocab. (54)	0.32	0.11	0.18	0.44	0.47	1.00				
47	Mem. Sent. (54)	0.31	0.11	0.09	0.40	0.47	0.42	1.00			
48	Inc. Words (54)	0.28	0.11	0.15	0.37	0.46	0.39	0.49	1.00		
49	Internalizing (54)	-0.06	-0.09	-0.03	0.02	-0.03	-0.03	-0.02	-0.04	1.00	
50	Externalizing (54)	-0.09	-0.15	-0.13	-0.03	-0.03	-0.05	-0.10	-0.05	0.59	1.00

Note. n=552. Only children from mothers who had not completed college were included here, and all non-missing cases for each pairwise correlation were included. Tables 2 and 3 include the full variable names and labels. The "G1" abbreviation stands for "grade 1," "15" stands for "age 15," and "54" stands for 54 months.

Table S2

Correlations Between All Analysis Variables for Children of Mothers Completed College

PANEL 1		1	2	3	4	5	6	7	8	9	10
Gratification Delay (54)											
1	Continuous	1.00									
2	<0.333 min.	-0.66	1.00								
3	0.333- 2 min.	-0.54	-0.11	1.00							
4	2 to 7 min.	-0.20	-0.12	-0.12	1.00						
5	7 min.	0.91	-0.48	-0.48	-0.55	1.00					
Related Measures											
6	Self-control (54)	0.17	-0.10	-0.17	0.01	0.17	1.00				
7	Attention (54)	0.15	-0.15	-0.02	-0.05	0.14	0.17	1.00			
8	Impulsivity (54)	-0.26	0.18	0.12	0.06	-0.24	-0.33	-0.29	1.00		
Outcome Measures											
9	Achievement (G1)	0.17	-0.16	-0.05	-0.03	0.15	0.25	0.25	-0.17	1.00	
10	Achievement (15)	0.15	-0.11	-0.08	-0.04	0.14	0.27	0.14	-0.15	0.59	1.00
11	Behavior (G1)	-0.05	0.06	0.04	-0.04	-0.03	-0.20	-0.08	-0.01	-0.02	-0.05
12	Behavior (15)	0.03	0.02	-0.00	-0.11	0.07	-0.18	-0.05	0.06	-0.06	-0.07
Demographic Controls											
13	Male	-0.03	0.04	-0.04	0.07	-0.05	-0.16	-0.01	0.22	-0.02	0.00
14	Black	-0.16	0.02	0.09	0.13	-0.16	-0.11	-0.05	0.02	-0.14	-0.18
15	Hispanic	-0.02	0.00	0.06	-0.06	0.01	-0.01	-0.01	0.01	0.02	-0.00
16	Other	-0.09	-0.02	0.12	0.05	-0.10	0.01	0.08	0.03	0.08	0.12
17	Age	0.03	-0.06	-0.01	0.06	0.00	0.05	0.14	-0.06	0.10	-0.08
18	Log of Income	0.14	-0.14	-0.03	0.00	0.11	0.15	0.01	-0.05	0.09	0.10
19	Mother's Age	-0.04	-0.00	0.01	0.04	-0.03	-0.08	-0.06	0.09	-0.09	0.04
20	Mother's Ed (yrs)	0.08	-0.06	-0.07	0.01	0.07	0.05	-0.02	-0.07	0.06	0.05
21	Mother PPVT	0.21	-0.10	-0.18	-0.09	0.24	0.16	0.02	-0.05	0.12	0.32
22	Site 1	0.06	-0.05	-0.05	0.01	0.06	-0.01	0.05	0.05	0.06	-0.06
23	Site 2	-0.02	0.01	0.04	-0.02	-0.02	-0.01	-0.01	0.01	0.12	0.15
24	Site 3	0.05	-0.02	-0.02	-0.02	0.04	-0.03	0.05	-0.06	-0.08	-0.13
25	Site 4	0.00	0.00	-0.03	0.04	-0.01	-0.00	0.11	-0.03	0.00	0.04
26	Site 5	-0.08	-0.00	0.09	0.05	-0.09	-0.03	0.07	-0.05	-0.02	-0.10
27	Site 6	0.04	-0.06	-0.03	0.04	0.02	0.10	0.02	0.01	0.02	0.06
28	Site 7	-0.02	0.06	-0.04	-0.00	-0.01	-0.02	-0.12	0.09	0.08	0.14
29	Site 8	0.07	0.05	-0.10	-0.10	0.10	0.05	-0.01	-0.05	-0.09	0.00
30	Site 9	-0.09	0.02	0.10	0.04	-0.10	-0.04	-0.03	0.07	-0.06	-0.10
31	Birthweight (g's)	-0.08	0.09	0.05	-0.06	-0.05	-0.04	-0.00	0.04	0.09	0.02
32	Bracken	0.14	-0.17	-0.01	0.01	0.11	0.26	0.22	-0.18	0.40	0.33
33	Bayley	0.26	-0.25	-0.01	-0.06	0.21	0.22	0.16	-0.27	0.24	0.22
34	Temperament	0.04	-0.03	0.04	-0.08	0.05	-0.04	0.02	-0.06	0.01	0.05
H.O.M.E. Controls											
35	Learn. Mater.	0.09	-0.12	-0.04	0.06	0.06	0.17	0.05	-0.15	0.13	0.13
36	Lang. Stim.	0.14	-0.19	0.02	0.00	0.11	0.15	0.00	-0.12	0.08	0.07
37	Phys. Env.	0.06	-0.07	-0.03	0.07	0.01	-0.03	0.02	-0.01	-0.06	-0.10
38	Responsivity	0.11	-0.11	-0.04	-0.04	0.12	0.06	0.04	-0.05	-0.09	-0.04
39	Academ. Stim.	0.13	-0.14	-0.04	0.06	0.07	0.06	-0.08	-0.10	0.04	-0.02
40	Modeling	0.06	-0.08	0.05	-0.05	0.06	0.08	0.01	0.03	0.03	0.06
41	Variety	0.12	-0.11	-0.06	0.02	0.10	0.14	-0.04	-0.09	0.02	0.04
42	Acceptance	0.08	-0.10	0.04	-0.07	0.09	0.17	0.06	-0.14	0.07	0.06

43	Respons. Emp.	0.15	-0.17	-0.01	-0.06	0.15	0.08	0.06	-0.07	-0.01	0.04
54-month Controls											
44	Letter Word (54)	0.16	-0.16	-0.04	-0.02	0.14	0.19	0.15	-0.14	0.57	0.39
45	App. Prob. (54)	0.26	-0.15	-0.19	-0.03	0.23	0.28	0.27	-0.28	0.48	0.50
46	Pic. Vocab. (54)	0.20	-0.13	-0.11	-0.06	0.19	0.21	0.20	-0.13	0.33	0.37
47	Mem. Sent. (54)	0.22	-0.19	-0.08	-0.06	0.22	0.23	0.17	-0.17	0.31	0.32
48	Inc. Words (54)	0.17	-0.07	-0.11	-0.07	0.16	0.11	0.07	-0.13	0.27	0.18
49	Internalizing (54)	-0.02	0.03	0.01	-0.02	-0.01	-0.15	-0.07	0.04	-0.00	-0.01
50	Externalizing (54)	-0.03	0.01	0.07	-0.03	-0.03	-0.37	-0.07	0.10	-0.03	-0.04

PANEL 2		11	12	13	14	15	16	17	18	19	20
11	Behavior (G1)	1.00									
12	Behavior (15)	0.56	1.00								
Demographic Controls											
13	Male	-0.03	-0.04	1.00							
14	Black	-0.07	-0.02	-0.01	1.00						
15	Hispanic	-0.00	-0.02	-0.02	-0.02	1.00					
16	Other	0.08	-0.02	0.03	-0.03	-0.03	1.00				
17	Age	0.02	0.02	-0.02	0.07	-0.02	-0.07	1.00			
18	Log of Income	-0.07	-0.03	-0.01	-0.03	-0.08	0.08	0.00	1.00		
19	Mother's Age	0.05	-0.04	0.01	-0.02	-0.12	-0.04	-0.03	0.22	1.00	
20	Mother's Ed (yrs)	-0.09	-0.10	-0.01	-0.01	-0.05	0.02	0.05	0.24	0.28	1.00
21	Mother PPVT	0.01	0.06	0.06	-0.16	-0.14	-0.18	-0.04	0.21	0.31	0.28
22	Site 1	0.07	0.02	0.03	-0.04	-0.04	-0.05	0.05	-0.06	-0.05	-0.02
23	Site 2	-0.04	-0.07	0.02	-0.05	0.05	-0.07	-0.17	0.13	0.01	-0.11
24	Site 3	0.08	-0.05	0.04	-0.05	0.16	-0.07	0.12	-0.19	-0.06	-0.01
25	Site 4	-0.03	-0.01	-0.06	0.03	0.01	0.04	0.04	0.12	0.05	-0.03
26	Site 5	0.00	0.05	0.04	0.01	-0.06	-0.03	0.23	-0.02	-0.01	-0.03
27	Site 6	-0.15	-0.07	-0.07	0.11	-0.07	0.03	-0.09	0.12	0.03	0.02
28	Site 7	-0.03	-0.03	0.01	-0.04	0.06	-0.02	0.05	-0.01	0.02	0.06
29	Site 8	0.10	0.11	-0.01	0.00	-0.02	0.16	-0.20	0.07	0.06	0.05
30	Site 9	0.04	0.04	0.04	0.04	-0.04	-0.05	0.03	-0.15	-0.06	-0.05
31	Birthweight (g's)	-0.00	0.13	0.15	-0.03	-0.03	-0.03	-0.07	-0.04	-0.03	-0.02
32	Bracken	0.00	-0.03	-0.19	-0.14	-0.14	0.02	-0.01	0.20	-0.04	0.05
33	Bayley	-0.05	-0.06	-0.20	-0.14	-0.05	-0.03	-0.03	0.15	-0.05	0.07
34	Temperament	0.10	0.12	-0.07	0.02	-0.06	0.02	-0.06	-0.03	-0.05	-0.00
H.O.M.E. Controls											
35	Learn. Mater.	-0.05	-0.01	-0.05	-0.18	-0.20	-0.07	0.04	0.09	0.02	0.05
36	Lang. Stim.	-0.08	-0.02	-0.02	0.04	-0.16	-0.07	-0.02	0.05	-0.03	0.05
37	Phys. Env.	-0.03	0.03	0.09	-0.03	0.05	0.05	-0.01	-0.06	-0.14	-0.02
38	Responsivity	-0.12	-0.07	-0.02	-0.00	-0.17	-0.08	-0.07	0.09	0.10	0.06
39	Academ. Stim.	0.02	0.05	-0.07	-0.03	-0.08	-0.05	-0.02	-0.01	-0.04	0.01
40	Modeling	-0.10	-0.05	-0.06	0.04	-0.05	-0.04	-0.06	0.09	-0.01	0.01
41	Variety	-0.08	-0.01	-0.00	-0.04	-0.12	-0.07	0.04	0.07	-0.01	0.08
42	Acceptance	-0.10	-0.06	-0.07	0.01	-0.06	0.07	0.02	0.12	0.04	0.09
43	Respons. Emp.	-0.07	-0.04	0.03	-0.06	-0.10	-0.05	-0.05	0.10	0.08	0.09
54-month Controls											
44	Letter Word (54)	-0.05	-0.03	-0.15	-0.02	-0.04	0.07	0.03	0.17	-0.12	0.07
45	App. Prob. (54)	0.01	-0.04	-0.10	-0.18	-0.01	0.00	0.01	0.11	0.01	0.13
46	Pic. Vocab. (54)	-0.09	0.02	0.12	-0.12	-0.12	0.01	0.04	0.20	0.02	0.13

47	Mem. Sent. (54)	-0.10	-0.08	-0.02	-0.06	-0.02	0.02	0.05	0.09	0.02	0.08
48	Inc. Words (54)	-0.09	-0.09	0.00	-0.08	-0.06	-0.05	0.03	-0.02	-0.03	0.07
49	Internalizing (54)	0.56	0.38	0.02	-0.10	-0.05	0.05	0.01	0.02	0.05	-0.05
50	Externalizing (54)	0.59	0.50	-0.10	-0.04	-0.00	0.00	-0.03	-0.05	0.04	-0.08

PANEL 3		21	22	23	24	25	26	27	28	29	30
21	Mother PPVT	1.00									
22	Site 1	-0.01	1.00								
23	Site 2	-0.07	-0.09	1.00							
24	Site 3	-0.08	-0.09	-0.12	1.00						
25	Site 4	0.01	-0.08	-0.10	-0.10	1.00					
26	Site 5	-0.01	-0.09	-0.12	-0.12	-0.11	1.00				
27	Site 6	0.08	-0.11	-0.14	-0.14	-0.12	-0.15	1.00			
28	Site 7	0.09	-0.08	-0.11	-0.11	-0.09	-0.11	-0.13	1.00		
29	Site 8	0.11	-0.10	-0.14	-0.14	-0.12	-0.14	-0.16	-0.13	1.00	
30	Site 9	-0.10	-0.07	-0.09	-0.09	-0.08	-0.10	-0.11	-0.08	-0.11	1.00
31	Birthweight (g's)	0.11	0.05	0.05	-0.05	-0.00	-0.06	-0.08	-0.03	0.10	0.05
32	Bracken	0.11	-0.05	0.08	-0.04	0.00	-0.06	0.06	0.00	-0.04	-0.05
33	Bayley	0.08	0.00	0.07	-0.10	0.09	-0.01	0.05	-0.06	-0.11	-0.06
34	Temperament	-0.09	0.03	-0.02	-0.00	-0.01	-0.07	0.02	-0.00	0.08	0.00
H.O.M.E. Controls											
35	Learn. Mater.	0.19	-0.02	-0.19	-0.01	-0.13	0.05	0.05	0.13	-0.09	0.05
36	Lang. Stim.	0.10	0.01	-0.15	-0.11	-0.22	0.04	0.17	0.09	-0.00	0.08
37	Phys. Env.	-0.14	-0.06	-0.08	-0.01	-0.02	0.14	0.03	-0.18	-0.07	0.17
38	Responsivity	0.09	-0.13	-0.01	-0.04	-0.02	-0.12	0.34	-0.27	-0.02	0.05
39	Academ. Stim.	0.07	0.03	-0.24	-0.05	-0.13	0.07	0.11	-0.04	-0.00	0.10
40	Modeling	0.10	-0.09	0.02	-0.11	-0.07	0.00	0.15	-0.04	0.00	-0.08
41	Variety	0.19	0.09	-0.20	-0.08	-0.09	-0.02	0.17	-0.02	0.02	0.06
42	Acceptance	0.08	-0.11	-0.13	-0.06	-0.01	-0.09	0.17	0.08	0.04	-0.08
43	Respons. Emp.	0.14	-0.02	-0.03	0.00	-0.04	-0.13	0.17	-0.03	-0.06	0.00
54-month Controls											
44	Letter Word (54)	0.08	0.02	-0.01	-0.08	-0.00	-0.05	0.12	0.06	-0.02	-0.08
45	App. Prob. (54)	0.26	-0.07	0.02	-0.02	0.02	-0.03	0.09	0.03	0.03	-0.10
46	Pic. Vocab. (54)	0.34	-0.06	-0.00	-0.03	-0.01	-0.03	0.06	0.03	0.12	-0.02
47	Mem. Sent. (54)	0.19	-0.07	-0.06	-0.08	0.05	0.05	0.07	0.08	0.02	-0.03
48	Inc. Words (54)	0.14	-0.01	-0.06	-0.03	-0.04	0.03	0.10	0.11	-0.10	0.08
49	Internalizing (54)	-0.04	-0.00	-0.05	-0.06	-0.03	0.12	-0.12	0.00	0.06	0.03
50	Externalizing (54)	-0.06	0.04	0.01	0.02	-0.02	0.04	-0.04	-0.04	-0.01	-0.00

PANEL 4		31	32	33	34	35	36	37	38	39	40
31	Birthweight (g's)	1.00									
32	Bracken	-0.03	1.00								
33	Bayley	-0.01	0.45	1.00							
34	Temperament	-0.01	-0.02	-0.05	1.00						
H.O.M.E. Controls											
35	Learn. Mater.	-0.03	0.31	0.22	0.02	1.00					
36	Lang. Stim.	0.13	0.21	0.15	0.05	0.53	1.00				
37	Phys. Env.	-0.10	0.02	0.01	0.01	0.16	0.09	1.00			
38	Responsivity	0.00	0.22	0.15	-0.04	0.28	0.28	0.15	1.00		

39	Academ. Stim.	-0.01	0.24	0.16	0.05	0.48	0.48	0.18	0.24	1.00	
40	Modeling	0.06	0.08	0.08	0.03	0.18	0.27	0.18	0.26	0.21	1.00
41	Variety	-0.03	0.18	0.27	-0.02	0.30	0.28	0.10	0.26	0.23	0.22
42	Acceptance	0.01	0.15	0.16	0.05	0.17	0.23	0.02	0.14	0.05	0.23
43	Respons. Emp.	-0.00	0.18	0.08	-0.10	0.30	0.43	0.17	0.68	0.21	0.27
54-month Controls											
44	Letter Word (54)	-0.01	0.55	0.25	0.04	0.21	0.20	-0.03	0.14	0.25	0.07
45	App. Prob. (54)	-0.02	0.44	0.42	0.02	0.19	0.08	0.00	0.09	0.12	-0.03
46	Pic. Vocab. (54)	0.10	0.30	0.36	-0.01	0.23	0.18	0.01	0.14	0.13	0.06
47	Mem. Sent. (54)	0.03	0.26	0.35	-0.03	0.15	0.13	-0.07	0.09	0.13	-0.05
48	Inc. Words (54)	0.03	0.15	0.19	-0.06	0.14	0.12	-0.01	0.09	0.10	-0.02
49	Internalizing (54)	0.02	0.04	-0.08	0.11	-0.00	-0.07	0.08	-0.08	0.04	-0.03
50	Externalizing (54)	0.05	-0.06	-0.06	0.06	-0.08	-0.06	0.00	-0.05	-0.03	-0.07

PANEL 5		41	42	43	44	45	46	47	48	49	50
41	Variety	1.00									
42	Acceptance	0.21	1.00								
43	Respons. Emp.	0.24	0.29	1.00							
54-month Controls											
44	Letter Word (54)	0.09	0.04	0.12	1.00						
45	App. Prob. (54)	0.12	0.07	0.08	0.45	1.00					
46	Pic. Vocab. (54)	0.25	0.07	0.15	0.32	0.44	1.00				
47	Mem. Sent. (54)	0.13	0.10	0.09	0.32	0.36	0.40	1.00			
48	Inc. Words (54)	0.10	-0.02	0.09	0.22	0.32	0.22	0.42	1.00		
49	Internalizing (54)	-0.05	0.00	-0.08	-0.02	-0.01	-0.12	-0.08	-0.02	1.00	
50	Externalizing (54)	-0.08	-0.07	-0.04	0.01	-0.01	-0.08	-0.07	0.00	0.57	1.00

Note. n=366. Only children from mothers completed college were included here, and all non-missing cases for each pairwise correlation were included. Tables 2 and 3 include the full variable names and labels. The "G1" abbreviation stands for "grade 1," "15" stands for "age 15," and "54" stands for 54 months.

Table S3

Delay of Gratification and Later Outcomes with Controls for Attention and Impulsivity Added (Lower-SES sample)

	Grade 1 Achievement	Age 15 Achievement	Grade 1 Behavior	Age 15 Behavior
	(1)	(2)	(3)	(4)
Delay of Gratification (categorical)				
<0.333 minutes	ref	ref	ref	ref
0.333- 2 minutes	0.166	0.200	0.096	-0.073
	(0.104)	(0.105)	(0.139)	(0.150)
2 to 7 minutes	0.186	0.271*	-0.016	-0.110
	(0.103)	(0.105)	(0.138)	(0.146)
7 minutes	0.206*	0.185*	0.006	-0.103
	(0.088)	(0.092)	(0.118)	(0.128)
Continuous Performance Task				
Sustained Attention	0.088*	0.026	-0.013	0.013
	(0.034)	(0.035)	(0.044)	(0.048)
Impulsivity	-0.072*	-0.062	0.011	-0.005
	(0.036)	(0.039)	(0.047)	(0.053)
p-value of test of equality of all categories	.113	.056	.868	.853
p-value of test of equality of 2^{nd}, 3^{rd} and 4^{th} categories	.914	.646	.714	.968
Child demographic and H.O.M.E. controls	Inc.	Inc.	Inc.	Inc.
Concurrent 54-month controls	-	-	-	-

Note. n= 552. Standard errors are in parentheses. These estimates compare to the middle column in each set of estimates in Table 4 (i.e., the estimates with controls measured prior to age 54 months), with standardized measures of age 54-month attention and impulsivity added. See Table 4 note for full explanation of model parameters.

* p< .05

Table S4

Delay of Gratification and Later Outcomes with Index of Self-Control Added (Lower-SES Sample)

	Grade 1 Achievement	Age 15 Achievement	Grade 1 Behavior	Age 15 Behavior
	(1)	(2)	(3)	(4)
Delay of Gratification (categorical)				
<0.333 minutes	ref	ref	ref	ref
0.333- 2 minutes	0.221*	0.252*	0.050	-0.108
	(0.103)	(0.102)	(0.133)	(0.146)
2 to 7 minutes	0.204*	0.305*	-0.018	-0.111
	(0.102)	(0.102)	(0.132)	(0.142)
7 minutes	0.261*	0.224*	0.040	-0.068
	(0.085)	(0.087)	(0.109)	(0.121)
Self-Control Composite	0.157*	0.114*	-0.303*	-0.212*
	(0.034)	(0.035)	(0.044)	(0.049)
p-value of test of equality of all categories	.018*	.011*	.943	.848
p-value of test of equality of 2^{nd}, 3^{rd} and 4^{th} categories	.800	.680	.861	.926
Child demographic and H.O.M.E. controls	Inc.	Inc.	Inc.	Inc.
Concurrent 54-month controls	-	-	-	-

Note. n= 552. Standard errors are in parentheses. These estimates compare to the middle column in each set of estimates in Table 4 (i.e., the estimates with controls measured prior to age 54 months), with standardized measures of age 54-month self-control added. See Table 4 note for full explanation of model parameters.

Table S5
Descriptive Characteristics of Supplemental Age 15 Behavioral Measures (Lower-SES Sample)

	M (SD)
Age 15 Behavioral Measures	
Stoplight- Brake Applications	4.85 (1.47)
Stoplight- Brake Time (ms)	900.26 (358.33)
Internalizing (self-report)	47.44 (10.38)
Externalizing (self-report)	50.15 (10.02)
Impulse Control	3.48 (0.89)
Risk Taking	6.81 (5.82)
Observations	478

Note. See Table 2 note. Mean values are presented in each cell, and standard deviations are in parentheses.

Table S6

Associations Between Age 54 month Measure of Delay of Gratification and Measures of Age 15 Risk Taking (Lower-SES Sample)

	Stoplight- Brake Applications			Stoplight- Time Waited			Risk Taking		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Delay minutes (categorical)									
<0.333 minutes	ref	ref	ref	ref	ref	ref	ref	ref	ref
0.333- 2 minutes	-0.085	-0.182	-0.179	0.073	0.096	0.076	-0.072	0.087	0.085
	(0.154)	(0.156)	(0.156)	(0.155)	(0.153)	(0.153)	(0.151)	(0.147)	(0.146)
2 to 7 minutes	0.234	0.143	0.106	0.111	0.179	0.188	-0.065	0.113	0.120
	(0.156)	(0.156)	(0.158)	(0.156)	(0.153)	(0.155)	(0.151)	(0.145)	(0.146)
7 minutes	0.094	0.032	0.011	0.067	0.089	0.107	-0.377*	-0.129	-0.129
	(0.123)	(0.132)	(0.133)	(0.124)	(0.129)	(0.130)	(0.120)	(0.123)	(0.124)
p-value of test of equality of all categories	.236	.254	.374	.906	.709	.677	.004*	.189	.174
p-value of test of equality of 2^{nd}, 3^{rd} and 4^{th} categories	.157	.131	.213	.949	.793	.765	.015*	.093	.084
Child demographic and H.O.M.E. controls	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.
Concurrent 54-month controls	-	-	Inc.	-	-	Inc.	-	-	Inc.

Note. n = 552. Standard errors are in parentheses. See Table 4 note.

* p< .05

Table S7

Associations Between Age 54 month Measure of Delay of Gratification and Measures of Age 15 Behavior Problems (Lower-SES Sample)

	Internalizing			Externalizing			Impulse Control		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Delay minutes (categorical)									
<0.333 minutes	ref	ref	ref	ref	ref	ref	ref	ref	ref
0.333- 2 minutes	0.107	0.196	0.182	-0.011	0.128	0.090	0.040	-0.050	-0.039
	(0.152)	(0.150)	(0.151)	(0.150)	(0.148)	(0.148)	(0.146)	(0.143)	(0.143)
2 to 7 minutes	-0.032	0.047	0.008	-0.124	0.039	0.020	-0.056	-0.150	-0.130
	(0.152)	(0.149)	(0.151)	(0.150)	(0.147)	(0.149)	(0.146)	(0.142)	(0.144)
7 minutes	-0.075	-0.024	-0.038	-0.156	-0.041	-0.068	0.087	-0.041	-0.024
	(0.121)	(0.127)	(0.128)	(0.120)	(0.125)	(0.126)	(0.116)	(0.120)	(0.122)
p-value of test of equality of all categories	.594	.425	.443	.503	.641	.676	.704	.752	.811
p-value of test of equality of 2^{nd}, 3^{rd} and 4^{th} categories	.400	.270	.270	.548	.438	.467	.545	.672	.695
Child demographic and H.O.M.E. controls	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.
Concurrent 54-month controls	-	-	Inc.	-	-	Inc.	-	-	Inc.

Note. n = 552. Standard errors are in parentheses. See Table 4 note.

* p< .05

Table S8

Associations Between Age 54 month Measure of Delay of Gratification and Disaggregated Measures of Age 15 Achievement and Behavior (Compare with Table 4)

	Math Achievement			Reading Achievement			Externalizing			Internalizing		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Delay minutes (categorical)												
<0.333 minutes	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
0.333- 2 minutes	0.305*	0.192	0.143	0.339*	0.222*	0.176	-0.173	-0.125	-0.158	-0.076	0.002	-0.030
	(0.122)	(0.109)	(0.105)	(0.128)	(0.112)	(0.107)	(0.152)	(0.148)	(0.131)	(0.150)	(0.148)	(0.138)
2 to 7 minutes	0.440*	0.287*	0.197	0.400*	0.257*	0.220*	-0.180	-0.114	-0.046	-0.146	-0.080	-0.049
	(0.124)	(0.110)	(0.107)	(0.129)	(0.112)	(0.108)	(0.151)	(0.145)	(0.129)	(0.149)	(0.145)	(0.136)
7 minutes	0.569*	0.222*	0.130	0.615*	0.208*	0.142	-0.307*	-0.196	-0.140	-0.032	0.033	0.060
	(0.098)	(0.093)	(0.090)	(0.103)	(0.095)	(0.092)	(0.120)	(0.123)	(0.109)	(0.119)	(0.123)	(0.116)
p-value of test of equality of all categories	.001*	.042*	.278	.001*	.070	.183	.084	.465	.503	.777	.861	.790
p-value of test of equality of 2nd, 3rd and 4th categories	.046*	.699	.778	.024*	.887	.715	.477	.770	.655	.688	.688	.596
Child demographic and H.O.M.E. controls	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.	-	Inc.	Inc.
Concurrent 54-month controls	-	-	Inc.	-	-	Inc.	-	-	Inc.	-	-	Inc.

Note. n = 552. Standard errors are in parentheses. See Table 4 note.

* p< .05

Data Analysis Files for

“Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Gratification Delay and Later Outcomes”

README

The attached files detail the data analysis process for "Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Gratification Delay and Later Outcomes." The primary dataset used for the study was the NICHD Study of Early Childcare and Youth Development. The study authors obtained a de-identified version of the dataset from Deborah Vandell at the University of California- Irvine, a principal investigator of the study. The data use agreement that the study authors signed prevented them from posting the data online. However, the data can be obtained from ICPSR website with submission of an application and fee (website listed below). It is our understanding that all variables used in the current study should be available in the ICPSR version of the data, but if any questions regarding the data arise, please contact the corresponding author (tyler.watts@nyu.edu).

https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/00233

The analyses described in the attached paper relied on data from Phases I, II, and IV of the study. To replicate all study tables, one should run the 3 attached Stata Do-Files in the order they are listed. The first Stata Do-File (titled "1. Marshmallow Data Set Up") includes code that cleans the study variables taken from the raw SECCYD files to create an analytic dataset. The second Stata Do-File (titled "2. Marshmallow Analysis") uses this analytic dataset to run the models displayed in Tables 2 through 7 of the main text and in the supplementary file tables. Both Do-Files contain extensive comments to assist with any replication efforts.

Finally, Table 1 of the main text also includes data from the Early Childhood Longitudinal Study Kindergarten Cohort (1998), which can be obtained online at the following website:

https://nces.ed.gov/ecls/dataproducts.asp

The authors downloaded the "Child Catalog" for Stata, and created a single dataset that included the variables shown in Table 1. The Stata Do-File called "3. Table 1 Set Up" contains code linking the variables shown in Table 1 to the raw data files from both the NICHD SECCYD and the ECLS-K data sets.

STATA DO-FILE: “1. Marshmallow Data Set Up”

```
/*

Before running any analyses, set a path directory for a single folder using the
"cd" command.  Within this folder, create two new folders: 1) "raw_data";
2) "data".  The "raw_data" folder should contain all raw datasets from the SECCYD.
Each section of this do-file lists the raw datasets that are used to create
the variables used in the analyses. The "data" folder will be used to save
altered datasets that contain cleaned variables.

This section of the file cleans the demographic measures.  Variables created
here include:

Table 3 Name                                            Syntax Name
------------                                            -------------
Male                                                    male
White                                                   dwhite
Black                                                   dblack
Hispanic                                                dhisp
Other                                                   dother
Birth Weight (g)                                        wtgms
Log of Family Income                                    logincome
Mother's Age at Birth                                   momage
Mother's Education                                      momed
Mother's PPVT                                           ppvt
Site (not listed in Table 3)                            site1 - site10

EXTRA Variables
Birthdate (for calculating age at 54 month interview)

Datasets used include:
demo0, demo1, demo6, demo15, demo24, demo36, demo54, fam36, and fam54
*/


use raw_data/demo0.dta, clear

*Merging on other datasets
foreach data in "demo1.dta" "demo6.dta" "demo15.dta" "demo24.dta" "demo36.dta" ///
"demo54.dta" "fam36.dta" "fam54.dta" {
                                            merge 1:1 id using raw_data/`data'
                                            drop _merge
                                            }

******  Gender  ***********
rename CSEX_M01 gender
codebook gender

gen male=.
replace male=1 if gender==1
replace male=0 if gender==2

gen female=.
replace female=1 if gender==2
replace female=0 if gender==1

tab1 male
*705 males; 52%

******  Ethnicity  ***********
rename CRACEM01 ethnicity
codebook ethnicity
tab MHISPM01

tab ethnicity CHISPM01
*Hispanic was not designated as an ethnic category, so Hispanics are classified
*across categories

gen dother=.
```

```
replace dother=1 if (ethnicity==1 | ethnicity==2 | ethnicity==5) & CHISPM01!=1
replace dother=0 if dother!=1 & ethnicity!=.

tab dother CHISPM01
tab dother ethnicity if CHISPM01==0

gen dblack=.
replace dblack=1 if ethnicity==3 & CHISPM01!=1
replace dblack=0 if dblack!=1 & ethnicity!=.

tab dblack CHISPM01
tab dblack ethnicity if CHISPM01==0

gen dwhite=.
replace dwhite=1 if ethnicity==4 & CHISPM01!=1
replace dwhite=0 if dwhite!=1 & ethnicity!=.

tab dwhite CHISPM01
tab dwhite ethnicity if CHISPM01==0

gen dhisp=.
replace dhisp=1 if CHISPM01==1
replace dhisp=0 if CHISPM01!=1 & CHISPM01!=.
tab dhisp

tab1 dother dblack dhisp dwhite
*the 1's add up to 1364 (i.e., no missingness)

*dother - 66; 4.85%
*dblack - 173; 12.68%
*dhisp- 83; 6.09%
*dwhite- 1,042; 76.39%

******  Birth Weight  ***********
rename BWTGMM00 wtgms


******Income*****************

foreach var in INCNTM01 INCNTM06 INCNTM15 INCNTM24 INCNTM36 INCNTM54 {
       gen missing`var'=.
       replace missing`var'= 1 if `var' == .
       }

misstable sum INCNTM*
tab1 missingINCNTM*

egen sum_missing_income = rowtotal (missingINCNTM01 missingINCNTM06 missingINCNTM15
missingINCNTM24 ///
missingINCNTM36 missingINCNTM54)

tab sum_missing_income

*restricting the average to people who have at least 2/6 observations*
egen incomeavg= rowmean (INCNTM01 INCNTM06 INCNTM15 INCNTM24 INCNTM36 INCNTM54) if
sum_missing_income < 5

sum incomeavg
*M= 3.46; SD= 2.70; N= 1299
gen logincome=log(incomeavg)
sum logincome
*M= .958; SD= .796; N= 1299

******Mother's Education***********
rename MEDUCM01 momed
sum momed
*M= 14.23 SD= 2.51; N= 1363

******Mother's PPVT***********
rename STDSCM36 ppvt
sum ppvt
```

```
*M= 99.01; SD= 18.35; N=1167

******Mother's Age**********
rename MAGE_M01 momage
sum momage
*M= 28.12; SD= 5.63; N= 1364

******Site**********
tab site, gen(site)

/*
. tab site, gen(site)

LOCATION OF |
       DATA |
 COLLECTION |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        150       11.00       11.00
          1 |        132        9.68       20.67
          2 |        133        9.75       30.43
          3 |        140       10.26       40.69
          4 |        123        9.02       49.71
          5 |        136        9.97       59.68
          6 |        136        9.97       69.65
          7 |        139       10.19       79.84
          8 |        144       10.56       90.40
          9 |        131        9.60      100.00
*/



*Saving dataset for first set of variables

keep id male dwhite dblack dhisp dother wtgms logincome momage momed ///
       ppvt site1-site10

save data/seccyd_1.dta, replace

*-------------------------------------------------------------------------------*
*-------------------------------------------------------------------------------*
*-------------------------------------------------------------------------------*


/* This section of the file cleans the remaining "Child Demographic and Home
Controls" listed in Table 3.

Table 3 Name                                        Syntax Name
------------                                        -------------
Child's Age at Delay Meas (mos)                     agemo
Bracken Standard Score (36 mos)                     bracken
Bayley (24 mos)                                     bayley
Child Temperament (6 mos)                           temperament_6
HOME Score
       Learning Materials                           hhlrnm36
       Language Stimulation                         hhlanm36
       Physical Environment                         hhphym36
       Responsivity                                 hhresm36
       Academic Stimulation                         hhacam36
       Modeling                                     hhmodm36
       Variety                                      hhvarm36
       Acceptance                                   hhaccm36
       Responsivity- Empirical Scale                hhrsem36



Datasets used include:
demo0, cout6, cout24, cout36, cout54, home36
*/

use raw_data/demo0.dta, clear
```

```
*Merging on other datasets
foreach data in "cout6" "cout24.dta" "cout36.dta" "cout54.dta" "home36.dta" {
                                            merge 1:1 id using raw_data/`data'
                                            drop _merge
                                            }


***** Age at Delay Measure ********

*start with birthdate
rename BRDATM00 birthdate
sum birthdate
*M= 11456.81 (stored as stata numeric daily date) N= 1364

*date of gratification delay measure
sum INTDT55E
*M= 13161.63 (stored as stata numeric daily date) N= 1038

gen agemo= INTDT55E - birthdate
*in days
replace agemo= agemo/30.42
*convert to months, using average number of days per month in a year
sum agemo
*M= 56.05; SD= 1.14; N= 1038

*****   Bracken Standard Score   ********
rename BKSTDO36 bracken
sum bracken
*M= 9.02; SD= 2.89; N= 1159

*****    Bayley   *************
rename MDI24O24 bayley
sum bayley
*M= 92.15; SD=14.64; N=1162

**** Child Temperament *********
rename TEMP_M06 temperament_6
sum temperament_6
*M= 3.18; SD= .40 N=1279


***** HOME SCORES ***************

*Making names lowercase
foreach var of varlist HHLRNM36- HHRSEM36 {
       rename `var' `=lower("`var'")'
       }

sum hhlrnm36 hhlanm36 hhphym36 hhresm36 hhacam36 hhmodm36 hhvarm36 ///
       hhaccm36 hhrsem36


/*
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    hhlrnm36 |      1179    7.157761    2.519644          0         11
    hhlanm36 |      1179    6.016964    1.135049          0          7
    hhphym36 |      1179    5.995759    1.282962          0          7
    hhresm36 |      1179    5.605598    1.362113          0          7
    hhacam36 |      1179    3.370653     1.22406          0          5
-------------+--------------------------------------------------------
    hhmodm36 |      1179    3.166243    1.131962          0          5
    hhvarm36 |      1179    6.754877    1.503296          1          9
    hhaccm36 |      1179     3.38592    .9192371          0          4
    hhrsem36 |      1179    5.429177    1.046798          0          6

*/

keep id agemo bracken bayley temperament_6 hhlrnm36 hhlanm36 hhphym36 ///
       hhresm36 hhacam36 hhmodm36 hhvarm36   hhaccm36 hhrsem36
```

```
save data/seccyd_2.dta, replace

*------------------------------------------------------------------------------*
*------------------------------------------------------------------------------*
*------------------------------------------------------------------------------*

/* This section of the file cleans the variables measuring 54-month cognitive
skills and behaviors, including the delay of gratification measure.

Table 2 Name                                                    Syntax Name
------------                                                    -------------
Delay of gratification (min. waited)                            dog_min
Delay of gratification (categories)
       7 minutes                                                d4
       2 to 7 minutes                                           d3
       0.333 to 2 minutes                                       d2
       < 0.333 minutes                                          d1


Table 3 Name                                                    Syntax Name
------------                                                    -------------
54 mos. WJ-R scores
       Letter-Word ID                                           lwid_ss_54
       Applied Problems                                         appld_ss_54
       Picture Vocabulary                                       picvo_ss_54
       Memory for Sentences                                     memse_ss_54
       Incomplete Words                                         incom_ss_54

54 mos. Child Behavioral Checklist
       Internalizing                                            internalizing_54
       Externalizing                                            externalizing_54

Datasets used:
cout54
*/

use raw_data/cout54.dta, clear


********* Delay of Gratification **************
sum DOG*

rename DOGPFO54 dog_pass
tab dog_pass
*N= 966, 514 passed (53.21%)

*make sure 1 = pass
assert DOGTWO54== 7 if dog_pass==1
*good, everyone who passed has max wait time: 7 min

*KEY INDEPENDENT VARIABLE (MINUTES WAITED):
rename DOGTWO54 dog_min
sum dog_min
*M= 4.47 SD= 3.01; N= 961

********** Delay Categories ********************

gen d1=.
replace d1= 1 if dog_min<= .333
replace d1= 0 if dog_min > .333 & dog_min!=.

gen d2=.
replace d2= 1 if dog_min<= 2 & dog_min > .333
replace d2= 0 if (dog_min > 2 | dog_min <= .333) & dog_min!=.

gen d3=.
replace d3= 1 if dog_min> 2 & dog_min< 7
replace d3= 0 if (dog_min >= 7 | dog_min <= 2) & dog_min!=.

gen d4=.
replace d4= 1 if dog_min>= 7 & dog_min!=.
```

```
replace d4= 0 if dog_min < 7

tab1 d1 d2 d3 d4

/*  All 1's equal 961 (i.e., 961 kids have data across the 4 measures)

d1: 180 coded "1"; 18.73%
d2: 129 coded "1"; 13.42%
d3: 138 coded "1"; 14.36%
d4: 514 coded "1"; 53.49%

*/


********** 54 month WJ-R Scores ***************

*** Letter-Word ID ***
rename WJLWSC54 lwid_ss_54
sum lwid_ss_54
*M=98.93 SD= 13.52; N= 1056

*** Applied Problems ***
rename WJAPSC54 appld_ss_54
sum appld_ss_54
*M= 102.94 SD= 15.63; N=1053

**** Picture Vocabulary ***
rename WJPVSC54 picvo_ss_54
sum picvo_ss_54
*M= 100.24 SD=15.03; N=1060

**** Memory for Sentences ***
rename WJMSSC54 memse_ss_54
sum memse_ss_54
*M= 91.74 SD= 18.49; N=1054

**** Incomeplete Words ***
rename WJIWSC54 incom_ss_54
sum incom_ss_54
*M= 96.67 SD= 13.63; N=1050


********** 54 month Child Behavioral Checklist ***************

*** Externalizing ***
rename BEX_TM54 externalizing_54
sum externalizing_54
*M=51.69 SD=9.39; N=1061

*** Internalizing ***
rename BIN_TM54 internalizing_54
sum internalizing_54
*M= 47.29 SD=8.88; N=1061


keep id dog_min d1 d2 d3 d4 lwid_ss_54 appld_ss_54 picvo_ss_54 ///
       memse_ss_54 incom_ss_54 externalizing_54 internalizing_54

save data/seccyd_3.dta, replace

*-------------------------------------------------------------------------------*
*-------------------------------------------------------------------------------*
*-------------------------------------------------------------------------------*

/* This section of the file cleans the key outcome variables: grade 1 and
age-15 achievement and behavior.

Table 2 Name                                            Syntax Name
------------                                            -------------
Outcome Measures- Grade 1
       Achievement Composite                            ach1
```

```
        Behavior Composite                                      beh1

Outcome Measures - Age 15
        Achievement Composite                                   ach15
        Behavior Composite                                      beh15

Variables used to make composite (not listed in Table 2):
        Letter-Word ID (grade 1)                                lwid_ss_1
        Applied Problems (grade 1)                              appld_ss_1
        Passage Comprehenion (age 15)                           passage_ss_15
        Applied Problems (age 15)                               appld_ss_15

        Internalizing (grade 1)                                 internalizing_1st
        Externalizing (grade 1)                                 externalizing_1st
        Internalizing (age 15)                                  internalizing_15
        Externalizing (age 15)                                  externalizing_15

*It should be noted that the age 15 measures of passage comprehension, applied
problems, internalizing and externalizing were used as the main dependent
variables in Table S8 (results for disaggregated outcome measures).

Datasets used:
coutg1, coutx5
*/

use raw_data/coutg1.dta, clear

merge 1:1 id using raw_data/coutx5.dta
drop _merge

*** Letter-Word ID (grade 1) ***
rename WJLWSC1S lwid_ss_1
sum lwid_ss_1
*M= 111.98; SD=15.79; N= 1025

*** Applied Problems (grade 1) ***
rename WJAPSC1S  appld_ss_1
sum appld_ss_1
*M= 110.80; SD=17.14; N= 1025

*** Passage Comprehension (age 15) ***
rename WJPCSCX5 passage_ss_15
sum passage_ss_15
*M= 107.71; SD=15.72; N= 887

*** Applied Problems (age 15) ***
rename WJAPSCX5 appld_ss_15
sum appld_ss_15
*M= 102.92; SD=14.22; N= 887

*** Internalizing (grade 1) ***
rename BIN_TM1S internalizing_1st
sum internalizing_1st
*M= 48.27; SD=8.94; N= 1028

*** Externalizing (grade 1) ***
rename BEX_TM1S externalizing_1st
sum externalizing_1st
*M= 48.64; SD=9.79; N= 1028

*** Internalizing (age 15) ***
rename BIN_TMX5 internalizing_15
sum internalizing_15
*M= 46.64; SD=9.86; N= 973

*** Externalizing (age 15) ***
rename BEX_TMX5 externalizing_15
sum externalizing_15
*M= 45.51; SD=10.46; N= 973

********** KEY OUTCOME VARIABLES: COMPOSITE SCORES **************
```

```
egen ach1 = rowmean(lwid_ss_1 appld_ss_1)
egen ach15 = rowmean(passage_ss_15 appld_ss_15)
egen beh1 = rowmean(internalizing_1st externalizing_1st)
egen beh15 = rowmean(internalizing_15 externalizing_15)

sum ach1 ach15 beh1 beh15

/*
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        ach1 |      1,025    111.3854    14.59709         59        152
       ach15 |        892    105.2876     13.7189         46        160
        beh1 |      1,028    48.45331    8.322566         32         77
       beh15 |        973    46.07451     9.10633         32         83
*/


keep id ach1 beh1 ach15 beh15 lwid_ss_1 appld_ss_1 passage_ss_15 appld_ss_15 ///
       internalizing_1st externalizing_1st internalizing_15 externalizing_15

save data/seccyd_4.dta, replace


*-------------------------------------------------------------------------------*
*-------------------------------------------------------------------------------*
*-------------------------------------------------------------------------------*

/* This section of the file cleans the variables used in the supplemental
analyses.  These variables appear primarily in the supplementary information
file.

Variable Name                                                 Syntax Name
------------                                                  -------------
Continuous Performance Task
       Sustained Attention                                    propcorrect
       Impulsivity                                            propincorrect

Self-Control Composite                                        selfcontrol54

Supplemental Age 15 Behavioral Measures (Table S5)
       Stoplight- Brake Applications                          int_brake
       Stoplight- Brake Time (ms)                             int_waittime
       Internalizing (self-report)                            internalizing_t
       Externalizing (self-report)                            externalizing_t
       Impulse Control                                        impulse_ctrl
       Risk Taking                                            risk_taking

Variables used for self-control composite:
       CBQ- Attentional Focusing (caregiv)                    cbqattention_cg
       CBQ- Inhibitory Control (caregiv)                      cbqinhibitory_cg
       CBQ- Attentional Focusing (mother)                     cbqattention
       CBQ- Inhibitory Control (mother)                       cbqinhibitory

Datasets used:
cout54, coutx5, cargiv54

Because id does not uniquely identify observations in the cargiv54 data file
(i.e., some children have multiple observations due to having multiple caregivers),
I start with that file and create a unique dataset before merging on the other
files.

*/

use raw_data/cargiv54.dta, clear

keep id ccid CBQAFA54 CBQICA54
egen miss= rowmiss(CBQAFA54 CBQICA54)
tab miss
sort id ccid miss
sum CBQAFA54 CBQICA54
/*
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    CBQAFA54 |        788    4.843954    1.012823       1.25          7
    CBQICA54 |        795    5.069679    1.048816        1.7          7
*/


/*For the CBQ items, every child has a response from only ONE caregiver.  I will
drop all ID's who are missing on the 2 CQB items */

drop if miss==2
codebook id
*803 id's; 803 unique values

*** CBQ- Attention- Caregiver ***
rename CBQAFA54 cbqattention_cg
sum cbqattention_cg
*M= 4.84; SD= 1.01; N=788

*** CBQ- Inhibitory Control- Caregiver ***
rename CBQICA54 cbqinhibitory_cg
sum cbqinhibitory_cg
*M= 5.07; SD= 1.05; N=795

drop ccid miss

****MERGING ON OTHER DATASETS ****

foreach data in cout54 coutx5 {
       merge 1:1 id using raw_data/`data'.dta
       drop _merge
       }

*** CBQ- Attention- Mother ***
rename CBQAFM54 cbqattention
sum cbqattention
*M=4.71; SD=.85; N=1023

*** CBQ- Inhibitory- Mother ***
rename CBQICM54 cbqinhibitory
sum cbqinhibitory
*M=4.66; SD=.78; N=1061

****** SELF-CONTROL COMPOSITE (taken from Duckworth et al., 2013; p. 848) ******
sum cbqattention cbqinhibitory cbqattention_cg cbqinhibitory_cg
egen selfcontrol54= rowmean(cbqattention cbqinhibitory cbqattention_cg ///
       cbqinhibitory_cg)
sum selfcontrol54
*M= 4.77; SD= .72; N=1083


******* Continuous Performance Task ***************
rename CPPCRC54 propcorrect
sum propcorrect
*M= 0.75; SD= .19; N=1002

rename CPPIRC54 propincorrect
sum propincorrect
*M= 0.08; SD= .12; N=1002


****** Supplemental Age 15 Behavioral Measures ******

*** Stoplight Task- Brake Applications ***
rename NBRKSCX5 int_brake
sum int_brake
*M= 4.95; SD=1.42; N=934

*** Stoplight Task- Wait Time ***
rename ATBYBCX5 int_waittime
sum int_waittime
```

```
*M=911.11; SD=349.50; N=923

*** Risk Taking ***
rename ANYR_CX5 risk_taking
sum risk_taking
*M=6.16; SD=5.67; N=954

*** Internalizing (self-report) ***
rename BIN_TCX5 internalizing_t
sum internalizing_t
*M= 47.29; SD=10.17; N=956

*** Externalizing (self-report) ***
rename BEX_TCX5 externalizing_t
sum externalizing_t
*M= 49.31; SD= 9.91; N=956

*** Impulse Control ***
rename MPLSCCX5 impulse_ctrl
sum impulse_ctrl
*M= 3.51; SD= 0.90; N=957


keep id propcorrect propincorrect selfcontrol54 int_brake int_waittime risk_taking ///
       internalizing_t externalizing_t impulse_ctrl cbqattention_cg cbqinhibitory_cg ///
       cbqattention cbqinhibitory

save data/seccyd_5.dta, replace


*------------------------------------------------------------------------------*
*--------------------   MERGING TOGETHER ALL FILES       ----------------------*
*------------------------------------------------------------------------------*


use data/seccyd_1.dta, clear

forvalues i = 2/5 {
       merge 1:1 id using data/seccyd_`i'.dta
       drop _merge
       }

order id dog_min d1 d2 d3 d4 ach1 beh1 ach15 beh15

*Labeling Variables
label var dog_min "Delay of Gratification"
label var d4 "7 minutes"
label var d3 "2 to 7 minutes"
label var d2 "0.333 to 2 minutes"
label var d1 "< 0.333 minutes"
label var ach1 "Achievement Composite - G1"
label var beh1 "Behavior Composite- G1"
label var ach15 "Achievement Composite- Age 15"
label var beh15 "Achievement Composite- Age 15"

label var male "Male"
label var dwhite "White"
label var dblack "Black"
label var dhisp "Hispanic"
label var dother "Other"
label var agemo "Child's Age at Delay Measure"
label var wtgms "Birth Weight (g)"
label var bracken "Bracken Standard Score"
label var bayley "Bayley"
label var temperament_6 "Child Temperament"
label var logincome "Log of Family Income"
label var momage "Mother's Age at Birth"
label var momed "Mother's Education"
label var ppvt "Mother's PPVT"

label var hhlrnm36 "HOME Learning Materials"
```

```
label var hhlanm36 "HOME Language Stimulation"
label var hhphym36 "HOME Physical Environment"
label var hhresm36 "HOME Responsivity"
label var hhacam36 "HOME Academic Stimulation"
label var hhmodm36 "HOME Modeling"
label var hhvarm36 "HOME Variety"
label var hhaccm36 "HOME Acceptance"
label var hhrsem36 "HOME Responsivity- Empirical"

label var lwid_ss_54 "Letter-Word ID 54"
label var appld_ss_54 "Applied Problems 54"
label var picvo_ss_54 "Picture Vocab 54"
label var memse_ss_54 "Memory for Sentences 54"
label var incom_ss_54 "Incomplete Words 54"
label var internalizing_54 "Internalizing 54"
label var externalizing_54 "Externalizing 54"

label var propcorrect "CPT Attention 54"
label var propincorrect "CPT Impulsivity 54"
label var selfcontrol54 "Self-Control Comp. 54"
label var int_brake "Stoplight- Brake App"
label var int_waittime "Stoplight- Brake Time"
label var risk_taking "Risk Taking"
label var internalizing_t "Internalizing (self)"
label var externalizing_t "Externalizing (self)"
label var impulse_ctrl "Impulse Control"

label var cbqattention_cg "CBQ- Attention (caregiv)"
label var cbqinhibitory_cg "CBQ- Inhibitory (caregiv)"
label var cbqattention "CBQ- Attention (mom)"
label var cbqinhibitory "CBQ- Inhibitory (mom)"

label var internalizing_1st "Internalizing (G1)"
label var externalizing_1st "Externalizing (G1)"
label var internalizing_15 "Internalizing (Age 15)"
label var externalizing_15 "Externalizing (Age 15)"

label var lwid_ss_1 "Letter-Word ID (G1)"
label var appld_ss_1 "Applied Problems (G1)"
label var passage_ss_15 "Passage Comprehension (Age 15)"
label var appld_ss_15 "Applied Problems (Age 15)"

label var site1 "Site 1"
label var site2 "Site 2"
label var site3 "Site 3"
label var site4 "Site 4"
label var site5 "Site 5"
label var site6 "Site 6"
label var site7 "Site 7"
label var site8 "Site 8"
label var site9 "Site 9"
label var site10 "Site 10"

************** CHECKING FOR STRANGE MISSING VALUES ***************
misstable sum

*Replacing any alternative missing values to "."
foreach var in momed memse_ss_54 incom_ss_54 appld_ss_1 {
       replace `var' =. if `var' >.
       }

************** SAVING FINAL DATASET FOR ANALYSIS ******************

save data/seccyd_marshmallow.dta, replace
```

STATA DO-FILE: “2. Marshmallow Analysis”

```
/* This file creates Tables 2 through 7 and all of the supplementary file
tables.  This do-file uses the "seccyd_marshmallow" dataset created by
the "1. Marshmallow Data Set Up" file.

Use the same path directory used in "1. Marshmallow Data Set Up.do".  Within
this folder, create two additional subfolders: 1) tables; 2) tables/app.  The
"tables" folder will be used to save tables containing analytic models, and the
"app" folder will be used to save supplemental analyses. */



use data/seccyd_marshmallow.dta, clear



*================================================================================*
*===============       Create sample variables               ==================*
*================================================================================*
gen momedlow=.
replace momedlow=1 if momed<=14
replace momedlow=0 if momed>14 & momed!=.

gen sample1=.
replace sample1=1 if dog_min!=. & (ach1!=. | ach15!=.) & (beh1!=. | beh15!=.)
*918

gen sample2=.
replace sample2=1 if momed<=14 & sample1==1
*552

gen sample3=.
replace sample3 = 1 if momed>14 & momed!=. & sample1==1
*366

gen highses=.
replace highses=1 if momed>14 & momed!=.
replace highses=0 if momed<=14
*482 coded as "1"

*================================================================================*
*============  Standardizing all Continuous Variables        ==================*
*================================================================================*
foreach var in dog_min  ach1 beh1 ach15 beh15 wtgms momage momed ppvt ///
        logincome temperament_6 bayley bracken hhlrnm36 hhlanm36 hhphym36 ///
        hhresm36 hhacam36 hhmodm36 hhvarm36 hhaccm36 hhrsem36 agemo ///
        memse_ss_54 incom_ss_54 picvo_ss_54 lwid_ss_54 appld_ss_54 ///
        internalizing_54 externalizing_54 internalizing_1st externalizing_1st ///
        lwid_ss_1 appld_ss_1 passage_ss_15 appld_ss_15 internalizing_15 ///
        externalizing_15 cbqattention_cg cbqinhibitory_cg propcorrect ///
        propincorrect cbqattention cbqinhibitory internalizing_t externalizing_t ///
        impulse_ctrl risk_taking int_brake int_waittime selfcontrol54 {

         egen z`var' = std(`var')

         }

*================================================================================*
*======================     Variable Lists           ============================*
*================================================================================*

*--------------------------- Basic background controls ------------------------*

global dem1 male dblack dhisp dother zagemo zlogincome zmomage zmomed zppvt ///
        site1 site2 site3 site4 site5 site6 site7 site8 site9

global dem2 male dblack dhisp dother zagemo zlogincome zmomage zppvt ///
        site1 site2 site3 site4 site5 site6 site7 site8 site9
```

```
*without mother's ed for interactions

*------------------------- Additional background controls --------------------*

global msr36 zwtgms zbracken zbayley ztemperament_6 zhhlrnm36 zhhlanm36 zhhphym36 ///
                      zhhresm36 zhhacam36 zhhmodm36 zhhvarm36 zhhaccm36 zhhrsem36


*------------------------- 54 month controls ---------------------------------*

global msr54 zlwid_ss_54 zappld_ss_54 zpicvo_ss_54 zmemse_ss_54 zincom_ss_54 ///
                       zinternalizing_54 zexternalizing_54


*=============================================================================*
*====================       TABLE 2           ================================*
*=============================================================================*
global des1 dog_min d4 d3 d2 d1 ach1 beh1 ach15 beh15

est clear
estpost sum $des1 if sample2==1
est store m1
estpost sum $des1 if sample3==1
est store m2

esttab * using tables/table2_des.csv,  main(mean) b(2) aux(sd) replace label nogaps

*Beta values
est clear
foreach var in zdog_min d4 d3 d2 d1 zach1 zbeh1 zach15 zbeh15   {
       reg `var' highses site1-site9 if sample1==1, robust
       est store m`var'
       }

esttab * using tables/table2_bvalues.csv, aux(se) b(2) r2 nogaps label replace noparen ///
       nostar

*P-Values
est clear
foreach var in $des1  {
       reg `var' highses site1-site9 if sample1==1, robust
       est store m`var'
       }

esttab * using tables/table2_pvalues.csv, main(p) b(3) r2 nogaps label replace noparen ///
       star(* 0.05)


*=============================================================================*
*====================       TABLE 3           ================================*
*=============================================================================*
global des2 male dwhite dblack dhisp dother agemo wtgms bracken bayley ///
       temperament_6 logincome momage momed ppvt hhlrnm36 hhlanm36 hhphym36 ///
       hhresm36 hhacam36 hhmodm36 hhvarm36 hhaccm36 hhrsem36 lwid_ss_54 ///
       appld_ss_54 picvo_ss_54 memse_ss_54 incom_ss_54 internalizing_54 ///
       externalizing_54

est clear
estpost sum $des2 if sample2==1 & d4==1
est store m1
estpost sum $des2 if sample2==1 & d4==0
est store m2

estpost sum $des2 if sample3==1 & d4==1
est store m3
estpost sum $des2 if sample3==1 & d4==0
est store m4

esttab * using tables/table3_des.csv, main(mean) b(2) aux(sd) replace label nogaps

**************************
```

```
* Beta Values and P-Values *
****************************

*LOW SES SAMPLE
*---------------*
*B Values
est clear
foreach var in male dwhite dblack dhisp dother zagemo zwtgms zbracken zbayley ///
       ztemperament_6 zlogincome zmomage zmomed zppvt zhhlrnm36 zhhlanm36 ///
       zhhphym36 zhhresm36 zhhacam36 zhhmodm36 zhhvarm36 zhhaccm36 zhhrsem36 ///
       zlwid_ss_54 zappld_ss_54 zpicvo_ss_54 zmemse_ss_54 zincom_ss_54 ///
       zinternalizing_54 zexternalizing_54  {

       reg `var' d4 site1-site9 if sample2==1, robust
       est store m`var'

       }

esttab * using tables/table3_bvalues1.csv, aux(se) b(2) r2 nogaps label replace noparen ///
       nostar

*P-values
est clear
foreach var in $des2  {
       reg `var' d4 site1-site9 if sample2==1, robust
       est store m`var'
       }

esttab * using tables/table3_pvalues1.csv, main(p) b(3) r2 nogaps label replace noparen ///
       star(* 0.05)


*HIGH SES SAMPLE
*---------------*
*Beta Values
est clear
foreach var in male dwhite dblack dhisp dother zagemo zwtgms zbracken zbayley ///
       ztemperament_6 zlogincome zmomage zmomed zppvt zhhlrnm36 zhhlanm36 ///
       zhhphym36 zhhresm36 zhhacam36 zhhmodm36 zhhvarm36 zhhaccm36 zhhrsem36 ///
       zlwid_ss_54 zappld_ss_54 zpicvo_ss_54 zmemse_ss_54 zincom_ss_54 ///
       zinternalizing_54 zexternalizing_54   {

       reg `var' d4 site1-site9 if sample3==1, robust
       est store m`var'

       }

esttab * using tables/table3_bvalues2.csv, aux(se) b(2) r2 nogaps label replace noparen ///
       nostar

*P-values
est clear
foreach var in $des2  {
       reg `var' d4 site1-site9 if sample3==1, robust
       est store m`var'
       }

esttab * using tables/table3_pvalues2.csv, main(p) b(3) r2 nogaps label replace noparen ///
       star(* 0.05)



*=============================================================================*
*====================       TABLE 7           ================================*
*=============================================================================*

est clear
estpost correlate zdog_min d1 d2 d3 d4 zselfcontrol zpropcorrect zpropincorrect ///
       zach1 zach15 zbeh1 zbeh15 $dem1 $msr36 $msr54 if sample1==1, matrix
est store m1
esttab * using tables/table7_correlation.csv, b(2) unstack not noobs nostar compress replace
```

```
***********************
** LOW SES- Table S1 ***
***********************
est clear
estpost correlate zdog_min d1 d2 d3 d4 zselfcontrol zpropcorrect zpropincorrect ///
       zach1 zach15 zbeh1 zbeh15 $dem1 $msr36 $msr54 if sample2==1, matrix
est store m1
esttab * using tables/app/tables1_lowses.csv, b(2) unstack not noobs nostar compress replace


***********************
** High SES- Table S2 **
***********************
est clear
estpost correlate zdog_min d1 d2 d3 d4 zselfcontrol zpropcorrect zpropincorrect ///
       zach1 zach15 zbeh1 zbeh15 $dem1 $msr36 $msr54 if sample3==1, matrix
est store m1
esttab * using tables/app/tables2_highses.csv, b(2) unstack not noobs nostar compress replace




*-------------------------------------------------------------------------------*
*-------------------                                 -------------------------------*
* ==================      ANALYTIC MODELS        ================================*
*-------------------                                 -------------------------------*
*-------------------------------------------------------------------------------*

*Getting correlation coefficients
corr zdog_min zach15 if sample2==1
*r= .29
corr dog_min ach15 if sample2==1
reg zach15 zdog_min if sample2==1
*b= .24; r2= 0.08; sqrt(0.08)= .28

*===============================================================================*
*===================       TABLE 4           ================================*
*===============================================================================*
capture log close
log using tables/table4.log, replace

*PANEL 1
foreach var in zach1 zach15 zbeh1 zbeh15 {
       sem (zdog_min -> `var') if sample2==1, method(mlmv) nolog

       sem (zdog_min $dem1 $msr36 -> `var') if sample2==1, method(mlmv) nolog

       sem (zdog_min $dem1 $msr36 $msr54 -> `var') if sample2==1, method(mlmv) nolog
}

*PANEL 2
foreach var in zach1 zach15 zbeh1 zbeh15 {
       sem (d2 d3 d4 -> `var') if sample2==1, method(mlmv) nolog
              test d2=d3=d4=0
              gen x4`var'1= r(p)
              test d2=d3=d4
              gen x4`var'1a= r(p)

       sem (d2 d3 d4 $dem1 $msr36 -> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen x4`var'2=r(p)
              test d2=d3=d4
              gen x4`var'2a=r(p)

       sem (d2 d3 d4 $dem1 $msr36 $msr54 -> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen x4`var'3=r(p)
```

```
               test d2=d3=d4
               gen x4`var'3a=r(p)
}

log close
est clear
estpost sum x4*
est store m1
esttab * using tables/table4_pvalues.csv, cells("mean(fmt(3))") replace label nogaps
drop x*


*================================================================================*
*===================       TABLE 5           ==================================*
*================================================================================*
capture log close
log using tables/table5.log, replace

*PANEL 1
foreach var in zach1 zach15 zbeh1 zbeh15 {
       sem (zdog_min -> `var') if sample3==1, method(mlmv) nolog

       sem (zdog_min $dem1 $msr36 -> `var') if sample3==1, method(mlmv) nolog

       sem (zdog_min $dem1 $msr36 $msr54 -> `var') if sample3==1, method(mlmv) nolog
}

*PANEL 2
foreach var in zach1 zach15 zbeh1 zbeh15 {
       sem (d2 d3 d4 -> `var') if sample3==1, method(mlmv) nolog
               test d2=d3=d4=0
               gen x5`var'1= r(p)
               test d2=d3=d4
               gen x5`var'1a= r(p)

       sem (d2 d3 d4 $dem1 $msr36 -> `var') if sample3==1, ///
               method(mlmv) nolog
               test d2=d3=d4=0
               gen x5`var'2=r(p)
               test d2=d3=d4
               gen x5`var'2a=r(p)

       sem (d2 d3 d4 $dem1 $msr36 $msr54 -> `var') if sample3==1, ///
               method(mlmv) nolog
               test d2=d3=d4=0
               gen x5`var'3=r(p)
               test d2=d3=d4
               gen x5`var'3a=r(p)
}
log close
est clear
estpost sum x5*
est store m1
esttab * using tables/table5_pvalues.csv, cells("mean(fmt(3))") replace label nogaps
drop x*


*================================================================================*
*===================       TABLE 6           ==================================*
*================================================================================*

*Creating Interaction Variables
foreach var in d2 d3 d4 {
       gen ses_`var' = `var' * highses
       }

gen ses_cont= highses * zdog_min

*PANEL 1
capture log close
log using tables/table6.log, replace
```

```
foreach var in zach1 zach15 zbeh1 zbeh15 {

	*NO CONTROLS
	sem (zdog_min highses ses_cont ///
	 -> `var') if sample1==1, method(mlmv) nolog

	*SIMPLE CONTROLS
	sem (zdog_min highses ses_cont ///
	$dem2 $msr36 -> `var') if sample1==1, method(mlmv) nolog


	*FULL CONTROLS
	sem (zdog_min highses ses_cont ///
	$dem2 $msr36 $msr54 -> `var') if sample1==1, method(mlmv) nolog

}

*PANEL 2
foreach var in zach1 zach15 zbeh1 zbeh15 {

	*NO CONTROLS
	sem (d2 d3 d4 highses ses_d2 ses_d3 ses_d4 ///
	 -> `var') if sample1==1, method(mlmv) nolog

	test ses_d2=ses_d3=ses_d4=0
	gen x6_`var'1=r(p)

	*SIMPLE CONTROLS
	sem (d2 d3 d4 highses ses_d2 ses_d3 ses_d4 ///
	$dem2 $msr36 -> `var') if sample1==1, method(mlmv) nolog

	test ses_d2=ses_d3=ses_d4=0
	gen x6_`var'2=r(p)

	*FULL CONTROLS
	sem (d2 d3 d4 highses ses_d2 ses_d3 ses_d4 ///
	$dem2 $msr36 $msr54 -> `var') if sample1==1, method(mlmv) nolog

	test ses_d2=ses_d3=ses_d4=0
	gen x6_`var'3=r(p)
}
log close
est clear
estpost sum x6*
est store m1
esttab * using tables/table6_pvalues.csv, cells("mean(fmt(3))") replace label nogaps
drop x6*


*------------------------------------------------------------------------------*
*------------------                              ------------------------------*
* ==================      SUPPLEMENTARY MODELS   ==============================*
*------------------                              ------------------------------*
*------------------------------------------------------------------------------*

*==============================================================================*
*====================    TABLE S3 (CPT ADDED)   ==============================*
*==============================================================================*

capture log close
log using tables/app/tables3.log, replace

foreach var in zach1 zach15 zbeh1 zbeh15 {
	sem (d2 d3 d4 zpropcorrect zpropincorrect $dem1 $msr36 -> `var') if sample2==1, ///
		method(mlmv) nolog
		test d2=d3=d4=0
		gen xs3`var'2=r(p)
		test d2=d3=d4
		gen xs3`var'2a=r(p)
}
```

```
log close
est clear
estpost sum xs3*
est store m1
esttab * using tables/app/tables3_pvalues.csv, cells("mean(fmt(3))") replace label nogaps


*================================================================================*
*=================  TABLE S4 (Self-Control ADDED)   ============================*
*================================================================================*

capture log close
log using tables/app/tables4.log, replace

foreach var in zach1 zach15 zbeh1 zbeh15 {
       sem (d2 d3 d4 zselfcontrol54 $dem1 $msr36 -> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen xs4`var'2=r(p)
              test d2=d3=d4
              gen xs4`var'2a=r(p)
}
log close
est clear
estpost sum xs4*
est store m1
esttab * using tables/app/tables4_pvalues.csv, cells("mean(fmt(3))") replace label nogaps

*================================================================================*
*=============  TABLE S5 (Supplemental Behavioral Outcomes)   ==================*
*================================================================================*

est clear
estpost sum int_brake int_waittime internalizing_t externalizing_t impulse_ctrl ///
       risk_taking
est store m1
esttab * using tables/app/tables5_des.csv, main(mean) b(2) aux(sd) replace label nogaps


*================================================================================*
*=============  TABLE S6 & S7 (Supplemental Behavioral Outcomes)   ============*
*================================================================================*
capture log close
log using tables/app/tables6&s7.log, replace

foreach var in zint_brake zint_waittime zinternalizing_t ///
       zexternalizing_t zimpulse_ctrl zrisk_taking {

       sem (d2 d3 d4 -> `var') if sample2==1, method(mlmv) nolog
              test d2=d3=d4=0
              gen xs6`var'1= r(p)
              test d2=d3=d4
              gen xs6`var'1a= r(p)

       sem (d2 d3 d4 $dem1 $msr36-> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen xs6`var'2=r(p)
              test d2=d3=d4
              gen xs6`var'2a=r(p)

       sem (d2 d3 d4 $dem1 $msr36 $msr54 -> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen xs6`var'3=r(p)
              test d2=d3=d4
              gen xs6`var'3a=r(p)
}

log close
est clear
```

```
estpost sum xs6*
est store m1
esttab * using tables/app/tables6&s7_pvalues.csv, cells("mean(fmt(3))") replace label nogaps
drop xs6*

*================================================================================*
*=============  TABLE S8 (Disaggregated Outcomes)   ============================*
*================================================================================*
capture log close
log using tables/app/tables8.log, replace

*Disaggregated Results
foreach var in zappld_ss_15 zpassage_ss_15 zexternalizing_15 zinternalizing_15 {
       sem (d2 d3 d4 -> `var') if sample2==1, method(mlmv) nolog
              test d2=d3=d4=0
              gen xs8`var'1= r(p)
              test d2=d3=d4
              gen xs8`var'1a= r(p)

       sem (d2 d3 d4 $dem1 $msr36 -> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen xs8`var'2=r(p)
              test d2=d3=d4
              gen xs8`var'2a=r(p)

       sem (d2 d3 d4 $dem1 $msr36 $msr54 -> `var') if sample2==1, ///
              method(mlmv) nolog
              test d2=d3=d4=0
              gen xs8`var'3=r(p)
              test d2=d3=d4
              gen xs8`var'3a=r(p)
}
log close
est clear
estpost sum xs8*
est store m1
esttab * using tables/app/tables8_pvalues.csv, cells("mean(fmt(3))") replace label nogaps
drop xs8*
```

STATA DO-FILE: “3. Table 1 Set Up”

```
/*In this section, I take measures from the NICHD SECCYD demographic file at
age 54 months to make comparisons with the ECLS-K dataset.  Some variables will
come from the "seccyd_marshmallow" dataset created in the "1. Marshmallow Data
Set Up.do" file.

Variables created here:

Table 1 Name                                          Syntax Name
------------                                          -------------
Mother Unemployed                                     mom_unemp54
Number of Children in Home                            childinhome54
Mother Married                                        married54

Mother's Education
       Did Not Complete High School                   momed_nohs
       Graduated from High School                     momed_hs
       Some College                                   momed_somecol
       Bachelor's Degree                              momed_bamore

Income to Needs Ratio
       Income/Needs <=1                               inc1
       Income/Needs >1 & <=2                          inc2
       Income/Needs >2 & <=3                          inc3
       Income/Needs >3 & <=4                          inc4
       Income/Needs >4                                inc5


*Not listed, but used to make "inc" dummies:
       Income at age 54 months                        income54

Datasets used include demo54, seccyd_marshmallow
*/

use raw_data/demo54.dta, clear

*making all variables lowercase
foreach var of varlist id- MOCCUM54  {
rename `var' `=lower("`var'")'
}

*** Income at age 54 Months ***
rename incntm54 income54
sum income54
*M= 3.59; SD=3.17; N=1073

*** Mother Unemployed ***
gen mom_unemp54=.
replace mom_unemp54=1 if mempsm54==0
replace mom_unemp54=0 if mempsm54==1
tab mom_unemp54
*297 coded as "1"; 28%

*** Number of Children in Home ***
rename chldnm54 childinhome54
sum childinhome54
*M= 2.27; SD= 0.98; N=1084

*** Mother Married ***
codebook mstatm54
*see CCDR212- mothers set to "1" are married and living with their partner
recode mstatm54 (2 3 4 5 6 = 0)
rename mstatm54 married54
tab married54
*835 coded to "1"; 77.03%

keep id income54 mom_unemp54 childinhome54 married54
```

```
save data/seccyd_6.dta, replace

*Merge Onto Full Dataset
use data/seccyd_marshmallow.dta, clear

merge 1:1 id using data/seccyd_6.dta
drop _merge

*Generating Sample Variables
gen sample1=.
replace sample1=1 if dog_min!=. & (ach1!=. | ach15!=.) & (beh1!=. | beh15!=.)
*918

gen sample2=.
replace sample2=1 if momed<=14 & sample1==1
*552

gen sample3=.
replace sample3 = 1 if momed>14 & momed!=. & sample1==1
*366

*** Mother Education Dummy Categories ***
gen momed_nohs=.
 replace momed_nohs=1 if momed<12
 replace momed_nohs=0 if momed>=12 & momed!=.

gen momed_hs=.
 replace momed_hs=1 if momed==12
 replace momed_hs=0 if momed!=12 & momed!=.


gen momed_somecol=.
 replace momed_somecol=1 if momed==14
 replace momed_somecol=0 if momed!=14 & momed!=.

gen momed_bamore=.
 replace momed_bamore=1 if momed>=16 & momed!=.
 replace momed_bamore=0 if momed<16

tab1 momed_*

*No high school - 139 coded "1"; 10.20%
*high school - 287 coded "1"; 21.06%
*some college- 455 coded "1"; 33.38%
*BA or more - 482 coded "1"; 35.36%

*** Income to Needs Dummies ***
gen inc1=.
replace inc1=1 if income54 <=1
replace inc1=0 if inc1!=1 & income54!=.

gen inc2=.
replace inc2=1 if income54 <=2 & income54 > 1 & income54!=.
replace inc2=0 if inc2!=1 & income54!=.

gen inc3=.
replace inc3=1 if income54 <=3 & income54 > 2 & income54!=.
replace inc3=0 if inc3!=1 & income54!=.

gen inc4=.
replace inc4=1 if income54 <=4 & income54 > 3 & income54!=.
replace inc4=0 if inc4!=1 & income54!=.

gen inc5=.
replace inc5=1 if income54 > 4 & income54!=.
replace inc5=0 if inc5!=1 & income54!=.

tab1 inc1 inc2 inc3 inc4 inc5

*inc1 - 127 coded "1"; 11.84%
*inc2- 204 coded "1"; 19.01%
```

```
*inc3- 231 coded "1"; 21.53%
*inc4- 180 coded "1"; 16.78%
*inc5- 331 coded "1"; 30.85%


est clear
estpost sum male dblack dhisp dwhite momage momed_nohs momed_hs momed_somecol ///
       momed_bamore inc1 inc2 inc3 inc4 inc5 mom_unemp54 childinhome54 married54 if sample2==1
est store m1
estpost sum male dblack dhisp dwhite momage momed_nohs momed_hs momed_somecol ///
       momed_bamore inc1 inc2 inc3 inc4 inc5 mom_unemp54 childinhome54 married54 if sample3==1
est store m2

esttab * using tables/table1_seccyd.csv, main(mean) aux(sd) b(2) replace label nogaps


*------------------------------------------------------------------------------*
*----------------------- ECLSK DATASET ----------------------------------------*
*------------------------------------------------------------------------------*

/*This section of the file cleans the ECLS-K dataset.  This dataset was downloaded
from https://nces.ed.gov/ecls/dataproducts.asp

The files were downloaded for Stata, and I combined the 6 files into one dataset
called "eclsk1998_full.dta"  Because the dataset is so large, I started by renaming
the variables used for Table 1, then I pared down the dataset to only include these
variables before cleaning.

The process used to create the "income to needs ratio" measure is described
below with the syntax.

Table 1 Name                                        Syntax Name
------------                                        -------------
Male                                                male
Black                                               dblack
Hispanic                                            dhisp
White                                               dwhite
Mother's Age at Child Birth                         momagebirth

Mother's Education
       Did Not Complete High School                 momed_nohs
       Graduated from High School                   momed_hs
       Some College                                 momed_somecol
       Bachelor's Degree                            momed_bamore

Income to Needs Ratio
       Income/Needs <=1                             inc1
       Income/Needs >1 & <=2                        inc2
       Income/Needs >2 & <=3                        inc3
       Income/Needs >3 & <=4                        inc4
       Income/Needs >4                              inc5

Mother Unemployed                                   dmom_employ_unemp
Number of Children in Home                          num_l18
Mother Married                                      parentmarried

*Not listed, but used to make "inc" dummies:
       Income                                       income_cat_imp
       Total num in household                       num_house
       Total adults in household                    num_o18

*Not listed, but used to make "mother's age at child birth":
       Child age                                    r1_kage

Child Weight (not listed)                           C1CW0

*/

use raw_data/eclsk1998_full.dta, clear

rename CHILDID id
```

```
rename GENDER gender
rename RACE race
rename P1HMAGE momage
rename WKMOMED momed
rename W1INCCAT income_cat_imp
rename P1HMEMP mom_employ
rename P1HTOTAL num_house
rename P1LESS18 num_l18
rename P1OVER18 num_o18
rename W1MOMAR parentmarried
rename C1CW0 c1cw0
rename R1_KAGE r1_kage

keep id gender race momage momed income_cat_imp mom_employ num_house num_l18 ///
       num_o18 parentmarried c1cw0 r1_kage

*** Male ***
codebook gender
replace gender=. if gender == -9
recode gender 2=0
tab gender
rename gender male
tab male
*10,950 coded "1"; 51.18%


*** Race ***
gen dwhite=.
replace dwhite=1 if race==1
replace dwhite=0 if race!=1 & race!=.
tab dwhite, missing

gen dblack=.
replace dblack=1 if race==2
replace dblack=0 if race!=2 & race!=.
tab dblack, missing

gen dhisp=.
replace dhisp=1 if race==3 | race==4
replace dhisp=0 if race!=3 & race!=4 & race!=.
tab dhisp, missing

tab1 dwhite dblack dhisp
*white= 11,788; 55.06%
*black= 3,224; 15.06%
*hisp= 3,826; 17.87%

*** Mother's Age at Child's Birth ***
tab momage
codebook momage
*1 case= 0; 3 cases=2; 2 cases=5; missing coded as negative values
drop if momage<18

gen momagebirth=.
replace momagebirth= momage - (r1_kage/12)
sum momagebirth
*M=27.70; SD=6.67; N=17,716

*** Mother Unemployed ***
codebook mom_employ
replace mom_employ=. if mom_employ<0

gen dmom_employ_unemp=.
replace dmom_employ_unemp= 1 if mom_employ==3 | mom_employ==4
replace dmom_employ_unemp= 0 if mom_employ!=3 & mom_employ!=4 & mom_employ !=.
tab dmom_employ_unemp
*5,779 coded "1"; 32.86%

*** Parent Married ***
tab parentmarried
replace parentmarried=. if parentmarried<0
```

```
recode parentmarried 2 = 0
tab parentmarried
*11,049 coded "1"; 71.22%


*** Mother Education ***
codebook momed
replace momed=. if momed<0

gen momed_nohs=.
replace momed_nohs=1 if momed==1 | momed==2
replace momed_nohs=0 if momed!=1 & momed!=2 & momed!=.
tab momed_nohs, missing

gen momed_hs=.
replace momed_hs=1 if momed==3
replace momed_hs=0 if momed!=3 & momed!=.
tab momed_hs, missing

gen momed_somecol=.
replace momed_somecol=1 if momed==4 | momed==5
replace momed_somecol=0 if momed_somecol!=1 & momed!=.
tab momed_somecol, missing

gen momed_bamore=.
replace momed_bamore=1 if momed>=6 & momed!=.
replace momed_bamore=0 if momed <6

tab1 momed_*
*No High School- 2,809; 14.25%
*High School - 5,954; 30.21%
*Some college- 6,337; 32.15%
*BA or more- 4,609; 23.39%

*** INCOME TO NEEDS RATIO ***

*Convert Income to Continuous Variable
tab income_cat_imp, missing
gen income=2500
forvalues i = 2/9 {
       local j = `i' - 1
       replace income= income + (5000 * `j') if income_cat_imp==`i'
       }
replace income= 62500 if income_cat_imp==10
replace income= 87500 if income_cat_imp==11
replace income= 150000 if income_cat_imp==12
replace income= 200001 if income_cat_imp==13
replace income=. if income_cat_imp==.
sum income
tab income
tab income_cat_imp

*Income to Needs Ratio Based on 1997 Categories
codebook num_house num_l18

gen famunit=.
replace famunit=num_house
replace famunit=9 if num_house>=9 & num_house!=.
sum famunit
*M= 4.52; SD= 1.33; N= 17,716

gen childinhome=.
replace childinhome= num_l18
replace childinhome= 8 if num_l18>=8 & num_l18!=.
*Federal pov threshold is capped at 8 children in home
sum childinhome
*M= 2.49; SD= 1.17; N= 17,716

/* These poverty line calculations come from the Census Bureau, which sets
poverty level threshholds each year based on income, family size, and the
number of children in the home.  We used 1997 for these calculations, and the
```

```
table containing these values can be found here:
https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-
thresholds.html
*/

gen incomeneeds1=.
replace incomeneeds1 =11021 if famunit == 2 & childinhome==1
replace incomeneeds1 =12919 if famunit == 3 & childinhome==1
replace incomeneeds1 =12931 if famunit == 3 & childinhome==2
replace incomeneeds1 =16825 if famunit == 4 & childinhome==1
replace incomeneeds1 =16276 if famunit == 4 & childinhome==2
replace incomeneeds1 =16333 if famunit == 4 & childinhome==3
replace incomeneeds1 =20255 if famunit == 5 & childinhome==1
replace incomeneeds1 =19634 if famunit == 5 & childinhome==2
replace incomeneeds1 =19154 if famunit == 5 & childinhome==3
replace incomeneeds1 =18861 if famunit == 5 & childinhome==4
replace incomeneeds1 =23053 if famunit == 6 & childinhome==1
replace incomeneeds1 =22578 if famunit == 6 & childinhome==2
replace incomeneeds1 =22123 if famunit == 6 & childinhome==3
replace incomeneeds1 =21446 if famunit == 6 & childinhome==4
replace incomeneeds1 =21045 if famunit == 6 & childinhome==5
replace incomeneeds1 =26586 if famunit == 7 & childinhome==1
replace incomeneeds1 =26017 if famunit == 7 & childinhome==2
replace incomeneeds1 =25621 if famunit == 7 & childinhome==3
replace incomeneeds1 =24882 if famunit == 7 & childinhome==4
replace incomeneeds1 =24021 if famunit == 7 & childinhome==5
replace incomeneeds1 =23076 if famunit == 7 & childinhome==6
replace incomeneeds1 =29811 if famunit == 8 & childinhome==1
replace incomeneeds1 =29274 if famunit == 8 & childinhome==2
replace incomeneeds1 =28804 if famunit == 8 & childinhome==3
replace incomeneeds1 =28137 if famunit == 8 & childinhome==4
replace incomeneeds1 =27290 if famunit == 8 & childinhome==5
replace incomeneeds1 =26409 if famunit == 8 & childinhome==6
replace incomeneeds1 =26185 if famunit == 8 & childinhome==7
replace incomeneeds1 =35719 if famunit == 9 & childinhome==1
replace incomeneeds1 =35244 if famunit == 9 & childinhome==2
replace incomeneeds1 =34845 if famunit == 9 & childinhome==3
replace incomeneeds1 =34190 if famunit == 9 & childinhome==4
replace incomeneeds1 =33289 if famunit == 9 & childinhome==5
replace incomeneeds1 =32474 if famunit == 9 & childinhome==6
replace incomeneeds1 =32272 if famunit == 9 & childinhome==7
replace incomeneeds1 =31029 if famunit == 9 & childinhome==8


gen incometoneeds=.
replace incometoneeds = income/incomeneeds1
sum incometoneeds
*M=3.29; SD= 2.75; N=13,779
*hist incometoneeds


/*The ECLS-K Base Year documentation also contains suggested thresholds, but they
have less specificity.  Will create an alternate version using these values to cross
check "incometoneeds"

Chart provided in Table 7-2 of https://nces.ed.gov/pubs2001/2001029rev.pdf

*/

gen incomeneeds2=.
replace incomeneeds2= 10973 if famunit==2
replace incomeneeds2= 13001 if famunit==3
replace incomeneeds2= 16655 if famunit==4
replace incomeneeds2= 19682 if famunit==5
replace incomeneeds2= 22227 if famunit==6
replace incomeneeds2= 25188 if famunit==7
replace incomeneeds2= 28023 if famunit==8
replace incomeneeds2= 33073 if famunit==9

gen incometoneeds2=.
replace incometoneeds2= income/incomeneeds2
```

```
sum incometoneeds2
*M=3.22; SD= 2.70; N=13,779
*very close to first calculation

corr incometoneeds incometoneeds2
*r=0.9998

*** Income to Needs Dummies ***
gen inc1=.
replace inc1=1 if incometoneeds <=1
replace inc1=0 if inc1!=1 & incometoneeds!=.

gen inc2=.
replace inc2=1 if incometoneeds <=2 & incometoneeds > 1 & incometoneeds!=.
replace inc2=0 if inc2!=1 & incometoneeds!=.

gen inc3=.
replace inc3=1 if incometoneeds <=3 & incometoneeds > 2 & incometoneeds!=.
replace inc3=0 if inc3!=1 & incometoneeds!=.

gen inc4=.
replace inc4=1 if incometoneeds <=4 & incometoneeds > 3 & incometoneeds!=.
replace inc4=0 if inc4!=1 & incometoneeds!=.

gen inc5=.
replace inc5=1 if incometoneeds > 4 & incometoneeds!=.
replace inc5=0 if inc5!=1 & incometoneeds!=.

tab1 inc1 inc2 inc3 inc4 inc5

*inc1 - 2,271 coded "1"; 16.48%
*inc2 - 3,497 coded "1"; 25.38%
*inc3 - 2,267 coded "1"; 16.45%
*inc4 - 2,293 coded "1"; 16.64%
*inc5 - 3,451 coded "1"; 25.05%


********** Producing Descriptive Table ****

svyset [pweight=c1cw0]


global des male dblack dhisp dwhite momagebirth momed_nohs momed_hs momed_somecol ///
       momed_bamore inc1 inc2 inc3 inc4 inc5 dmom_employ_unemp num_l18 parentmarried

est clear
foreach var in $des {
       svy: mean `var'
       est store m`var'
       estadd sd
       }

esttab * using tables/table1_eclsk.csv,  replace b(%12.2f) aux(sd) nostar wide compress
```