IRVINE NETWORK ON
INTERVENTIONS IN DEVELOPMENT

Greg J. Duncan
Principal Investigator
gduncan@uci.edu

Andrea Karsh
Project Coordinator
akarsh@uci.edu

March 17, 2012

Understanding Variation in the Impacts of Human Capital Interventions on Children and Youth

Greg J. Duncan and Deborah Lowe Vandell

University of California, Irvine

Acknowledgment

The authors are grateful to the NICHD-supported Irvine Network on Interventions in Development (HD065704 P01) and for comments from other Network members (George Farkas, Christopher Carpenter, Marianne Bitler, Margaret Burchinal, Sara Wakefield, Tim Bruckner, Thurston Domina, AnneMarie Conley, Andrew Penner and Candice Odgers), members of the Network's advisory board (Howard Bloom, Daniel Bolt, Douglas Clements, Robert Crosnoe, Kenneth Dodge, Jacquelynne Eccles, Dale Farran, Daniel Keating, Susana Loeb, Jens Ludwig, Wayne Osgood, Sean Reardon and Jeffrey Smith), and from Dorothy Duncan, Pamela Morris and Robert Siegler.

Understanding Variation in the Impacts of Human Capital Interventions on Children and Youth

Abstract

Policy researchers have long known that the impacts of policies and programs differ across individuals. Recent efforts have been directed toward quantifying the scope of treatment effect heterogeneity. But little energy has been expended to conceptualize and test hypotheses about the nature of such treatment effect heterogeneity. With regard to human capital skill and behavior interventions directed at children and adolescents, we posit that a human development perspective provides a useful framework for understanding major sources of treatment effect heterogeneity. The cognitive and socio-emotional development of children and adolescents follows a predictable pattern of stages in which windows for profitable interventions widen or narrow. Moreover, there are predictably persistent differences in individual characteristics such as temperament and family circumstances (e.g., socioeconomic status) that can help explain within-stage variation in intervention impacts. We join others in arguing that key to successful human capital interventions is the "fit" between a given policy intervention and the developmental stage and context of the child or youth participants. Our framework generates many concrete predictions regarding the kinds of human capital interventions that are likely to work for which kinds of children and youth and when during their development.

Understanding Variation in the Impacts of Human Capital Interventions on Children and Youth

I. Introduction

Although policy researchers have long recognized that the impacts of human capital policies and programs differ across individuals, disciplinary barriers prevent many from understanding the sources of this variation. Economists and, increasingly, sociologists bring strong experimental and non-experimental methods to their policy research, and have developed methods for quantifying the scope of heterogeneous treatment impacts (Bloom & Micholopolous, 2011; Imbens & Angrist, 1994; Xie, Brand & Jann., 2011). But most economic theories make few concrete predictions regarding the causes of that heterogeneity or the processes producing the impacts inside of the black policy box. And while sociologists offer a sophisticated conception of the many contexts (e.g., neighborhoods, schools) in which children develop, they rarely link those conceptions to the circumstances of individual children within a given context that lead some to profit much more than others from interventions. We posit that discoveries from neuroscience and psychology regarding child and adolescent development provide a useful framework in which the "fit" between a given intervention and children's developmental needs and capacities explains substantial variation in the impacts of human capital policies and programs.

Evidence on treatment impacts in the Infant Health and Development Program (IHDP) provides some motivation for our paper. Beginning shortly after a child's birth, the IHDP offered a package of services including free, full-day, cognitively-enriching child care between ages one and three to a randomly chosen subset of nearly 1,000 low birth weight infants in eight sites scattered around the country (Gross, Spiker & Haynes, 1997). To be eligible for the program, infants had to have weighed less than 2,500 grams (5.5 pounds) at birth, but eligibility was not further restricted by family income, race or ethnicity. This led to the enrollment of a very demographically heterogeneous set of children and families. The study design permitted the investigators to contrast effects of the intervention on children who were at risk because of family circumstance (low income versus not) and individual circumstances (very low birth weight versus low birth weight).

IHDP treatment impacts varied markedly across these different subgroups of participants (Figure 1). Among children whose mothers reported that family income in the child's first year of life was below 180 percent of the poverty line, low-income children in the treatment group outscored their control-group counterparts by .82 sd on the Stanford-Binet IQ mental subscale at age three – a difference of more than 13 IQ points.[1] For children in high-income families, the treatment impact was much less -- .18 sd, which is less than three IQ points and significantly less than the 13-point treatment impact for low-income children. Thus, if "disadvantage" is defined by an indicator of socioeconomic status, IHDP treatment impacts heavily favored disadvantaged infants. But alternative definitions of "disadvantage" can lead to opposite conclusions: children disadvantaged by being born with a "very low" birth weight (less than 1500 grams or 3.3 lbs) profited much less (.39 sd vs. .95 sd) from the IHDP intervention than "advantaged" heavier babies.

[Insert Figure 1 here]

Although these patterns of impact heterogeneity fail to support a simple conclusion such as "the IHDP program worked best for disadvantaged children," it is not difficult to generate post-hoc explanations for them. For example, the program's early education focus may compensate for the lower levels of academic stimulation in poor relative to higher-income families, while very low birth weight babies face neurological challenges that the program was unable to address. But after-the-fact rationales are a poor substitute for a conceptual framework that generates *a priori* testable hypotheses regarding program treatment effect heterogeneity. Creating such a framework for human capital interventions directed at children and youth is the goal of this paper.

More generally, we seek to identify ways in which educators, intervention scientists and policymakers can make use of neurologically and developmentally informed principles to increase the likelihood that educational programming and behavioral interventions will achieve their intended goals of supporting positive development in children and youth. Key to successful interventions, we argue, is an understanding of how the cognitive and socio-emotional development of all children and adolescents follows a predictable pattern of stages in which windows for profitable interventions widen or narrow. Moreover, successful interventions may also rest on understanding predictable variation *within* a given stage in family circumstances and in the timing of normative development as children reach important developmental milestones – e.g., the ability to focus attention on learning tasks for young children and puberty for adolescents. Thus, a thorough understanding of both across- and within-stage variation in children's capacities and circumstances should inform the design of effective skill-building interventions and their evaluations.

To generate hypotheses regarding the likely impacts of these policies across and within children's developmental stages, we follow past work in focusing on the *congruence* ("fit") between the developmental needs of children and youth and the design and nature of intervention policies (Clements & Sarama, 2011; Eccles, et al., 1993). Children profit most from interventions that are well matched to their developmental stage and individual circumstances. We elaborate our ideas with a review of the features of child and adolescent development that have direct implications for optimizing the fit between the child and the skill or behavioral intervention under consideration.

In the final section of the paper we apply our framework to a number of rules of thumb and puzzles emerging in the human capital intervention literature. First is the idea of "skill begets skill" -- should we always expect that children with the highest levels of cognitive and socio-emotional skills would profit most from human capital interventions? Second is the proposition that earlier is better when it comes to human interventions. How does this square with an approach that tailors an intervention to the fit between program offerings and the needs and abilities of the children or adolescents? We close with some ideas about how the impacts of interventions might be enhanced through better matches between programs and the developmental needs of children and youth.

II. Matching interventions to normative development

Normative development is defined as age-related changes in behavior or functioning that occur under ordinary conditions and follow a predictable sequence. These changes can have important implications for the effectiveness of human capital interventions. Large bodies of

research literature in neuroscience and developmental psychology describe normative development in physical, cognitive-linguistic and socio-emotional domains (Steinberg, Vandell, & Bornstein, 2011). We argue that these normative changes provide a general framework in which interventions may prove to be more or less effective, depending on how well the intervention's design is matched to a child's capacities and capabilities.

The importance of understanding differences in functioning across child and adolescence stages can be seen in the pattern of age-related impacts averaged across four random-assignment welfare and anti-poverty policies conducted in the 1990s, all of which lasted between two and three years and all of which increased parental employment and family income (Morris, Duncan & Clark-Kauffman, 2005). The impacts of these programs on children's school achievement varied markedly by age (Figure 2). Treatment-group children between the ages of two and five when the programs began, most of whom would be making the transition into elementary school during or shortly after the programs were in operation, scored significantly higher on achievement tests than their control group counterparts. The achievement of children age six to nine did not appear to be affected by the programs. But the achievement of children who were age 10 and 11 when the programs began seemed to be hurt by the programs' efforts to increase parental employment and family income.

[Insert Figure 2 here]

As do Morris et al. (2005), we believe that salient features of children's normative development can account for many of these differences. In the case of children who were transitioning into elementary school, employment and income supports for parents fit well with the children's full-day, reliable and structured school environments. However, qualitative research revealed that adolescents with younger siblings were burdened by child care responsibilities when their mothers increased their employment (Gennetian et al., 2004). Indeed, additional analyses of the experimental data showed that negative impacts were concentrated among adolescents with younger siblings.[2] Taken together, it appears that work-focused anti-poverty programs fit the developmental needs and family circumstances of younger children much better than older children.

In the following sections, we examine the implications of normative development for the design of a variety of targeted interventions during age periods from infancy through adolescence.

Development and the Brain

Normative psychological and behavioral development is linked to functional and structural changes in the brain, and the timing of the development of different regions of the brain is associated with certain physical, cognitive, and social competencies. Of special importance during the first three years of life are the areas associated with vision, hearing, motor development and balance. During early and middle childhood, brain development is concentrated in the frontal lobes, an area related to higher-level information processing skills. The prefrontal cortex, an area that promotes higher-level cognitive and emotional functions, has a long developmental trajectory, and is not fully formed until late adolescence. Of course, development of these areas does not occur in isolation but is part of a broader, integrated neural system.

Across each of these regions, brain development depends heavily on two complementary processes: *synaptogenesis*, or the creation of connections or synapses between neurons that

increases the efficiency of communication between brain cells, and *synaptic pruning*, or the elimination of unused synapses. Contemporary neuroscience research has shown that experience plays a powerful role in sculpting the detailed architecture of the brain (Nelson & Sheridan, 2011). Relatively few aspects of physical, cognitive, linguistic or socio-emotional development reflect a maturational unfolding of a fixed genetic plan. It is the interplay of biology and experience that drives most development.

Moreover, the timing of an intervention may matter a great deal. Changing developmental trajectories is easier during periods of dramatic growth than when behaviors or skills are firmly established (Caspi & Shiner, 2006; Nelson & Sheridan 2011). Neuroscientists have used the term *developmental plasticity* to refer to periods in which substantial change is possible. But, as Nelson and Sheridan (2011) argue, developmental plasticity is a two-edged sword because children are more receptive to experiences that change trajectories in both positive and negative directions. In periods of developmental plasticity, children are more receptive to growth-promoting experiences *and* more vulnerable to harsh, inhospitable conditions.

Not all periods of plasticity occur early in life since neuroscience research has also demonstrated that the kinds of learning and behaviors targeted by human capital interventions involve the use and coordination of a wide range of neurological processes and functions (Howard-Jones, Washbrook & Meadows 2012). Classroom learning cannot take place before children have mastered an ability to control their impulses and engage in learning tasks.

Birth to Three

At birth, human infants enter the world with rudimentary capacities – sucking and rooting reflexes, auditory capabilities that are sensitive to sound frequencies in the range of the human voice, a visual system that is sensitive to the human face, and a communication system (crying) that can signal hunger and pain and is useful for eliciting the care of adults.

From these initial building blocks, the next three years are marked by dramatic gains in physical, social, cognitive, and language development. By 18 months, children have formed strong emotional or attachment bonds to their parents or primary caregivers, which function as sources of emotional security and a foundation for other healthy relationships later in life. By age three, considerable growth has occurred in children's vocabulary, language, memory and attention.

The importance of these early building blocks can be seen in studies of early deprivation and later remediation. Children who receive the implants at earlier ages typically show better speech perception, speech production, vocabulary, grammar, and reading comprehension (Connor, Craig, Raudenbush, Heavner & Zwolan, 2006). There is less improvement in speech production and vocabulary following implantation at older ages, and none at all in vocabulary after 3.5 years or in speech production after age seven. In this case, earlier implants best match children's speech and language needs and competencies.

A second example is based on horrific policies implemented by Romanian dictator Nicolae Ceaușescu that sent hundreds of infants to orphanages and exposure to extreme deprivation. With 30 infants per caregiver, children were confined to cribs for much of the day, where they had little contact and few interactions with adults. Nor did they have stimulation from toys or other children. Feedings were administered using propped bottles.

The fall of the Ceaușescu regime in 1989 and news stories about the plight of these children led to a brief period in which hundreds of children were adopted by families in the UK, Canada, and the United States. The varying adoption ages provided a natural experiment for studying impacts of the timing and duration of early deprivation on children's development. Rutter and colleagues (2007) studied 165 adopted children and found that shortly after their adoptions, the children showed severe physical, cognitive and social deficits. Weight, height and head circumference were far below norms. Repetitive, autistic-like behaviors were evident in many. When tested on a battery of social and cognitive assessments, the children displayed severe cognitive deficits and had great difficulty regulating their own emotional reactions and reading the emotions of others. These problems were most serious for children who had spent at least two years in orphanages before they were adopted.

After both younger and older children were adopted, some aspects of development, such as weight and height, gradually recovered to a considerable degree. But recovery in the socio-emotional and cognitive domains was more differentiated. For example, when children reached the age of 11, their scores on an inattention/hyperactivity scale as reported by parents and teachers were significantly higher if they were adopted from Romanian orphanages when they were older than six months old relative to either Romanian children adopted before six months or a comparison group of children adopted in the UK as infants (Figure 3). Clearly, then, the adoption "treatment" fit best when the Romanian children had spent less than six months in an orphanage.

[Figure 3 here]

The Preschool Years (three to five years)

The preschool years are marked by impressive gains in children's language and cognition, which are the building blocks for later academic skills (National Research Council, 1998; Sarama & Clements, 2009). Changes also occur during this period in children's behavioral regulation. Three-year-olds have great difficulty suppressing their first impulses in games like Simon Says or when slowly tracing a pattern with their finger, while six-year-olds can perform these tasks with ease (Kochanska, Aksan, Penney & Doobay, 2007; Rothbart, Posner & Kieras,2006).

Gains are not uniform across children, though. At kindergarten entry, children from low-income families and of less educated parents score significantly lower, as a group, than children from more advantaged families on tests of behavioral regulation, vocabulary, numeracy, and pre-reading skills (Duncan & Magnuson, 2011). Accordingly, a major thrust of educational interventions in the preschool period has been to improve the school readiness skills of poor children.

We posit that the success of these interventions will be greater if the programming is closely attuned to children's competencies and skills. For example, Ramani and Siegler (2008) argue that understanding the position of and distances between numbers on the number line is key for children's mastery of addition and subtraction. In a series of random assignment experiments, they show that a simple Chutes and Ladders-type board game is very well suited to promoting preschoolers' understanding of the number line. Specifically, preschool children counting out their moves in a linear version of the board game with numbers in the squares show greater mastery of the number line and arithmetic two months later than children playing an

identical game with colors instead of numbers, or even than children playing a circular version of a numbered board.

Middle Childhood (6 to 12 years)

Freud (1910) mistakenly referred to the middle childhood years (ages 6 to 12 years), situated between early childhood and adolescence, as the "latency period" in which relatively little happens. We now know that a lot happens during this time. In terms of brain development, the frontal lobes, which are responsible for critical thinking and problem solving, undergo a growth spurt between six and eight years. This is also the period when most children begin their formal education.

These changes set the stage for the growth in academic skills that occurs during middle childhood. At the beginning of formal schooling (age five in the U.S.) children may be able to write their name, recognize some letters and count to ten. Six years later, most are able to read and comprehend chapter books, compose multi-paragraph essays, and solve complex mathematical problems. These changes are accompanied by expectations of specific academic materials to be mastered at each grade level. These competencies are cumulative, build on earlier skills, and set the stage for instruction and skills for the next year.

As shown in Figure 4, initial gains in both math and literacy across the kindergarten year are very large, amounting to more than one standard deviation (Hill, Bloom, Black & Lipsey, 2008).[3] By third grade, gains are less than half as large, and they continue to decline throughout the high school years. These patterns suggest that opportunities for academic interventions may be particularly plentiful in the early grades. Indeed, interventions such as Tennessee Star, which decreased class sizes, show much larger impacts in kindergarten and first grade than later (Finn & Achilles, 1999*)*.

[Insert Figure 4 here]

Why do achievement gains decline across middle childhood? One possible contributor, consistent with our "fit" argument, is that instruction in many elementary schools does not keep pace with children's cognitive and academic capacities. In particular, a focus on basic skills and lack of attention to conceptual understanding, reasoning and problem solving can be problematic for children in the older grades. A longitudinal study of 1,000 children enrolled in more than 400 school districts found that across all elementary school grades, teachers focused almost exclusively on teaching basic skills rather than providing conceptually focused instruction and provided very little evaluative feedback of student performance (Pianta, Belsky, Houts, Morrison, NICHD ECCRN, 2007).

Adolescence

Normative development during the adolescent period (ages 13 to 20) poses many challenges – but also opportunities – for educators and other intervention designers. Chief among the developmental events are the onset and timing of puberty, which produces a host of physical, hormonal, psychological and social changes (Steinberg et al., 2011). Puberty brings with it renegotiations in parent-child relationships as adolescents seek greater autonomy and independence from parental control, but there is evidence that adolescents still need and want their parents' support and affirmation. Developing a good balance between autonomy and connectedness is viewed by scholars as a critical challenge during this developmental period (Cooper, Grotevant & Condon, 1983).

Adolescents' needs can also conflict with the structure of education interventions. For example, Eccles et al. (1993) argue that the primary/middle-school educational model is inferior to an integrated K-8 structure because middle schools are poorly matched to the emerging developmental demands of children as they transition to adolescence.

Transitioning adolescents are in special need of close relationships with adults outside of their homes, and yet promotion to middle school involves changing from elementary schools' greater focus on personal aspects of schooling to a middle school structure in which teachers are subject-specific, see many students in one day, and focus on impersonal tasks rather than personal relationships (See Weiss & Kipnes, 2006, for a review of evidence on middle-school structure). Adolescents have fragile and malleable self-perceptions, but most middle school teachers have tougher grading standards and higher expectations than elementary school teachers. Adolescents' heightened concern about their status relative to peers is ill served by the ways in which middle schools track students by ability level. An adolescent's need for more complex academic tasks is often met by more rote teaching styles. Needs for self-determination, participation in rule making and emotional support are met by increased middle-school regimentation and rigid disciplinary policies.

Blyth et al. (1983) tested some of these ideas by taking advantage of the fact that building space constraints led Milwaukee public schools to operate schools spanning kindergarten through eighth grade as well as elementary (K-6th grade) and junior high (grades 7-9) schools. When compared with students in K-8 schools, children forced to change schools between grades six and seven reported significantly lower grades, more anonymity and less participation in extracurricular activities. Girls also reported lower self-esteem, with drops in self-esteem for those in middle schools corresponding to the transitions into and out of middle school (Figure 5). The transition into high school was less problematic for the K-8 than for middle school students.

[Insert Figure 5 here]

Corroborating evidence comes from a study that randomly assigned teachers to an intervention explicitly designed to improve the fit between the nature of teacher-student interactions and adolescents' developmental, intellectual and social needs (Allen, Pianta, Gregory, Mikami, & Lun, 2011). Teachers of a variety of middle and high school courses were coached extensively on improving the quality of their ongoing classroom interactions with students. Scores on state standardized tests rose significantly more for students in the treatment relative to the control classroom in the year following the intervention training.

A final hallmark of normative development during adolescence is the increasing importance of relationships with peers (Steinberg et al., 2011). The hypersensitivity of adolescents to peer approval can produce virulent consequences among groups of deviant youth. Studies of the micro-dynamics of deviant groups show how talk of deviance often elicits approving reactions from peers, which in turn leads to more such talk (Dishion & Andrews, 1995). The resulting reinforcing process has been called "deviancy training" and can subvert well-intentioned programs that provide services to groups of deviant youth (Dishion, McCord & Poulin, 1999; Dodge, Dishion & Lansford, 2006). The problem appears to be most acute in the case of moderately deviant youth – young adolescents in the early stages of manifesting serious behavior problems who lack the family or other supports that might otherwise keep their behavior in check.

Poulin et al. (2001) conducted one such study, in which they evaluated the effects of an ambitious intervention program that brought deviant teens together for sessions that emphasized prosocial goals and self-regulation. Teachers blind to treatment status reported *more* problem behavior three years later for youth in the treatment group than in a comparison group, while treatment-group youth also reported more tobacco use than control-group youth.

Thus, human capital intervention policies addressing problem behaviors risk these unintended (iatrogenic) consequences by providing their treatments to groups of problematic children or youth (Dodge et al., 2006). This is most obvious in juvenile justice programs such as detention centers, prisons and boot camps. But it is also common in the educational system, when children who display conduct problems or are diagnosed with behavioral or emotional disorders are assigned to self-contained classrooms or separate schools. Avoiding situations that foster iatrogenic behaviors can be expensive (e.g., providing individual rather than group treatment) or costly in other ways (e.g., if non-deviant students are prevented from learning by the presence of classmates with conduct problems). On the other hand, evaluation studies have identified effective programs that provide their treatments to entire schools or classes, or that isolate, use adult mentors effectively, or treat behavior-problem "hotspots" in inexpensive ways that avoid promoting the formation of deviant groups (Dishion, Dodge & Lansford, 2006).

III. Matching policy to individual and family differences

Beyond the lessons from normative development are a set of principles with implications for understanding heterogeneous program impacts *within* developmental stages. Some relate to persistent individual differences such as temperament that are formed early in life and change little as children grow up. Others relate to recurrent group differences defined by the contexts in which a child develops. Among the many contexts (e.g., family, neighborhood, cultural) that may matter, we discuss the one likely to be most important – the family.

Individual differences

Developmental psychologists define *individual differences* as within-stage variations that persist over time and, for the most part, do not develop in an orderly sequence. For example, children do not routinely and systematically change from calm to frenetic, shy to outgoing, or happy to sad. Instead, these temperamental qualities are believed to reflect persistent biological predispositions in how children approach and react to the world, although they may be manifested in different ways during different developmental periods. Some individual differences are even apparent in the fetal period. Higher levels of fetal activity predict precocious development of motor skills, stronger reactions to being physically restrained at ages one and two, and more advanced language and play during the toddler period (DiPietro, Bornstein, Hahn, Costigan, & Achy-Brou, 2007).

Some children may lack the capacity or competencies to profit from certain interventions. Figure 1 suggests that the IHDP enrichment program was not well matched to the needs of very low birth weight children in the program. Although the cognitive developmental trajectories of babies close to the 2,500-gram low birth weight threshold differ little from those of normal birth weight babies, children with much lower birth weight face greater neurological challenges that the IHDP treatment appeared less able to address (Duncan & Sojourner, 2011). As we explain below, this would appear to be a case of "skill begets skill" (Cunha, Heckman &

Schennach2010): children coming into the program with normal cognitive capacities seem best matched to take advantage of what the program had to offer.

Vygotsky's (1978) conceptualization of the *Zone of Proximal Development* provides a useful way to think about the match between a child's individual capabilities and any given education intervention. The Zone of Proximal Development refers to the space (or zone) between what a child can do alone without assistance (a mastered skill) and what the child can do with the assistance of an adult or more skilled peer. Vygotsky argued that this zone defines potential learning and development. Initially, children may not be able to accomplish a task (read a book, solve a math problem, ride a bike, recite a story) even with help and prompts. As they get older, children may be able to master tasks such as these if provided lots of help, referred to as "scaffolding" (Bruner, Jolly & Sylva, 1976). Key features of effective scaffolding are that they are tuned to children's developing skills and that adults (or curricula) provide only as much support or assistance as children need. Once skills are mastered, effective instruction moves on to the next task that requires substantial scaffolding to accomplish.

Nonexperimental studies suggest that mismatches may be quite common. Xue and Meisels (2004) found that children who entered kindergarten with higher rather lower literacy skills profited more the most from integrated language arts instruction and the least when the emphasis was on phonics in their kindergarten classroom. For students in third grade, a pivotal year in which children are expected to be able to read longer and comprehend more complex texts, Connor et al. (2006) found that initially poor readers profited the most from teacher-managed strategies, whereas children with good initial reading comprehension enjoyed greater reading achievement growth when more classroom time was spent in child-managed strategies.

One of the most startling hypotheses about treatment heterogeneity associated with individual differences comes from the field of neurodevelopment, and centers on children who begin life with difficult temperaments characterized by high reactivity (Ellis, Boyce, Belsky, Bakermans-Kranenburg & Van Ijzendoorn, 2011). Medical studies have shown that highly reactive young children exposed to adverse home or child care environments have the highest rates of respiratory illness. Unexpectedly, though, highly reactive children exposed to high-quality environments have the *lowest* rates of illness.

Might this interaction apply to a broader set of contexts and outcomes? A great deal of empirical evidence supports the prediction that when exposed to harsh and unsupportive environments, highly reactive children exhibit more problematic behavior. But what about the complementary hypothesis that "highly reactive" children in unusually positive environments fare *even better than* their more even-keeled peers? Pluess and Belsky (2009) find this very pattern with child care quality in non-experimental data from the NICHD Study of Early Child Care and Youth Development. Children rated by their mothers as having difficult temperaments show more problem behavior when exposed to low quality child care than children with normal temperaments. Yet when child care quality is high, children with difficult temperaments display the best behavior of all (Figure 6). Bakermans-Kranenburg and Van Ijzendoorn (2011) show how these patterns of differential sensitivity appear to be connected to dopamine-related genes.

[Insert Figure 6 here]

The example of high reactivity raises the unsettling possibility that treatment effects may be so heterogeneous that while most children may benefit from a given treatment, some children may be harmed by them. Vaccine developers face this dilemma – how frequent do adverse

reactions need to be before it is decided that a given vaccine is not in the public interest? Such decision making requires knowledge of the distribution of treatment effects and some risk function that weighs the beneficiaries against the individuals harmed by the intervention.

Family environments

Another source of within-stage variation in human capital treatment impacts are the predictable differences in the environments provided to children in different socioeconomic strata and cultural circumstances. Since children depend so heavily on their families, variation in family environments can enhance or retard development as well as the effectiveness of human capital interventions (Bronfenbrenner & Morris, 2006). Children who routinely experience activities offered by an intervention that match those that are provided at home as part of their normal lives are less likely to profit from them than children who would otherwise lack access to such experiences.

A case in point is the language environment of young children, which differs dramatically across socioeconomic lines. Parents with more education and income are more likely to support their children's academic success by providing rich language and literacy environments and engaging children in learning activities (Bradley & Corwyn, 2002). Hart and Risley (1995) tracked home language practices for children between ages one and three and found that on average, professional parents uttered 11 million words to their toddlers, as compared with 6 million for working class families and 3 million for welfare-recipient families. Higher SES parents also read more to their children (Phillips, 2011), use teaching strategies that mimic formal instructional techniques, such as asking questions and offering feedback rather than issuing directives (Laosa, 1983), and spend much more time with them in "novel" contexts – other than at home, at school, or in the care of another parent or a day care provider (Phillips, 2011).

The curricula of some early childhood education programs are designed to address some of these gaps by providing to low-SES children some of the academic supports found in higher SES families. The Infant Health and Development Program mentioned in the introduction was based on a coordinated educational curriculum of learning games and activities used both during home visits and at the program's early childhood education centers. It is hardly surprising that Figure 1 shows much stronger cognitive gains for children from low- than high-income families – the IHDP program was a better "fit" for the low-SES children because it provided the children with academic supports that were not otherwise available to them.

Seemingly inconsequential differences in family environment like the kinds of games children play may contribute to school readiness. Ramani and Siegler (2008) report that children from middle-income families are twice as likely as children in low-income families to report playing board games and less likely to report playing video games. These differences may well contribute to the large income-based gaps in math proficiency observed in kindergarten. Indeed, Ramani and Siegler find that low-income children who did report board play at home had higher math proficiency than their video-game-playing counterparts.

IV. Some implications

Our goal in this paper has been to articulate a set of principles for understanding why well-intentioned education-oriented intervention programs and policies, directed at children in

the infant-toddler, preschool, middle childhood, and adolescent stages of development, have the effects, non-effects and, in some cases, perverse effects that they do. We posit that developmental theory provides a useful framework for understanding major sources of treatment heterogeneity. The normative cognitive and socio-emotional development of all children and adolescents follows a predictable pattern in which windows for profitable interventions widen or narrow. Moreover, there are systematic within-stage variations associated with family, neighborhood and cultural factors and with individual differences in the timing or age at which children achieve important developmental milestones. Key to successful interventions, we believe, is the "fit" between a given policy intervention and the developmental stage and context of the child or youth participants. Table 1 summarizes some of our examples of interventions that fit developmental needs quite well or badly.

In this final section we apply our policy/child fit framework to two current issues in the intervention literature: i) Do "skills beget skills"? and ii) Is it true that "early is better" with respect to the timing of these interventions? We close with a discussion of how the structure of interventions might be improved to better account for heterogeneous treatment effects.

Fit and "skill begets skill"

Cunha et al. (2010) develop a model of the production of human capabilities that includes the cumulative role of cognitive and noncognitive skills, as well as skill investments made by families, preschool programs and schools, in producing adult human capital. Two features of their model drive their conclusion that "skill begets skill." First, they assume that higher-order skills are built up from lower-order ones (they call this "self-productivity"). Second, they assume that the productivity of any given human capital investment is greatest for children who bring the most skills into the intervention ("complementarity"). In terms of treatment effect heterogeneity, complementarity means that intervention impacts ought to be larger for children who start programs with higher as opposed to lower skills.

A great virtue of the Cunha et al. (2010) model is that it provides a rigorous framework for conceptualizing and estimating the roles that cognitive skills, noncognitive skills and human capital investments play in the production of human capital at different points across childhood, adolescence and early adulthood. They estimate the parameters of their model using data on children observed every two years, beginning at birth, in the National Longitudinal Survey of Youth. The survey provides measures of an assortment of cognitive (mostly reading and math achievement) measures, as well as noncognitive (mostly temperament and externalizing behavior problems) measures. Their human capital investment measures are limited to parental practices and include items such as reading to and praising children and making museum trips. Child care or schooling inputs were not measured well in the data and were not included in their analyses. Their results suggest that parental investments in children's cognitive skills are most productive early in childhood, while parental investments in their children's behavior is most productive in adolescence.

As suggested in our review of the developmental literature, abundant evidence from many developmental fields supports the hypothesis of self-productivity. Brain circuits wired sequentially are the most efficient (Knudsen, Heckman, Cameron & Shonkoff, 2006). Reading proficiency begins with letter recognition and phonemic skills such as knowledge of beginning and ending word sounds (National Research Council, 2008). Understanding the number line

undergirds addition and subtraction (Ramani and Siegler, 2008), while mastering fractions facilitates later proficiency in algebra (Siegler et al., forthcoming).

What about the idea that the higher the skills of the child or youth coming into a program, the greater the benefit from the program? This is certainly in accord with the IHDP-based evidence in Figure 1 showing that the more cognitively capable heavier babies profited more from the intervention than did the very low birth weight babies. But our view that treatment heterogeneity is a function of the "fit" between the given intervention and the developmental stage or individual circumstances of the treated children suggests many ways in which the complementarity hypothesis might not be confirmed. The IHDP intervention produced much greater gains for children from low- than high-income families, perhaps in part because the curriculum best matched the circumstances of children with neither very low nor particularly high skill levels.

Occam's Razor would value their simpler model and its sharper, more readily testable set of predictions. Based as they are on the fit between program characteristics and the multi-faceted nature of child and adolescent development, our more complicated models provide the sharpest predictions when the features of a given intervention are either well or ill matched to the salient features of the developmental stage or individual circumstances of the child. Examples are summarized in Table 1.

Is earlier better?

The economic case for the profitability of programs begun early as opposed to later in life faces significant hurdles. First, even with a modest 3% interest rate, the time value of money makes dollars invested at age 1 nearly twice as expensive as dollars invested at age 20. Second, the high staff-to-child ratios of programs directed at very young children lead to high costs relative to programs serving children only a few years older. Nevertheless, Knudsen et al. (2006) propose that the greater malleability of children's development early in life, coupled with the tendency for skill to beget skill, leads to a monotonic decline in the profitability of human capital investment with age. Does our developmental-based theory of fit between intervention and child yield the same prediction?

A neuro-developmental perspective reveals that while growth in brain functioning and very basic skills and behaviors is often astonishingly rapid in the first few years of life, other windows for skill development open up later in childhood and in adolescence (Howard-Jones et al., 2012). Growth in concrete academic skills is particularly rapid early in the school years, but only after children develop a degree of self-regulation that can sustain engagement in school-level curricula. Puberty ushers in a host of neurological changes that both increase risk of sensation-seeking activities and provide a degree of integrated, executive-type functioning that supports mastery of college-level knowledge.

Although many interventions directed at adolescents are unsuccessful, it is unclear whether they fail because adolescents are less able to change, or merely because they are less willing to change because the intervention designs are mismatched to their developmental needs. We have already reviewed evidence showing that even intensive, well-intentioned programs directed at deviant teens will fail if they bring those teens together in ways that provide them with opportunities for unsupervised interaction (Dodge et al., 2006).

The same may hold for disadvantaged youth in skill building interventions such as the

Job Training Partnership Act (JTPA; Orr et al., 1996). Male youth (age 16-22 and not enrolled in school) offered the chance of JTPA training earned less than their control-group counterparts; for female youth, the difference was positive but very small and statistically insignificant. Were youth training programs ineffective because youth were too old to respond to skill-building interventions? A simple age story is inconsistent with evidence that impacts on older groups offered JTPA training – adult women and men (age 22 and older) — were positive and statistically significant. In the case of adult women, the program impact amounted to an earnings gain of close to 10% (Orr et al., 1996, Exhibit 4.4). Program impacts were particularly large for single adult women with young children, which suggests that their motivation to provide for their families may have matched best to the JTPA training opportunities.

The classroom training offered by JTPA to disadvantaged youth may also have led to some of the kinds of iatrogenic effects detailed in Dodge et al. (2006). Mentoring models for youth such as Big Brothers Big Sisters, with one-to-one, sustained matches between adolescent and adult mentors, have proved more successful in random-assignment evaluations (Grossman & Rhodes, 2002), as have small high schools that emphasize close interactions between teachers and students (Bloom, Thompson & Unterman 2010). Thus, it may indeed be possible to develop successful human capital interventions for older children and youth, but only if those interventions are tailored to their developmental needs.

Tailoring interventions to children's individual circumstances

If the effectiveness of interventions depends on the match between program, developmental stage and individual differences, how best to allocate scarce human capital intervention resources to the children and youth who might enroll in them? Froelich (2008) and Smith and Staghøj (2009) outline five possible rules for assigning individuals to programs. First is what they term *profiling*, in which the programs are targeted to individuals whose outcomes are predicted to be the worst in the absence of program help. Examples include many training programs in the U.S. and Europe that enroll individuals predicted to be at greatest risk of long-term unemployment.

Second, programs could be *targeted* toward individuals predicted to profit the most from them. This can be quite different from profiling if, say, individuals at greatest risk of long-term unemployment share conditions (e.g., depression, very weak skills) that render them incapable of profiting from the intervention. Smith and Staghøj (2009) present evidence that at least in the case of training programs for the unemployed, program impacts are typically not largest for those at greatest risk of unemployment. Our theory of stage/policy match is an obvious example of targeting, in which program impacts are presumed to depend on the Vygotsky-esque (1978) match between a child's capabilities and the treatment.

Third, individuals or families can *self-select* into programs, as with universal but voluntary pre-kindergarten programs. In the case of New Hope, a Milwaukee-based work support program, parents were offered a menu of work supports that included an earnings supplement, health insurance, and child care subsidies, as well as a community-service job if needed. Few parents took advantage of all of the benefits all of the time. Instead, they matched benefits to their current needs and values (Gibson & Weisner, 2002). The benefits to children were apparent in the impact evaluation: teachers ranked children in New Hope families significant higher in both achievement and positive behavior (Duncan, Huston & Weisner,2007).

A fourth set of rules are *deterministic,* based, say, on demographic characteristics such as

age or formal education. Fifth, with *caseworker discretion*, enrollment decisions are left up to an administrator. The sixth is largely unsystematic *random* assignment.

The logical extreme of our argument is a one-to-one match between the circumstances of a given child and the chosen human capital intervention – the Oxford don. For inventions involving significant staff, one-to-one matches can easily become prohibitively expensive. Computer-based interventions have greater potential for one-to-one matching; achievement test designers long ago learned of the time and attention-saving virtues of tests that quickly gear questions to the ability levels of the test taker.

Computer-based instruction programs attempt to tailor instruction to the achievement levels of individual students. One thoughtful if not always effective example is *Cognitive Tutor*, a series of programs designed to bolster reading and math skills at both the elementary and the secondary school level. The program is based on a theory of learning that focuses on motivation and speeding up the retrieval of knowledge needed to accomplish the task at hand. To increase motivation, word problems are personalized to reflect individual students' interests and include the names of peers. And, to the extent possible, challenges are tailored to a student's achievement level. Rigorous evaluations show stronger evidence for gains in high school than in elementary school, but more general comparative studies of computer-based instruction programs often produce disappointing results (Campuzano, Dynarski, Agodini, & Rall, 2009)

Good teachers are well attuned to student differences and structure their classrooms and instruction accordingly. A vivid qualitative example is found in Tracy Kidder's book *Among Schoolchildren* (1989), which chronicles the school year of Christine Zajac, a fifth-grade teacher in the mixed-income Holyoke public school system.[4] On the first day of school, Chris let her students choose their desks within the large square arrangement she had set up. On the second day, her keen sense of observation led her to make some strategic adjustments that were designed to help struggling students and avoid some of the negative peer influences she had come to fear over the years:

> *She put four desks in the middle of the square, so that each of those four had space between it and any other desk. These were Chris's "middle person desks," where it was especially hard to hide...*
>
> *She had caught Courtney not paying attention several times yesterday. Courtney was small and doll-like, with the mobile rubbery face... "If school doesn't become important to her, and she doesn't do better at it, she'll have a boyfriend at fourteen and a baby at sixteen." Courtney got a middle person spot...*
>
> *She sent handsome, enthusiastic Felipe to a spot between Margaret and Alice. Felipe seem to be very talkative and excitable. He was probably used to being the center of attention. Chris guessed, "He's easily influenced by the people around him. If he sits between twits, he'll be a twit." Placed between two obviously well-mannered children, Felipe might be an asset....*
>
> *Several children seemed quick academically, especially Alice and Judith, a Puerto Rican girl with long, dark, curly hair and penetrating eyes....Chris moved Judith next to Alice. "Judith's exceptional, and I want Alice to get to know an Hispanic kid who's at her level." Maybe Judith and Alice would become friends* (Kidder, 1989, pp. 13-16).

Schools often adopt programs designed to tailor education to children's grouped circumstances. Special education, gifted and talented, and English Language Learner classes are nearly universal examples in U.S. elementary schools, while various forms of tracking (honors and Advanced Placement classes) are examples in high schools. At the same time, schools sometimes adopt practices such as combining grades that might be seen as reducing matches between instruction and student ability.

Implications for intervention design

A child/policy match perspective generates a number of important implications for the design of human capital policy evaluations. Most obviously, the more the distribution of treatment effects matters, the larger the sample size needed to detect effects for subgroups hypothesized to benefit differentially by the given treatment. And if one wants to test hypotheses that different treatments best fit the needs of different kinds of children or youth, then the sample sizes increase yet more, although theory may not require all combinations of treatments and subgroups. Increasing the scale of interventions in this way may create painful choices for intervention funders – is it better to test a smaller number of interventions at a scale that addresses important fit issues, or it is more important to estimate average treatment effects for a larger number of interventions?

Designing interventions so that their data can be pooled is one possible way of coping with some of these tradeoffs. The welfare-to-work program impacts on child achievement shown in Figure 2 were estimated from data pooled across four different work-support programs (Morris et al., 2005). Collectively, the power afforded by pooling proved sufficient to detect medium-sized impacts within age groups.

More generally, we believe that the idea of the match between a program and the developmental and individual circumstances of the children and youth they target is a very fruitful one. Because it involves the characteristics of both the children *and* the program, it is more complicated than proposition such as "early is better" or "skill begets skill," which focus only on the treated individuals. And it can risk *post hoc* situations in which a successful program is presumed to be a good match. When used well, knowledge of neuroscience and human development can generate testable *a priori* predictions and perhaps lead to more productive designs and implementations of skill-building human capital interventions.

Table 1: Examples of Matches and Mismatches of Intervention Programs Across and Within Developmental Stages

	Matches	**Mismatches**
Across Stage		
Early Childhood	Adoptions from orphanages in the first six months of life provide vital caregiver attachment experiences for infant	Orphanage adoptions after six months appear too late to completely remediate cognitive and socioemotional functioning
	Cochlear implants best promote speech production and vocabulary among younger children	Head Start impacts appear to fade out most quickly if coupled with low-quality primary school
	Board games appear well suited for number line development in 4 year olds	
Middle Childhood	Small class sizes appear particularly well-matched to student needs in kindergarten and first grade	Whole-class instruction and lack of emphasis on concepts appears ill-matched to 4th and 5th graders' needs and capabilities
	Work support programs boost achievement for children making the transition into school	Separate classes or other programming that congregate anti-social children may induce harmful iatrogenic effects
Adolescence	Interventions explicitly focused on promoting teacher-adolescent student interaction can boost achievement	K-5/Middle schools changes appear ill-matched to various needs in early adolescence
	Small high schools, if well-designed, and some targeted interventions can promote teacher/student interactions and thus meet adolescents' needs	Programs that provide group or class treatment of deviant youth can induce harmful iatrogenic effects
		Work support programs can burden adolescents with child care responsibilities that interfere with school work
Within stage		
	For children with difficult temperaments, high quality environments appear particularly effective	For children with difficult temperaments, low quality environments appear particularly harmful
	Intensive early childhood enrichment programs appear to remediate home environments of low- but not high-income children	Intensive early childhood enrichment programs might not be able to compensate for the cognitive deficiencies associated with very low birth weight

REFERENCES

Allen, J., Pianta, R., Gregory, A., Mikami, A., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. Science, 333: 1034-1037. doi: 10.1126/science.1207998

Bakermans-Kranenburg, M., & Van Ijzendoorn, M. (2011). Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. Development and Psychopathology, 23, 39–52. doi: 10.1017/S0954579410000635

Bloom, H. S., Thompson, S. L., & Unterman, R. (2010). Transforming the high school experience: How New York City's new small schools are boosting student achievement and graduation rates. Retrieved March 11, 2012, from http://www.mdrc.org/publications/560/full.pdf

Bloom, H. S., & Micholopolous, C. (2011). When is the story in the subgroups? Strategies for interpreting and reporting intervention effects for subgroups. Prevention Science. doi: 10.1007/s11121-010-0198-x.

Blyth, D., Simmons, R., & Carlton-Ford, S. (1983). The adjustment of early adolescents to school transitions. Journal of Early Adolescence, 3(1-2), 105-120. doi: 10.1177/027243168331008

Bradley, R. H., & Corwyn, R. (2002). Socioeconomic status and child development. Annual Review of Psychology, 53(1), 371-399. doi: 10.1146/annurev.psych.53.100901.135233

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). Effectiveness of reading and mathematics software products: Findings from two student cohorts (NCEE 2009-4041). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Caspi, A. & Shiner, R. (2006). Personality development. In N. Eisenberg (Ed.). Handbook of Child Psychology, Vol 3. Social Emotional, and Personality Development. Hoboken: Wiley and Sons.

Clements, D. & Sarama, J. (2011). Early childhood mathematics intervention. Science, 333, 968-970. doi: 10.1126/science.1204537

Connor, C. M., Craig, H. K. Raudenbush, S. W., Heavner, K., & Zwolan, T. A. (2006). The age at which young deaf children receive cochlear implants and their vocabulary and speech: Is there an added value for early implantation? Ear and Hearing, 27, 628-644. doi: 10.1097/01.aud.0000240640.59205.42

Cunha, F., Heckman, J.J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. Econometrica, 78(3), 883-931. doi: 10.3982/ECTA6551

DiPietro, J. A., Bornstein, M. H., Hahn, C. S., Costigan, K., & Achy-Brou, A. (2007). Fetal heart rate and variability: Stability and prediction in early childhood. Child Development, 78, 1788-1798. doi: 10.1111/j.1467-8624.2007.01099.x

Dishion, T., Dodge, K., & Lansford, J. (2006). Findings and recommendations: A blueprint to minimize deviant youth influence in youth interventions and programs. In K. Dodge, T.

Dishion, & J. Lansford, (Eds). Deviant peer influences in programs for youth: Problems and solutions (pp. 366-394). New York: Guilford Press.

Dishion, T., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. American Psychologist, 54(9), 755-764.

Dodge, K., Dishion, T., & Lansford, J. (2006). Deviant peer influences in programs for youth: Problems and solutions. New York: Guilford Press.

Duncan, G., Huston, A., & Weisner, T. (2007). Higher Ground: New Hope for the Working Poor and Their Children. New York: Russell Sage Foundation.

Duncan, G. & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills and behavior problems. Pp. 47-70 in G. Duncan & R. Murnane (Eds.), Whither opportunity: Rising inequality, schools, and children's life chances (pp. 47-69). New York: Russell Sage.

Duncan, G. & Sojourner, A. (2011). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps? University of California, Irvine working paper.

Eccles, J., Midgley, C., Wigfield, A., Buchanan, C., Reuman, D., Flanagan, C.,…Iver, D. M. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and families. American Psychologist, 48, 90-101. doi: 10.1037/0003-066X.48.2.90

Ellis, B. J., Boyce, T., Belsky, J., Bakermans-Kranenburg, C. M., & Van Ijzendoorn, M. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. Development and Psychopathology, 23, 7-28. doi: 10.1017/S0954579410000611

Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. Educational Evaluation and Policy Analysis, 21(2), 97-109. doi: 10.3102/01623737021002097

Freud, S. (1910). Three contributions to the sexual theory. Journal of Nervous and Mental Disorders. (Monograph Series No. 7). 1 - 91. (Translated by A. A. Brill).

Froelich, M. (2008). Statistical treatment choice: An application into active labour market programmes. Journal of the American Statistical Association, 103, 547-558.

Gennetian, L., Duncan, G., Knox, V., Vargas, W., Clark-Kauffman, E., & London, A. (2004). How welfare policies affect adolescents' school outcomes: A synthesis of evidence from experimental studies. Journal of Research on Adolescence, 14(4), 339-423. doi: 10.1111/j.1532-7795.2004.00080.x

Gibson, C., &. Weisner, T. S. (2002). Rational and ecocultural circumstances of program take-up among low-income working parents. Human Organization 61(2), 154-166.

Gross, R. T., Spiker, D., & Haynes, C. (1997). Helping low birth weight, premature babies: The Infant Health and Development Program. Stanford, CA: Stanford University Press.

Grossman, J., & Rhodes, J. E. (2002). The test of time: Predictors and effects of duration in youth mentoring relationships. American Journal of Community Psychology, 30(2), 199-219. doi: 10.1023/A:1014680827552

Hart, B. & Risley, T. (1995). Meaningful differences in the everyday experiences of young American children. Baltimore: Brookes.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. Child Development Perspectives, 2, 172-177. doi: 10.1111/j.1750-8606.2008.00061.x

Howard-Jones, P., Washbrook, E. & Meadows, S. (2012). The timing of educational investment: A neuroscientific perspective. Developmental Cognitive Neuroscience, 25, S18-S29.

Imbens, G. & Angrist, J. (1994). Identification and estimation of local average treatment effects. Econometrica, 62(2), 467-475.

Kidder, T. J. (1989). Among Schoolchildren. Boston: Houghton Mifflin Company.

Knudsen E., Heckman, J., Cameron, J., & Shonkoff, J. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. PNAS, 103, 10155-62. doi: 10.1073/pnas.0600888103

Laosa, L. (1983). School, occupation, culture and family. In E. Sigel & L. Laosa (Eds.), Changing families (pp.79-135). New York: Plenum Press.

Morris, P., Duncan, G. J., & Clark-Kauffman, E. (2005). Child well-being in an era of welfare reform : The sensitivity of transitions in development to policy change : Development psychology and public policy. Developmental psychology, 41(6), 919–932. doi: 10.1037/0012-1649.41.6.919

National Research Council (1998). Preventing Reading Difficulties in Young Children. Washington D.C.: National Academies Press.

Nelson, C. A., & Sheridan, M. A. (2011). Lessons from neuroscience research for understanding causal links between family and neighborhood characteristics and educational outcomes. Pp. 27-46 in G. Duncan & R. Murnane (Eds.), Whither opportunity: Rising inequality, schools, and children's life chances (pp. 27-46). New York: Russell Sage.

Orr, L., Bloom, H., Bell, S., Doolittle, F., Lin, W., & Cave, G. (1996). Does training for the disadvantaged work? Evidence from the National JTLA study. Washington, D.C.: Urban Institute Press.

Phillips, M. (2011). Parenting, time use and disparities in academic outcomes. In G. Duncan & R. Murnane (Eds.), Whither opportunity: Rising inequality, schools, and children's life chances (pp. 207-228). New York: Russell Sage.

Pianta, R. C., Belsky, J., Houts, R., Morrison, F., & NICHD Early Child Care Research Network. (2007). Opportunities to learn in America's elementary classrooms. Science, 315, 1795-1796. doi: 10.1126/science.1139719

Pluess, M., & Belsky, J. (2009). Differential susceptibility to rearing experience: The case of childcare. Journal of Child Psychology and Psychiatry, 50, 396–404. doi: 10.1111/j.1469-7610.2008.01992.x

Poulin, F., Dishion, T. J., & Burraston, B. (2001). Three-year iatrogenic effects associated with aggregating high-risk adolescents in cognitive-behavioral preventive interventions. Applied Developmental Science, 5, 214–224.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. Child Development, 79(2), 375 – 394. doi: 10.1111/j.1467-8624.2007.01131.x

Sarama, J., & Clements, D. (2009). Early childhood mathematics education research: Learning trajectories in young children. New York: NY Routledge.

Siegler, R., Duncan, G., Davis-Kean, P., Duckworth, K., Claessens, A., Engel, M…& Chen, M. (Forthcoming). Early predictors of high school mathematics achievement. Psychological Science.

Smith, J. & Staghøg, J. (2009). Using statistical treatment rules for assignment of individuals in labor market programs, Working Paper, Department of Economics, University of Michigan.

Steinberg, L., Vandell, D. L., & Bornstein, M. (2011). Development: Infancy through adolescence. Belmont CA: Wadsworth Cengage Learning.

Stevens, S., Sonuga-Barke, E., Kreppner, J., Beckett, C., Castle, J., Colvert, E., …& Rutter, M. (2008). Inattention/overactivity following early severe institutional deprivation: Presentation and associations. Journal of Abnormal Child Psychology 36, 385–398. doi: 10.1007/s10802-007-9185-5

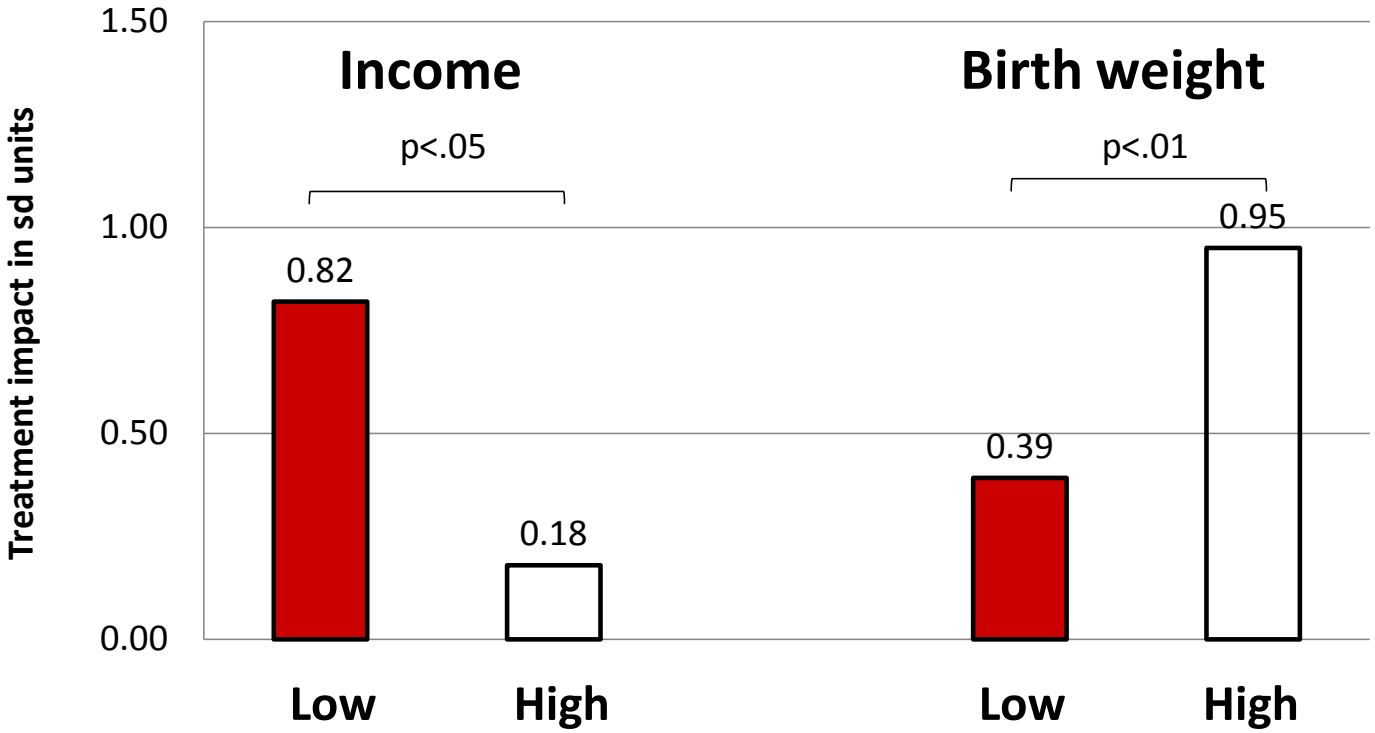
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge MA: Harvard University Press.

Weiss, C. C. & Kipnes, L. (2006). Reexamining middle school effects: A comparison of middle grades students in middle schools and K–8 schools. American Journal of Education,112, 239-272. doi: 10.1086/498996

Xie, Y., Brand, J & Jann, B. (2011). Estimating heterogeneous treatment effects with observational data. Working paper, Population Studies Center, University of Michigan.

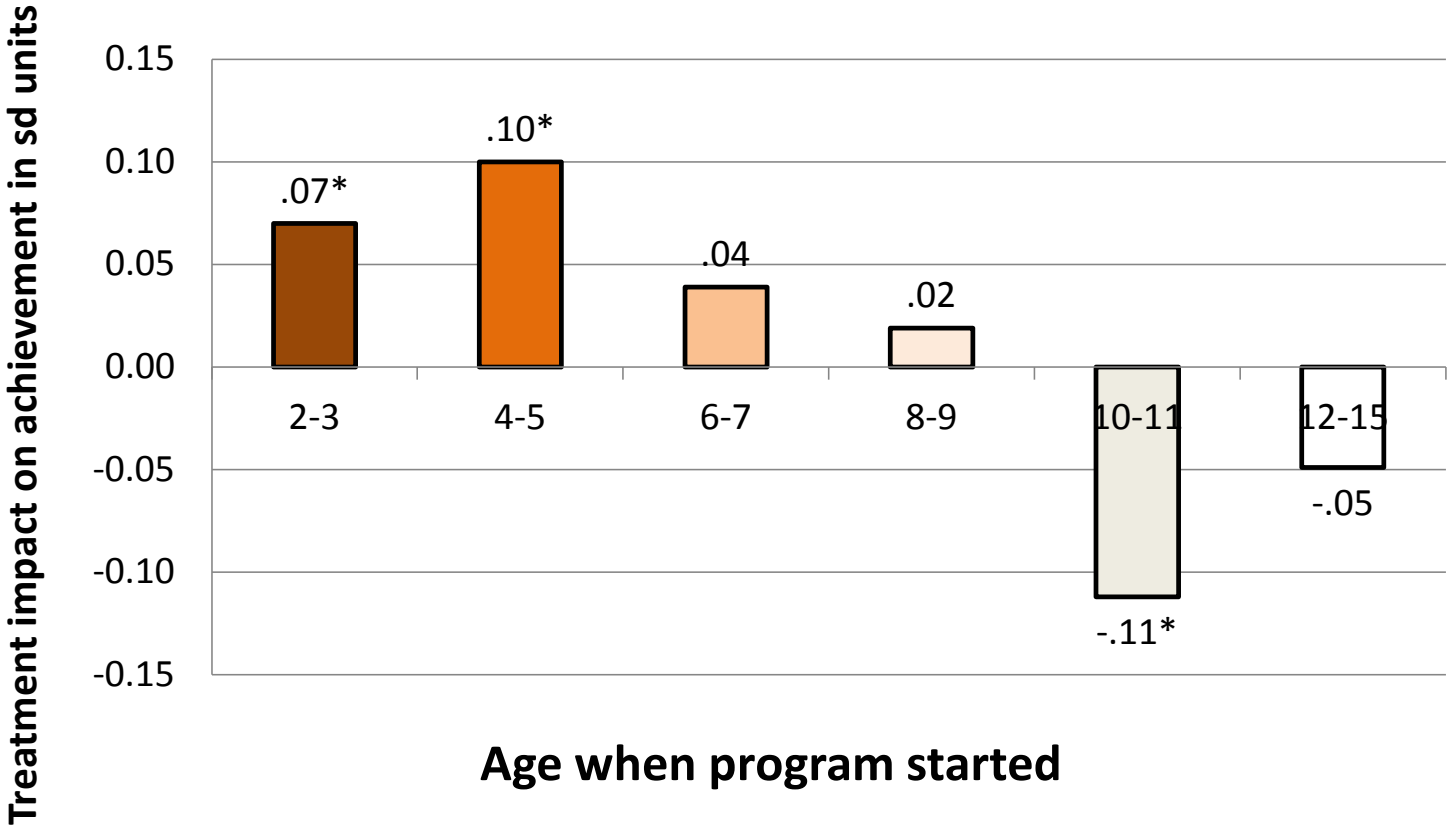
Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the Early Childhood Longitudinal Study: Kindergarten Class of 1998-1999. American Educational Research Journal, 41, 191-229. doi: 10.3102/00028312041001191

Figure 1: Impacts of the IHDP treatment on age-3 IQ by income and birth weight



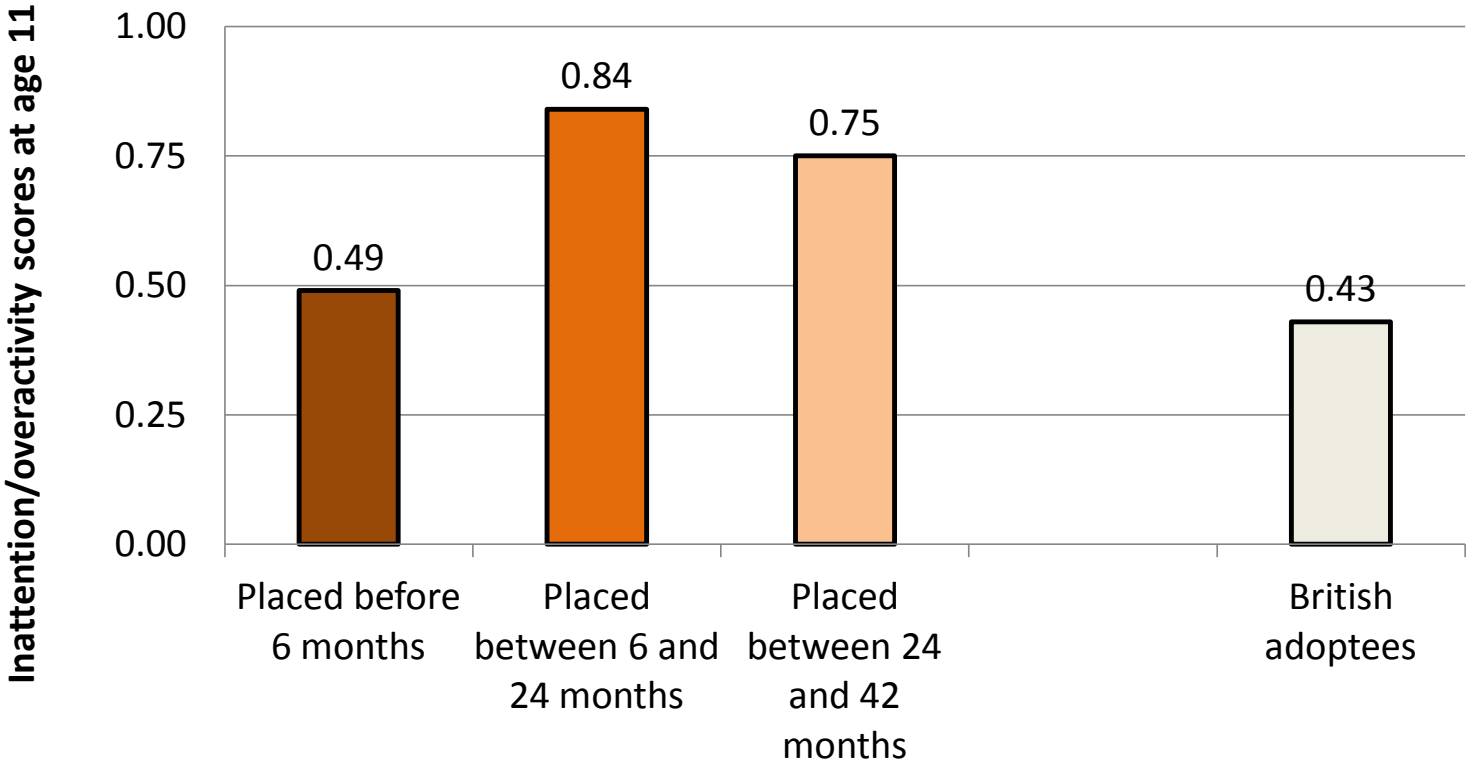
All models also condition on child gender, birth weight, gestational age at birth, neonatal health index and site indicators.

Figure 2: Impacts of Earnings Supplement Programs on School Achievement, by Age of Child



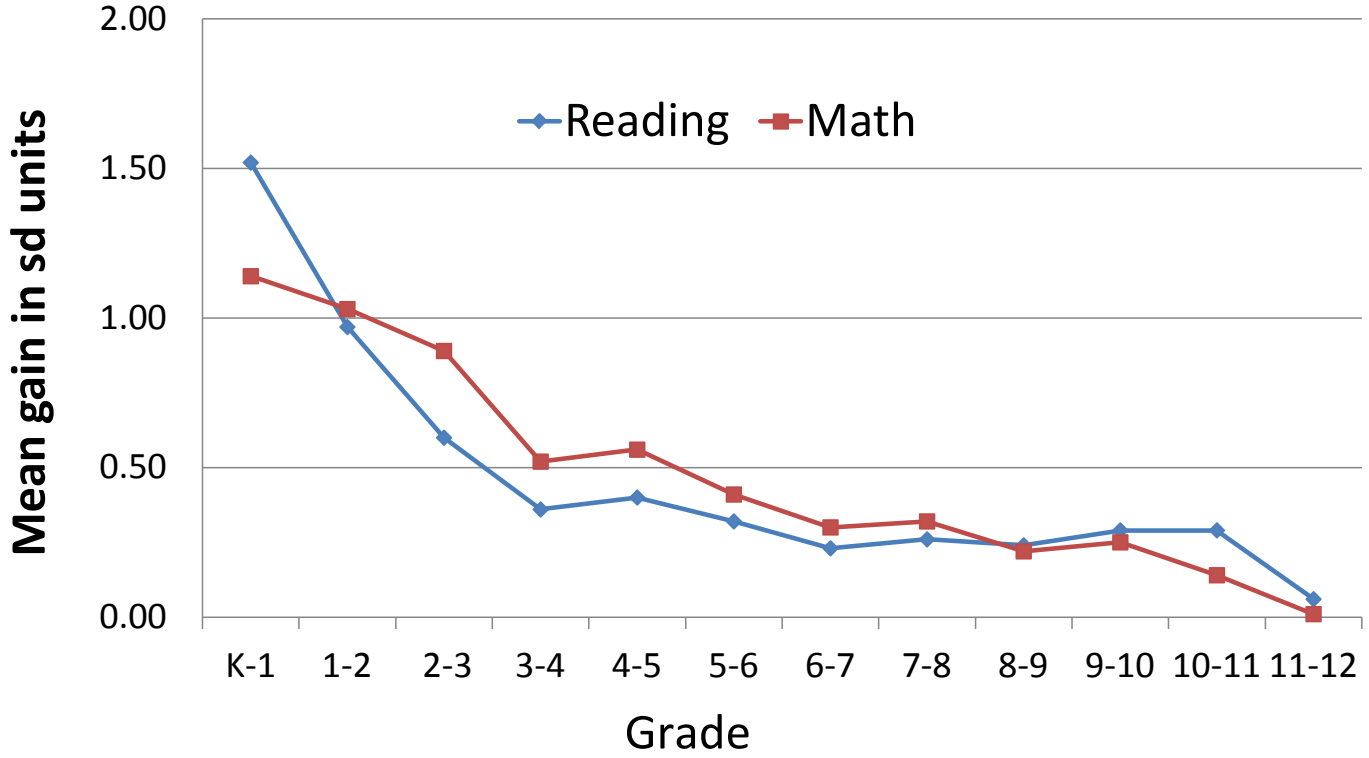
Note: * p<.05; Source: Morris et al. (2005).

Figure 3: Age 11 inattention/overactivity scores for adopted Romanian and British children



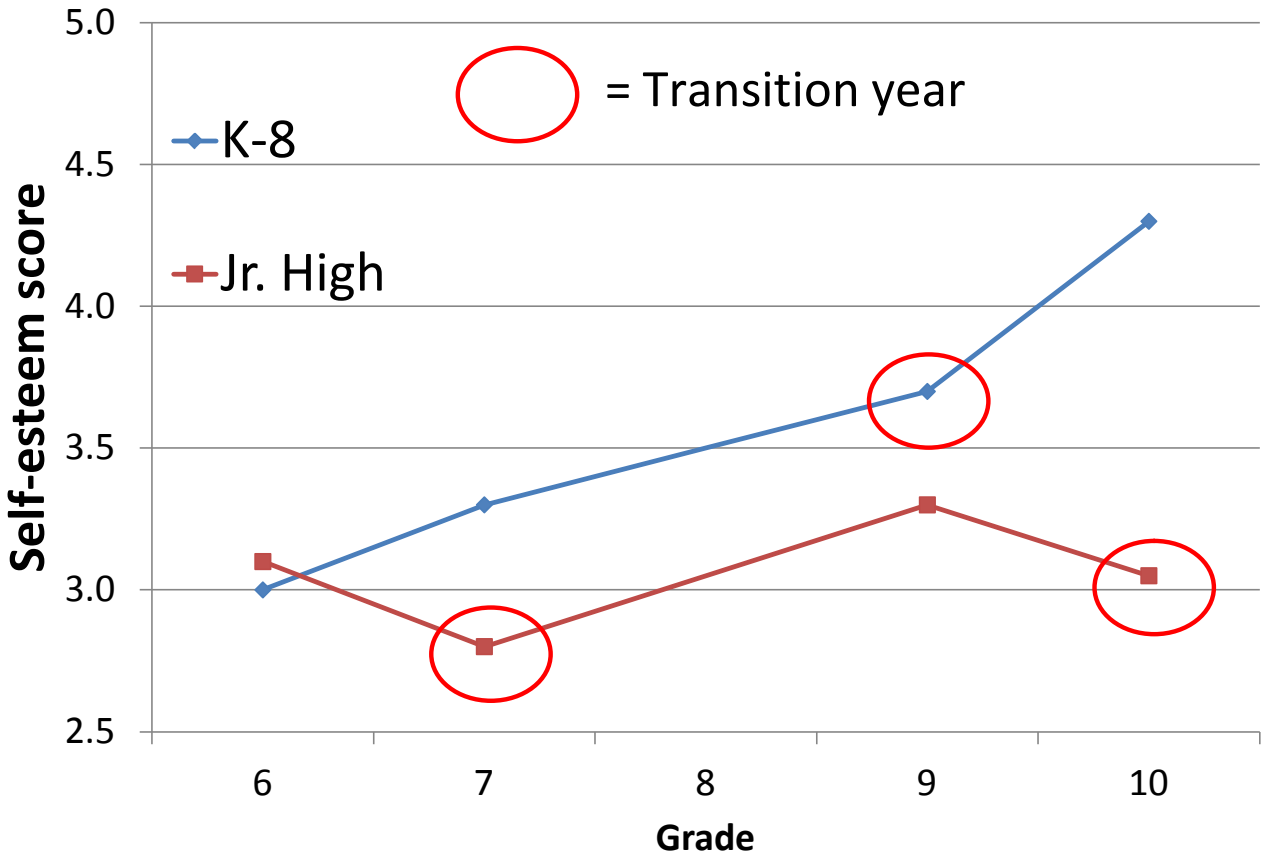
Note: Scores for "Placed before 6 months" group are statistically different from the other two placement groups., but not statically difference from the British adoptee group. Source: Stevens et al. (2008).

Figure 4: Average annual gain in math and reading achievement across grades



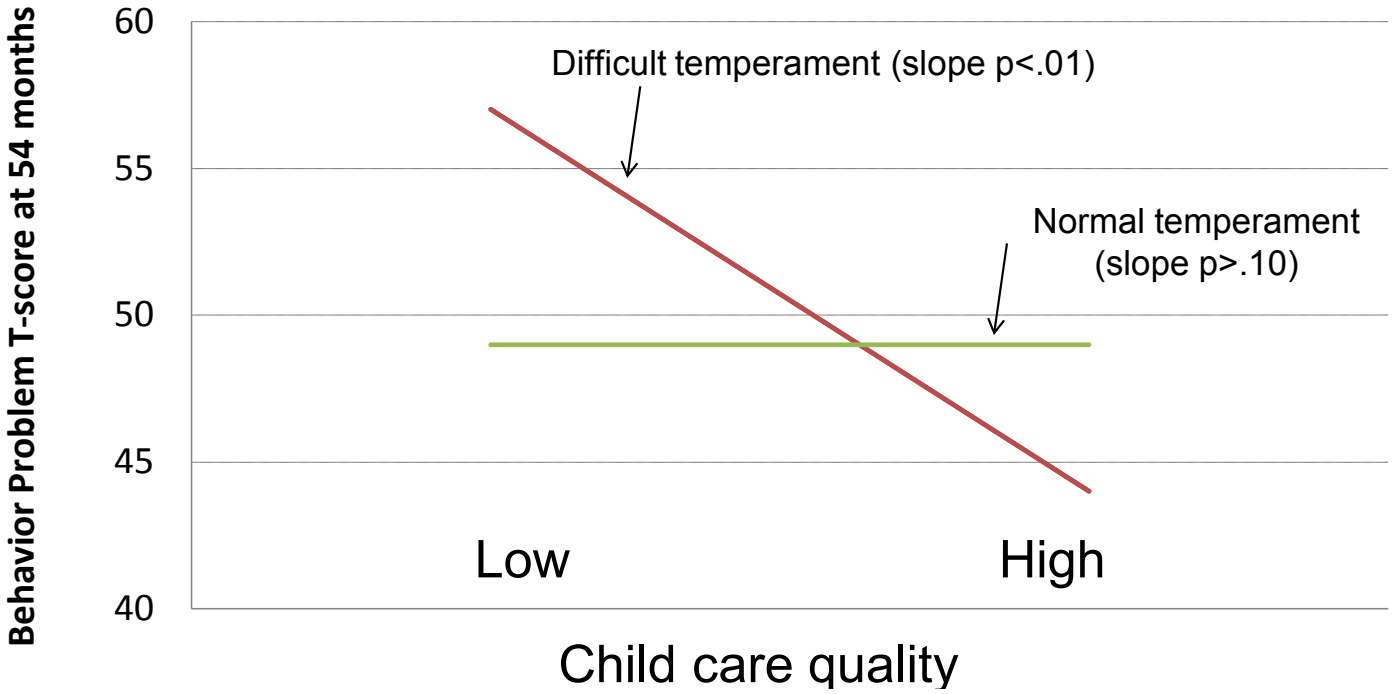
Source: Hill et al. (2008).

Figure 5: Female self-esteem by school structure



Source: Blyth et al. (1983)

Figure 6: Interaction of temperament and child care quality on behavior problems



Models controls for a number of child and parental characteristics, and child care type of hours. Source: Pluess and Belsky (2009), Figure 1.

ENDNOTES

[1] This estimate comes from Duncan and Sojourner (2011) and is based on weights designed to match the demographic characteristics of the IHDP sample to those of all U.S. births.

[2] A vivid example from the qualitative research concerned an adolescent whose mother needed to start work at six o'clock in the morning (Gennetian et al., 2004). The adolescent was responsible for taking her younger sibling to school in the morning, but doing that meant that the adolescent herself could not make it to her own school on time. Her no-excuses high school sent her to detention every day.

[3] Figure 3 is based on a meta-analysis of studies estimating gains in math and reading achievement between kindergarten and the end of high school.

[4] Kidder (1989).