Intelligence 53 (2015) 202–210

Contents lists available at ScienceDirect

Intelligence

The environment in raising early intelligence: A meta-analysis of the fadeout effect

John Protzko

University of California, Santa Barbara, United States

ARTICLE INFO

Article history:
Received 23 May 2015
Received in revised form 15 October 2015
Accepted 16 October 2015
Available online xxxx

Keywords:
Intelligence
Intervention
Cognitive development

ABSTRACT

Many theories about the role of the environment in raising IQ have been put forward. There has not been an equal effort, however, in experimentally testing these theories. In this paper, we test whether the role of the environment in raising IQ is bidirectional/reciprocal. We meta-analyze the evidence for the fadeout effect of IQ, determining whether interventions that raise IQ have sustained effects after they end. We analyze 7584 participants across 39 randomized controlled trials, using a mixed-effects analysis with growth curve modeling. We confirm that after an intervention raises intelligence the effects fade away. We further show this is because children in the experimental group lose their IQ advantage and not because those in the control groups catch up. These findings are inconsistent with a bidirectional/reciprocal model of interaction. We discuss explanations for the fadeout effect and posit a unidirectional–reactive model for the role of the environment in the development of intelligence.

1. Introduction

What role does the environment play in the development of early intelligence? Such a question has sparked the interest and ire of scientists for decades. The question has no easy answer and the methods used to solve it have been wide and varied. Here we investigate this question by quantitatively analyzing the existence of the fadeout effect—the finding that after an intervention raises the intelligence of children the effects appear to fade away once the intervention ends. The existence and details of the fadeout effect allow us to understand the causal role of the environment in the development of intelligence.

Among the main schools of thought over the role of the environment and the development of intelligence, we first focus on two. One is the little-to-no effect school which posits that the environment either has no effect on the development of intelligence, or that only a restricted environment can suppress intelligence (e.g. Scarr, 1992; Herrnstein & Murray, 1994). Regarding the fadeout effect, it seems contrary to such theories to initially admit an increase in intelligence that would then fade away. If the environment cannot improve intelligence, there is nothing there to fade. Showing early interventions can indeed raise IQ would be a first step to negating such theories. If the IQ gains were the result of teaching to the test or test familiarity, we would expect a fade to occur because the control group, who become more exposed and familiar with the IQ tests, catch up.

Theories of reciprocal interactions posit a dynamic interplay between the environment and a child's intelligence: intelligence feeds into the environments children are in which scaffold and help develop future intelligence (e.g. Gottlieb, 1983; van der Maas et al., 2006). One such model is the probabilistic epigenetic model (Gottlieb, 1983). This theory—as directed towards the environment and IQ—posits early neural systems in the prefrontal cortex (PFC) used in habituation are with the child at birth. Enriched environments strengthen these neural connections; these stronger connections further develop early executive function and self-regulatory abilities through PFC development. Enriched environments further strengthen PFC connections by providing supportive and stimulating environments; these increased PFC connections directly confer increases in fluid and general intelligence. This way an increase in supportive environments leads to reciprocal interactions, causing further gains in IQ directly and through accessing environments that further enhance neural and behavioral responses. Such models of reciprocal effects are able to account for a large amount of evidence concerning developing neural pathways in regions of interest for IQ, and can accommodate much of the evidence of the role of the environment in IQ (Blair, 2010).

The problem is such fully-reciprocal models would not accord with the fadeout effect. Children who participated in Head Start preschools, for example, left the program with higher IQs; by the end of first grade, they scored no higher than if they had not gone (Puma et al., 2010). Children who participated in the Perry Preschool Project had higher IQs at the intervention's end; the gains faded away six years later (Schweinhart & Weikart, 1980). We see similar fadeout effects of IQ for almost every intervention when researchers followed their participants.

Therefore, the existence and details of the fadeout effect are of great importance in testing the causal role of the environment in the raising of IQ. Previous investigations into the fadeout effect, however, have largely been qualitative and consisted of demonstrating how a few big-name studies failed to have permanent IQ gains (Herrnstein & Murray, 1994;

http://dx.doi.org/10.1016/j.intell.2015.10.006
0160-2896/

Howe, 1997). One concern with any such qualitative review of confirming studies is that the authors may only be selecting those studies that support their argument and ignoring those that run counter to it. Some have criticized previous investigations into the fadeout effect for cherry picking studies (e.g. Devlin, 1997) while others defended the results (e.g. Gottfredson, 1997). In addition, the fadeout effect is not as prominent for academic outcomes like grades, with the effects from many early interventions lasting (Barnett, 2011; but see Bailey et al., under review). Despite such lasting effects of many early interventions for academic achievement, IQ continues to fadeout (Darlington et al., 1980). An investigation into the lasting effects of Head Start showed that for most outcomes, "Fade-out is more apparent than real (except for IQ)" (Barnett, 2002, p. 2).

To investigate whether the fadeout effect for IQ is real and not a matter of cherry-picking big-name studies, we present this analysis of all of the early interventions that attempted to raise IQ and followed their participants afterwards. With this research, we examine the fadeout effect to see if interventions on average really do fade (for IQ). We identify specific elements of early interventions associated with larger effects and slower declines, using a longitudinal meta-analysis of randomized controlled trials.

2. Methods

2.1. Inclusion criteria and literature search

To be included in this study an intervention has to meet the following criteria: i) the participants are drawn from a general, nonclinical population; ii) the study employs an individual-level randomized controlled design; iii) the outcome variable is a widely accepted measure of IQ; iv) the intervention includes at least two IQ measurements after it ends; v) the intervention starts before the children enter kindergarten (is an early intervention). We include a study regardless of whether it is published. The reason for including only randomized controlled trials (RCTs) will be explained in more detail in the Discussion.

Each study comes from cross-referencing meta-analyses and reviews of early interventions (e.g. Jester & Guinagh, 1983; Herrnstein & Murray, 1994; Protzko, Aronson, & Blair, 2013) and also a search of the literature using Google Scholar and PSYCHInfo, using keywords such as ~random, IQ, cognitive. Every study that meets all of the requirements was then subject to exhaustive backward and forward searches.

We code all studies into effect sizes based on the post-intervention differences in IQ scores, using the sample standard deviations where available. In cases where no standard deviation (SD) data are available, we contact the study authors for the data. If the authors or the data remain unavailable, we impute the SDs using the value from the standardization sample (most commonly 15 or 16).

The purpose of this investigation is to test whether the increase in intelligence from a targeted intervention lasts or if recursive processes maintain or even increase the effects. To test these theories we aggregate all attempts to raise intelligence. While it could be useful to investigate whether type of intervention (nutritional, educational, training etc.) moderate the findings, there is not enough studies per category to allow for such an investigation (see Table 1 for all studies included). This meta-analysis does not just look at cognitive training studies, but any type of intervention (nutritional, educational, training, etc.) which has attempted to raise intelligence and followed the participants after the intervention ended. Only this way can the fadeout effect but put to experimental test.

2.2. Statistical tests

The distribution of long-term follow-up assessments on interventions is sporadic; with some studies followed for decades (e.g. Schweinhart, Barnes, Weikart, Barnett, & Epstein, 1993) and others for just one year or two after the intervention ends (e.g. Puma et al., 2010). The best way to analyze this data is using growth curve analysis with meta-analytic weights. This is referred to as a mixed-effects model in the meta-analysis literature (Hedges & Olkin, 1985). The intercept of each study is when an intervention ends and the time variable is years from the end of the intervention. Allowing both the slope and intercept to vary allows one to analyze what aspects of the interventions produce higher intercepts or different rates of decline.

The basic idea behind a meta-analysis is to use a weighted regression on a number of effect sizes (Card, 2011). In this way, larger studies have less error and produce more accurate estimates of a true effect size. Growth curve modeling is a longitudinal data analytic procedure where an average growth curve is fit to the trends of many different participants. One advantage of growth curve modeling is it is well suited to missing data and measurements taken at different times. All analyses are run in STATA version 13.1.

The full model for this analysis involves the following variables: delay from the end of the intervention, age at when the intervention began, duration of the intervention, and an interaction of age and duration with time to investigate different slopes. All weights were calculated using the following formula, consistent with meta-analytic procedures for investigating standardized mean-differences across studies (e.g. Hedges, 1981; Card, 2011):

$$w = \frac{1}{SE^2} = \frac{1}{\frac{n_E + n_C}{n_E n_C} + \frac{ES^2}{2(n_E + n_C)}}$$

There is a major theoretical issue when dealing with this data. The research question is specifically: Will the salutary effects of an early intervention persist or do they fadeout?" This is a question different from traditional meta-analyses which asks: *Is* a certain type of intervention effective?" So in this instance what should be done with ineffective experiments? If we are interested in asking 'are early interventions effective?' we should keep such studies. Removing ineffective interventions could ignore possible sleeper effects where the intervention is not effective at raising the IQ at first—but then the effects occur after a delay. A quick inspection of Fig. 1 indicates there is little reason to believe in such sleeper effects.

The question we are asking, however, is: do the IQ benefits of early interventions last? This implicitly assumes that the intervention worked in the first place. In the interest of transparency, we run two analyses: one with the full model and one with interventions only if their earliest effect size was greater than .2 (less than this indicates a small effect unlikely to be statistically significant; Cohen, 2009). Commonalities between both models will help converge on what may be happening to the participants after an early intervention ends. We start with a full model with all of the variables of interest; we then remove non-significant variables only if doing so improves model fit (examined through a likelihood-ratio test).

There are a number of studies that contribute multiple effect sizes. In order to consider this nesting, we first run the analysis clustering the errors by which study they come from. Under the all-in model this was not able to converge, as there was not enough variability to nest the errors. Including these studies could possibly introduce bias into the analysis. As such, we run both analyses with a binary variable for each study that contributes more than two effect sizes to the total.

One possibility suggested to us was to use a survival analysis instead of the meta-analytic growth curve modeling. Our major concern in using survival analysis is that such an analysis requires a binary event to mark the end of survival (e.g. death, attrition, relapse); we cannot identify any such event in the analysis of the fadeout effect. One could possibly use when an effect size reaches 0, but it is rare for any study to follow data through to 0 after it has already reached statistical nonsignificance. Alternately, one could use lack of statistical significance as the event; however statistical significance is deeply flawed (e.g.

Table 1
Studies used in the analysis of the fadeout effect of IQ.

Study		Test	Age at start	ES1	ES2	ES3	ES4	ES5	ES6	ES7
Pregnant Supplemen-tation (Helland et al., 2003; Helland et al., 2008)	LCPUFA supplementation	K-ABC	−.5[a]	.44 (3.5)	.25 (6.5)					
Pregnant supplementation (Harrell et al., 1955)	Ascorbic acid	Terman-Merril	−.42	.26 (3.5)	.4 (4.5)					
	B-complex	S-B	−.42	.51 (3.5)	.72 (4.5)					
	Thiamine		−.42	.36 (3.5)	.39 (4.5)					
Nurse visiting program (Olds, Henderson, Chamberlin, & Tatelbaum, 1986; Olds, Henderson, & Kitzmna, 1994)	*During Pregnancy only*	*CFIT, SB*	*−.35*	*−.05 (2)*	*.15 (3)*	*.16 (4)*				
	Through age 2		*−.35*	*.19 (2)*	*.11 (3)*	*.17 (4)*				
Abecedarian (Ramey & Campbell, 1991; Campbell & Ramey, 1994; Ramey et al., 2000; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002)	Intense Early Education + School	S-B, WISC-R, WAIS-R	0	.217 (8)	.403 (12)	.28 (15)	.407 (21)			
	Intense Early Education		0	.731 (5)	.598 (6.5)	.232 (8)	.525 (12)	.34 (15)	.279 (21)	
	School-age only		*5*	*−.297 (8)*	*−.12 (12)*	*−.14 (15)*	*−.089 (21)*			
Project CARE (Wasik, Ramey, Bryant, & Sparling, 1990)	Intense Early Education With Development Center	S-B	.083	1.1 (3.1)	.4 (4.1)					
	Intense Early Education Without Development Center		*.083*	*−.7 (3.1)*	*−.6 (4.1)*					
Parent Child Center (Bridgeman, Blumenthal, & Andrews, 1991)	Detroit Parent-Center Wave 1	S-B	.17	.88 (3)	1.14 (4)					
Early Head Start (Love et al., 2005; Vogel, Xue, Moiduddin, Carlson, & Kisker, 2010)	*Early Head Start Home Care*	*PPVT*	*.42*	*.09 (3)*	*.13 (10)*					
	Early Head Start Center Care		*.42*	*.09 (3)*	*−.03 (10)*					
	Early Head Start Home and Center Care		.42	.23 (3)	.02 (10)					
Ypsilanti IEP (Lambie, Bond, & Weikart, 1974; Epstein & Weikart, 1979)	Intense Early Education	S-B	.583	.206 (3.6)	−.01 (5.9)	.05 (6.9)				
Gordon PEP (Jester & Guinagh, 1983)	*Intense Early Education First Year only*	*S-B/WISC-R*	*1*	*.06 (4)*	*.57 (5)*	*.64 (6)*	*.2 (7)*	*.2 (11)*		
	Intense Early Education First 2 years in Program		1	.40 (4)	.78 (5)	.53 (6)	.85 (7)	.6 (11)		
	Intense Early Education first and third year in program only		*1*	*−.1 (3)*	*.13 (4)*	*.22 (5)*	*.12 (6)*	*.04 (10)*		
	Intense Early Education 3 years in program		1	.59 (4)	.57 (5)	.59 (6)	.59 (7)	.81 (11)		
	Intense Early Education second year only		*2*	*−.1 (3)*	*−.44 (4)*	*−.2 (5)*	*.12 (6)*	*.18 (10)*		
	Intense Early Education Last 2 years		2	.70 (4)	.75 (5)	.51 (6)	.59 (7)			
	Intense Early Education Last Year Only		3	.33 (4)	.43 (5)	.33 (6)	.51 (7)	.36 (11)		
Reading (Whitehurst et al., 1988)	Encour-aging Reading	PPVT	2.33	.58 (2.4)	.01 (3.2)					
Dialogic Reading (Huebner, 2000)	Dialogic Reading	PPVT	2.39	.27 (2.6)	0 (2.8)					
Preschool (Deutsch, 1971)	Wave 1	S-B	3	.81 (4.5)	.3 (7.5)					
	Wave 2		*3*	*.01 (4.5)*	*.04 (7.5)*					
	Wave 3		3	.77 (4.5)	.42 (7.5)					
Perry Curriculum Development Program (Weikart, Epstein, Schweinhart, & Bond, 1978a, Weikart, Bond, & McNeil, 1978b)[b]	Cognitive Program	S-B	3	.25 (4.6)	.27 (5.6)	−.14 (6.6)	.21 (7.6)	.32 (9.6)		
	Language Program		3	1.29 (4.6)	.43 (5.6)	.22 (6.6)	.35 (7.6)	.5 (9.6)		
Preschool (Weikart et al., 1978b; Schweinhart & Weikart, 1980)	Ypsilanti Perry Preschool Program	S-B	3	.98 (4.6)	.45 (5.6)	.42 (6.6)	.1 (7.6)	.08 (8.6)	.04 (9.6)	.13 (14.6)
Head Start Impact Study (Puma et al., 2010)	*3 year olds*	*PPVT*	*3*	*.07 (5)*	*.01 (6)*	*.09 (7)*				
	4 year olds		*4*	*.11 (5)*	*.06 (6)*	*.13 (7)*				
Dialogic Reading (Whitehurst, Arnold, Epstein, & Angell, 1994)	*at School*	*PPVT*	*3.46*	*.15 (3.6)*	*−.23 (4.1)*					
	at Home and School		3.46	.24 (3.6)	.01 (4.1)					
Preschool (Karnes, Zehrbach, & Teske, 1974)	Karnes Program	S-B	4	.48 (4.8)	.538 (5.8)	.269 (6.8)	.289 (7.8)	.163 (8.8)		
Early Training Project	Three year program	S-B	4	.81 (7)	.42 (8)	.21 (8.8)	.11 (10.9)			

Table 1 (*continued*)

Study		Test	Age at start	ES1	ES2	ES3	ES4	ES5	ES6	ES7
Summer School (Gray & Klaus, 1970)	Two year program		5	.86 (7)	.52 (8)	.51 (8.8)	.33 (10.9)			
Head Start (Abbott-Shim, Lambert, & McCarty, 2003)	Head Start	PPVT	4.58	.24 (5.3)	.32 (5.6)					
Cognitive Training (Rueda, Checa, & Combita, 2011)	*Effortful Control Training*	*K-BIT Matrices*	*5.39*	*−.17 (5.5)*	*.18 (5.7)*					

Note: Table ordered by age at when the intervention started, effect sizes and the age at follow-up in parentheses. Data in italics were dropped for the effective-only model. Ages are rounded for table fit. [a] = note that ages are in year format; negative numbers indicate that the intervention was started before birth (e.g. –.5 indicate that the intervention started 6 months before birth). One Study (Miller & Bizzell, 1983, 1984) was removed due to concerns about whether the control group was a control or comparison group. [b] = Age 15 & Age 23 data included participants in a wave of the study who were not randomly assigned to groups; all were placed in the experimental groups. Therefore, further datapoints are not analyzed here.

Cohen, 1994) and the purpose of meta-analysis is to bypass the use of significance testing within individual studies. In short, we are trying to understand the change in effect size over time; growth curve modeling is a superior way to do this in the context of this study.

The Results section is in two parts. The first part explains the different results of the two models (all-in, effective-only); the second part is the general results from this analysis. We find 23 studies that were fully randomized control trials that attempted to raise IQ and followed the subjects after the interventions ended. These 23 studies yield 39 different interventions averaging 3.3 follow-up observations across an average of 7584 participants. In the effective-only model, we use 17 studies yielding 26 effect sizes averaging 3.4 follow-up observations across 1758 participants (see Table 1).

Data analysis ran as follows: we started by fitting a model with the following theoretically-relevant variables: delay since end of intervention, delay^2 (to take into account the non-linearity of a fadeout), age when the intervention began, age interacted with delay, the duration of the intervention, duration interacted with delay. We then removed non-significant terms provided doing so resulted in an increase in model fit (as indicated by a non-significant change in likelihood-ratio test). Once we had a final model, we include random intercepts (if it increases model fit), then include random slopes (if it increases model fit).Then, since multiple studies contributed multiple effects, we test whether nesting the effects in studies improves model fit. Finally, we test whether allowing the intercept and slope to correlate improves model fit.

3. Results

Under the full data model, interventions were successfully able to increase intelligence by the time the intervention ended (ES = .368, 95%CI = .212 to .523). Over time, the salutary effects an early intervention declined to nothing (b = −.109/year,[i] 95%CI = −.199 to −.019; quadratic b = $9.86 * 10^{-6}$, 95%CI = $8.8 * 10^{-6}$ to $1.09 * 10^{-6}$). Interventions that started earlier in a child's life were no more effective than those which started later in a child's life, nor did they affect how long the effects lasted (both *ps* > .15) and as such were dropped from the model. Duration played no appreciable role in explaining the fadeout effect and was also removed from the model.

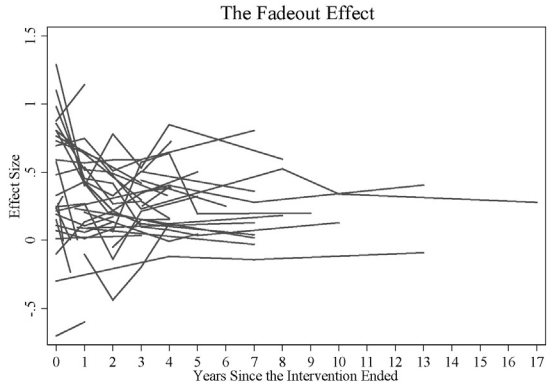

Fig. 1. Effect sizes over time after an intervention has ended, all studies.

Under the effective-only data model where we only include studies that would have reasonably raised IQ, interventions were successfully able to increase intelligence by the time the intervention ended (ES = .523, 95%CI = .451 to .666). Over time, the salutary effects an early intervention declined to nothing (b = −.132/year, 95%CI = −.243 to −.021; quadratic b = $1.27 * 10^{-5}$, 95%CI = $1.1 * 10^{-5}$ to $1.45 * 10^{-5}$).[ii,iii] Age when the intervention started and duration of the intervention played no appreciable role in explaining the fadeout effect and were thus removed from the model (see Fig. 3).

3.1. General results

3.1.1. Declining effect sizes

Under both models, we see a significant decline in the magnitude of effect sizes. Therefore, we believe it safe to say that the fadeout effect for IQ is a real phenomenon and intervention effects will decline to zero after the intervention ends.

3.1.2. Age

Under the all-in model, interventions that started earlier in a child's life were actually less effective than interventions that started later in a child's life. Studies which started later, however, did fade faster than those which started later. This later is better effect was not seen in the effective-only data. This is important because some interventions may have trained children on individual skills that were only age

[i] All coefficients involving time are presented multiplied by a constant to express them in change over years

[ii] At one reviewer's request, we ran a supplementary analysis with imputing the median SD (11.762) instead of the SD from the population (15 or 16). This required imputing the median SD into the following four studies where the standardization sample was imputed for the purpose of calculating effect size (Gray & Klaus, 1970; Karnes et al., 1974; Lambie et al., 1974; and Abbott-Shim et al., 2003). Results of this model were the same and all coefficients within the confidence intervals of the previous model: constant = .609, b_{delay} = −.136/year, $b^2_{delay} = 1.17 \times 10^{-5}$, all *ps* < .021.

[iii] At another reviewer's request, we ran a supplementary analysis with imputing the population SD (15 or 16, depending on the test) instead of the SD from the sample. This required imputing 15 or 16 as the pooled standard deviations for all studies. This altered the magnitude of effect sizes and thus, the weights. Results of this model were the same as the previous model, although the intercept was smaller than that in the original model: constant = .407, b_{delay} = −.133/year, $b^2_{delay} = 8.25 \times 10^{-6}$; all *ps* < .016. This reduced intercept is to be expected as a larger SD would produce a smaller effect size with no correction to the difference between experimental and control groups.

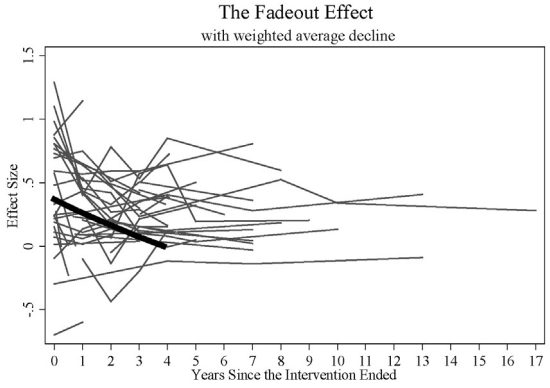

Fig. 2. Average declines to zero in all interventions.

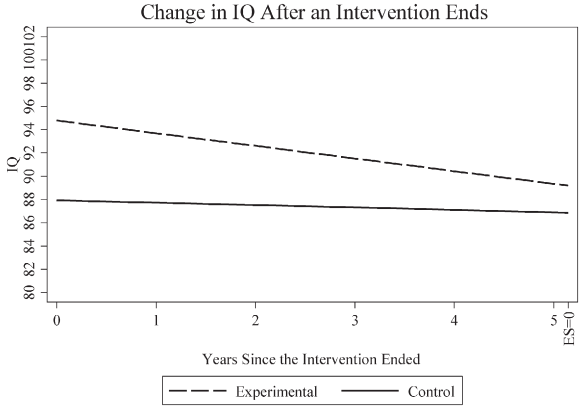

Fig. 4. IQ scores decline after an intervention ends.

appropriate. Since age does not play a role in the fadeout of effective interventions, however, we do not believe such age-appropriate training is a cause of the fadeout effect (see also the sections on 'Lost knowledge' and 'The changing *g* hypothesis' in the Discussion).

To our knowledge, however, one reciprocal model of IQ and the environment does posit that early environmental effects would be shorter-lived than environmental effects later in life (Dickens, Turkheimer, & Beam, 2011). What is apparent is that there is no evidence for the notion that earlier interventions are more effective than later interventions. This mirrors some of the experimental work on sending children to preschool, sending a child when they are younger is of no greater benefit to IQ than sending a child to preschool when they are a little older (Behrman, Cheng, & Todd, 2004; Protzko et al., 2013).

3.1.3. Duration

Duration of the intervention had no significant effect on either the magnitude of effect sizes or on how long the effects lasted.

We can now ask how the salutary effects fade by examining the pattern of IQs of the experimental and control groups. There are two ways we can see the effects disappear. Either the control group would catch up and we would see an increase in their scores over time, or the experimental group's scores would decline. If the gains fade because the control group catches up, we would have to posit a new interpretation of the fadeout effect that does not include a loss of ability. If the experimental group declines relative to the control group, we would have evidence that the fadeout occurs because of a genuine loss of added ability in the experimental group.

We approach this test using the same studies and methods for the analysis of the fadeout effect. First off, the scores from the control group stayed stable over time ($b = -.206$, 95%CI $= -1.149$ to .737). The experimental group, however, exhibited a steep decline over time ($b = -1.09$, 95%CI $= -2.149$ to $-.026$). This means that the fadeout effect occurs because children in the experimental group lose the IQ gains they received (see Fig. 4).

We can now quantitatively affirm that the fadeout effect is real. We verify that the fadeout happens because children in the experimental groups lose the IQ gains they acquired. The effects of early interventions fade because children lose the IQ gain the intervention gave them—not because the control group caught up (Fig. 4). It would be troubling if the only reason scores increased across these interventions were because children were getting used to the IQ tests or testing situation. If this practice effect were the operator, we would see the control group's IQ rise as they became more practiced with IQ testing. That does not happen. We can therefore reject practice effects as the cause of the fadeout effect.

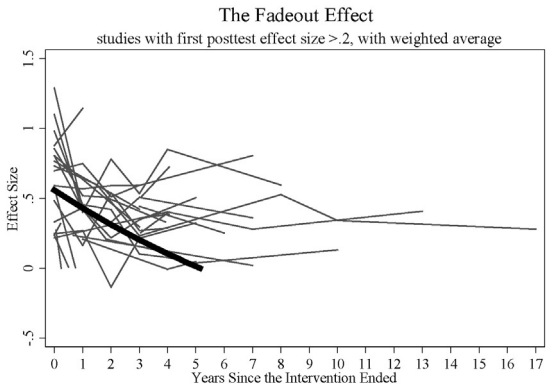

Fig. 3. Average decline to zero in effective interventions.

4. Discussion

Previous investigations into the fadeout effect have been few and qualitative in method (Herrnstein & Murray, 1994; Howe, 1997). As a reminder, the conclusions we draw are based on the experimental literature from attempts to raise IQ. This analysis provides a firm quantitative framework for examining the fadeout effect. Here we see that interventions that raise the IQ of young children decline after the intervention ends.

4.1. Previous theories of the fadeout effect

Before presenting our own theory of the fadeout effect, we briefly address previous theories, showing that most of them do not adequately explain the data. We present these arguments as the proponents of such views hold them, not as our own beliefs.

4.1.1. Raising IQ/raising g

The fading of IQ gains could occur because the gains were, in the first place, never real. They were illusory or instances of teaching to the test. In this view, an IQ test reflects two elements: a latent ability *g* which is responsible for performance, and extraneous variance that contributes to item- or subtest-level performance. In effect, the interventions did not affect the underlying latent variable, and thus the scores were doomed to fade (Howe, 1997, p. 60). Others have noted, because the gains fade, the results from such interventions are not appreciable or real (Jensen, 1985; te Nijenhuis, Jongeneel-Grimen, & Kirkegaard, 2014).

This theory does not adequately explain the findings here for multiple reasons. First, and most crucially, it offers no explanation of why *g*-effects *should be lasting*. Second, it offers no explanation why such non-*g* effects would behave differently than the *g* effects. Third, evidence from an early intervention of over 1000 low-birthweight infants has been shown to raise *g* when the intervention ended and yet

still exhibit the fadeout effect (Protzko, under review). Fourth, an extension of this theory could see increases in non-*g* variance as testing effects and exposure to the IQ tests and material (te Nijenhuis, van Vianen, & van der Flier, 2007). Were this the case, we would expect the fadeout effect to occur because the children in the control group would see rising scores over time as they gain familiarity with and exposure to IQ tests. No such pattern exists (see Fig. 4). This also stands as evidence that the effects of these interventions are not instances of test sophistication, as that would show a pattern of control groups "catching up" to the experimental groups. This pattern is also predicted by other models of intellectual development (Dickens & Flynn, 2001). Therefore, for the multiple reasons above, we have good reason to believe these instances are not teaching to the test or instances of effects on test scores and not *g*. In addition, such an argument has difficulty explaining the effects of physiological interventions on IQ (such as supplementation of healthy mothers (Helland, Smith, Saarem, Saugstad, & Drevon, 2003; Helland et al., 2008)). It has even more difficulty explaining the fadeout of such interventions as well.

4.1.2. Lost knowledge

One way theorists have approached the fadeout effect is to deny the validity of IQ tests and claim that performance on such tests simply reflects acquired knowledge with no underlying ability. Under this view, the fadeout effect is seen simply as a loss of memory for the knowledge directly taught during these interventions (Zigler & Seitz, 1982; Howe, 1997, p. 53). This view of intelligence and IQ tests has long been refuted (see Jensen, 1998 for a historical review). In addition, it cannot explain instances of raising IQ from interventions such as double-blind supplementation in infants (Helland et al., 2003, 2008). Also, it cannot account for instances where the fadeout effect does not occur for academic achievement (see Zigler & Seitz, 1982; Barnett, 2002), subject matter that *is* learned knowledge and skills.

4.1.3. Back into poverty

Another proposed reason for the fadeout is that children from many of these interventions come from substandard environments. Children who attend Head Start preschools end up in middle schools in poorer neighborhoods that have lower levels of academic achievement (Lee & Loeb, 1995; Currie & Thomas, 2000). Therefore, the reason that the fadeout effect occurs is because they are unable to use their newfound intelligence, as the environment offers no scaffolding or opportunities (Zigler & Seitz, 1982). While the majority of studies included in this meta-analysis use children that come from poverty, critically, the work on LCPUFA supplementation (Helland et al., 2003, 2008) did not specifically use children from poverty yet still experienced the fadeout effect.

4.1.4. The changing g hypothesis

As suggested by one reviewer, the fadeout effect occurs because the nature of intelligence changes. Therefore, there is a different intelligence and the effects *should not* carry over. This hypothesis has a number of empirical and philosophical problems which render it an unlikely source of the results demonstrated here.

The first problem with the changing *g* hypothesis is that it must somehow account for the high correlation of intelligence at different ages (e.g. Jensen, 1998). The correlation of intelligence, even in childhood, is often over .7 from year to year; this is true when compared at the manifest level of IQ scores (Jensen, 1998) or at the hierarchical level using latent variables (e.g. Žebec, Demetriou, & Kotrla-Topić, 2015; Protzko under review). If the latent *g*, or intelligence, is constantly in flux, how do we propose to explain such strong stability? The most common answer is that earlier processes are causally relevant for later processes (Demetriou, Spanoudis, & Shayer, 2014). That is, earlier intelligence or processes involved in intelligence cause later intelligence or similar processes; hence, the correlation. This, however, could not be used to explain away the fadeout effect. In fact, this causal interpretation would predict that the decline effect should *not* happen, as increases to intelligence at time 1 would cause higher intelligence at time 2. Either this causal account must be kept and the results presented here act as evidence *against* such a causal account, or the causal account must be adjusted with some other explanation for the stability despite the changing nature of intelligence must be put forward.

Second, if the nature of intelligence changed over development, thus leading to a washing out of the intelligence gains, one would expect a faster decline than is observed. Though present, the loss of ability in the experimental group is gradual and takes place over years. Were the nature of intelligence to change (ignoring the previous concern of causality between ages or developmental processes), we would expect a rapid loss of intelligence in the experimental groups, not the gradual decline we instead observe.

Third, the changing *g* hypothesis must explain how intelligence deeply changes and yet interventions that occur in the womb or before first year of life is over can still have effects four (Harrell, Woodyard, & Gates, 1955) and five years later (Helland et al., 2003). It would seem that either the nature of intelligence would have to be stable from birth until at least these ages and then start changing, or the changing *g* hypothesis is unlikely the reason for the fading. Since the stability of IQ increases throughout development (e.g. Jensen, 1998), it could be argued that the nature of intelligence is less stable from birth until at least these ages. This further suggests that the changing *g* hypothesis is not driving the fadeout effect.

Though the empirical problems may not be insurmountable for future research and theorists, the largest problem of the changing *g* hypothesis is the operationalism that begins to creep in. Even in developmental work on the nature of intelligence, a unitary *g* fits developmental models better than a new *g* at each age (Žebec et al., 2015). Positing a different latent intelligence at each age brings with it the problem of interpretability. Suppose a researcher wishes to see if a given environmental intervention raises intelligence in six-year-olds. At the end of the experiment, what can the researcher conclude? Under the changing *g* hypothesis, the researcher could not conclude that they have raised intelligence, merely that they raised the current manifestation of intelligence at age six (also note that 'age 6' is only an average of the ages of participants in such a study). If the *construct* changes between ages, then we cannot have a scientific or theoretical dialog since no two researchers would be discussing the same *construct*. There is no problem with the manifestation of the construct being different, but then one must contend with the stability of the construct. One way this hypothesis could be salvaged given the empirical and philosophical issues would be to posit that all of the interventions here did not have an effect on the core of intelligence or the part of intelligence that is causally efficacious for later intelligence. This line of argument has already been dealt with in the Raising IQ/raising *g* section.

4.1.5. Genetic set point

One suggested possibility has been that humans have a genetic set point for their intelligence. Interventions may be able to move intelligence up or down, but they will always return to their set point. This is indeed a theoretical possibility, especially given that the majority of variance in intelligence is accounted by genetic differences (e.g. Bouchard & McGue, 1981). The evidence here, however, may not accord with such a hypothesis. First, a genetic set-point (in our opinion) would predict a much faster "return" to the set point than the fadeout seen here occurring over years. Second, and most importantly, genes become more responsible for variation in intelligence across development, with small effects at the earliest ages and increasing effects over development (e.g. Briley & Tucker-Drob, 2013). This would suggest we should see an age effect on rate of fading, specifically, interventions which began earlier should fade at a different rate than those which began later in a child's life. As we see no such pattern, a genetic set point is unlikely to be the reason for the effects found here.

4.1.6. Selective attrition

The idea behind selective attrition is that the fadeout effect could occur because of a methodological flaw in the data points—participant dropout. For this to be the case to match the data here, children who were of *higher* intelligence would have to selectively drop out of the study *disproportionately in the experimental groups*. This is unlikely. In addition, in longitudinal panel studies, individuals of *lower* intelligence are more likely to drop out of successive waves, not those of higher ability (see Salthouse, 2013, for a brief review). Therefore, if selective attrition was happening within each study, the continued presence of the fadeout effect would be robust.

4.1.7. Publication bias

Another possibility is that the results here are the results of publication bias. We agree that publication bias undoubtedly exists—though we argue any such bias would be *against* the fadeout effect. Interventions that had no effect on intelligence are less likely to be subject to follow-up with IQ as a dependent variable. Therefore, they would not be identified for inclusion here. This may bias the full-data results but would likely have no effect on the effective-only analysis. Additionally, among studies that do find a significant effect at immediate posttest, if they do not find an effect on IQ at follow-up, they would be less likely to publish that null finding. Studies continuing to find a significant effect may be over-identified here, thus making the continued existence of the fadeout effect robust.

5. The current theory

What does all of this tell us about the role of the environment in raising IQ, and how do we understand the implications and explanation of the fadeout effect? First, this analysis shows that the relationship between the environment and IQ is unlikely to be fully-reciprocal, or at least any reciprocal interactions are not as strong as theorized from longitudinal investigations. A direct implication of a fully-reciprocal interaction is that any change made at any part to the system creates a lasting change—regardless of the point of entry (see Fig. 5).

For a relationship to be fully-reciprocal, both sides of the equation must hold. On the left side of the figure, an increase in IQ causes an increase in environmental demands (E) which leads to further gains in IQ and so on. On the right, an increase in E causes an increase in IQ that causes a further increase in E and so on. We see from our results that it is unlikely that either side of the equation withstands falsification. This does not mean, however, that reciprocal interactions cannot be occurring between intelligence and environments. It simply suggests that the strength of these interactions is much weaker than a driving force in intellectual development.

The most telling example is as follows: Supplementing fetuses in a double-blind way with long-chain polyunsaturated fatty acids causes them to have higher IQs by the time they are about four years old (Helland et al., 2003). Keeping in mind Fig. 5, this study creates a change to the IQ of the children with no concomitant change to the environment (because it was double blind). By the time the children were seven, however, the effects had completely faded away (Helland et al., 2008). This negates the likelihood that the interaction of IQ and the environment is fully-reciprocal. Those children, with their higher IQs, did not create or seek out environments that were more demanding *which then fostered further IQ gains*. Full reciprocity breaks down. We argue that the fadeout effect falsifies such a model of intellectual development. Therefore, it appears that the relationship between the environment and IQ is more likely to be unidirectional on the part of the environment, since increases in IQ do not confer further IQ-increasing environmental demands.

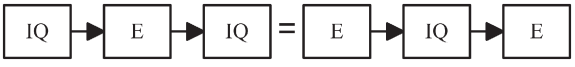

Fig. 5. Implication of a fully-reciprocal interaction between IQ and environmental effects (E).

One reviewer suggested, however, that the infants could have had an effective IQ that was immense at the end intervention; it was simply untestable because the children were infants. While possible, this is untestable. In addition, if these infants, as they aged, sought out new and more cognitively demanding environments, the fadeout effect should not occur unless either: a) the environments they sought out were unable to support or further raise intelligence; b) the increase in intelligence was not enough to *alter* selection or self-selection effects in the environment; or c) increased intelligence is not a causal factor in self-selection into new environments and thus, the children did not seek out these new environments.

These possibilities, that children either do not seek out scaffolding-enough environments or simply do not seek them out, are all compatible with the fadeout effect. There is not enough evidence in other literature to figure out which may be occurring. We think the second option, the increase in intelligence is not large enough to alter selection, is unsupported by the data here. As the analysis is a multilevel model, it allows examination of an intercept*slope interaction. This means we can test whether the magnitude of effect at the end of the programs is related to the rate of decline. It may be that the intelligence gains on average are not strong enough to alter selection effects of the environment. Therefore, we would expect to see that interventions that had larger effects at the programs' end had a longer effect on IQ. This was not the case ($r = -.298$, 95%CI $= -.613$ to .099). Therefore, even interventions with larger effects faded at the same rate as those with smaller effects (in the effective-only analysis). This plausibly leaves that either the newly-more-intelligent children are not selecting new environments, or they are selecting into new environments that are not strong enough to scaffold their new intelligence. Both possibilities can be made to be consistent with the current state of the literature.

Not selecting/being selected into new environments: The fact that children who were made more intelligent should select into new environments is based on the (untested) assumption that intelligence is causally relevant for selection into demanding environments. The closest data available show that traits such as need for cognition are barely correlated with intelligence (e.g. Fleischhauer et al., 2010), yet could also predict selection into more demanding environments. In addition, we know that many other traits are important for academic achievement, such as grit and self-regulation (e.g. Duckworth & Seligman, 2005). It is therefore possible that people select more cognitively demanding environments for reasons other than their intelligence and that intelligence is not a causal factor in those decisions.

On the other hand, assuming intelligence is causally relevant for selection into more cognitively demanding environments, it could be possible that children are selecting into new environments, it is just that the environments are not cognitively demanding *enough* to scaffold their newfound intelligence. Evidence for this comes from causal estimates on the effects of gifted education. Gifted and Talented programs in the United States would count as 'more cognitively demanding' programs over traditional public education. Entry into these programs is often determined via a cutoff score on an academic achievement test. It is therefore possible to test students a couple of points above or below the cutoff, students who are ostensibly of the same intelligence where one was accepted and the other rejected. This allows a test of the causal effect of such programs on academic outcomes (unfortunately, no work on the effects on IQ exists yet). This regression-discontinuity design shows little to no effect of gifted education in Western countries on SAT scores (Dobbie & Fryer, 2011; Abdulkadiroglu et al., 2011) or academic achievement (Bui et al., 2011), despite positive effects on graduation rates and exposure to better performing teachers and peers. This could suggest that the range of new environments, even in the Western extreme of suddenly being placed in a gifted program or school, is not strong enough to scaffold the newfound increase in intelligence. We leave it to future work to explicitly test these two possibilities.

We argue that the fadeout effect is incompatible with reciprocal models of the environment and intelligence. Instead, changes in the environment can increase intelligence, yet no reciprocal effect occurs. We term this pattern of interaction a *unidirectional–reactive model*. Such a model posits that the trait (IQ) reacts to the demands of the environment; growing when the demands increase and shrinking when the demands decrease.

At the macro-level, increased demands to the system cause an increase in performance of the system. Interventions raise IQ. Once those demands are reduced, the IQ adapts to the new, lesser demands. This behavior is not consistent with a reciprocal interaction but is explained by a unidirectional interaction of adaptation to environmental demands.

The strength of this interpretation of the evidence presented here comes not only from the meta-analytic results—but also from the fact that the only studies included are RCTs. Children in all of these studies began school either after their intervention ended, during their intervention, or as part of their intervention, for example. This represents an increased cognitive demand; one that could be argued should work against the unidirectional–reactive model. The effects of going to school, however, are present *for both groups*, *experimental and control*. Therefore, in these studies, such an increase in cognitive demand from schooling plays no part in the unidirectional–reactive model for the fadeout effect. The effects cancel out in between-group comparisons because they are seen in both of the randomly assigned groups. Consistent with this unidirectional-reactive model, quasi-experimental data (e.g. Cahan & Cohen, 1989) suggests that going to school a year early because of birthday cutoffs increases the IQ of children who are only a couple of months apart in age, yet a full year apart in schooling.

The strength of random assignment also buys ignorability of appeals to genetic differences. Genetic effects, often ignored in traditional developmental research, present a major confounding factor for longitudinal studies. Specifically, it is easy and defensible to posit traditional longitudinal models of reciprocal interactions are contaminated by genetic effects. What appears to be an environmental effect is due to the unfolding of genetic processes. In randomized controlled trials as in those presented here, individuals in the experimental and control groups are experiencing the same unfolding, the same genetic effects, the same selection into environments due to different genetic profiles. So for every person with a genetic advantage for intelligence who was intervened on, there is a corresponding individual with that same advantage who was not. Any differences between the groups cannot be accounted for by genetic differences.

6. Conclusion

Before we summarize the implications of this study, it is of the utmost importance to state what these results do not imply. We are in no way suggesting that environmental effects do not alter IQ in the first few years (see Figs. 2 & 3: intercepts). The existence of the fadeout effect does not mean that interventions to raise intelligence are ineffective or doomed to fail or pointless. Such an interpretation is not warranted by the data. The interventions show a strong effect on IQ that does not immediately snap back but instead gradually fades over years. This fadeout occurs because those children whose ability was increased lose their abilities once returned to their previous environment. Because the effects of an intervention are not permanent in no way means the intervention 'failed'. The permanency of the effects rested on an untested assumption, one that we show is false. There is no reason to believe that results from an intervention to raise intelligence should be permanent.

The warranted implications of this study are as follows. The interaction between environmental effects and IQ is unlikely to be a fully-reciprocal one. Such a model, while able to account for a large amount of longitudinal correlational data, does not withstand experimental testing. Instead, the pattern of interaction is more of adaptation to environmental demands. Under increased demands from the environment (due to an intervention) we can raise IQ. Once those demands are removed, the system adapts to the new, reduced demands. This is more consistent with a pattern of unidirectional adaptation.

Following participants from experimental trials is of the utmost importance for testing not only the durability of intelligence gains, but also for the testing of theories of development and the role of the environment in intelligence. As shown here, the fadeout effect is real, but the fade is slow and occurs over years. Therefore, we recommend that researchers in adolescence, adulthood, and old age, follow participants for extended period of time. It is not enough to explore durability over weeks, months, or even a year. Extended follow-up will further shed light on these processes in adolescence and adulthood.

The environment can cause increases in intelligence. We believe, however, that there is no reason to assume that these increases *should be permanent*. Reciprocal theories of intelligence and the environment make this permanency explicit. Unfortunately, the data presented here fails to support such strong reciprocal interactions. We argue that this assumption of permanency is misplaced; a unidirectional model of the environment and intelligence makes no such assumption. The intelligence of children will react (within reason) to the demands of the environments they are placed in, for good and for ill.

Acknowledgements

We would like to thank Tricia Voda, Laura Scully, Josh Aronson, Scott Barry Kaufman, Clancy Blair, Lisa Suzuki, and Susan Neumann for their helpful comments on an earlier draft of this paper and Irina Petkova for her work verifying the coding of studies. We would like to thank Marc Scott for his help with the analysis. Thanks go to Jelte Wicherts, Drew Bailey, and the other members of the ISIR conference for their helpful comments and criticisms which were incorporated into this paper. We would also like to thank Eric Turkheimer for his discussion on reciprocal models, Jim Flynn and Andreas Demetriou for their insightful and helpful reviews, and Doug Detterman and the anonymous reviewers for their help in improving this paper. We would finally like to thank Ingrid Helland, for sharing her data.

References[iv]

*Abbott-Shim, M., Lambert, R., & McCarty, F. (2003). A comparison of school readiness outcomes for children randomly assigned to a head start program and the program's wait list. *Journal of Education for Students Placed at Risk, 8*, 191–214.

Abdulkadiroglu, A., Angrist, J.D., & Pathak, P.A. (2011). The elite illusion: Achievement effects at boston and new york exam schools (No. w17264). *National Bureau of Economic Research*.

Bailey, D., Duncan, G.J., Odgers, C., & Yu, W. (2015). *Persistence and fadeout in the impacts of child and adolescent interventions.* (under review).

Barnett, W.S. (2002). The battle over Head Start: What the research shows. *Presentation at a Science and Public Policy Briefing Sponsored by the Federation of Behavioral, Psychological, and Cognitive Sciences* (pp. 999).

Barnett, W.S. (2011). Effectiveness of early educational intervention. *Science, 333*(6045), 975–978.

Behrman, J.R., Cheng, Y., & Todd, P.E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *The Review of Economics and Statistics, 86*(1), 108–132.

Blair, C. (2010). Fluid cognitive abilities and general IQ. *The Handbook of Life-Span Development*.

Bouchard, T.J., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*(4498), 1055–1059.

*Bridgeman, B., Blumenthal, J.B., & Andrews, S.R. (1991). *Parent–child development centers: Final evaluation report.* Unpublished manuscript Princeton, NJ: Educational Testing Service.

Briley, D.A., & Tucker-Drob, E.M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24*, 1704–1713.

Bui, S.A., Craig, S.G., & Imberman, S.A. (2011). Is gifted education a bright idea? assessing the impact of gifted and talented programs on achievement (No. w17089). *National Bureau of Economic Research*.

[iv] *indicates studies that are included in the meta-analysis.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development*, *60*, 1239–1249.

*Campbell, F.A., & Ramey, C.T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development*, *65*, 684–698.

*Campbell, F.A., Ramey, C.T., Pungello, E.P., Sparling, J.J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the abecedarian project. *Applied Developmental Science*, *6*, 42–57.

Card, N. (2011). *Applied meta-analysis for social science research.* New York, NY: Guilford Press.

Cohen, B.H. (2009). When the use of p-values actually makes sense. *Paper presented at the annual meeting of the American Psychological Association, Toronto*.

Cohen, J. (1994). The earth is round (p<. 05). *The American Psychologist*, *49*(12), 997–1004.

Currie, J.D., & Thomas, D. (2000). School quality and the longer-term effects of Head Start. *Journal of Human Resources*, *35*, 755–774.

Darlington, R.B., Royce, J.M., Snipper, A.S., Murray, H.W., & Lazar, I. (1980). Preschool programs and later school competence of children from low-income families. *Science*, *208*(11), 202–204.

Demetriou, A., Spanoudis, G., & Shayer, M. (2014). Inference, reconceptualization, insight, and efficiency along intellectual growth: A general theory. *Enfance*, *3*, 365–396.

*Deutsch, M. (1971). Regional research and resource center in early childhood. *Final Report*.

Devlin, B. (1997). *Intelligence, genes and success: Is it all in the genes?: Scientists respond to the bell curve.* Springer.

Dickens, W.T., & Flynn, J.R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, *108*(2), 346–369.

Dickens, W.T., Turkheimer, E., & Beam, C. (2011). The social dynamics of the expression of genes for cognitive ability. *The dynamic genome and mental health: The role of genes and environments in youth development* (pp. 103–127).

Dobbie, W., & Fryer, R.G., Jr. (2011). Exam high schools and academic achievement: Evidence from New York city. *National Bureau of Economic Research Working Paper Series* (No. 17286).

Duckworth, A.L., & Seligman, M.E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.

*Epstein, A.S., & Weikart, D.P. (1979). *The Ypsilanti-Carnegie Infant Education Project: Longitudinal follow-up.* Ypsilanti, MI: The High/Scope Press.

Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin*, *36*(1), 82–96.

Gottfredson, L.J. (1997). Mainstream science on intelligence. *Intelligence*, *24*, 13–23.

Gottlieb, G. (1983). The psychobiological approach to developmental issues. In P.M. Mussen (Ed.), *Handbook of child psychology* (pp. 1) (4th ed.). New York: John Wiley & Sons.

*Gray, S.W., & Klaus, R.A. (1970). The early training project: A seventh-year report. *Child Development*, *41*, 909–924.

*Harrell, R.F., Woodyard, E., & Gates, A.I. (1955). *The effect of mothers' diets on the IQ of offspring.* New York, NY: Teachers College.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Hedges, L.V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*(2), 107–128.

*Helland, I.B., Smith, L., Biomen, B., Saarem, K., Saugstad, O.D., & Drevon, C.A. (2008). Effect of supplementing pregnant and lactating mothers with n–3 very-long-chain fatty acids on children's IQ and body mass index at 7 years of age. *Pediatrics*, *122*, e472–e479.

*Helland, I.B., Smith, L., Saarem, K., Saugstad, O.D., & Drevon, C.A. (2003). Maternal supplementation with very long-chain n − 3 fatty acids during pregnancy and lactation augments children's IQ at 4 years of age. *Pediatrics*, *111*, e39–e44.

Herrnstein, R., & Murray, C. (1994). *The bell curve.* New York, NY: Free Press.

Howe, M.J.A. (1997). *IQ in question: The truth about IQ.* London: SAGE.

*Huebner, C.E. (2000). Promoting toddlers' language development through community-based intervention. *Journal of Applied Developmental Psychology*, *21*, 513–535.

Jensen, A.R. (1985). Compensatory education and the theory of IQ. *Phi Delta Kappan*, *67*, 554–558.

Jensen, A.R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

*Jester, R.E., & Guinagh, B.J. (1983). The Gordon parent education infant and toddler program. *The consortium for longitudinal studies, As the twig is bent…Lasting effects of preschool programs* (pp. 103–132).

*Karnes, M.B., Zehrbach, R.R., & Teske, J.A. (1974). The Karnes' preschool program: Rationale, curricula offerings and follow-up data. In S. Ryan (Ed.), *A report on longitudinal evaluations of pre-school programs. Vol. 1*, Washington, DC: Department of Health, Education, and Welfare.

*Lambie, D.Z., Bond, J.T., & Weikart, D.P. (1974). Home teaching with mothers and infants. *Monographs of the high/scope educational research foundation, No. 2*. Ypsilanti, MI: The High/Scope Press.

Lee, V.E., & Loeb, S. (1995). Where do head start attendees end up? One reason why preschool effects fade out. *Educational Evaluation and Policy Analysis*, *17*, 62–82.

*Love, J.M., Kisker, E.E., Ross, C., Raikes, H., Constantine, J., Boller, K., ... Vogel, C. (2005). The effectiveness of early head start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*, *41*(6), 885–901.

Miller, L.B., & Bizzell, R.P. (1984). Long-term effects of four preschool programs: Ninth- and tenth-grade results. *Child Development*, *55*, 1570–1587.

Miller, L.R., & Bizzell, R.P. (1983). Long-term effects of four preschool programs: Sixth, seventh and eighth grades. *Child Development*, *54*, 727–741.

te Nijenhuis, J., Jongeneel-Grimen, B., & Kirkegaard, E.O. (2014). Are headstart gains on the *g* factor? A meta-analysis. *Intelligence*, *46*, 209–215.

te Nijenhuis, J., van Vianen, A.E.M., & van der Flier, H. (2007). Score gains on g-loaded tests: no g. *Intelligence*, *35*, 283–300.

*Olds, D.L., Henderson, C.R., Chamberlin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics*, *78*, 65–78.

*Olds, D.L., Henderson, C.R., & Kitzmna, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 to 50 months of life? *Pediatrics*, *93*, 89–98.

(under review) Protzko, J. (2015). *Does the raising IQ/raising g theory explain the fadeout effect?*

Protzko, J., Aronson, J., & Blair, C. (2013). How to make a young child smarter: Evidence from the database of raising intelligence. *Perspectives on Psychological Science*, *8*, 25–40.

*Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., ... Spier, E. (2010). Head Start impact study. *Final report. Administration for children & families*.

*Ramey, C.T., & Campbell, F.A. (1991). Poverty, early childhood education, and academic competence: The Abecedarian experiment. In A.C. Huston (Ed.), *Children in poverty: Child development and public policy* (pp. 190–221).

Ramey, C.T., Campbell, F.A., Burchinal, M., Skinner, M.L., Gardner, D.M., & Ramey, S.L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science*, *4*, 2–14.

*Rueda, M.R., Checa, P., & Combita, L.M. (2011). Enhanced efficiency of the executive attention network after training in preschool children: Immediate and after two months effects. *Cognitive Developmental Neuroscience*, *2*, S192–S204.

Salthouse, T.A. (2013). Selectivity of attrition in longitudinal studies of cognitive functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbt046.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, *63*, 1–19.

*Schweinhart, L.J., & Weikart, D.P. (1980). Young children grow up: The effects of the preschool program on youths through age 15. *Monographs of the high/scope educational research foundation no. 7*. Ypsilanti, MI: High/Scope Press.

Schweinhart, L.J.H.V., Barnes, D.P., Weikart, W.S., Barnett, A.S., & Epstein, A.S. (1993). Significant benefits: The High/Scope Perry preschool study through Age 27. *Monographs of the High/Scope Educational Research Foundation, number 10*. Ypsilanti, Michigan: High/Scope Press.

Van Der Maas, H.L., Dolan, C.V., Grasman, R.P., Wicherts, J.M., Huizenga, H.M., & Raijmakers, M.E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861.

*Vogel, C.A., Xue, Y., Moiduddin, E.M., Carlson, B.L., & Kisker, E.E. (2010). Early head start children in grade 5: Long-term followup of the Early Head Start Research and Evaluation Project Study Sample (no. 4e81525b2ed848808383ce904c03ba2b). *Mathematica Policy Research*.

*Wasik, B.H., Ramey, C.T., Bryant, D.M., & Sparling, J.J. (1990). A longitudinal study of two early intervention strategies: Project CARE. *Child Development*, *61*, 1682–1696.

*Weikart, D.P., Bond, J.T., & McNeil, J.T. (1978b). *The Ypsilanti Perry preschool project: Preschool years and longitudinal results through fourth grade.* Ypsilanti, MI: High/Scope Press.

*Weikart, D.P., Epstein, A.S., Schweinhart, L.J., & Bond, J.T. (1978a). The Ypsilanti preschool curriculum demonstration project: Preschool years and longitudinal results. *Monographs of the High/Scope Educational Research Foundation*. Ypsilanti, MI: High/Scope Press.

*Whitehurst, G.J., Arnold, D.S., Epstein, J.N., & Angell, A.L. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, *30*, 679–689.

*Whitehurst, G.J., Falco, F.L., Lonigan, C.J., Fischel, J.E., DeBaryshe, B.D., Valdez-Menchaca, M.C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*(4), 552–559.

Žebec, M.S., Demetriou, A., & Kotrla-Topić, M. (2015). Changing expressions of general intelligence in development: A 2-wave longitudinal study from 7 to 18 years of age. *Intelligence*, *49*, 94–109.

Zigler, E., & Seitz, V. (1982). Social policy and intelligence. *Handbook of human intelligence* (pp. 586–641).