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/© 2015 Published by Elsevier Inc.

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 nonsignificant 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.