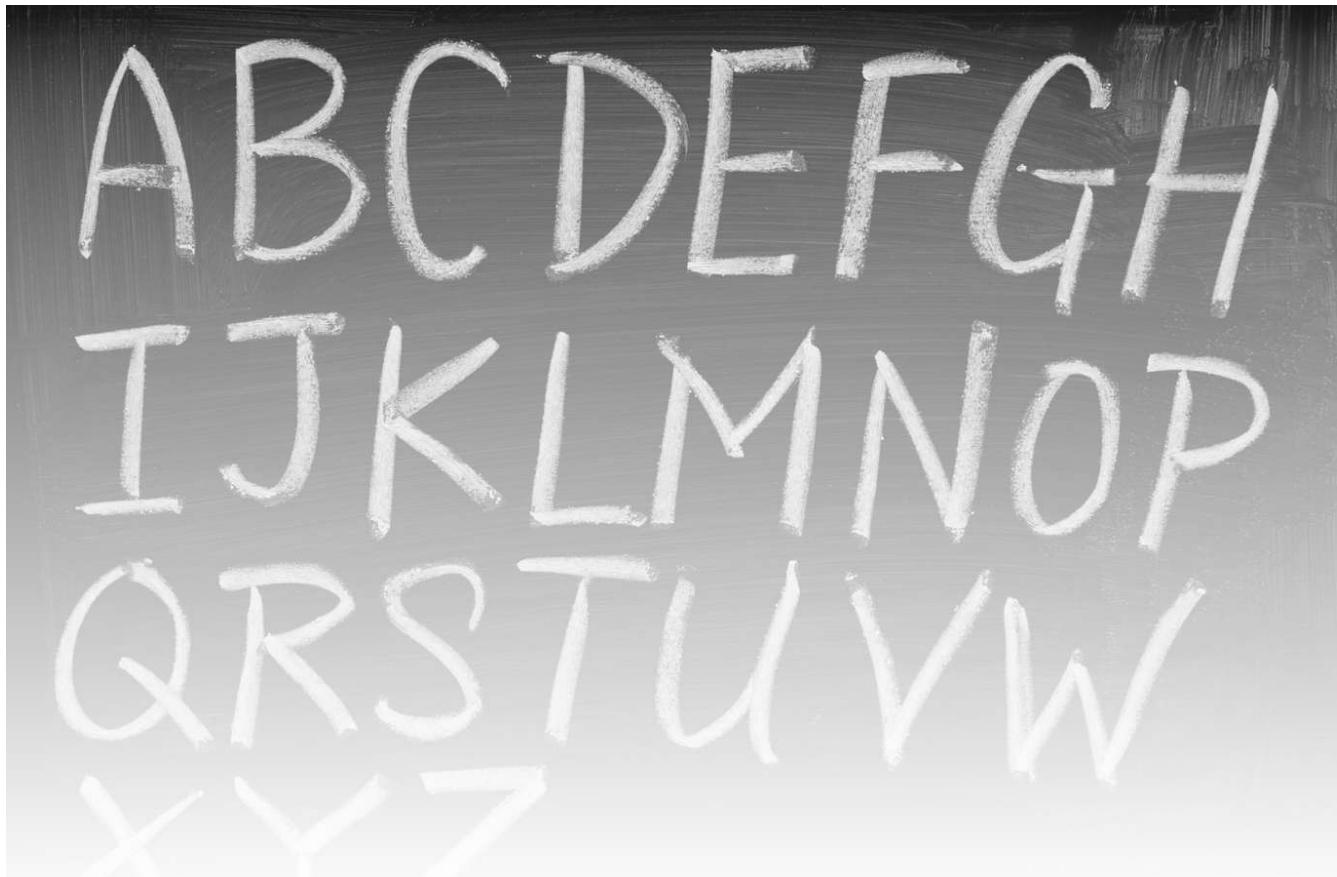


Why Do Most Education Interventions Fade Out Over Time?

There is evidence both to explain and complicate the “fadeout effect”

By DREW BAILEY, TYLER WATTS, AND EMMA HART



FOR MANY YEARS, education researchers have puzzled over a common phenomenon known as the “fadeout effect”—that is, that the benefits of education interventions, such as early childhood enrichment programs, often diminish or disappear.

Our research team has written extensively about fadeout over the past decade. We have focused on how this issue pertains to early childhood programs, though we have now studied the dynamics of fadeout across a range of interventions, including K-12 curricular reforms, charter schools, and even adolescent substance-use prevention programs. The fadeout issue has received substantial attention in early childhood education research, in part because landmark studies set the expectation that impacts from early childhood interventions should last—but many newer studies have found that initial benefits from early childhood education programs do not. We have spent the past 10 years trying to figure out why.

Although we have learned a lot about fadeout, many important questions remain unresolved. Here, we

present responses to some frequently asked questions about fadeout in an imaginary dialogue between our research team and an interested skeptic who suspects that concerns about fadeout are overblown. Many of these questions reflect conversations we have had with audiences at conferences and seminars, for which we are grateful.

I recognize that the effects of many education interventions don't last in the long term. But who would have expected otherwise? If an exercise intervention for weight loss ended and the treatment group stopped exercising, would we be surprised to see that they regained their weight? Why would education be any different?

Fair point! We agree that researchers and advocates shouldn't be surprised to see evidence of fadeout following an education intervention. However, those working in education consistently use the prospect of long-lasting effects in their grants and papers to motivate all kinds of education investments.

The promise of positive long-term effects is not just used as a pitch in grant salesmanship; it also reflects theoretical beliefs about skill building in education. The reason we as a society are drawn to education as a tool for social change is that we believe investments in education work differently from other types of interventions that will obviously produce transient effects.

As the Chinese philosopher Lao Tzu said, "Give a man a fish and you feed him for a day. Teach a man to fish and you feed him for a lifetime." Childhood is viewed as an important period during which investments can yield long-term payoffs.

Correlational studies showing long-term relationships between skills measured early in life and later outcomes seem consistent with skill-building theories and are often cited to support the possibility that programs will produce beneficial long-run effects. For example, the correlational finding that early mathematics skills are strong predictors of later school outcomes is often used as evidence that early interventions targeting mathematics will produce the desired long-lasting effects.

So, we cannot have it both ways. We cannot regularly argue for education investments on the expectation that they will produce positive long-lasting effects, then act as if fadeout was inevitable all along when we observe it.

OK, but isn't the term "fadeout" too negative for what is really going on most of the time? In the case of, say, a pre-K intervention, kids who received the program and kids who did not receive the program keep learning after the program ends. The kids who received the program don't actually *forget* what they learned during pre-K (for example, they don't forget how to identify letters or count). Instead, the kids who did not receive the program eventually catch up. So, isn't "catch-up" a better term? And isn't "catch-up" a good and equitable result, since the lower-achieving kids who did not attend pre-K eventually learn more?

This is a common perspective among early-childhood-education researchers. We think that the line of reasoning that reframes fadeout as socially desirable "catch-up" is misleading and obscures what is really going on.

The easiest way to understand the issue is to imagine a randomized controlled trial in which children are randomly assigned to either an education intervention or a control group. As any basic research methods textbook argues, a randomized controlled trial produces a control group that can be understood as an approximation of the *counterfactual condition*. In other words, the outcome for the control group is what we would have observed for the treatment group *had the treatment not been administered*. So, if we observe children's learning two years after a given intervention, and the control group has "caught up" to the learning levels of the treatment group, we should understand this to mean that the treatment group now has the same level of skills that they would have had if the intervention had never been administered.

If we want education interventions to have long-lasting effects for socially progressive reasons (for example, closing achievement gaps related to socioeconomic status), control group children "catching up" is not desirable. In the context of many education interventions that are targeted toward children at risk for underachievement compared with some other group (for example, poor versus non-poor children; children who are struggling to read versus children who are reading at grade level), "catch-up" implies that both groups are now lagging behind their higher-achieving peers. Indeed, both groups of children are at the same level as before the intervention started—which is presumably the problem that motivated the intervention in the first place.

The term "fadeout" is most often associated with research on early childhood interventions, but in fact the fadeout phenomenon occurs far more generally. Researchers have also observed fadeout in studies of adults, where members of the control group would not be expected to demonstrate gains in the targeted skills. Tailoring our definitions and explanations of fadeout too closely to early childhood contexts risks missing important insights into why it happens and its policy implications.

What about spillover effects on peers? Is it possible that an intervention has fading effects because its benefits are spilling over to the kids who did not receive the intervention?

This is an interesting question. If that were the case, the control group would no longer approximate the "counterfactual condition" that we described above, as the control group outcomes would be affected by the presence of the treatment. Certainly, spillovers in education are frequent. For example, peers affect children's learning and behavior, and siblings affect each other directly and, through their parents, indirectly. Attributing fadeout completely to spillover would lead to a drastically different understanding of fadeout. One clever method for probing this issue was John Protzko's approach of looking separately at the age-normed IQ test scores for the treatment and control groups in a 2015 meta-analysis of interventions that boosted children's IQ. If spillover was driving fadeout, one might expect the control group's IQ scores—something that is generally fairly stable—to increase in the years following these interventions. However, while Protzko found a decrease in the treatment group's scores following intervention, he didn't find increases in control-group IQ scores. Thus, while spillovers are certainly possible and worthy of study in education policy evaluation, they are probably not a major force driving fadeout.



ISTOCK PHOTO

If learning slows after an intervention ends, shouldn't we focus on the environments kids encounter after the intervention? Why are we saying that the effects of an intervention fade out when the problem actually happens after the intervention is over? If I run a successful pre-K program, and I find a large benefit for the pre-K group over the control group at the end of the school year, but the benefit fades during kindergarten, shouldn't we change what happens in kindergarten?

This is an appealing and intuitive idea that is especially popular in early childhood education research. In fact, many research teams have looked for evidence that early childhood intervention effects are more likely to last when the intervention is followed by higher-quality learning environments. The strongest example comes from Rucker C. Johnson and C. Kirabo Jackson, who in 2019 found that the effects of Head Start persisted more for students who subsequently attended better-funded schools during their K-12 education. However, the idea that fadeout happens more slowly when students later attend better education environments is far from a rule. For example, Pedro Carneiro and colleagues conducted an experiment in which students were randomly assigned to classrooms across the elementary school years. They found that the benefits of effective teachers faded as much for students who had more effective teachers in the following school year as for students who did not. And a meta-analysis we conducted with other members of our team found that early intervention effects were not more likely to persist when followed by higher-quality learning environments.

While it is probably true that subsequent environments help program effects endure in some cases, there is no guarantee that high-quality learning environments will disproportionately benefit the students that received an intervention. Indeed, particularly in the early school years, teachers' conceptions of "high-quality learning environments" probably include practices that benefit less-prepared students more (would we really want kindergarten teachers targeting instruction to the top 10 percent of achievers in the class?). Thus, subsequent school quality may have the opposite effect: Early intervention impacts may well persist *more* when children later enter lower-quality learning environments where control-group students catch up less readily.

Intervention effects might well be sustained at higher rates if we restricted the subsequent learning opportunities of control-group children in our studies (that is, tracking these students into lower-quality environments, beginning at kindergarten entry), but this is not a practical or ethically viable solution.

OK, but it seems as if you're painting with too broad a brush. Don't a lot of education programs produce positive long-lasting effects? Isn't it just the bad or ineffective interventions that generate effects that fade?

Our thinking on this has evolved over the years, but our work now suggests that fadeout is fairly ubiquitous across education interventions. In our recent meta-analysis of approximately 85 diverse randomized controlled trials, we found that only one intervention characteristic was a strong and consistent predictor of the follow-up impact size in the two years after programs ended: the size of the effect at the end of the intervention. Surprisingly, information about intervention features, skills targeted, and study characteristics provided little additional predictive power. In other words, knowing how big the effect of the intervention was at the end of the program was a good predictor of the follow-up effect a few years after interventions ended, but other characteristics of interventions and programs told us almost nothing about how much effects would fade or persist. While the persistence of effects varied considerably (and impacts did persist to some extent, on average), theoretically salient factors could not explain this variance. Instead, we found that effects faded fairly consistently across cognitive interventions, social-emotional interventions, early childhood interventions, and adolescent interventions.

Labeling a particular intervention as "good" or "bad" requires value determinations beyond the scope of this conversation. But in our own work, we have seen that even interventions developed from careful research and implemented with high fidelity can yield impacts that fade in the long term.

Perhaps we just have to give up on long-term effects. Why not just focus on the "here and now"? The short-term effects of an intervention may be important enough without the promise of long-lasting effects. In the case of pre-K, maybe it's OK if the effects don't last, so long as the kids were well taken care of that year.

We agree that making children better off, even temporarily, is a good thing. However, in a world with many potential investments that could target major societal challenges, such as intergenerational poverty, decisionmakers must make hard choices. Should pre-K be on the list of "worthwhile" investments? We think that long-run impacts on socially important outcomes such as education attainment and earnings should absolutely factor into such decisions, because it is the gains on those outcomes that often drive positive investment returns.



ISTOCK PHOTO

But you just told me that impacts on child skills fade out in the years after intervention programs, and that this pattern is ubiquitous. Isn't fadeout incompatible with positive long-term effects on life-course outcomes like education attainment and earnings? Doesn't fadeout mean that the intervention failed?

Maybe not! So far, we have focused on the dynamics of fadeout on skills measured consistently at the end of an intervention and in the years that follow. However, it is worth thinking more holistically about how interventions may shape other dimensions of a child's development despite fadeout on the discrete skills the intervention targeted. An intervention may succeed in other ways, even potentially shaping adult functioning, regardless of initial fadeout. Examples might include:

- 1) **Skill transfer.** The effects faded out completely on one measure—let's say early literacy skills—but the intervention continues to affect long-term outcomes, perhaps because phonics training built a child's higher-order reading comprehension skills or boosted their interest in reading.
- 2) **Unmeasured mediators.** The effects faded out completely on an achievement measure, but the intervention's longer-term effects operate primarily through pathways not captured by test scores, often called "noncognitive skills."
- 3) **Fadeout coupled with persisting effects.** The effects faded to statistical non-significance on a focal measure, but they persist at some non-zero level that may still justify the cost of the intervention and may still transfer to other skills.

Whether programs can generate adult effects despite fadeout on child skills is a question that has long puzzled researchers. As optimists about the potential for education interventions to improve lives, we are most intrigued by questions about the potential long-term benefits of such interventions.

Many in the research sphere hold to the notion that education interventions commonly produce “sleeper effects”—instances where short-run impacts on some targeted skills fade out completely yet are followed by positive long-run impacts on adult outcomes. Perhaps the most famous example is the Perry Preschool Project, where impacts on children’s IQ scores faded to nearly zero in the years following the end of treatment, yet children who received the intervention showed benefits on a range of economically relevant outcomes in adulthood, including elevated earnings and employment status and lower crime rates.

To better understand the very long-term effects of education programs, our team has conducted a meta-analysis examining randomized controlled trials that have measured adult effects. We identified 29 programs that evaluated end-of-treatment and adult outcomes and found that, despite medium-term fadeout on child skills, these programs did produce positive long-term impacts on adult functioning, though effects are fairly small on average (~0.05 SD).

Yet, isolating *how* programs generate long-term effects (either through the mechanisms named above or others) presents significant challenges. Returning to the case of Perry Preschool, researchers have toiled to identify how the program could have generated long-term impacts despite fadeout on IQ. As found by research teams led by Sneha Elango and Remy Pages, measured social-emotional mediators (such as



The Perry Preschool Project, conducted from 1962 to 1967 in Ypsilanti, Michigan, sought to determine the effect of early education on the cognitive abilities of underserved children. The work of renowned child psychologists Jean Piaget and Constance Kammi, pictured with two participants, laid the groundwork for the study.

externalizing behaviors, task orientation, and psychological maturity) and cognitive mediators (such as scores on IQ tests and early standardized language and math assessments) fall far short of accounting statistically for the full effect of Perry and other early programs on adult outcomes. Some economists use the term “social policy dark matter” to describe the mechanisms behind persisting effects.

One hypothesis—probably the most popular in the field—posits that the unexplained mediators are “noncognitive.” The best evidence for this assertion is that (1) increased test score impacts often fade rapidly following education interventions, and (2) certain education inputs that generate beneficial long-run effects, such as effective teachers and some preschool programs, sometimes affect medium- and long-term outcomes related to children’s behavior, such as school attendance and crime rates. These findings have led some researchers, including Raj Chetty and his colleagues, to assert that the impacts of education interventions on noncognitive skills may not fade out.

Wait! You never mentioned this. Now you’re saying that fadeout is really just an issue for cognitive skills? Then why don’t we just generate more programs that target social-emotional skills?

Not so fast! The notion that impacts on noncognitive outcomes do not fade out does not hold up to scrutiny. Changing human personality durably is famously difficult. Given the significant enthusiasm for this theory, our team examined the persistence of effects on cognitive and noncognitive skills using our meta-analytic sample of 85 diverse randomized controlled trials. We found that similarly sized intervention impacts on cognitive and noncognitive skills at the end of treatment faded at strikingly similar rates. It is therefore unlikely to be the case that long-term impacts of education programs on adult outcomes are solely driven by fully enduring impacts on noncognitive skills.

Then where do you think adult impacts come from? You can’t really believe that “social policy dark matter” is something we can take to the bank!

We think a good tentative hypothesis is that there is no single ingredient comprising social policy dark matter just waiting to be discovered. We doubt that there are “silver bullet” skills that interventions should target to produce long-lasting change in key outcomes.

Rather, the kinds of education inputs that generate positive effects in the short and long term (for example, high-quality preschool programs and high-quality teachers) affect children through a complex set of pathways such as those detailed in items 1 to 3 above, rippling out from the skills most strongly affected by the intervention to other aspects of children’s lives. For example, an early literacy curriculum may affect children’s early phonics skills but might also improve their love of reading, vocabulary, performance in other academic subjects, and their relationships with peers, teachers, and parents. These complex and interdependent processes will ring true to developmental psychologists, who view child development as highly multifaceted (and who frequently object to the lumping of skills into broad categories, such as “non-cognitive skills”).

This story offers hope in the promise of education programs, and it is completely compatible with fadeout. Like ripples in a pond, intervention impacts on targeted child skills will transfer to other aspects

of a child's life, the effects of which may then ripple further and feed back into the initially targeted domains. In the end, the child may be different in ways that would be difficult to predict based on an intervention's initial effects. Our team calls this account of long-term impacts on children's outcomes Large Interconnected Network Theory, or LINT.

Zooming back in on Perry Preschool, we think impacts in that study are compatible with LINT. Impacts on IQ scores were not statistically significant a few years after the intervention ended, but impacts on achievement tests were non-trivial throughout childhood and adolescence. The story of complete fadeout followed by emergence of large impacts in adulthood is an oversimplification of Perry's results. If LINT is right, it may be difficult to make reliable predictions in advance about differences in patterns of fadeout and persistence on a range of medium-term outcomes; however, larger and broader short-term intervention impacts should be more likely to transfer into longer-term impacts. Of course, there are almost certainly specific combinations of skills, populations, treatments, and settings that are more likely to generate long-term impacts, but the absence of a large database of randomized interventions with long-term follow-up results presents challenges to testing precise theories about how impacts emerge.

This is sounding more hopeful! So maybe most education interventions are actually successful even when we find fadeout because the effects are just transferring to other domains. Again, why are we worried about fadeout?

It may be tempting to conclude that any investment in children is a good investment, but we are far from a formalized understanding of whether, and to what degree, LINT-like processes occur. Homing in on what's going on is essential to forming an intervention science that can help education policymakers make difficult investment choices that reliably benefit children.

Unfortunately, long-term follow-up evaluation of education programs rarely happens, and when it does, researchers find substantial variation in the estimated effects on long-term outcomes. In the very short term, deciding between two otherwise identical programs on the basis of their end-of-treatment impacts on test scores may be the best we can do. But this approach is not ironclad—at least today. In the aforementioned meta-analysis of 29 interventions that followed up on long-term adult outcomes, our team found that end-of-treatment test-score impacts sometimes predicted adult impacts, but these associations were not always robust.

One cannot draw solid conclusions that might shape policymaking from a sample as small as 29 randomized controlled trials, but this constitutes the best data currently available. To make progress, the field needs more long-term follow-up assessments. Over time, we hope that researchers, policymakers, and funders will embrace the need for a well-funded, robust study of the long-run effects of education inputs. Meanwhile, we have encouraged funders and policymakers to think more about how to design and select education interventions that are more likely to produce positive longer-term impacts. For example, is a program effectively teaching fundamental skills that have been ignored by state curricula, or is it simply targeting the same skills a few months earlier than the status quo?

OK, let's say you have convinced me that fadeout is common across most education interventions. Where does that leave people trying to design effective programs in the real world? Are we just shooting in the dark, hoping that we can somehow activate this LINT process that you describe? Surely, there are steps we can take to maximize our chances of success.

Increasing the money and effort spent on long-term follow-up of education intervention evaluations should be a major priority. Following up with participants many years after the end of every evaluation would not be feasible. But forward-thinking researchers (if properly incentivized by funders and policy-makers) can plan to link student data to large administrative data sets many years later at relatively little cost. If we had 10 times more causally informative evaluations with long-term follow-up, we could learn a lot about these questions.

When researchers use cheaper and easier non-experimental approaches to identify skills to target intervention, we should demand more of them: Are their predictions holding up in stronger causal studies of skill dynamics? Are they able to help us identify regularities in how intervention impacts persist into adulthood in the extant experimental literature? Answering these questions puts developmental theories to work to improve children's lives.

As we have studied fadeout over the last 10 years, we have noticed a reflexive squeamishness by researchers in education policy or child development. The economist James J. Heckman and his colleagues at the nonprofit Heckman Equation initiative have asserted on social media that "fadeout is a myth." While well meaning, this pithy slogan avoids grappling with the complexities of the phenomenon: Fadeout is ubiquitous, but it often occurs along with persistence, and sometimes long-term impacts on economic outcomes are large enough that public investments in education programs are fully offset by gains in the income earned by their beneficiaries in adulthood and associated reductions in government spending.

Ignoring fadeout will lead to its repeated rediscovery by anyone curious enough to follow up on education interventions. If evaluations are not large enough to detect realistically sized positive long-term impacts, skeptics will continue to cite nonsignificant medium-term impacts as evidence of program ineffectiveness. At the same time, advocates and many researchers will cite selected (and thus upwardly biased) impacts as evidence of program effectiveness. We have spent enough time in this spin cycle. We can only hope to make progress in understanding human development and education interventions if we learn from our current body of evidence and use it to generate new hypotheses about skill building that take the reality of fadeout into account. **E**

Drew Bailey is a professor in the School of Education at the University of California, Irvine. Tyler Watts is an associate professor of psychology and education in the Department of Human Development at Teachers College, Columbia University. Emma Hart is a postdoctoral research fellow in the Lynch School for Education and Human Development at Boston College.

This article appeared at EducationNext.org on February 12, 2026.