



What a Modern Perry Preschool Project Has Taught Us about Early Childhood Investment

John A. List

The University of Chicago · NBER

This Research Road

All started with a plea for help



Chicago Height's 2006 Request

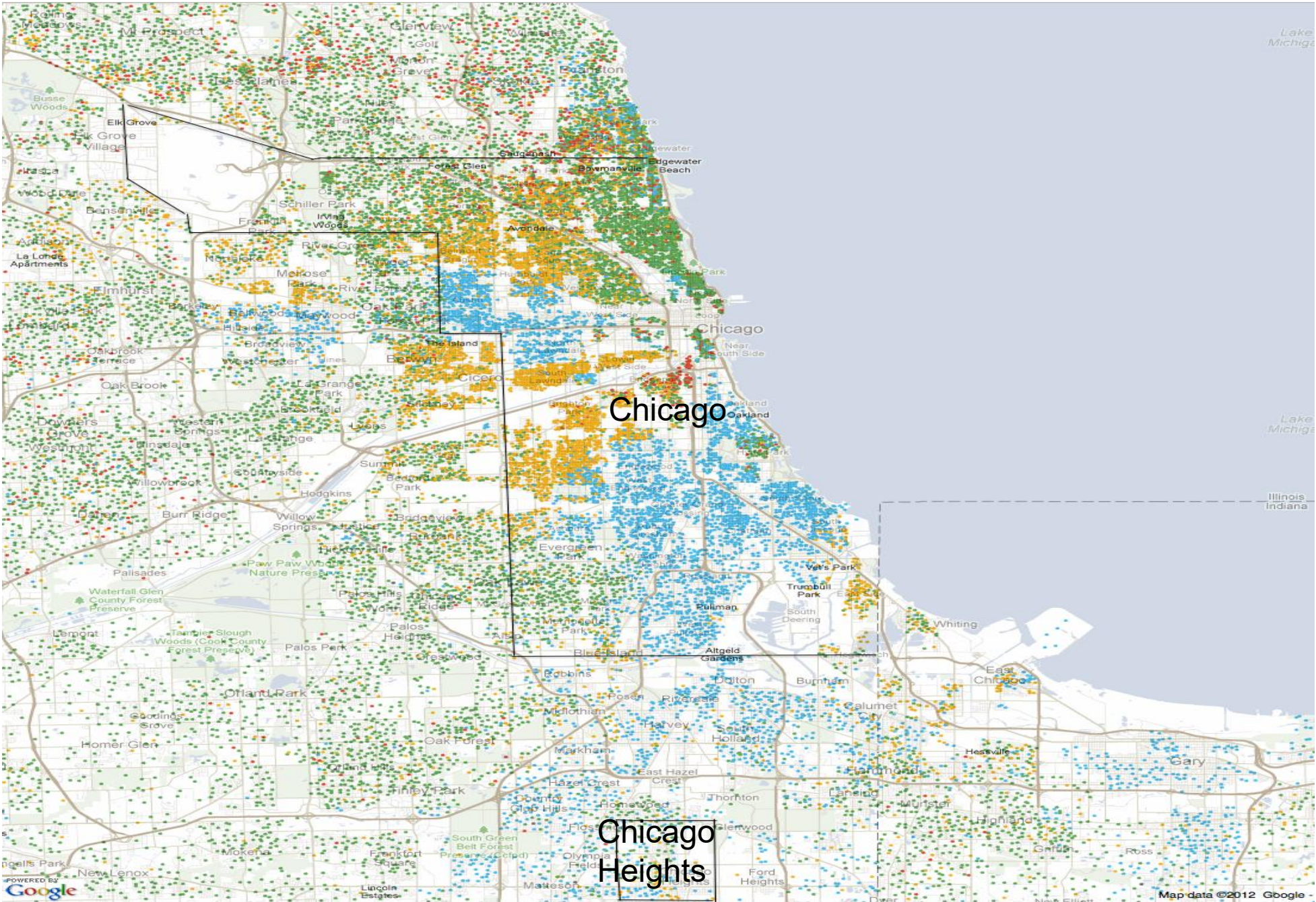
Can you help us
in any way possible?



Distribution of racial and ethnic groups in 2010

One dot = 100 people

- White
- Black
- Hispanic
- Asian
- Native American
- Other



Where We Started: High School

“Chicago Heights Miracle”

Three numbers I have believed about this program

2008, mid-stream, I told the press the \$400,000 experiment had kept 25–40 potential dropouts on track for a diploma: “call it \$10,000–\$16,000 per graduate” “Per student, that’s pretty expensive.”

Then, with two years of data, the paper projected \$1,200 per additional graduate, driven by threshold students, +14 p.p. on track in year 1 and +12 p.p. in year 2, exactly where the model said to look.

Then year 4 arrived: the on-track gains halved, then vanished, no impact on graduation, even for threshold students.

Same program, three horizons, three answers. The estimand depends on when you stop measuring!

Levitt, List, and Sadoff, “The Short and Long-Term Effects of Performance-Based Incentives on Educational Achievement: Evidence from a Randomized Experiment,” NBER Working Paper No. 22107.

THE STARTING POINT

Perry showed how preschool could transform lives — but left some big questions open

The Perry Preschool Project (1962–67) randomized 123 disadvantaged children in Ypsilanti and tracked them for decades. Graduation, earnings, and crime all moved. Rates of return of 7–10% reshaped policy.

But Perry was a product of its era: one treatment, one site, small sample, and 40 years of waiting before the evidence crystallized.

It could tell us that early investment works. It could not tell us why it works, which components matter, or how to deliver them at scale.

WHY

does it work? Mechanisms were a black box

WHICH

components matter? Everything was bundled

HOW

does it scale? One site, 123 children

My Idea: Let's Start Our Own Pre-K!

- Started building a Pre-K in 2008
- 3-5 year olds as the target population
- 3 major goals

Perry and I asked the same question. We just had in mind different estimands

Perry estimated whether the bundle works; I wanted to estimate the objects that decide when, why, and at what scale it works



The system we take to the data

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale } n: F(n), f_n, S(n), E(n), C(n)$$

The child's system

What builds skill — the arguments of f

- θ — the child's skill stock at age t : the state of the system
- f — the technology of skill formation, varying with age t
- P — parental investment: the home's inputs, beliefs, and norms
- S — schooling: what the classroom delivers
- C — the child's own investment: the effort the child brings to f
- E — the environment: other children's skills enter through $\sum w_{ij}\theta_j$

Measurement, outcomes, scale

What we observe — and what it buys

- Z — test scores: what we measure, not the object itself
- λ — the measurement map from θ to Z ; ε — the noise around it
- Y — adult outcomes: college, work, conduct
- g — the outcome map: which coordinates of θ the world prices
- n — scale: an argument of the whole system, not an afterthought

At scale: capital F is who you serve, little f is what you deliver, S is what you can staff, E is what surrounds the child in equilibrium, and $C(n)$ is what it costs — every one of them moves when n does.

Same question — a different estimand

What Perry identifies

ATE and a few CATEs, all reduced-form

- The return to one bundle: f , P , S , C , and E enter jointly; subgroup cuts (gender, baseline IQ) slice the same composite, never separate the inputs
- One site, one situation: a single draw of F and f
- Scores stand in for skills: λ assumed to be the identity map
- Y arrives after a 40-year wait; $n = 123$ is silent on scale

Same question — a different estimand

What Perry identifies

ATE and a few CATEs — all reduced-form

- The return to one bundle: f , P , S , C , and E enter jointly — subgroup cuts (gender, baseline IQ) slice the same composite, never separate the inputs
- One site, one situation: a single draw of F and f
- Scores stand in for skills: λ assumed to be the identity map
- Y arrives after a 40-year wait; $n = 123$ is silent on scale

What CHECC is built to identify

The objects behind the number

- $E: \sum w_{ij}\theta_j$ — peers and neighborhoods, randomized and measured
- P — parental capital: built early, pays off in bad states
- C — the child's own effort: chooseable, priced — and inside the score Z
- λ — the measurement map, tested against a second technology (EEG)
- g — which coordinates of θ carry into Y
- n — scale as an argument of the system, not an afterthought

Perry answers “does it work?” CHECC is designed to answer why, through what, for whom — and at what scale.

THE DESIGN

Who was randomized to what: nearly 3,000 children in two phases

Phase 1 · 2010–12

1,434 children randomized across five arms

Pre-K · Tools of the Mind	n = 164
Pre-K · Literacy Express	n = 170
Parent Academy · Cash	n = 186
Parent Academy · College	n = 186
Control	n = 728

THE DESIGN

Who was randomized to what: nearly 3,000 children in two phases

Phase 1 · 2010–12

1,434 children randomized across five arms

Pre-K · Tools of the Mind

n = 164

Pre-K · Literacy Express

n = 170

Parent Academy · Cash

n = 186

Parent Academy · College

n = 186

Control

n = 728

Phase 2 · 2012–14

~1,500 children randomized to CogX dosages

CogX · 12 months

n = 150

CogX · 9 months

n = 150

KinderPrep · 2-month summer

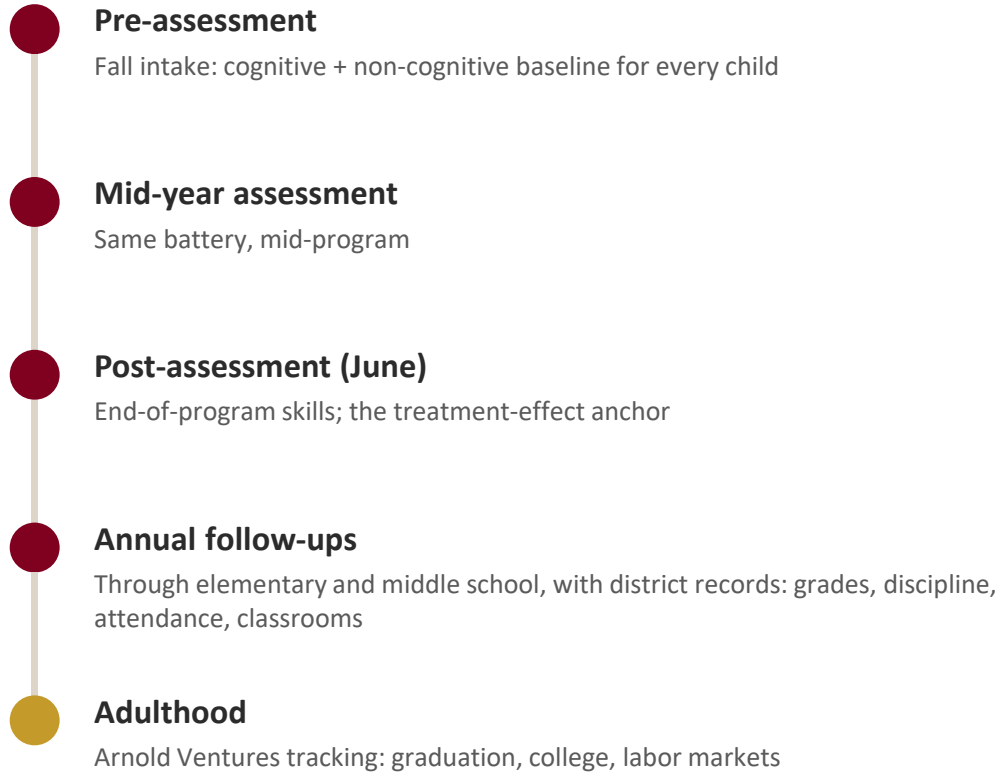
n = 150

Control

n = 1,054

What we measured, and when

The cadence



Cognitive battery (normed)

- PPVT — Peabody Picture Vocabulary Test (receptive vocabulary)
- Woodcock-Johnson III: Letter-Word Identification, Spelling, Applied Problems, Quantitative Concepts

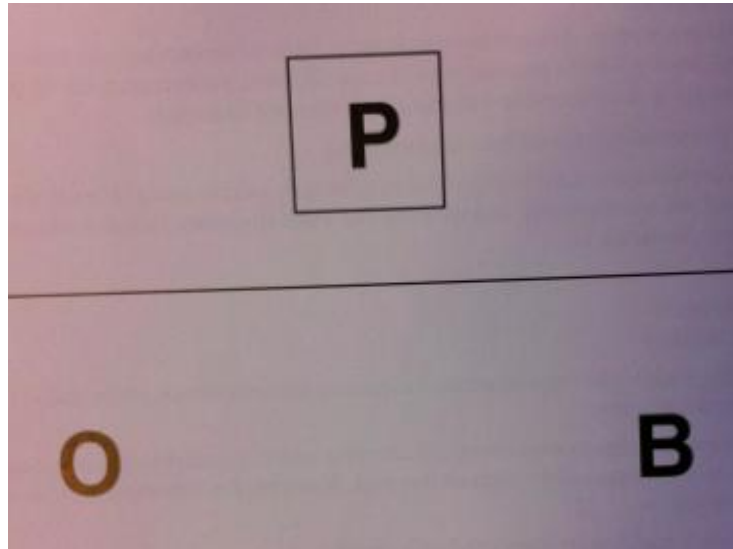
Non-cognitive & behavioral

- Executive function: inhibitory control (spatial conflict), operation span; PSRA self-regulation assessment
- Incentivized behavioral-economics experiments — 50+ studies on the platform: patience, sharing, cheating, discrimination, competitiveness
- Later layers: parental beliefs & time diaries, classroom peers, EEG brain activity

COGNITIVE:

Letter-Word/Receptive Vocabulary

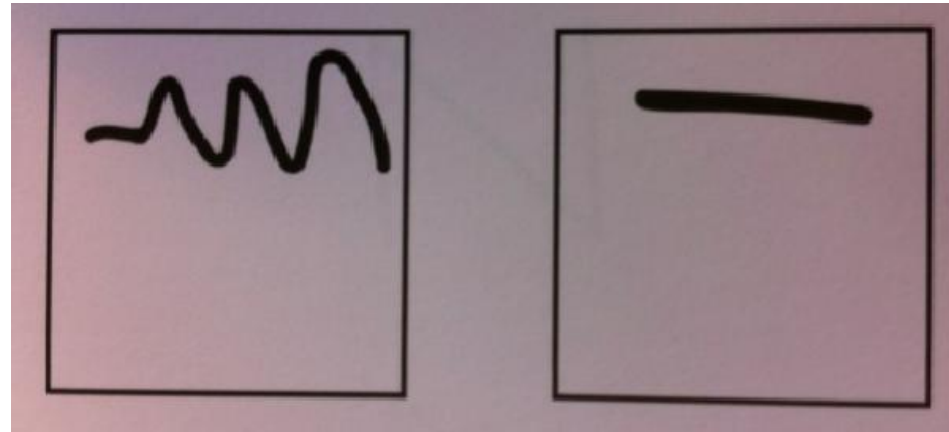
- In this part, each child was asked to name letters out loud from a list.
- “What letter/word is this?” – Letter-word



COGNITIVE

Spelling

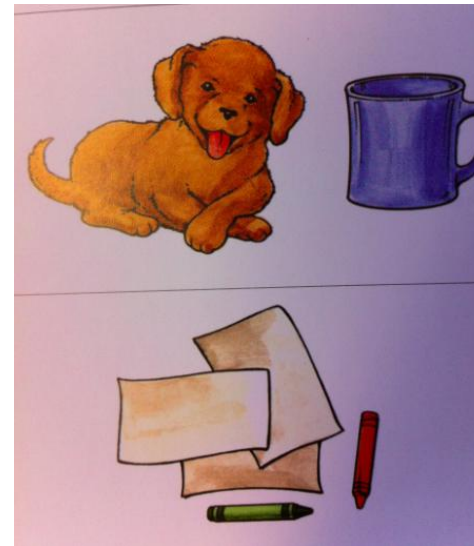
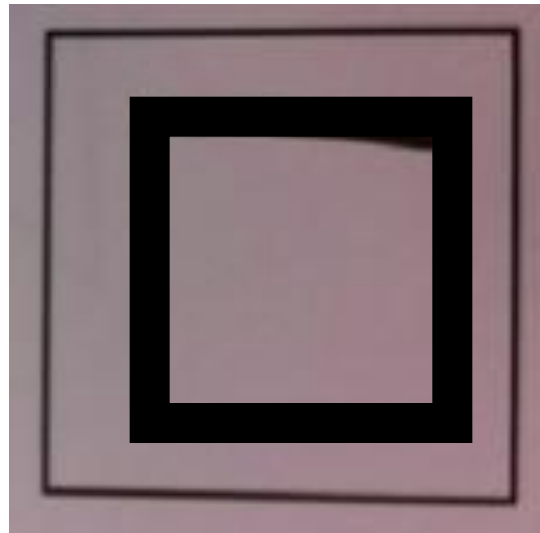
- In this part, each child was asked to draw shapes and letters.
- “Can you draw a line like this one”?
- Can you write the Letter “A”?



COGNITIVE

Applied Probs. / Quant.Concepts

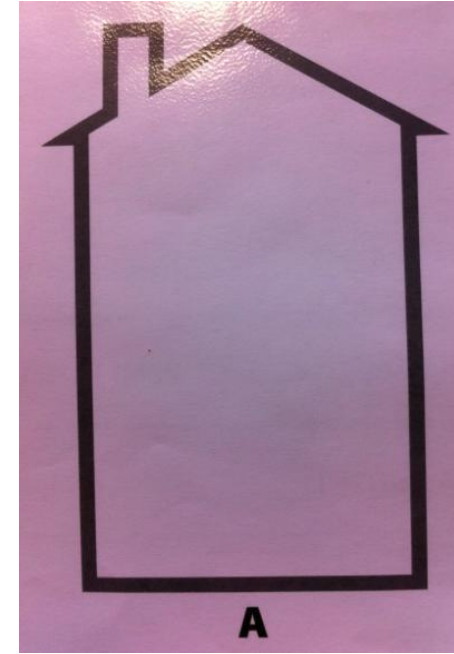
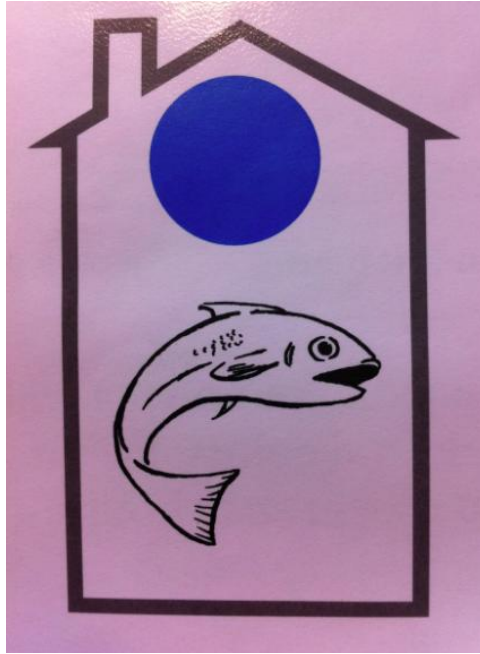
- In this part, each child was asked to count, solve problems, and identify shapes.
- “What shape is this?” “How many dogs are in this picture?”



NON-COGNITIVE

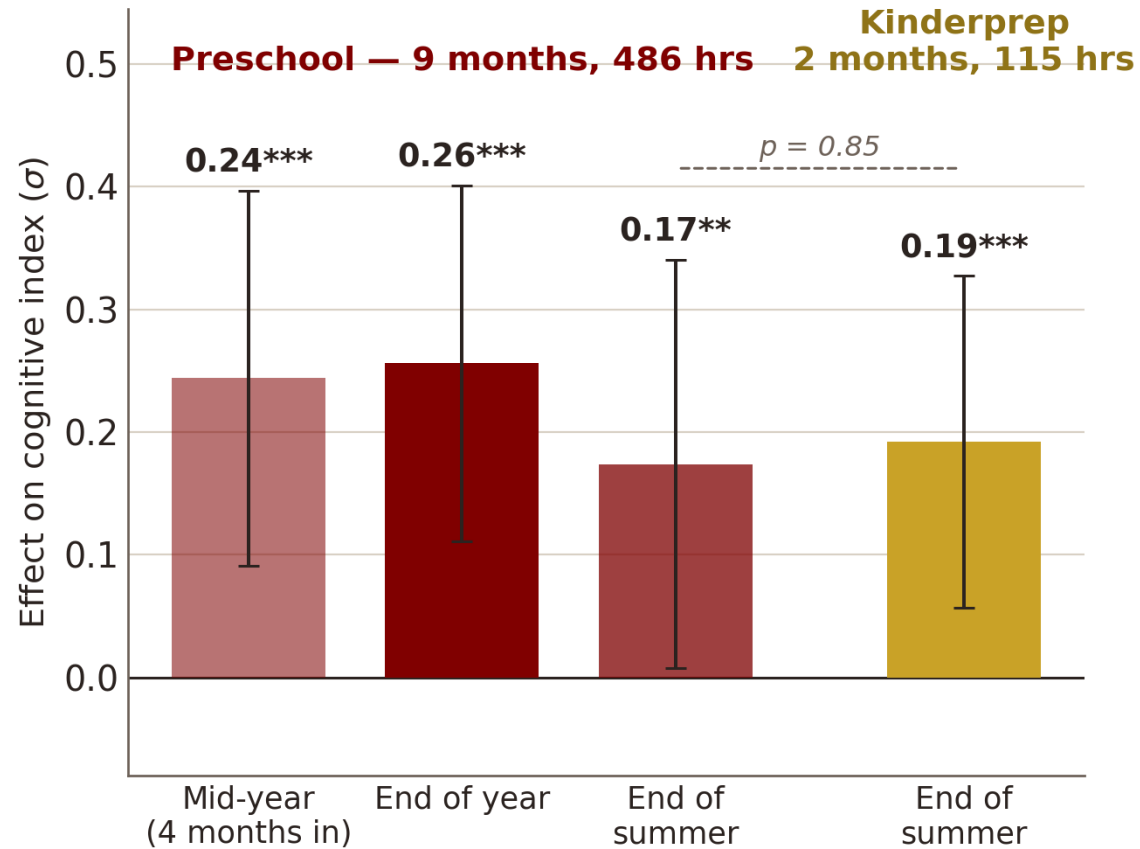
Operation Span (Clancy 1)

- “What color is in this house”?
- “What color WAS in this house”?



CORE RESULTS · THE EVIDENCE

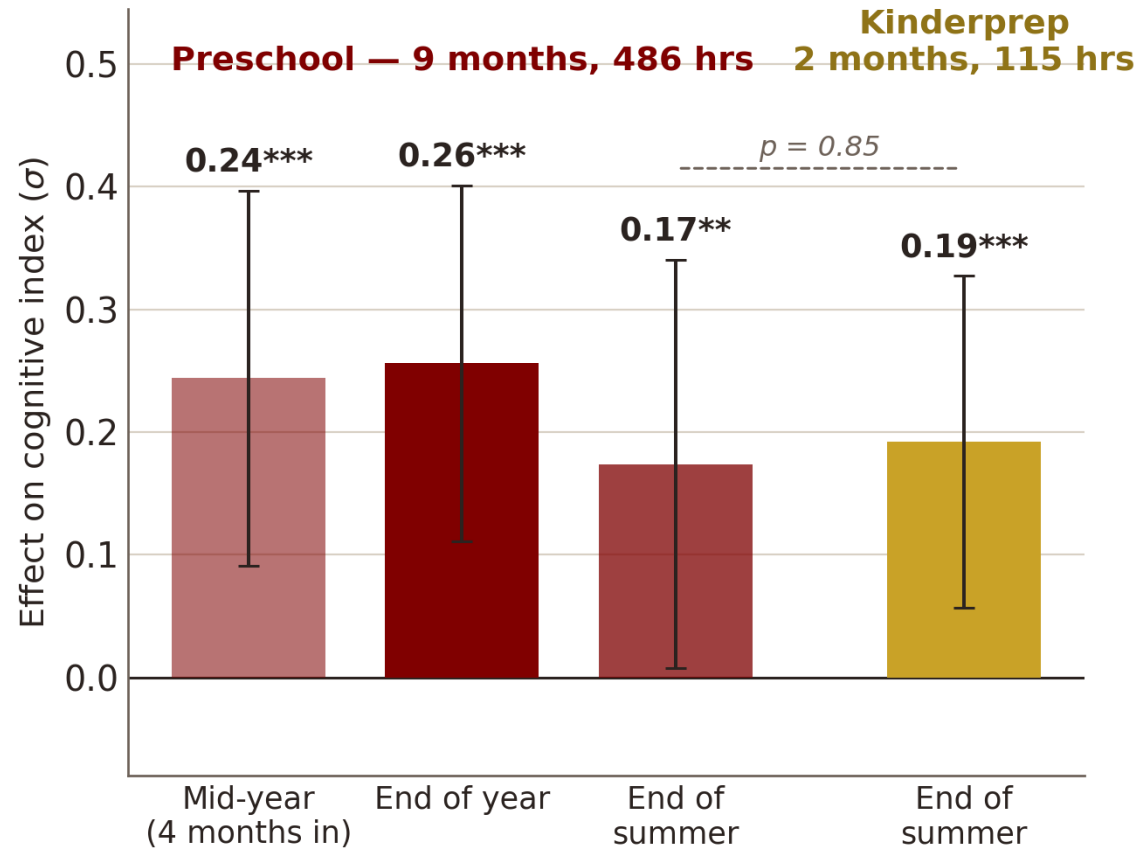
A · Cognitive effects by arm and timing



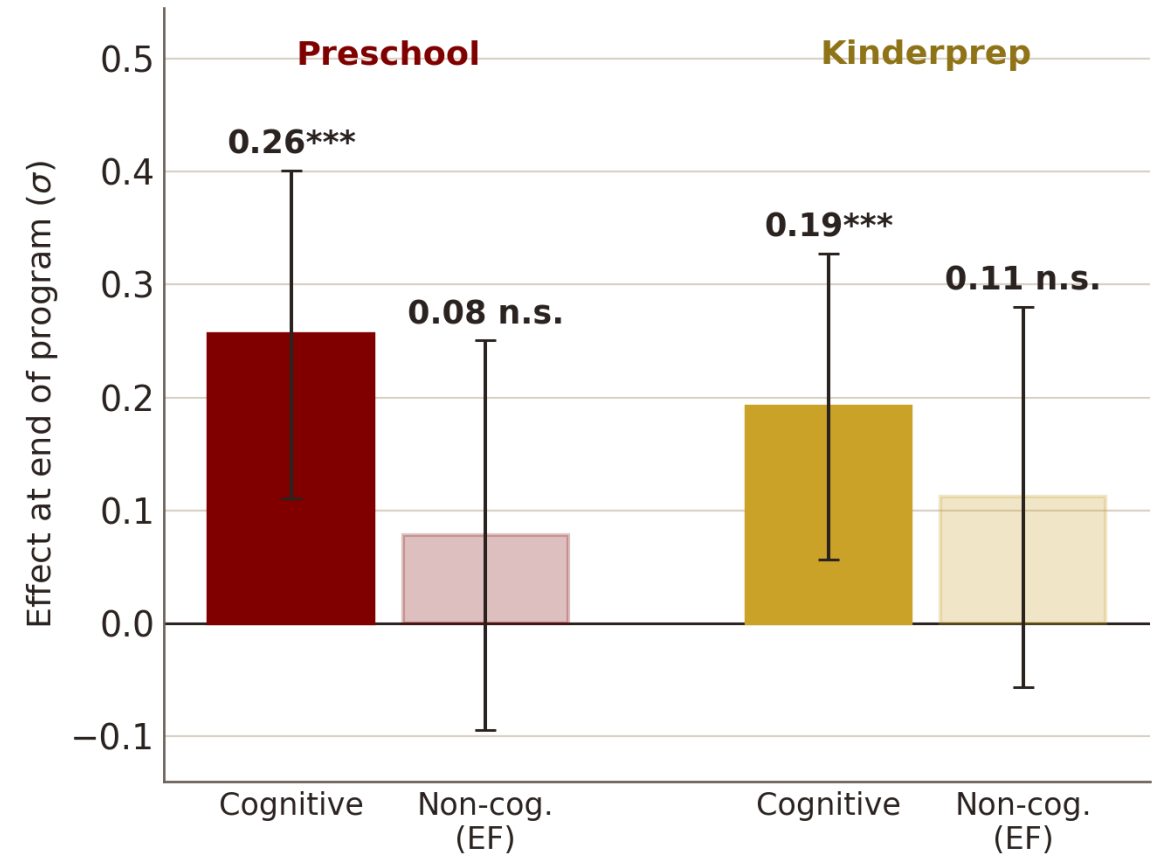
Fryer, Levitt, List & Samek (2020), "Introducing CogX: A New Preschool Education Program Combining Parent and Child Interventions." ITT estimates, full sample (Table B4); whiskers are 95% CIs. ** $p < 0.05$, *** $p < 0.01$.

CORE RESULTS · THE EVIDENCE

A · Cognitive effects by arm and timing



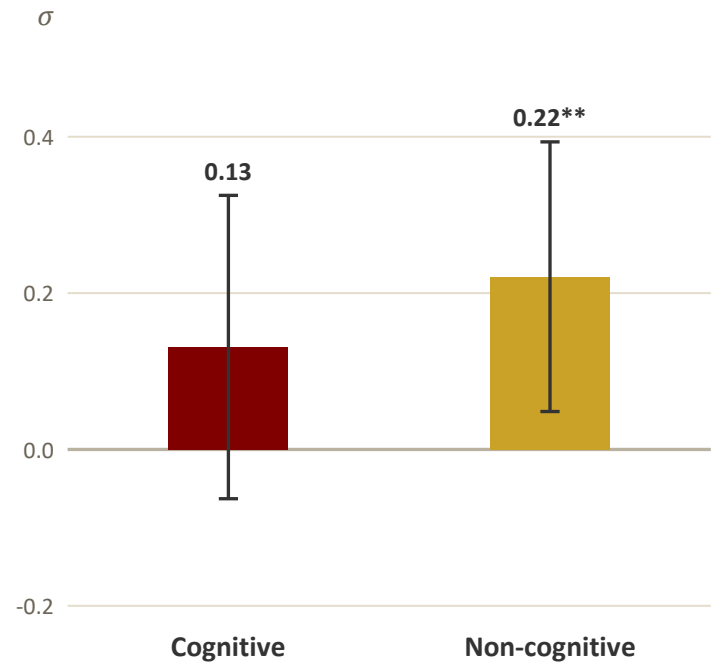
B · Cognitive vs. non-cognitive, end of program



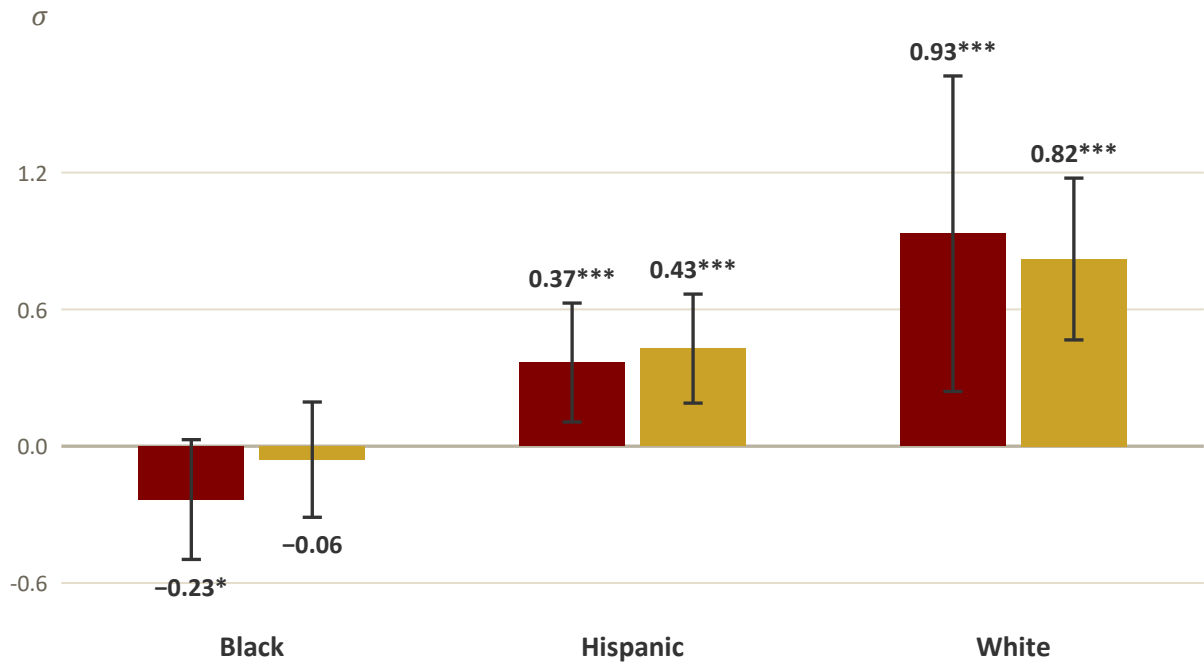
Fryer, Levitt, List & Samek (2020), "Introducing CogX: A New Preschool Education Program Combining Parent and Child Interventions." ITT estimates, full sample (Table B4); whiskers are 95% CIs. ** $p < 0.05$, *** $p < 0.01$.

The Parent Academy moved skills — just not for everyone

What the program moves



For whom



Program Gains, Short Run



$\sim 0.25\sigma$

cognitive gains from a year of preschool



Parents move

Parent Academy shifted beliefs and investments; large gains for Hispanic and White children



Not uniform

Near-zero Parent Academy effects for Black children; heterogeneity is a first-order finding, not a footnote

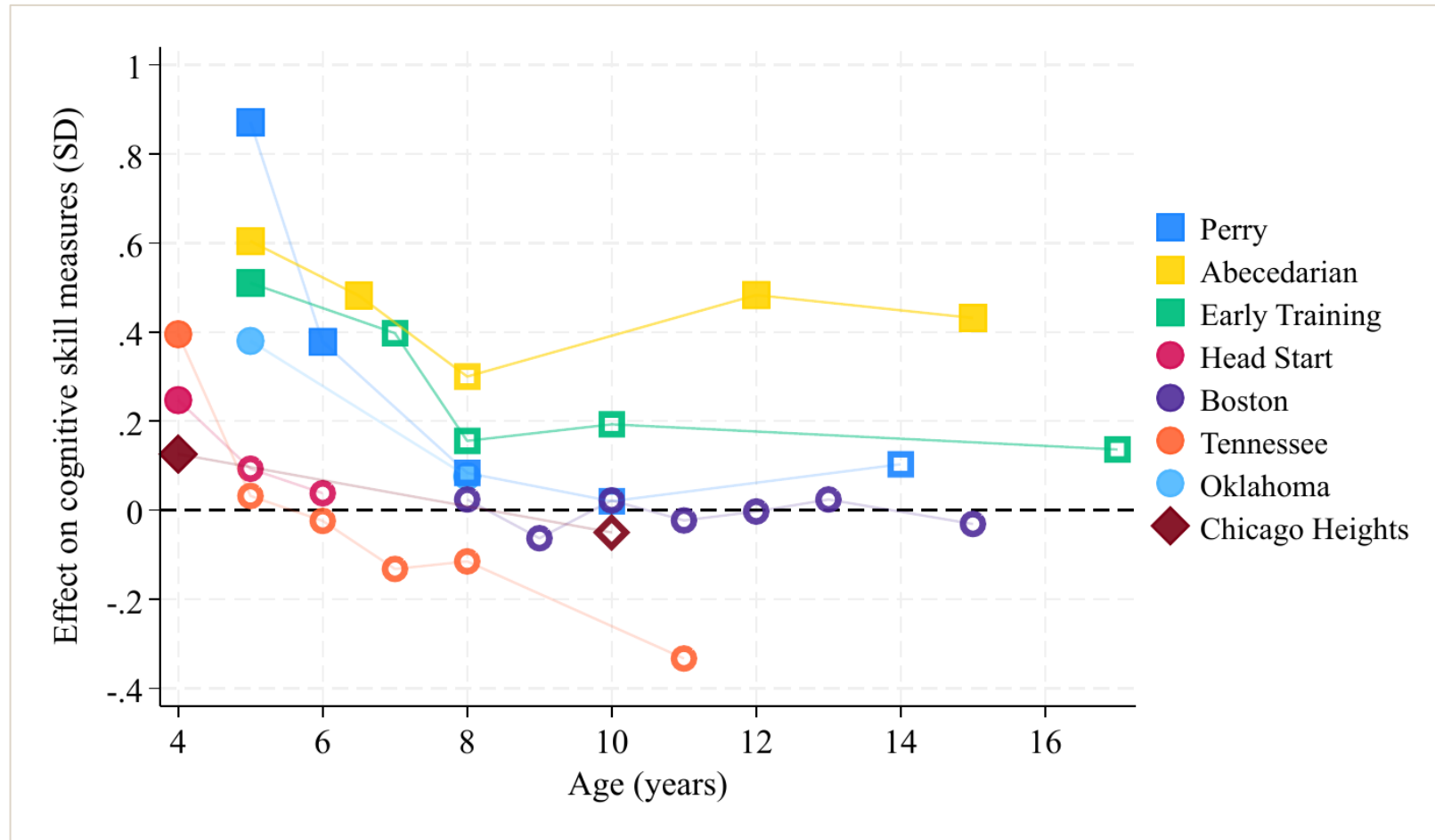


ONE PUZZLE IN THE FIELD

Decorated programs lift test scores. A few years later, the gains are gone. Decades later, the adults are different.

Is fade-out real skill loss? Or, an artifact of what, and whom, and how we measure? In the model's language: is it θ , f , E , or λ or a combination?

Here today, gone tomorrow: gains fade by formal schooling



Across five decades of decorated programs, initial cognitive gains erode by age 8: critics' exhibit A against early education.

CHECC: A Fade-Out Experiment Too

Two stages of randomization. First, children were randomized within CHECC. Then, with School District 170, the same children were randomized to elementary classmates, year after year.

List & Uchida, “Here Today, Gone Tomorrow? Toward an Understanding of Fade-out in Early Childhood Education Programs” (NBER WP 33027, 2024).

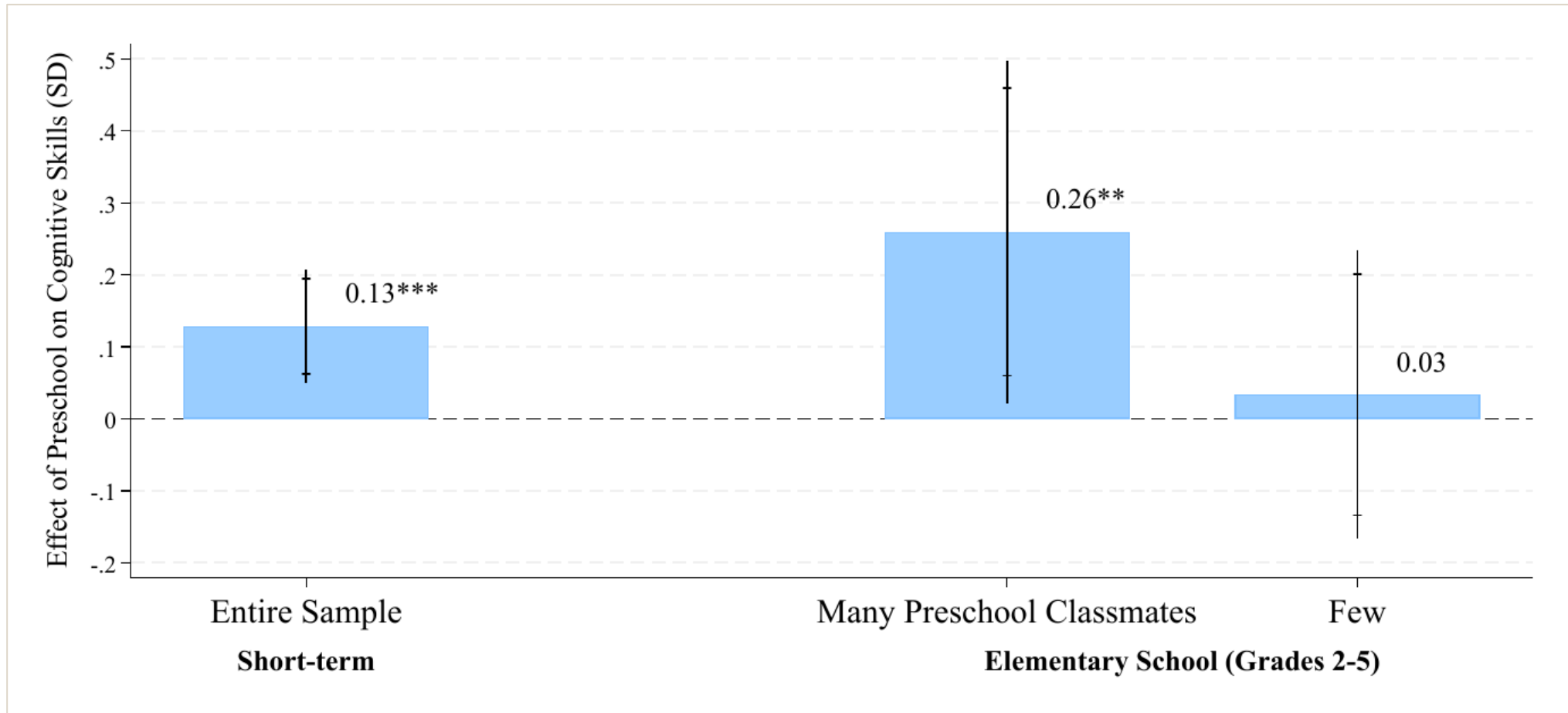


Few treated peers



Many treated peers

Fade-out is not fate: treated classmates preserve the effect



0.13 σ at program end grows to **0.26 σ** with many Preschool classmates — and vanishes (**0.03**) with few. Peers reverse fade-out.

Fade-out lives in the classroom, not the child

The payoff: with enough treated classmates, the classic fade-out effect is muted. Skills persist when the classroom sustains them; gains erode when treated children are scattered among untreated peers.

Fade-out is partly an equilibrium property of the environment we send children into,, which means it is a policy choice.

In the past fade-out has been used to argue against early childhood spending, our results provide an argument for pre-k schooling for all.

List & Uchida, “Here Today, Gone Tomorrow? Toward an Understanding of Fade-out in Early Childhood Education Programs” (NBER WP 33027, 2024).



Few treated peers

Gains erode: the textbook fade-out pattern re-emerges



Many treated peers

Fade-out is muted: skills persist through elementary school

Peers enter the technology through E

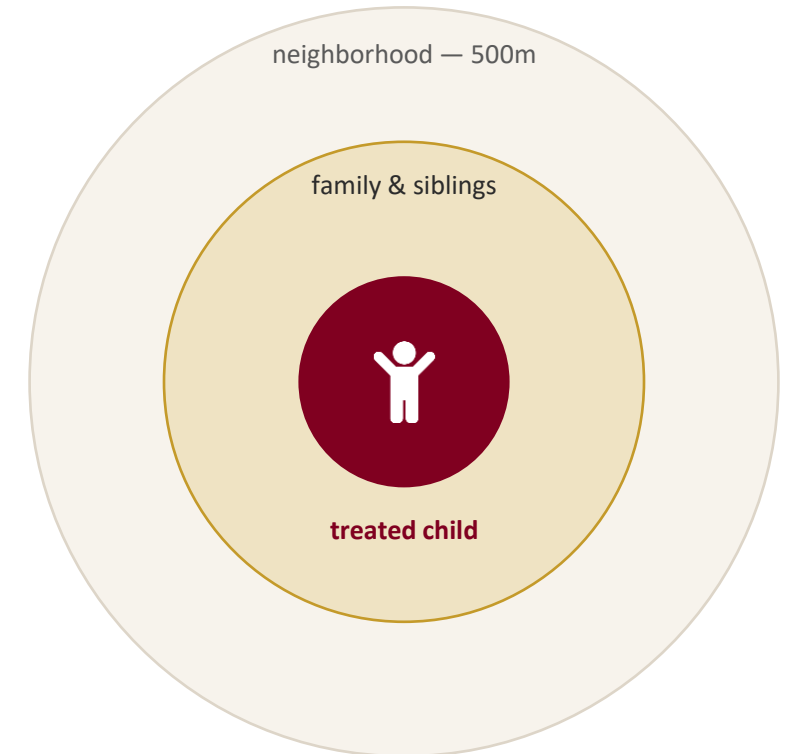
$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale n: } F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: peers	Persistence of θ depends on classroom composition	0.25 σ preschool gains persist when classmates are treated, erode when scattered (List & Uchida)
?	—	—
?	—	—
?	—	—
?	—	—
?	—	—

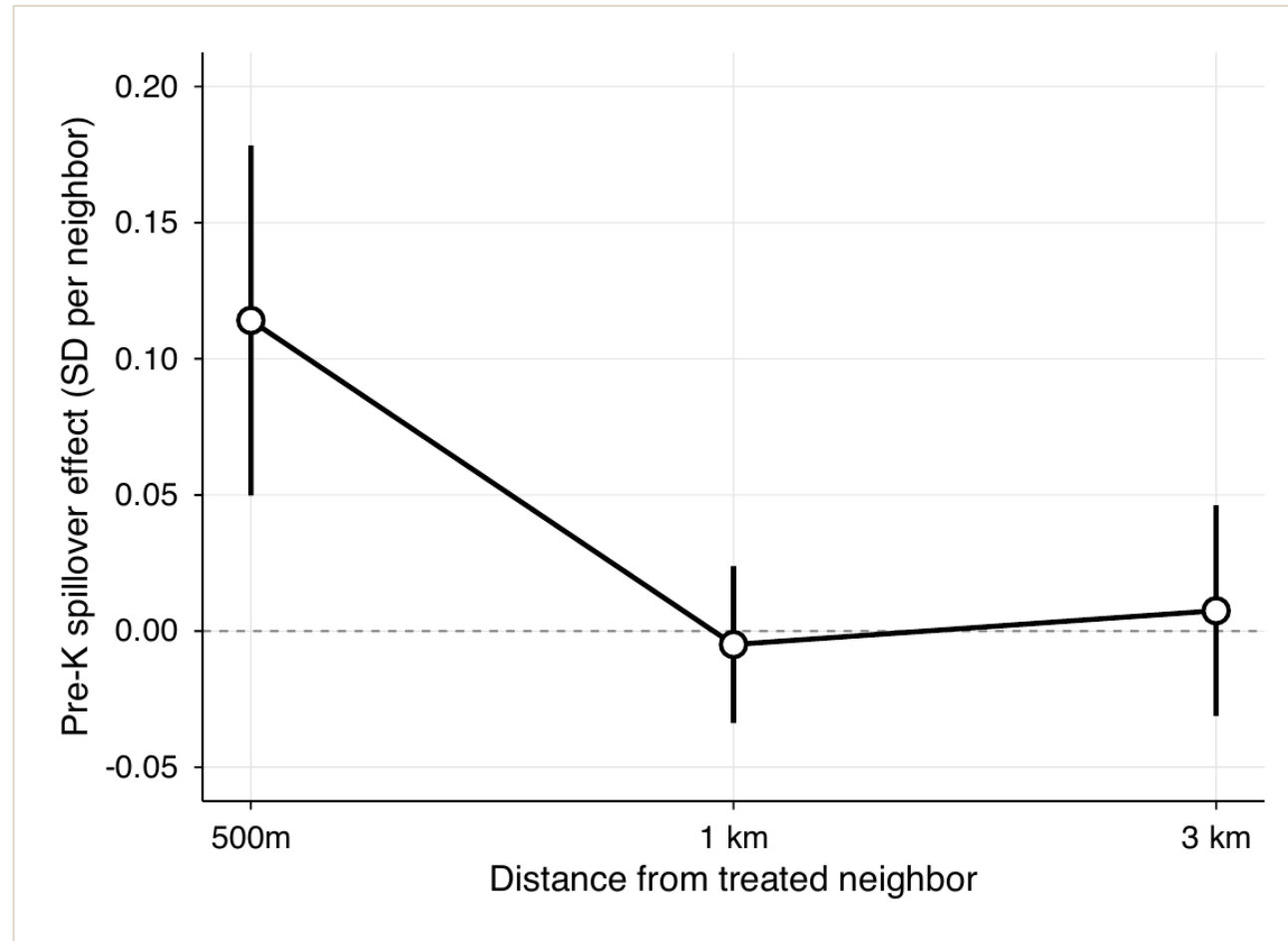
Spillovers in different places: neighborhoods

Because CHECC randomized within a community, we can ask what happens in terms of spillovers.

List, Momeni, Vlassopoulos & Zenou, “Neighborhood Spillover Effects of Early Childhood Interventions” (CEPR DP 18134)



Education is contagious, but only within 500 meters!



0.114 σ per treated Pre-K neighbor within 500m, and zero at 1km and 3km.

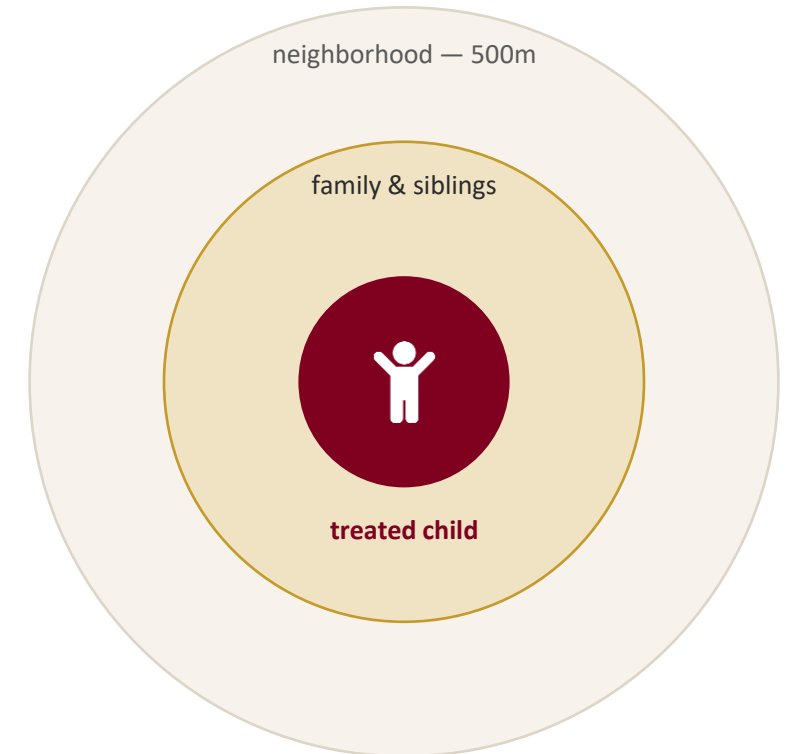
Spillovers change the arithmetic of social returns

Cognitive spillovers of 0.114σ per treated neighbor (mean treated neighbor is ~ 1)

Highly localized, a pattern consistent with local child-to-child (and parent to parent!) interaction rather than broad diffusion.

Siblings of treated children benefited too, echoing the Perry intragenerational findings.

Evaluations that stop at the participant systematically understate the social return.



List, Momeni, Vlassopoulos & Zenou, “Neighborhood Spillover Effects of Early Childhood Interventions” (CEPR DP 18134)

Same object, second network: $\Sigma w\theta$ identified twice

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale } n: F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: $\Sigma w_{ij}\theta_j$	Others' skills enter your production — classrooms and neighborhoods	Gains persist with treated classmates, erode when scattered (List & Uchida); spillovers to nearby non-participants $\approx 38\%$ of the total effect, within 500m (List, Momeni, Vlassopoulos, Zenou)
?	—	—
?	—	—
?	—	—
?	—	—
?	—	—

Same structural term, two networks, two designs — randomized w in classrooms, natural w in neighborhoods. Identified twice, same sign: the spillover sum is real, and the participant is the wrong unit of account.

Latent human capital: the skills were never gone

When Covid shuttered schools, it created the schooling shock no one would ever randomize. CHECC's decade-old randomization was suddenly live again.

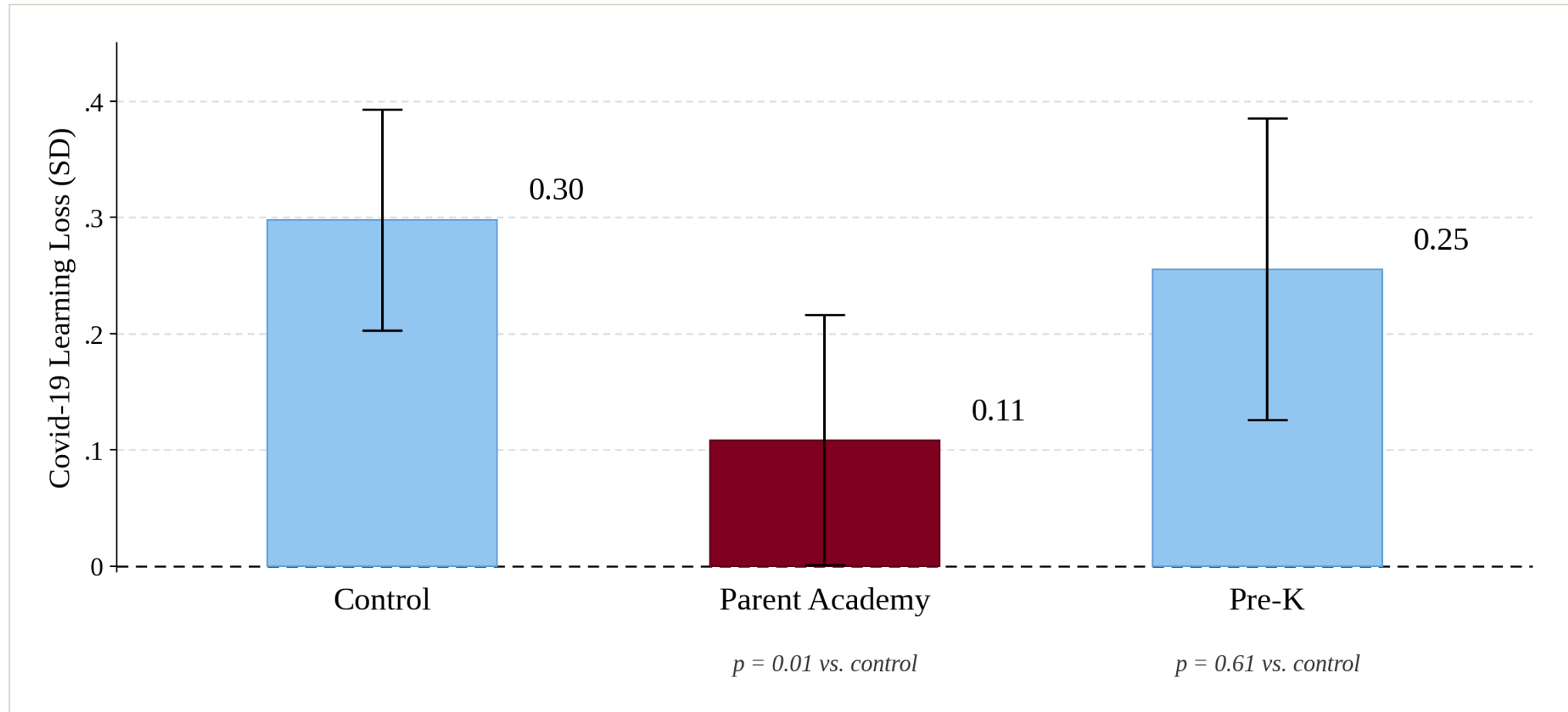
Were any of the CHECC treatment kids different from control?

Jha, List, Royer & Samek, "Latent Human Capital: How Early Investments in Parents Buffer Against Future Schooling Shocks" Coming soon!



Insurance, not just investment?

When schools closed, the decade-old treatment effect reappeared



0.30 σ loss for controls falls to **0.11 σ** under Parent Academy ($p = 0.01$), ten years on. Pre-K: no shield (**0.25 σ** , $p = 0.61$).

Latent human capital: the skills were never gone

Children whose parents received early investment training weathered the shock far better

The early treatment built durable parental capacity that switched on precisely when the school input switched off.

Fade-out on test scores is not evidence the capital is gone. It can sit latent (and buffer) until circumstances call on it.

Jha, List, Royer & Samek, “Latent Human Capital: How Early Investments in Parents Buffer Against Future Schooling Shocks” Coming soon!



Insurance, not just investment

Early parental programs act like an insurance policy on children's human capital — a value entirely invisible to standard test-score evaluations.

P is a stock: built early, it pays off when S fails

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale n: } F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: $\Sigma w_{ij}\theta_j$	Others' skills enter your production — classrooms and neighborhoods	Gains persist with treated classmates, erode when scattered (List & Uchida); spillovers to nearby non-participants $\approx$ 38% of the total effect, within 500m (List, Momeni, Vlassopoulos, Zenou)
P built early	Parental capital persists; substitutes for S when schools fail	Covid closures: control children lose 0.30σ of state test scores, Parent Academy children lose 0.11σ ($p = 0.01$): 8–10 years post-treatment (Jha, List, Royer, Samek)
?	—	—
?	—	—
?	—	—
?	—	—

C is movable by the designer, when effort responds to incentives!

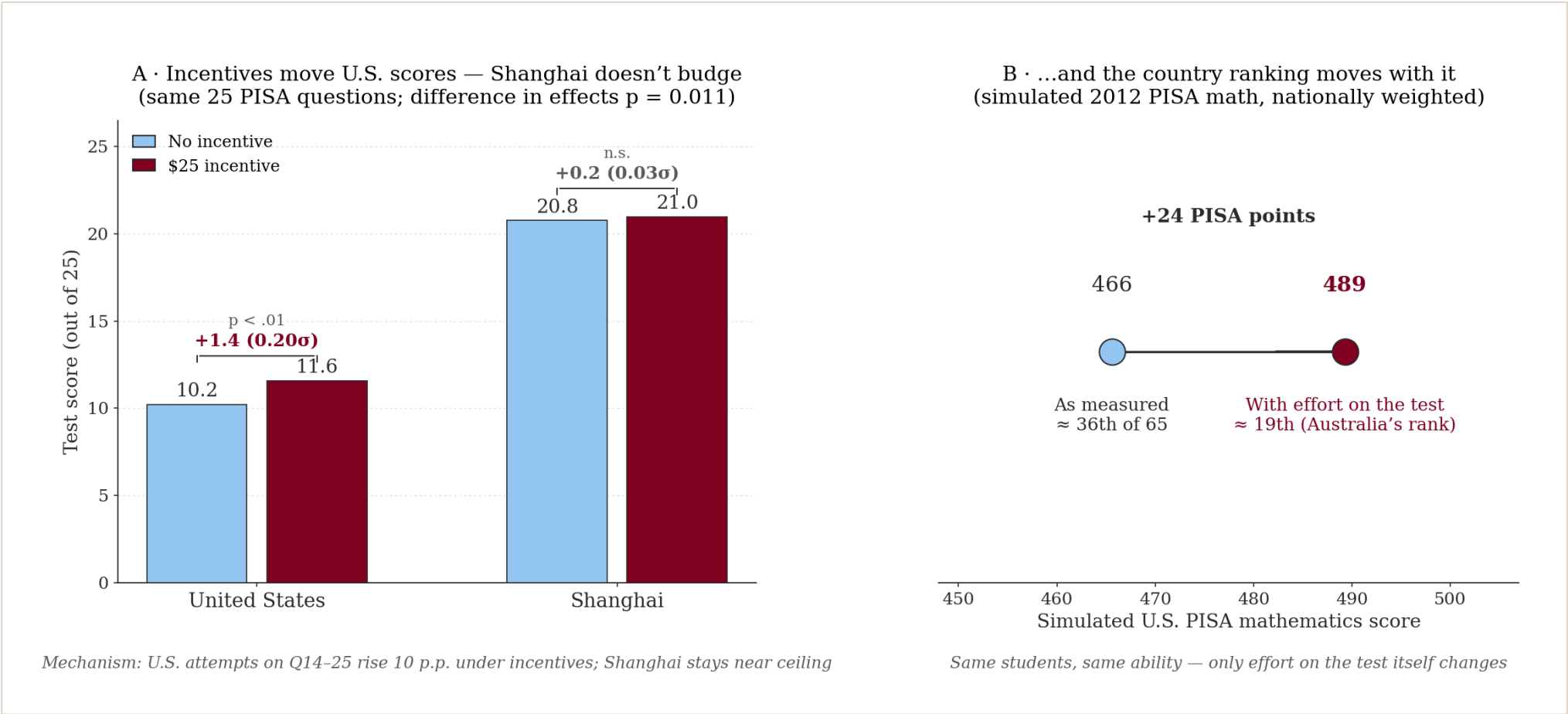
Morning of the test:

Offer surprise “clawback” incentives just before a PISA test: U.S. versus Shanghai students

What happens?

Gneezy, List, Livingston, Qin, Sadoff & Xu, AER: Insights 1(3), 2019.

Test scores measure trying: worth 24 points and 17 ranks



+0.20σ U.S. score gain under a \$25 incentive ($p < .01$); Shanghai: **+0.03σ** n.s. S
imulated PISA: **+24 points** : from 36th to ≈19th of 65, with ability unchanged.

C is chooseable: when effort predictably responds to incentives

Student incentives

Financial and non-financial rewards move student effort, but only when delivered immediately. Loss-framed, prepaid rewards lift math $\sim 0.12\sigma$; introduce any delay and the motivating power vanishes.

Levitt, List, Neckermann & Sadoff, AEJ: Economic Policy, 2016.

No lasting crowd-out

Incentives move incented tests, and any crowd-out of intrinsic motivation is temporary: treated students outperform a year later.

List, Livingston & Neckermann, Economics of Education Review, 2018.

In the model: C can be changed by the designer, the child prices effort, it enters the technology f , and it enters the score Z . Not all of θ shows up unless C does.

C is a choice: the child supplies the effort

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale n: } F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: $\sum w_{ij}\theta_j$	Others' skills enter your production — classrooms and neighborhoods	Gains persist with treated classmates, erode when scattered (List & Uchida); spillovers to nearby non-participants $\approx 38\%$ of the total effect, within 500m (List, Momeni, Vlassopoulos, Zenou)
P built early	Parental capital persists; substitutes for S when schools fail	Covid closures: control children lose 0.30σ of state test scores, Parent Academy children lose 0.11σ ($p = 0.01$) — 8–10 years post-treatment (Jha, List, Royer, Samek)
C: child's effort	The child's own investment: effort enters f, and it is in the score Z	Loss-framed, prepaid rewards lift math $\sim 0.12\sigma$; delay kills the effect (Levitt, List, Neckermann, Sadooff); no lasting crowd-out (List, Livingston, Neckermann); \$25 lifts U.S. scores 0.20σ , Shanghai nil (Gneezy et al.)
?	—	—

Row three filled: effort responds to price, decays with delay — and it lives inside Z. That contamination is exactly where the measurement map comes in next.

Take the Students to the EEG to Explore Executive Function

We built a space-themed EEG lab in a Chicago Heights storefront and measured brain activity of 72 CHECC children (36 preschool, 36 control).

The task: press a button for every animal except the dog. Dog trials force the brain to slam the brakes: inhibitory control, the core of executive function.

The measure: N2, the electrical signature of that braking.

If Preschool children show greater N2 on braking trials , then their inhibitory system is more engaged

Reall that in our measures behavioral EF tests (Fryer et al. 2020) barely moved in pre-school treatment.

Do the hands hide what the cortex is doing?

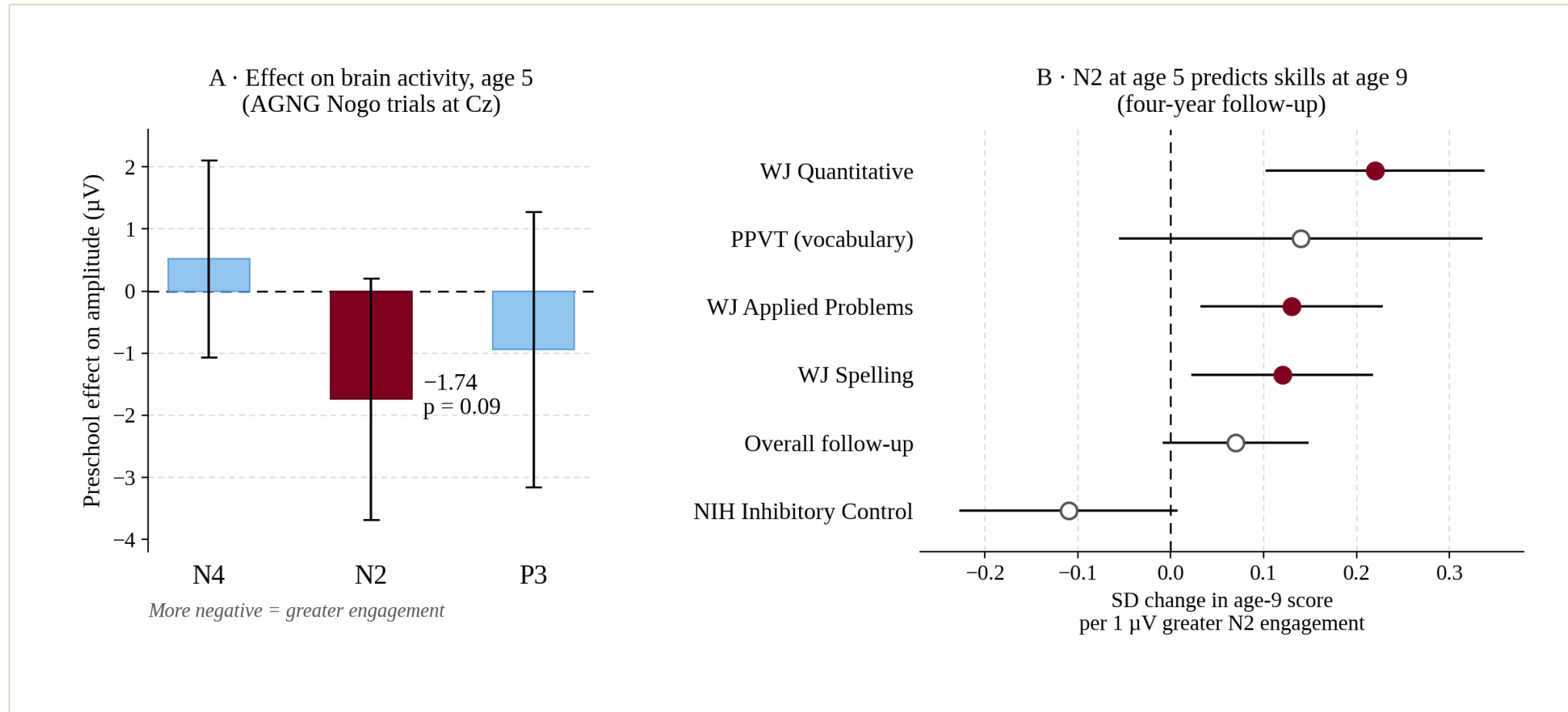
Ye, Samek, Yoder, Decety, Hortaçsu & List, “Early Childhood Programs Change Test Scores but Do They Change Brain Activity?” (working paper).



Why it matters

- Behavioral tests confound skill with motor control and test-day effort. Neural measures see through that.
- EEG biomarkers are candidate statistical surrogates (Athey et al.) : forecasting long-run effects without a 40-year wait.
- This study combines EEG and behavioral tasks to evaluate a preschool intervention.

The N2 moved at age 5, & predicted skills at age 9



1.74 μV deeper N2 under preschool ($p = 0.09$); each μV predicts **+0.22 σ** in age-9 math. Behavioral EF tests saw **nothing**.

Ye, Samek, Yoder, Decety, Hortaçsu & List (working paper), Tables 3–4. $n = 35$ (EEG), 22 (follow-up); whiskers are 95% CIs; filled markers $p < 0.05$.

λ is not the identity map: Z is a proxy, θ is the object

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale } n: F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: $\Sigma w_{ij}\theta_j$	Others' skills enter your production: classrooms and neighborhoods	Gains persist with treated classmates, erode when scattered (List & Uchida); spillovers to nearby non-participants $\approx$ 38% of the total effect, within 500m (List, Momeni, Vlassopoulos, Zenou)
P built early	Parental capital persists; substitutes for S when schools fail	Covid closures: control children lose 0.30σ of state test scores, Parent Academy children lose 0.11σ ($p = 0.01$) — 8–10 years post-treatment (Jha, List, Royer, Samek)
C: child's effort	The child's own investment: effort enters f, and it is in the score Z	Loss-framed, prepaid rewards lift math $\sim 0.12\sigma$; delay kills the effect (Levitt, List, Neckermann, Sadoff); no lasting crowd-out (List, Livingston, Neckermann); \$25 lifts U.S. scores 0.20σ , Shanghai nil (Gneezy et al.)
$\lambda < 1, \varepsilon \neq 0$	Z can fade or stall while θ moves: measurement is not the object	Behavioral EF gains small & n.s. (Fryer et al. 2020); preschoolers show greater N2 during Go-Nogo ($p < 0.1$, $n = 72$); EEG predicts skills 4 yrs later beyond baseline behavioral measures (Ye et al.)
?	—	—
?	—	—

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children. Each experiment identifies an object in the framework.

Social preferences

θ

Randomize early education, then elicit distributive choices: preschool makes children markedly more egalitarian in their fairness views, while the parenting arm tilts them toward efficiency.

In the model: fairness views are coordinates of ϑ , and f moves them. Education changes values, not just skills.

Cappelen, List, Samek & Tungodden, "The Effect of Early Childhood Education on Social Preferences," Journal of Political Economy 128(7), 2020.

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Time preferences

θ

1,200+ children ages 3–12 making incentivized intertemporal choices: patience develops substantially with age, with marked differences across racial groups.

In the model: the discount rate is a coordinate of ϑ with its own developmental path.

Andreoni, Kuhn, List, Samek, Sokal & Sprenger, "Toward an Understanding of the Development of Time Preferences: Evidence from Field Experiments," Journal of Public Economics 177, 2019.

Social preferences

θ

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Risk & ambiguity

 θ

1,400+ children ages 3–15: the famous gender gap in risk aversion is absent in childhood and emerges only in early adolescence, and math skill predicts risk-taking.

In the model: risk attitudes sit in ϑ , and the gender coordinate switches on in adolescence, not at birth.

Andreoni, Di Girolamo, List, Mackevicius & Samek, “Risk Preferences of Children and Adolescents in Relation to Gender, Cognitive Skills, Soft Skills, and Executive Functions,” J. Economic Behavior & Organization 179, 2020.

Social preferences

 θ

Time preferences

 θ

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Preferences → outcomes

g

Skills load on three factors, cognition, executive function, preferences. It is EF and early impatience, not cognition, that predict disciplinary referrals years later.

In the model: this identifies g : which coordinates of ϑ actually carry into Y . Not the ones tests reward.

Castillo, List, Petrie & Samek, "Detecting Drivers of Behavior at an Early Age: Evidence from a Longitudinal Field Experiment," Journal of Political Economy 132(12), 2024.

Social preferences θ **Time preferences** θ **Risk & ambiguity** θ

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Theory of mind

 λ

Disadvantaged children's low theory-of-mind performance improves sharply when other children's presence is made salient. It is partly attention, not capacity. And ToM competence predicts less sharing: strategic minds hold back.

In the model: partly a λ result, measured ToM confounds attention with capacity. The map, not the mind.

Charness, List, Rustichini, Samek & van de Ven, "Theory of Mind Among Disadvantaged Children: Evidence from a Field Experiment," JEBO 166, 2019; Cowell, Samek, List & Decety, "The Curious Relation Between Theory of Mind and Sharing in Preschool Age Children," PLoS ONE 10(2), 2015.

Social preferences

 θ

Time preferences

 θ

Risk & ambiguity

 θ Preferences $\rightarrow$ outcomes

g

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Dishonesty at home

P

Put parents in a die-roll task with real stakes: they cheat, but markedly less when their children are watching, and especially in front of daughters.

In the model: P transmits norms, not just skills: the home writes ϑ 's moral coordinates.

Houser, List, Piovesan, Samek & Winter, "Dishonesty: From Parents to Children," *European Economic Review* 82, 2016.

Social preferences θ **Time preferences** θ **Risk & ambiguity** θ **Preferences $\rightarrow$ outcomes** g **Theory of mind** λ

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Discrimination

 θ_0

In incentivized allocation tasks, discriminatory behavior is detectable already among preschoolers, before formal schooling begins.

In the model: the initial condition ϑ_0 already contains social preferences over groups. Policy starts before school does.

List, List & Samek, "Discrimination Among Pre-School Children: Field Experimental Evidence," Economics Letters 157, 2017.

Social preferences

 θ

Time preferences

 θ

Risk & ambiguity

 θ

Preferences → outcomes

 g

Theory of mind

 λ

Dishonesty at home

 P

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Health behaviors

f

Small incentives lift healthy snack choice from 17% to roughly 75%, with evidence of habit formation after incentives end.

In the model: incentives are an input in f , and habit formation is $\partial f / \partial \theta$ at work on behavior.

List & Samek, “The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption,” Journal of Health Economics 39, 2015.

Social preferences	θ
Time preferences	θ
Risk & ambiguity	θ
Preferences → outcomes	g
Theory of mind	λ
Dishonesty at home	P
Discrimination	θ_0

CHECC breathes: θ is wider than test scores

Because children came to us daily, CHECC doubled as a behavioral laboratory. We elicited preferences, social behavior, and moral behavior from thousands of children — and each experiment identifies an object in the framework. Watch the badges: this is the census of θ and its neighbors.

Social preferences

θ

Preschool makes children more egalitarian; the parenting arm shifts them toward efficiency.

Cappelen, List, Samek & Tungodden, "The Effect of Early Childhood Education on Social Preferences," Journal of Political Economy 128(7), 2020.

Time preferences

θ

1,200+ children ages 3–12: patience develops substantially with age, with marked racial differences.

Andreoni, Kuhn, List, Samek, Sokal & Sprenger, "Toward an Understanding of the Development of Time Preferences: Evidence from Field Experiments," Journal of Public Economics 177, 2019.

Risk & ambiguity

θ

The gender gap in risk aversion is absent in childhood — it emerges in early adolescence.

Andreoni, Di Girolamo, List, Mackevicius & Samek, "Risk Preferences of Children and Adolescents in Relation to Gender, Cognitive Skills, Soft Skills, and Executive Functions," J. Economic Behavior & Organization 179, 2020.

Preferences → outcomes

g

EF and early impatience — not cognition — predict disciplinary referrals years later.

Castillo, List, Petrie & Samek, "Detecting Drivers of Behavior at an Early Age: Evidence from a Longitudinal Field Experiment," Journal of Political Economy 132(12), 2024.

Theory of mind

λ

Disadvantaged children's low ToM improves sharply when other children's presence is made salient — and ToM predicts less sharing, not more.

Charness, List, Rustichini, Samek & van de Ven, "Theory of Mind Among Disadvantaged Children: Evidence from a Field Experiment," JEBO 166, 2019; Cowell, Samek, List & Decety, "The Curious Relation Between Theory of Mind and Sharing in Preschool Age Children," PLoS ONE 10(2), 2015.

Dishonesty at home

P

Parents cheat less when their children are watching — especially daughters. Honesty transmission starts at home.

Houser, List, Piovesan, Samek & Winter, "Dishonesty: From Parents to Children," European Economic Review 82, 2016.

Discrimination

θ_0

Discriminatory behavior is detectable in incentivized tasks already among preschoolers — before formal schooling begins.

List, List & Samek, "Discrimination Among Pre-School Children: Field Experimental Evidence," Economics Letters 157, 2017.

Health behaviors

f

Small incentives lift healthy snack choice from 17% to ~75%, with evidence of habit formation.

List & Samek, "The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption," Journal of Health Economics 39, 2015.

Many parts of “skills” carry into Y

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale } n: F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: $\Sigma w_{ij}\theta_j$	Others' skills enter your production: classrooms and neighborhoods	Gains persist with treated classmates, erode when scattered (List & Uchida); spillovers to nearby non-participants $\approx$ 38% of the total effect, within 500m (List, Momeni, Vlassopoulos, Zenou)
P built early	Parental capital persists; substitutes for S when schools fail	Covid closures: control children lose 0.30σ of state test scores, Parent Academy children lose 0.11σ ($p = 0.01$) — 8–10 years post-treatment (Jha, List, Royer, Samek)
C: child’s effort	The child’s own investment: effort enters f, and it is in the score Z	Loss-framed, prepaid rewards lift math $\sim 0.12\sigma$; delay kills the effect (Levitt, List, Neckermann, Sadoff); no lasting crowd-out (List, Livingston, Neckermann); \$25 lifts U.S. scores 0.20σ , Shanghai nil (Gneezy et al.)
$\lambda < 1, \varepsilon \neq 0$	Z can fade or stall while θ moves: measurement is not the object	Behavioral EF gains small & n.s. (Fryer et al. 2020); preschoolers show greater N2 during Go-Nogo ($p < 0.1$, $n = 72$); EEG predicts skills 4 yrs later beyond baseline behavioral measures (Ye et al.)
g: $\theta \rightarrow Y$	Which coordinates of θ cash out: and which choices they drive	1σ of EF $\rightarrow -1.9$ pp referral probability, 1σ of patience $\rightarrow -2.3$ pp (14.7% base); cognition: nil (Castillo, List, Petrie, Samek); preferences & beliefs, not ability, drive STEM sorting (Fiala et al.)
The living lab filled g: EF and preferences, not solely test scores, are what carry into outcomes.		

Scaling lives inside the model, not after it

Five vital signs, five objects. A voltage drop is not one wedge: each vital sign attacks a different part of the model, and only the supply-side trap runs through S.

False positives live in the estimate, not the technology. Unrepresentative samples shift the population $F(n)$ you draw from, so heterogeneity bites. Unrepresentative situations change f itself. Costs hit $C(n)$ directly.

CHECC's additional ledger entry: the spillover vital sign can run positive. Peer density $E(n)$ rises with coverage, fade-out mutes when classmates are treated, and neighborhood effects grow with saturation. Whether scaling drops or gains voltage is an empirical horse race, object by object.

The design implication: Option C Thinking (List, 2024, Nature) backward induct from the scale you intend, choose samples, situations, and dosages so the pilot estimates the policy parameter, not the pilot parameter.

False positives

$\hat{\Delta}$ — the inference, not the technology

Unrepresentative sample

$F(n)$ — who you draw shifts; heterogeneity bites

Unrepresentative situation

f_n — the technology changes with delivery

Supply side & cost trap

$S(n)$ quality falls; $C(n)$ marginal costs rise

Spillovers & GE

$E(n)$ — CHECC: this one can run positive

Scaling is an argument at the beginning, not the end

$$\theta_{i,t+1} = f_t(\theta_{i,t}, P_{i,t}, S_{i,t}, C_{i,t}, E_{i,t}) \quad Z_{i,t} = \lambda_t \theta_{i,t} + \varepsilon_{i,t} \quad Y_i = g(\theta_{i,T}) \quad \text{at scale } n: F(n), f_n, S(n), E(n), C(n)$$

Object	Economic content	Result it organizes
E: $\Sigma w_{ij}\theta_j$	Others' skills enter your production: classrooms and neighborhoods	Gains persist with treated classmates, erode when scattered (List & Uchida); spillovers to nearby non-participants $\approx 38\%$ of the total effect, within 500m (List, Momeni, Vlassopoulos, Zenou)
P built early	Parental capital persists; substitutes for S when schools fail	Covid closures: control children lose 0.30σ of state test scores, Parent Academy children lose 0.11σ ($p = 0.01$): 8–10 years post-treatment (Jha, List, Royer, Samek)
C: child's effort	The child's own investment: effort enters f , and it is in the score Z	Loss-framed, prepaid rewards lift math $\sim 0.12\sigma$; delay kills the effect (Levitt, List, Neckermann, Sadoff); no lasting crowd-out (List, Livingston, Neckermann); \$25 lifts U.S. scores 0.20σ , Shanghai nil (Gneezy et al.)
$\lambda < 1, \varepsilon \neq 0$	Z can fade or stall while θ moves: measurement is not the object	Behavioral EF gains small & n.s. (Fryer et al. 2020); preschoolers show greater N2 during Go-Nogo ($p < 0.1$, $n = 72$); EEG predicts skills 4 yrs later beyond baseline behavioral measures (Ye et al.)
$g: \theta \rightarrow Y$	Which coordinates of θ cash out: and which choices they drive	1σ of EF $\rightarrow -1.9$ pp referral probability, 1σ of patience $\rightarrow -2.3$ pp (14.7% base); cognition: nil (Castillo, List, Petrie, Samek); preferences & beliefs, not ability, drive STEM sorting (Fiala et al.)
n: scale	Five vital signs, five objects: inference on Δ , sample $F(n)$, technology f_n , inputs $S(n)$ & costs $C(n)$, environment $E(n)$	Dosage: 2 summer months $\approx$ an academic year ($\sim 0.25\sigma$); whether $V(n)$ rises or falls is $S(n)$ vs $E(n)$ — an empirical horse race (List, JPE 2026; Nature 2024)

The map is complete: environment E on two networks, parents P , the child's effort C , measurement λ , outcome map g , and scale n . One system, a dozen papers, no black box.

It's not over at five: skill compounds, and efficiency divides

Why doesn't adolescent Johnny do well in school?

Cotton et al.: model of time allocation and academic efficiency.

Black and Hispanic students are as motivated as White/Asian peers, but academic efficiency runs $\sim 1\sigma$ (Black) and $\sim 0.7\sigma$ (Hispanic) lower, conditional on school, family, and neighborhood. Schools act on efficiency: the best-to-worst-resourced district gap predicts a 0.6σ drop.

Cotton, Hickman, List, Price & Roy, "Why Don't Struggling Students Do Their Homework? Disentangling Motivation and Study Productivity as Drivers of Human Capital Formation," Journal of Political Economy, 2026.

Joensen et al.: skill begets skill; 18 classroom hours lift cognition $+0.12\sigma$ (Math) and EF $+0.13\sigma$ (Goals). Seven years on, +6.3 p.p. 4-year college enrollment off a 25% base.

Joensen, List, Samek & Uchida, "Using a Field Experiment to Understand Skill Formation During Adolescence" (coming soon)

Teacher Incentives, effort, and the anatomy of gaps could fill another table!

The lab also ran on the adults and institutions around the children: identifying the inputs P and S, the outcome map g, and the distribution F. The census continues.

Teacher incentives S Loss-framed prepaid bonuses raise math achievement ~0.12σ; conventional gain-framed bonuses do little. Fryer, Levitt, List & Sadoff, “Enhancing the Efficacy of Teacher Incentives through Framing: A Field Experiment,” AEJ: Economic Policy 14(4), 2022.	Racial Excellence gaps F The racial excellence gap among top achievers appears by 3rd grade and grows; average-gap policies barely move it. Karna, Lee, List, Simon & Uchida, “Measuring Racial Educational Disparities Over Time Amongst Top Achievers,” Nature 639, 2025.	Gender gaps start at home P Differential early parental investments help explain where the math and literacy gender gaps originate. Chuan, List, Samek & Samujjwala, “Parental Investments in Early Childhood and the Gender Gap in Math and Literacy,” AEA Papers & Proceedings 112, 2022.	Preferences → pipelines g Early gender differences in preferences and beliefs — not ability — drive high-school STEM sorting and college major gaps. Fiala, Humphries, Joensen, Karna, List & Veramendi, “How Early Adolescent Skills and Preferences Shape Economics Education Choices,” AEA Papers & Proceedings 112, 2022.
Beliefs are the root P Low-SES parents underestimate the returns to investment; beliefs are malleable, and intensive programs move investments and skills. List, Pernaudet & Suskind, “Shifting Parental Beliefs About Child Development to Foster Parental Investments and Improve School Readiness Outcomes,” Nature Communications 12, 2021.	First Gen Excellence gaps F The First-gen excellence gap among top achievers appears by 1st grade and grows Karna, List, Simon & Uchida, “Th First-Gen Excellence Gap, Nature Human Behavior, forthcoming. .		

Same question, different machine

Dimension	Perry (1962–67)	CHECC (2010–2014; experiments until 2026)
Sample	123 children	~3,000 children
Treatment arms	1 bundled (preschool + home visits)	4 arms: Preschool, Parent Academy, Kinderprep, Control
Mechanism data	None at treatment	Preferences, parental beliefs, EEG brain activity
Spillovers	Not measured	Neighborhood (500m) + classroom peer effects
Parental channel	Undecomposed home visits	Separate Parent Academy arm isolates the channel
Scaling evidence	None built in	Kinderprep dosage, district-led replications, voltage framework
Later-life arc	40-year wait	Tracked in real time; adolescent interventions extend the arc

Standard evaluations get the returns wrong in both directions



They overestimate when...

- Test-score gains fade with untreated peers and get read as durable skill
- Endogenous compliance inflates factorial treatment estimates
- Test-day effort, not skill, moves the measured outcome
- Pilot inputs are the best inputs: $S(n)$ degrades as supply stretches at scale



...and underestimate when

- Spillovers to neighbors and siblings are left out of the ledger
- Latent parental capital only pays off when shocks arrive
- Neural changes precede, and predict, what behavioral tests eventually detect
- Peer density $E(n)$ strengthens with coverage; universal provision mutes fade-out

One platform, six lessons

FOR POLICYMAKERS

Buy the system, not the pilot.

Fund at coverage levels that activate peer effects: universal provision changes efficiency itself, not just equity. Never cancel on 3rd-grade fade-out: latent capital and spillovers are off the test. Commission evidence at the scale you intend.

FOR ACADEMICS

Build platforms, not projects.

A randomization is an asset that appreciates: a 2010 coin flip became a 2026 paper when Covid hit. One-shot RCTs estimate a number; platforms identify the objects — f , P , E , λ , n .

FOR PRACTITIONERS

The classroom is tunable.

Dosage is a lever: two summer months $\approx$ an academic year, and gains front-load in four months. Composition sustains skills. Reward immediately; frame teacher bonuses as losses; teach EF, not just content.

FOR PARENTS

You are the durable input.

The binding constraint is your belief that investment matters, it does. And, for longer than you think: parental capital built at age 4 buffered a shock at age 14. And children copy what they watch: honesty, patience, sharing.

FOR EXPERIMENTALISTS

Design for the violations.

Randomize environments, not just treatments. Measure the non-participants; if in the design, SUTVA violations are findings, not nuisances. Put the lab where people live, and fifty studies ride one platform at near-zero marginal cost.

FOR METHODS PEOPLE

Respect the measurement map.

$Z \neq \theta$: effort, motor control, and $\lambda < 1$ all sit between skill and score. Bring MHT corrections, compliance-aware factorial estimators, and surrogate indices.

Where to next: when λ becomes an input, f changes

In a state-wide test across 1-3 year olds in Connecticut.....



Where to next: when λ becomes an input, f changes

Our Model: Teacher allocates rival directed speech and nonrival ambient words across children under an information friction. Think mixed private/public good model.

Treatment operates on three margins: it moves the classroom from inside its production possibilities frontier to the frontier, expands the frontier, and moves the allocation along it toward the inclusive point.

The arc closes where it began: the deck opened with λ as a ruler. Where we are headed, λ is an input the teacher holds, and every classroom is a living lab.

*List, Pernaudet, Royer, & Suskind, Making the Invisible Visible:
Toward an Understanding of the Effects of Classroom Interactions.*

There is a New Landscape Emerging

- Yesterday's economist visited the pin factory to scribe general economic theories.
- Today's economist models natural disruptions that affect the pin factory and occasionally generates data from the pin factory to test economic theories.
- Tomorrow's economist (ALL OF US) will surgically control the assignment mechanism, in an ethically responsible way, to provide causal evidence of the inner workings of the pin factory *and* simultaneously generate data necessary to understand if learnings from that environment will transfer to other environments, whether spatially, temporally, or scale differentiated.

The platform, the mechanisms, and scaling

Fryer, Levitt & List (2015), “Parental Incentives and Early Childhood Achievement: A Field Experiment in Chicago Heights,” NBER WP 21477.

Fryer, Levitt, List & Samek (2020), “Introducing CogX: A New Preschool Education Program Combining Parent and Child Interventions,” NBER WP 27913.

List & Uchida (2024), “Here Today, Gone Tomorrow? Toward an Understanding of Fade-out in Early Childhood Education Programs,” NBER WP 33027.

Jha, List, Royer & Samek (2026), “Latent Human Capital: How Early Investments in Parents Buffer Against Future Schooling Shocks.”

Ye, Samek, Yoder, Decety, Hortaçsu & List, “Early Childhood Programs Change Test Scores but Do They Change Brain Activity?” Working paper.

List, Momeni, Vlassopoulos & Zenou (2023), “Neighborhood Spillover Effects of Early Childhood Interventions,” CEPR DP 18134 / NBER WP 28283; R&R, AEJ: Economic Policy.

Castillo, List, Petrie & Samek (2024), “Detecting Drivers of Behavior at an Early Age: Evidence from a Longitudinal Field Experiment,” J. Political Economy 132(12).

Chuan, List & Samek (2021), “Do Financial Incentives Aimed at Decreasing Interhousehold Inequality Increase Intrahousehold Inequality?” J. Public Economics 196.

Joensen, List, Samek & Uchida, “Using a Field Experiment to Understand Skill Formation During Adolescence.”

Cotton, Hickman, List, Price & Roy, “Why Don't Struggling Students Do Their Homework? Disentangling Motivation and Study Productivity as Drivers of Human Capital Formation,” J. Political Economy.

List (2026), “The Economics of Scaling Early Childhood Programs: Lessons from the Chicago School,” J. Political Economy 134(1).

List (2024), “Optimally Generate Policy-Based Evidence Before Scaling,” Nature 626, 491–499.

List, Petrie & Samek (2023), “How Experiments with Children Inform Economics,” J. Economic Literature 61(2).

Heckman, Moon, Pinto, Savelyev & Yavitz (2010), “The Rate of Return to the HighScope Perry Preschool Program,” J. Public Economics 94.

The living lab

Cappelen, List, Samek & Tungodden (2020), “The Effect of Early Childhood Education on Social Preferences,” J. Political Economy 128(7).

Andreoni, Kuhn, List, Samek, Sokal & Sprenger (2019), “Toward an Understanding of the Development of Time Preferences: Evidence from Field Experiments,” J. Public Economics 177.

Andreoni, Di Girolamo, List, Mackevicius & Samek (2020), “Risk Preferences of Children and Adolescents in Relation to Gender, Cognitive Skills, Soft Skills, and Executive Functions,” J. Economic Behavior & Organization 179.

Charness, List, Rustichini, Samek & van de Ven (2019), “Theory of Mind Among Disadvantaged Children: Evidence from a Field Experiment,” J. Economic Behavior & Organization 166.

Cowell, Samek, List & Decety (2015), “The Curious Relation Between Theory of Mind and Sharing in Preschool Age Children,” PLoS ONE 10(2).

Houser, List, Piovesan, Samek & Winter (2016), “Dishonesty: From Parents to Children,” European Economic Review 82.

List, List & Samek (2017), “Discrimination Among Pre-School Children: Field Experimental Evidence,” Economics Letters 157.

List & Samek (2015), “The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption,” J. Health Economics 39.

Fryer, Levitt, List & Sadoff (2022), “Enhancing the Efficacy of Teacher Incentives through Framing: A Field Experiment,” AEJ: Economic Policy 14(4).

Levitt, List, Neckermann & Sadoff (2016), “The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance,” AEJ: Economic Policy 8(4).

List, Livingston & Neckermann (2018), “Do Financial Incentives Crowd Out Intrinsic Motivation to Perform on Standardized Tests?” Economics of Education Review 66.

Gneezy, List, Livingston, Qin, Sadoff & Xu (2019), “Measuring Success in Education: The Role of Effort on the Test Itself,” AER: Insights 1(3).

Karna, Lee, List, Simon & Uchida (2025), “Measuring Racial Educational Disparities Over Time Amongst Top Achievers,” Nature 639.

Chuan, List, Samek & Samujjwala (2022), “Parental Investments in Early Childhood and the Gender Gap in Math and Literacy,” AEA Papers & Proceedings 112.

Fiala, Humphries, Joensen, Karna, List & Veramendi (2022), “How Early Adolescent Skills and Preferences Shape Economics Education Choices,” AEA Papers & Proceedings 112.

List, Pernaudet & Suskind (2021), “Shifting Parental Beliefs About Child Development to Foster Parental Investments and Improve School Readiness Outcomes,” Nature Communications 12.