

# Education, Data, and Policy: Reflections on the Past, Present, and Future

James J. Heckman  
Department of Economics  
& Center for the Economics of Human Development (CEHD):  
The University of Chicago

Inaugural UNESCO Conference on Education Data and Statistics  
(2024)

February 7th, 2024





CENTER FOR THE ECONOMICS  
OF HUMAN DEVELOPMENT  
THE UNIVERSITY OF CHICAGO

[cehd.uchicago.edu](http://cehd.uchicago.edu)



**It is an honor to be asked to address you  
today.**

*Education...beyond all other devices of human origin, is a great equalizer of conditions of men –the balance wheel of the social machinery...It does better than to disarm the poor of their hostility toward the rich; it prevents being poor.*

—Horace Mann (1838)

[Link to Figures](#)

- Common frameworks of measurement and analysis: Examples

① Mincer equation:  $\underbrace{\ln Y}_{\text{earning}} = \alpha + \beta \underbrace{S}_{\text{schooling}} + \gamma \underbrace{E}_{\text{work experience}} + U$

② Occupational and educational choice, and choices of schooling, location, occupation, etc.

③ IGE:  $Y_C = \beta_0 + \beta Y_P + \varepsilon_C$

④ Peer networks and peer effects (social interactions and learning)

- Develop tools and models to analyze existing data and devise new methods to collect new data on dimensions of learning not yet explored.

# Life Cycle Learning



# I. The Skills That Matter

- Many major economic and social problems can be traced to low levels of skill and ability in the population.
- Inventories of skills essential in designing policy.
- Abilities and skills are **multiple** in nature.
- Considerable public policy discussion focuses on cognitive ability (IQ) or PISA scores.
- Yet **noncognitive** abilities are also very important. Especially important is the ability to cope with change.
- Need to develop useful inventories of skill

- Achievement tests widely used to screen, sort and monitor the success of schools and society. PISA tests; No Child Left Behind; Iowa Tests; NAEP tests; and the GED (General Educational Development); LSAT; SAT.
- They miss noncognitive (personality) traits.

- These are sometimes called “soft skills” or “character” traits.
- Perseverance, conscientiousness, motivation, willful planning.
- They are predictive of a range of important behaviors.
- Ignoring personality and character is a dangerous practice and can lead to costly mistakes in assessing and addressing social problems and in evaluating the success or failure of economic and social policies.

- Measurement systems overly reliant on standardized tests miss a great deal.
- Hard evidence on the importance of “soft skills” and that they can be measured and not just by psychometric tests, but by measures of behavior controlling for situation and ability, e.g., absences, expulsions, disciplinary actions, etc.

## Focus on Skills that Produce Outcomes and Behavior

- Not just specific outcomes
- Cognitive traits ( $\theta^C$ )
  - **a** **crystallized** and **fluid intelligence**
  - **b** different age profiles for their development
- Socioemotional traits: personality traits and preference parameters ( $\theta^S$ )
- Biological traits  $\theta^H$  (mental and physical health)
- Each vector of traits evolves with age, but levels of each trait are positively (but not perfectly) autocorrelated over age.

- A common low-dimensional set of capabilities (traits) predicts a variety of outcomes.
- Comparative advantage is an empirically important feature of economic and social life.
- The same bundle of personal traits has different productivity in different tasks.
- There is a causal effect of schooling on these capabilities.
- There is substantial *heterogeneity* in these latent traits.

These capabilities or traits have direct **causal** effects on

- Wages (controlling for schooling);
- Schooling;
- Performance on achievement tests;
- Crime;
- Compliance with health protocols;
- Smoking;
- Adult health outcomes (mental and physical);
- Teenage pregnancy

and many other aspects of social and economic life.



## The GED

- An achievement test that high school dropouts can take to certify that they are the equivalents (in cognition) of high school graduates.

## Lessons from the GED

- More than academic achievement is required for success in life.
- Personality—“character”—can be measured.
- Many inventories developed.

- Movements for “accountability” in education often create perverse incentives.
- Tests and test certification can create and conceal problems.
- Uncritical reliance on tests to screen students into schools warps educational goals, stifles creativity and does not even predict even success in school all that well.

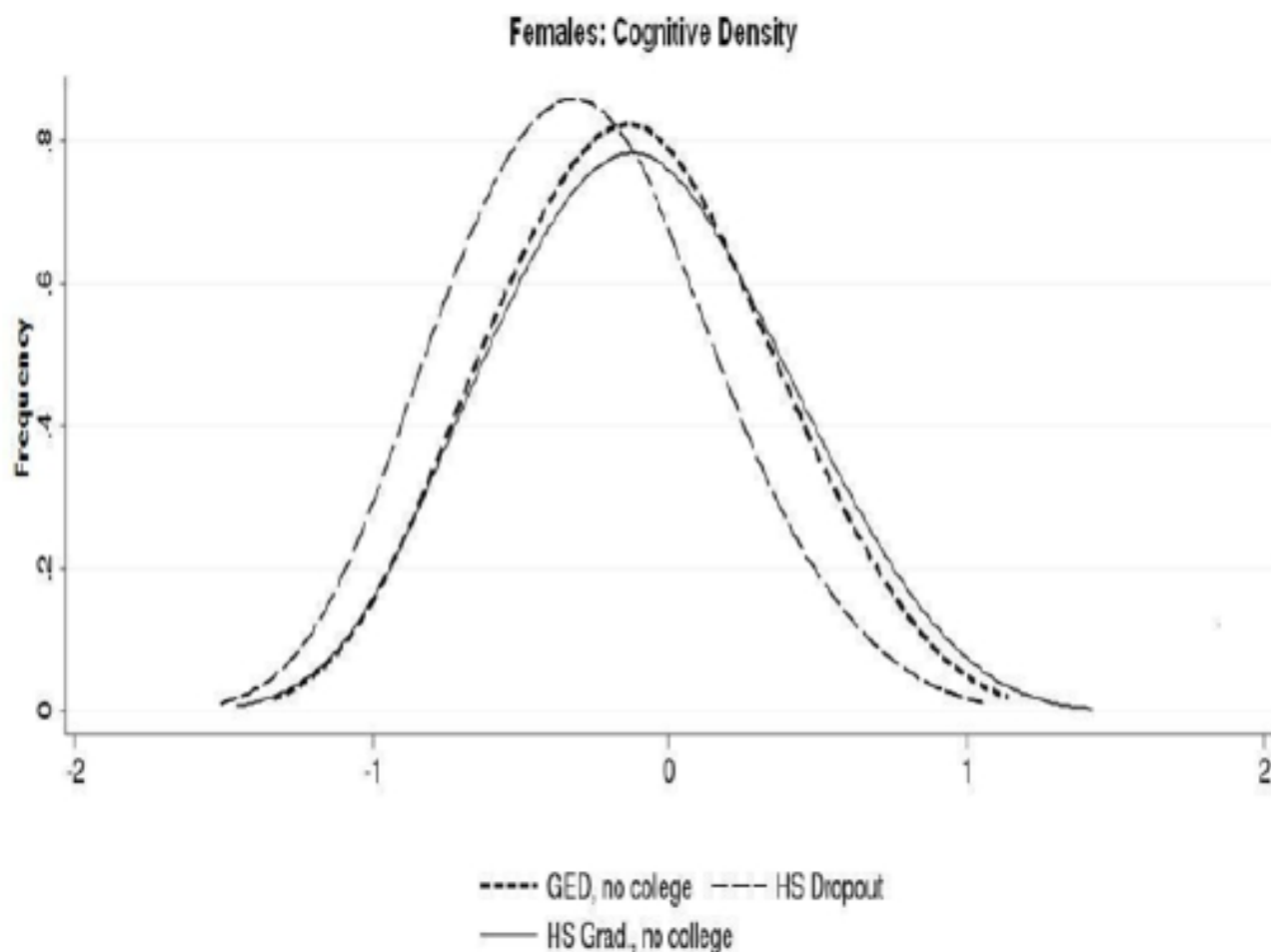
## Horace Mann, on tests:

*“Hence to value schools, by length instead of quality, is a matchless absurdity. Arithmetic, grammar, and the other rudiments, as they are called, comprise but a small part of the teachings in a school. The rudiments of feeling are taught not less than the rudiments of thinking. The sentiments and passions get more lessons than the intellect. Though their open recitations may be less, their secret rehearsals are more.”*

–Horace Mann (1867, p.420)

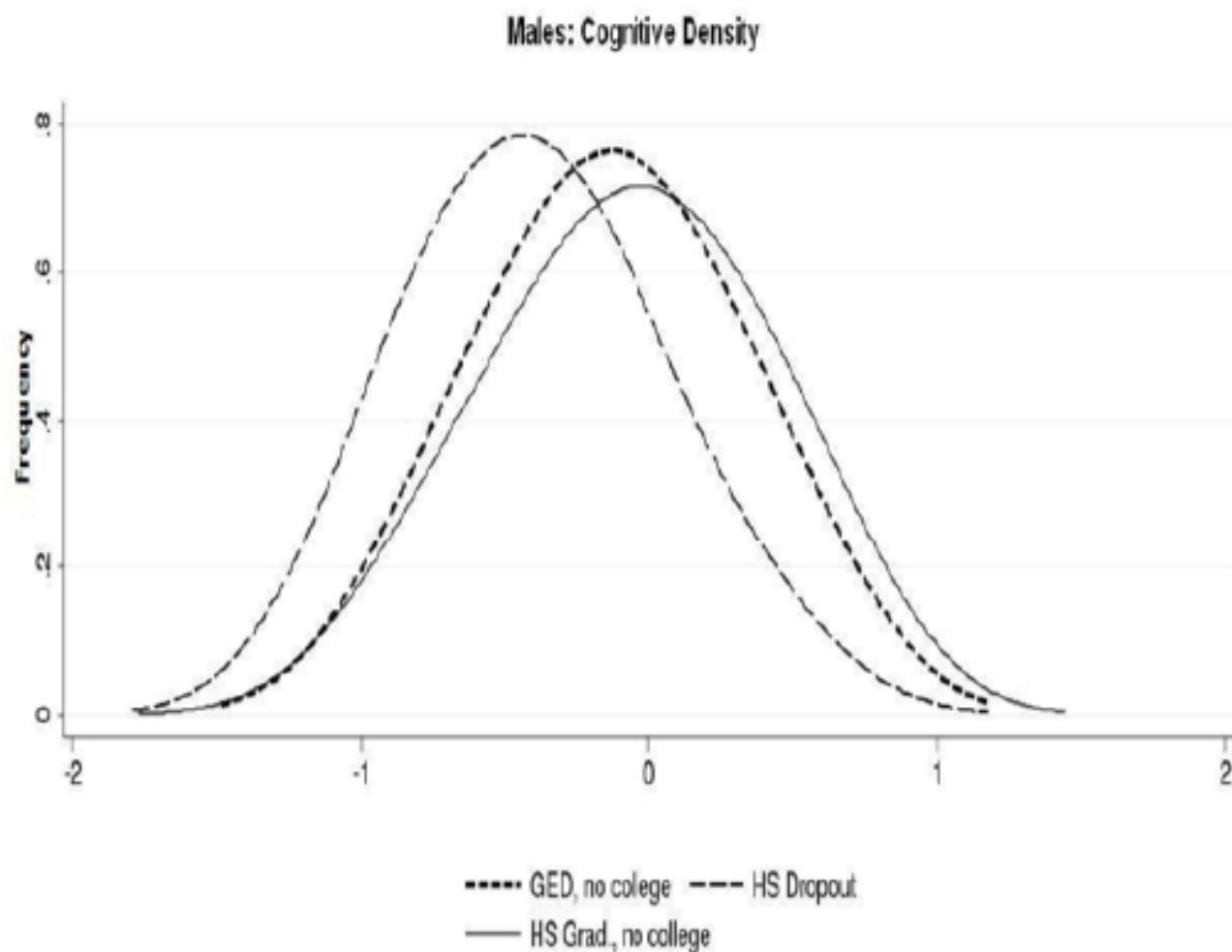
**Cognitive Skills: Comparable for GEDs and High School  
Grads Who Do Not Go on to College**

**Figure 1:** Cognitive ability by educational status (no college sample, all ethnic groups)



Source: : Heckman, Humphries, Urzua, and Veramendi (2010)

**Figure 2:** Cognitive ability by educational status (no college sample, all ethnic groups)



Source: : Heckman, Humphries, Urzua, and Veramendi (2010)

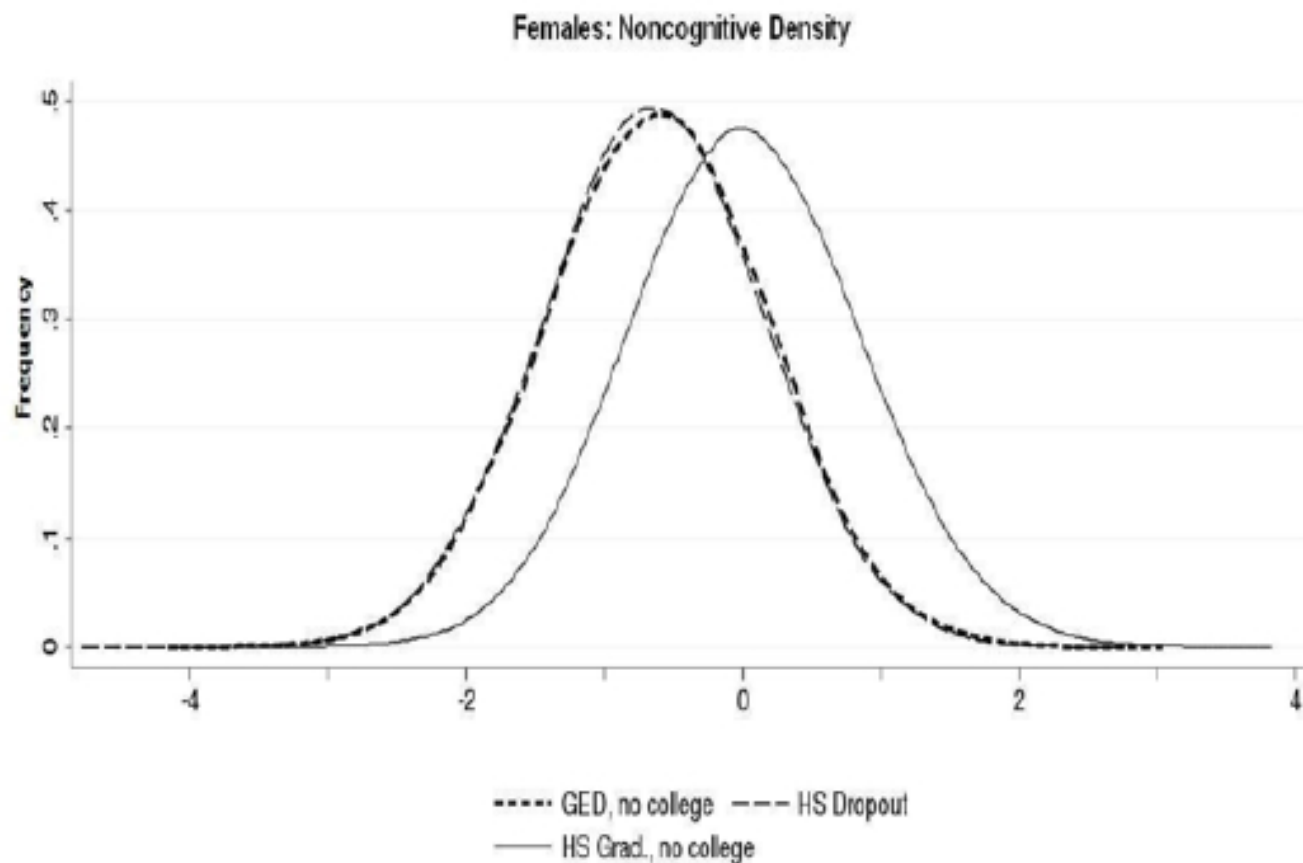
Heckman

## Noncognitive Skills: GEDs Resemble Dropouts

- They lack noncognitive or “soft” skills.

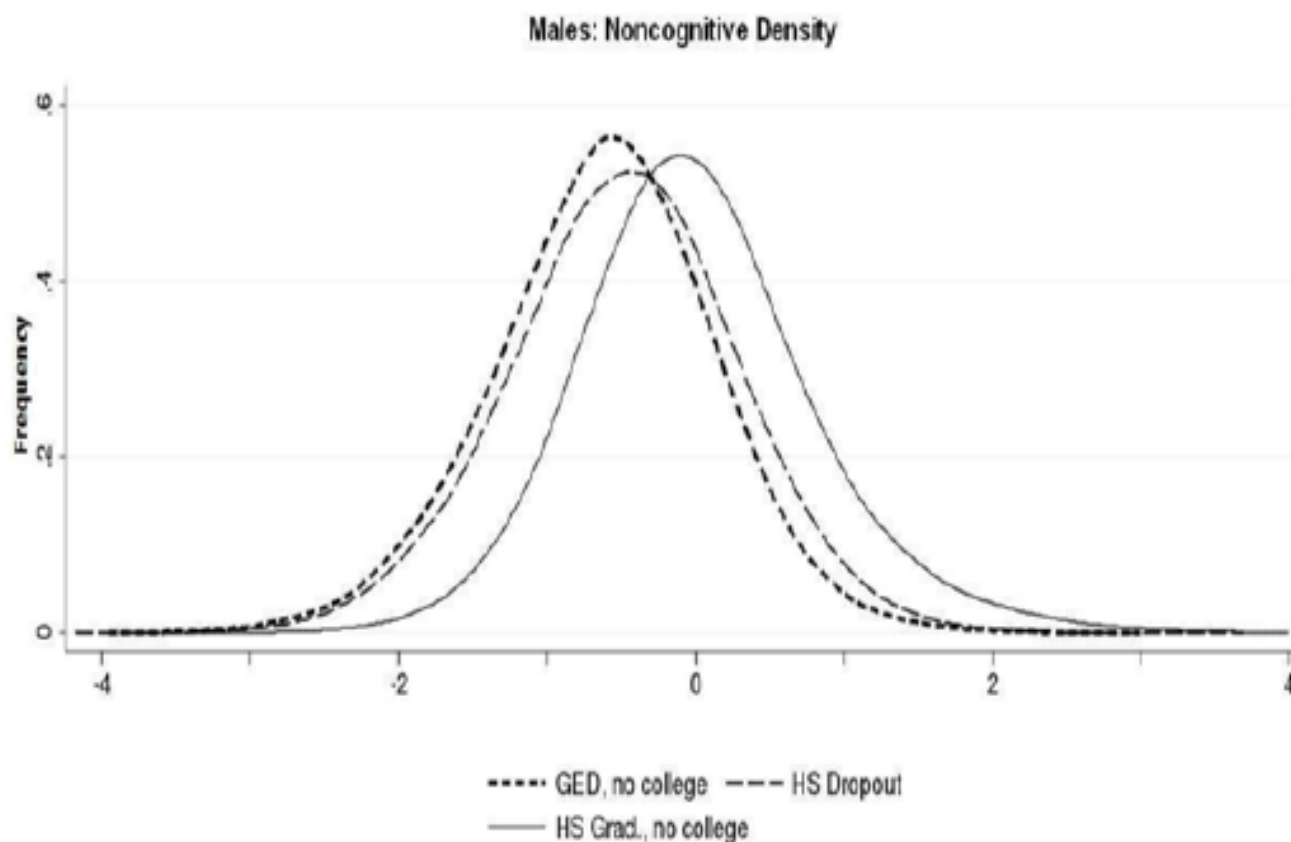


**Figure 3:** Cognitive ability by educational status (no college sample, all ethnic groups)



Source: : Heckman, Humphries, Urzua, and Veramendi (2010)

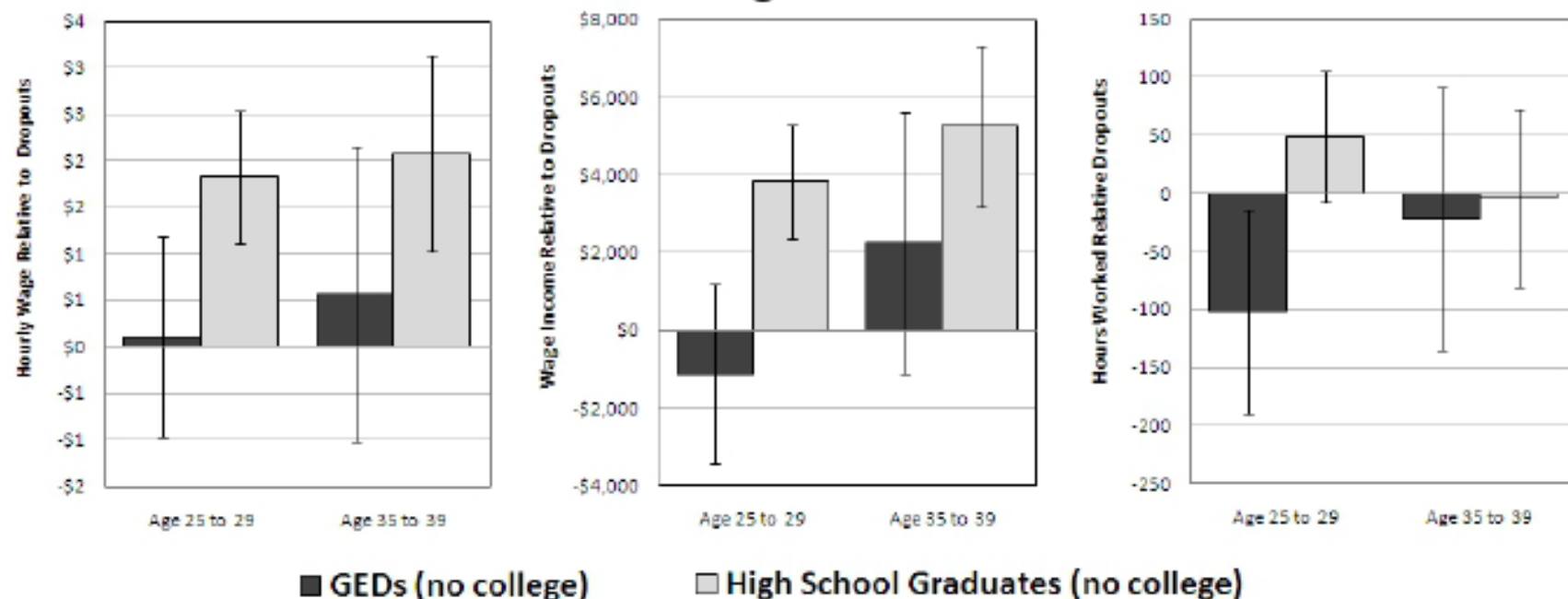
**Figure 4:** Cognitive ability by educational status (no college sample, all ethnic groups)



Source: : Heckman, Humphries, Urzua, and Veramendi (2010)

**Figure 5:** Ability-adjusted economic gaps relative to dropouts: GEDs and high school graduates

Male ability-adjusted economic gaps relative to dropouts: GEDs and high school graduates



## II. A Life Cycle Perspective

- Recent research documents the importance of early life parental, social, and environmental factors as predictors and causes of adult abilities, earnings, wages, preferences, health and personality.
- A developmental approach to understanding the evolution of skills over the life cycle.

## Modern Understanding of Skill Formation: Capabilities are Skills

- Skills are capacities to function and to flourish (Sen, 1986).
- A core low-dimensional set of skills explains a variety of diverse socioeconomic outcomes including health, earnings, schooling and occupational success.
- Cognitive and socioemotional skills are both important causal determinants with equal strength for many socioeconomic outcomes.
- An emerging literature shows that biological capacities are also powerfully predictive and causal of many socioeconomic outcomes.

- Sensitive periods for development:
  - Earlier for cognitive skills
  - Later for social and emotional skills
  - Occur at different stages for different biological capabilities.
  - Crucial to measure the nurturing environments at these sensitive ages.

## Critical and Sensitive Periods

- Sensitive and critical periods in the development of capabilities have been documented extensively for many organisms.



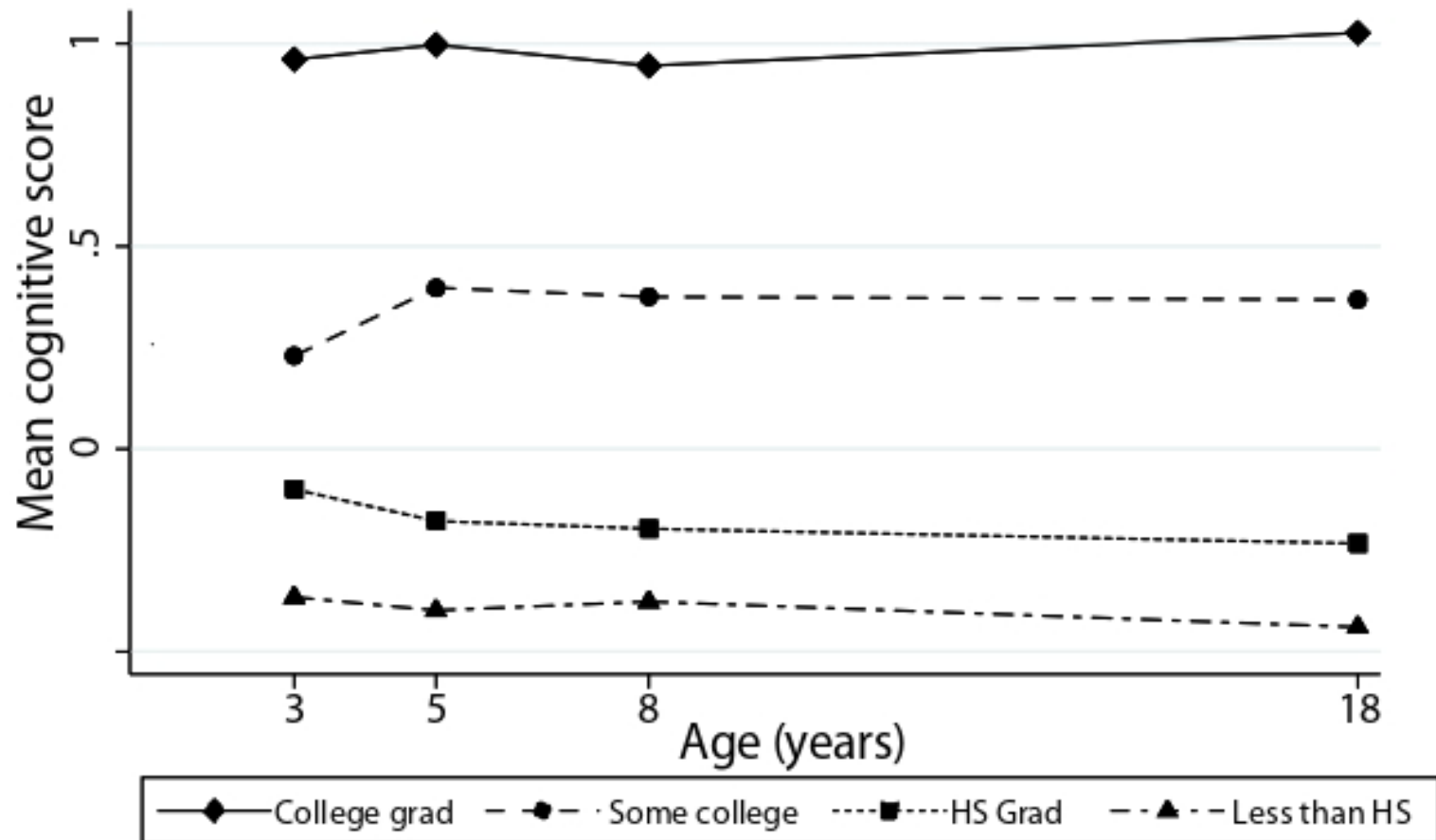
## Early Life Conditions Matter

- Fetal origins or Barker (2001) hypothesis.
- Fetal environment affects adult outcomes.  
(Gluckman and Hansen, 2005, 2006)
- Fetal alcohol (Nilsson, 2008); smoking.
- Related animal evidence.

## Skills Evolve and are Malleable

**For both cognitive and socioemotional skills, ability gaps across child socioeconomic groups open up at early ages and persist. For health traits, gaps tend to widen with age.**

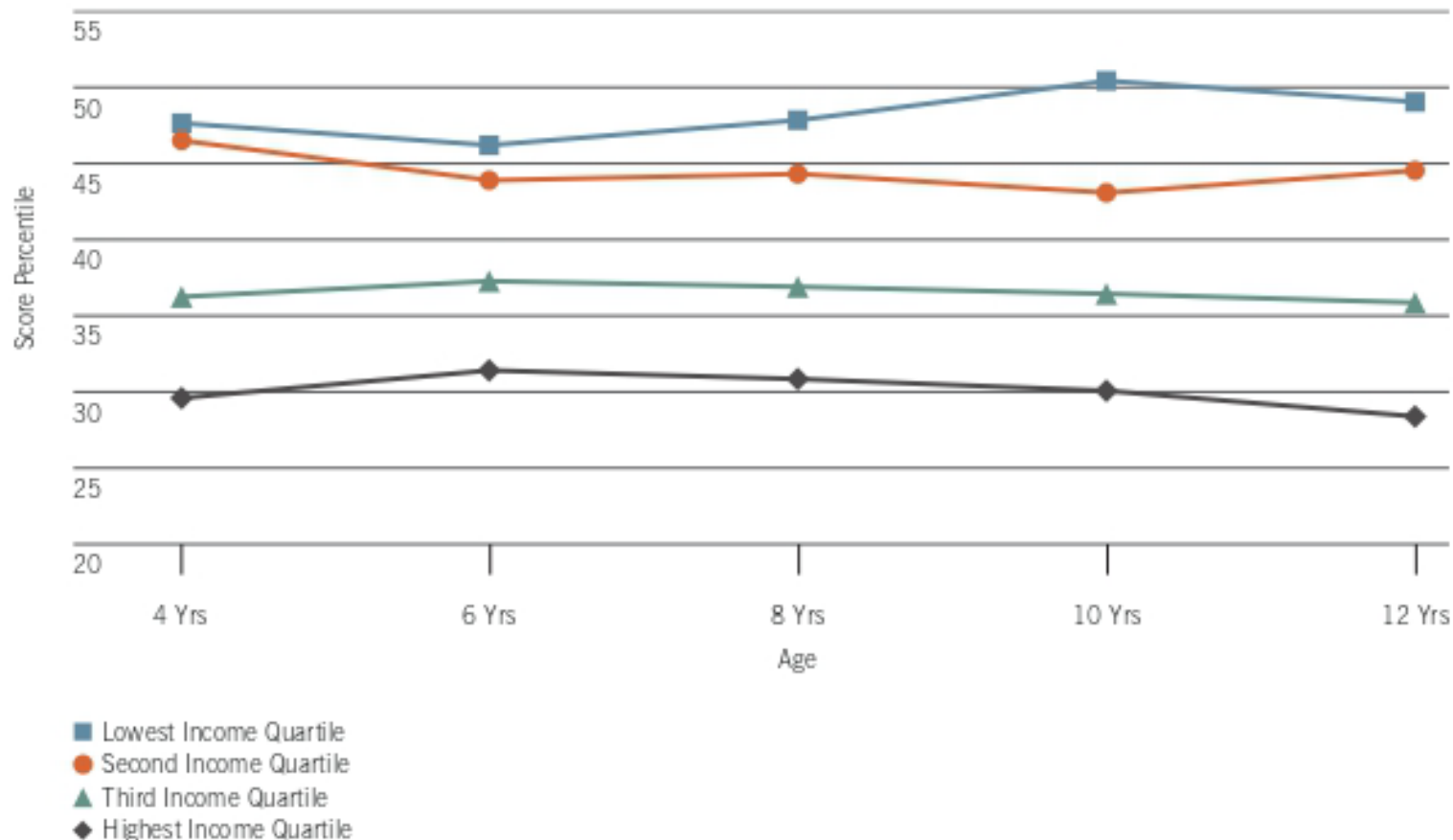
## Trend in mean age-adjusted cognitive score by maternal education



Each score standardized within observed sample. Using all observations and assuming data missing at random. Source: Brooks-Gunn et al. (2006).

- Controlling for early family environments using conventional statistical methods substantially narrows the gaps.
- Comparable phenomena in the evolution of gaps in behavioral problems.

## Average percentile rank on age-adjusted anti-social behavior score, by income quartile



## **These gaps have counterparts in gaps in family investments and environments.**

- Investment in children varies substantially by family type.
- Differences are persistent over the age of the child.
- Since Coleman (1966), we have known that family life is more important in predicting educational outcomes than schooling inputs.

## Family Environments

- In many other countries, a divide is opening up between the advantaged and the disadvantaged in the quality of early family environments.
- Those born into disadvantaged environments are receiving relatively less stimulation, child development resources, and access to health care than those from advantaged families.
- Fosters the persistence of inequality across generations.
- This is an important aspect of parenting and child welfare that should be measured.



## Early Life Experiences Are Not the Full Story

- There is also evidence of resilience and recovery at later ages.
- Early conditions are not fully determinative.
- Executive functioning and self-control emerge in adolescent years.
- Later life experiences are also important.
- The central economic question is what is the cost of remediation?
- How important are experiences and investments at various stages of the life cycle?

## **Enriched Early Investments Compensate In Part For the Risks Arising from Being Born into Disadvantaged Environments**

- A major channel through which early interventions work is through socioemotional skills.
- Enhancing parental knowledge and participation enhances child performance at school and elsewhere.
- Measure adversity: EDI indicator – school readiness.

# Modeling Human Capability Formation

## Capability Formation Process

- The capability formation process is governed by a multistage technology.
- Each stage corresponds to a period in the life cycle of the person.

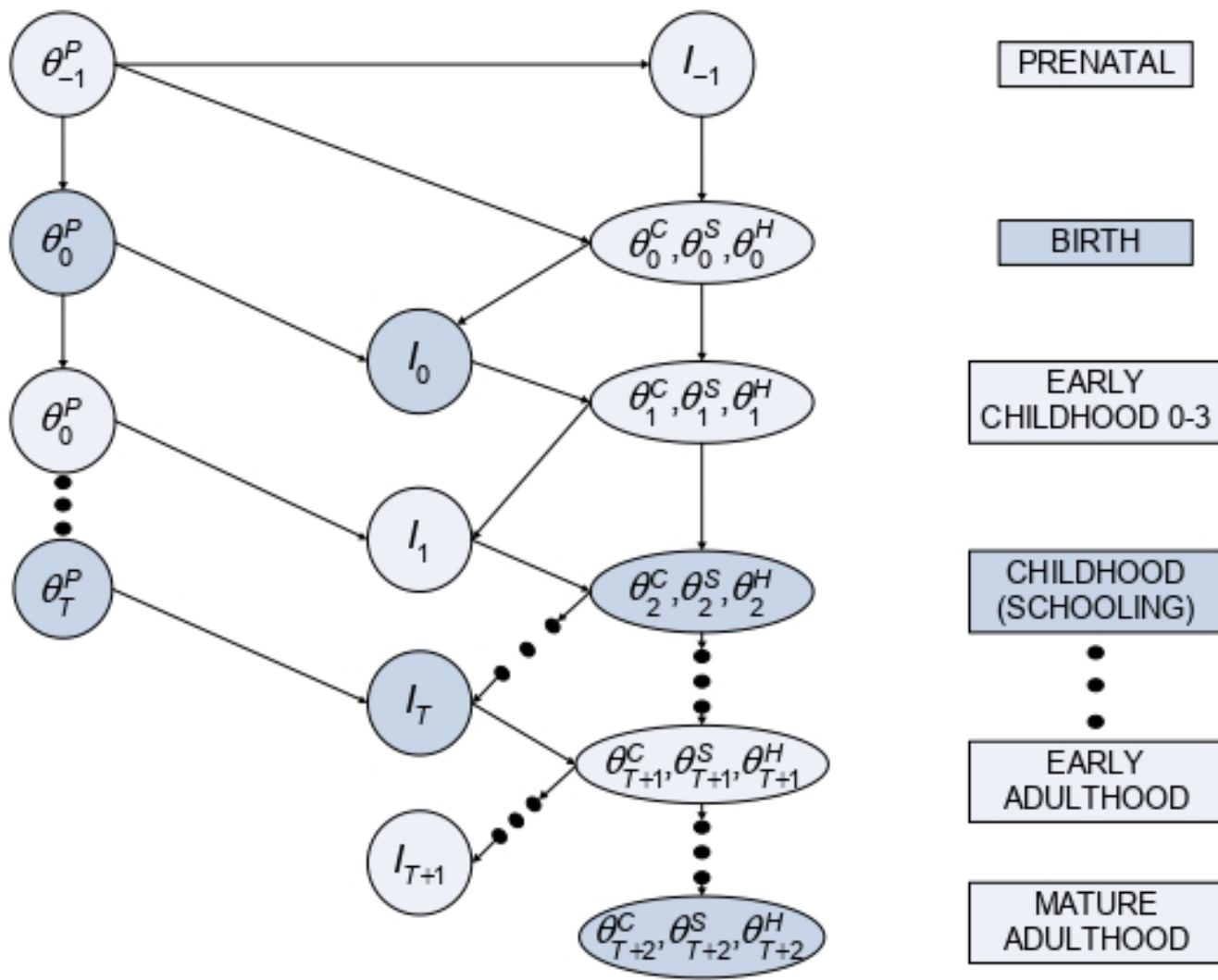
# A Life Cycle Framework for Organizing Studies and Integrating Evidence: $T + 1$ Periods of Life Cycle

$\theta_t = (\theta_t^C, \theta_t^S, \theta_t^H)$  capacities at  $t$

$l_t$ : investment at  $t$  (schooling; parenting)

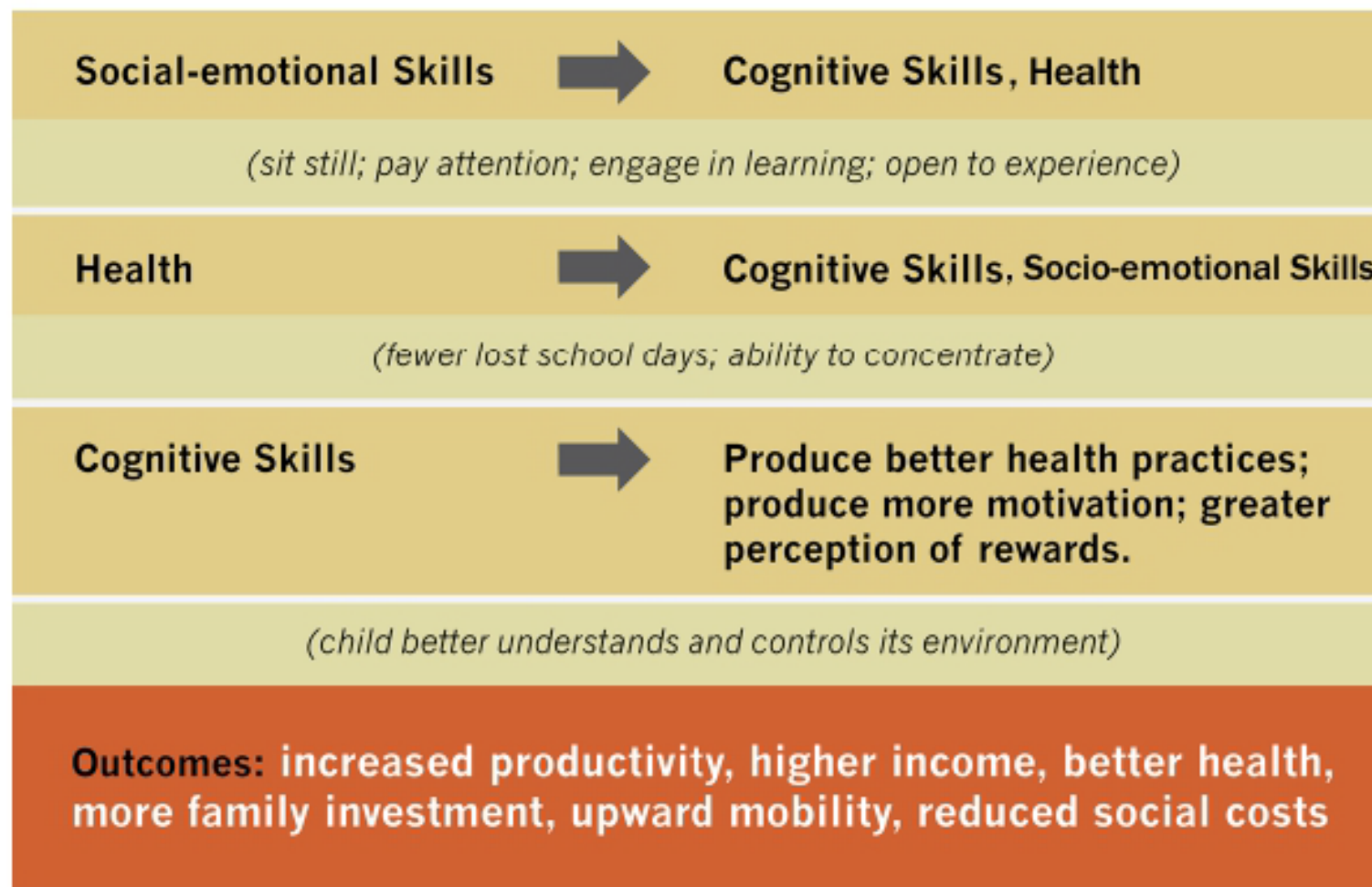
$\theta_t^P$ : parental capabilities at age  $t$

$$\theta_{t+1} = f_t(\theta_t, l_t, \theta_t^P)$$



# Dynamic Synergy of Skill Formation

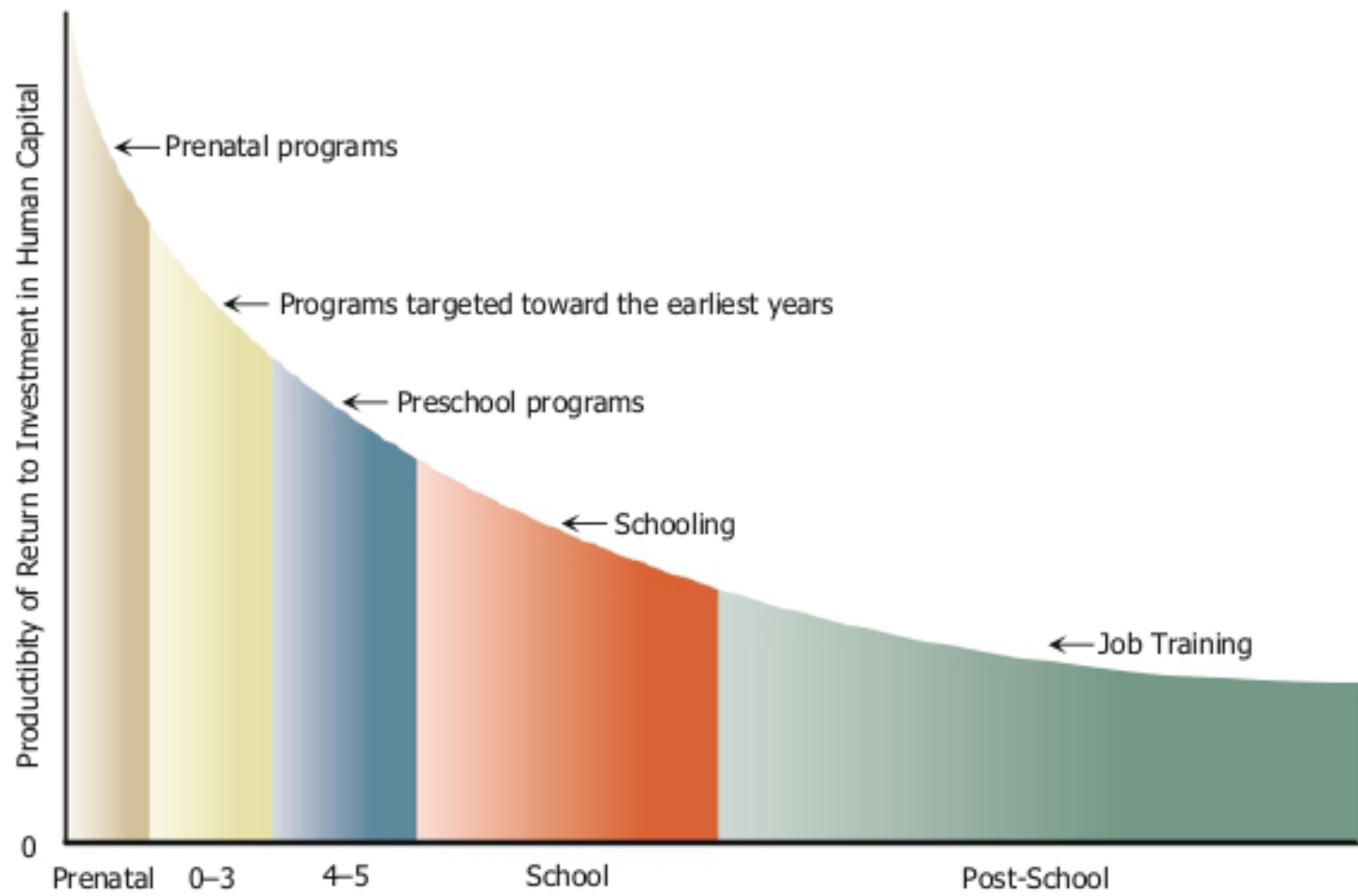
Figure 7: The Importance of the Early Years: Skills Beget Skills





**Dynamic Complementarity  
High Returns to Early Investment**

# Productivity Per Unit Value Invested



Source: Heckman (2008)



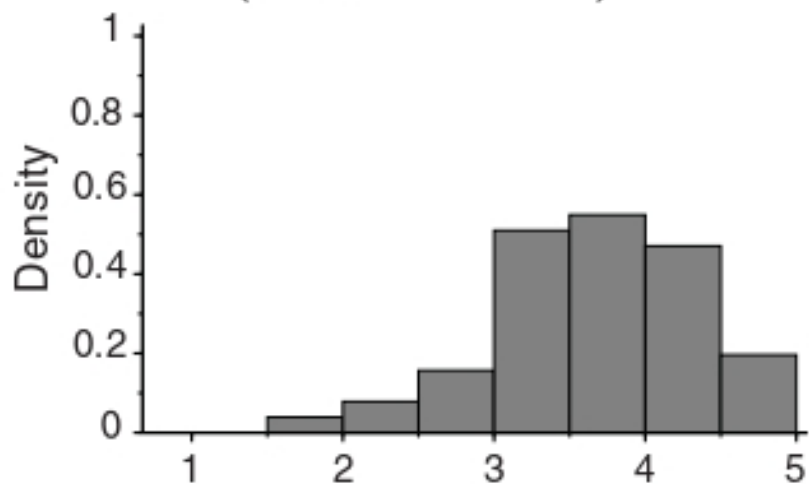
# Perry Preschool Project

**Starts at Age 3**  
**2 hrs a Day – Two Years**  
**Curriculum: Plan, Do, Review**

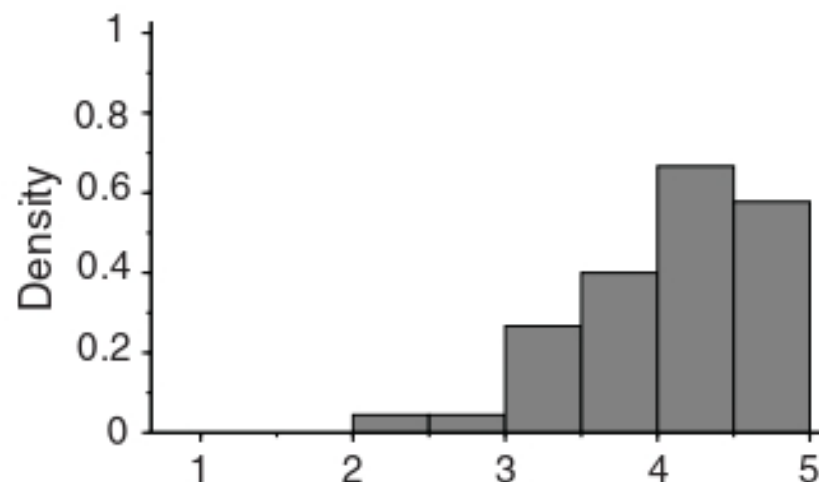
**10% Rate of Return Per Dollar Invested**  
**Noncognitive Factors Were Greatly Enhanced**

**Figure 8:** Perry Preschool Program: Histograms of Indices of Personality Skills and CAT Scores

Panel A. Externalizing behavior,  
control  
(Bigger is better)



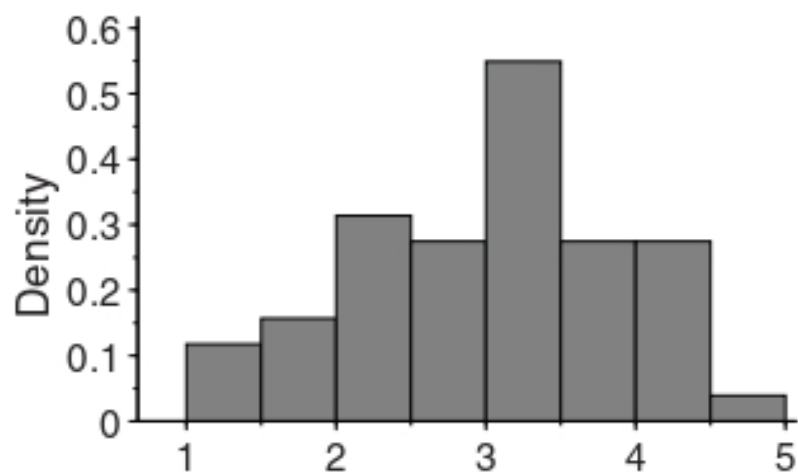
Panel B. Externalizing behavior,  
treatment  
(Bigger is better)



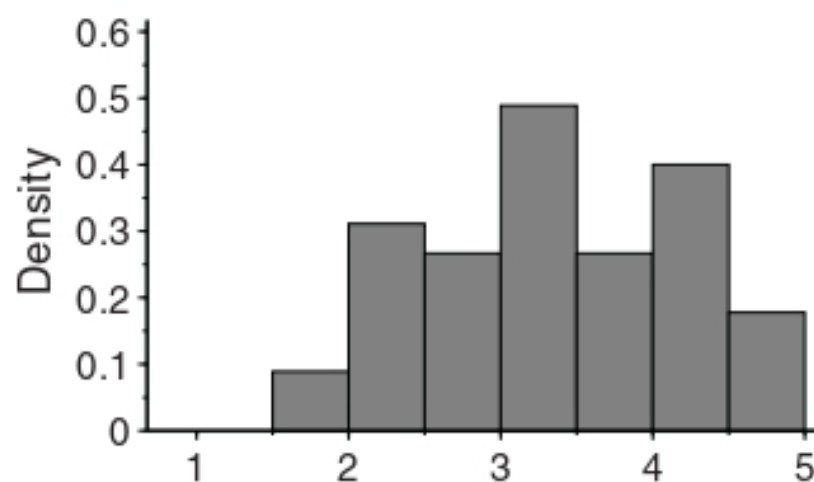
Source: Heckman et al. (2013).

**Figure 8:** Perry Preschool Program: Histograms of Indices of Personality Skills and CAT Scores, Cont'd

Panel C. Academic motivation, control



Panel D. Academic motivation, treatment



Source: Heckman et al. (2013).

## Abecedarian

- Abecedarian: a comprehensive approach.
- Starts earlier (at birth) and continues through age 5.
- 8 hours per day.
- Boosts IQ.



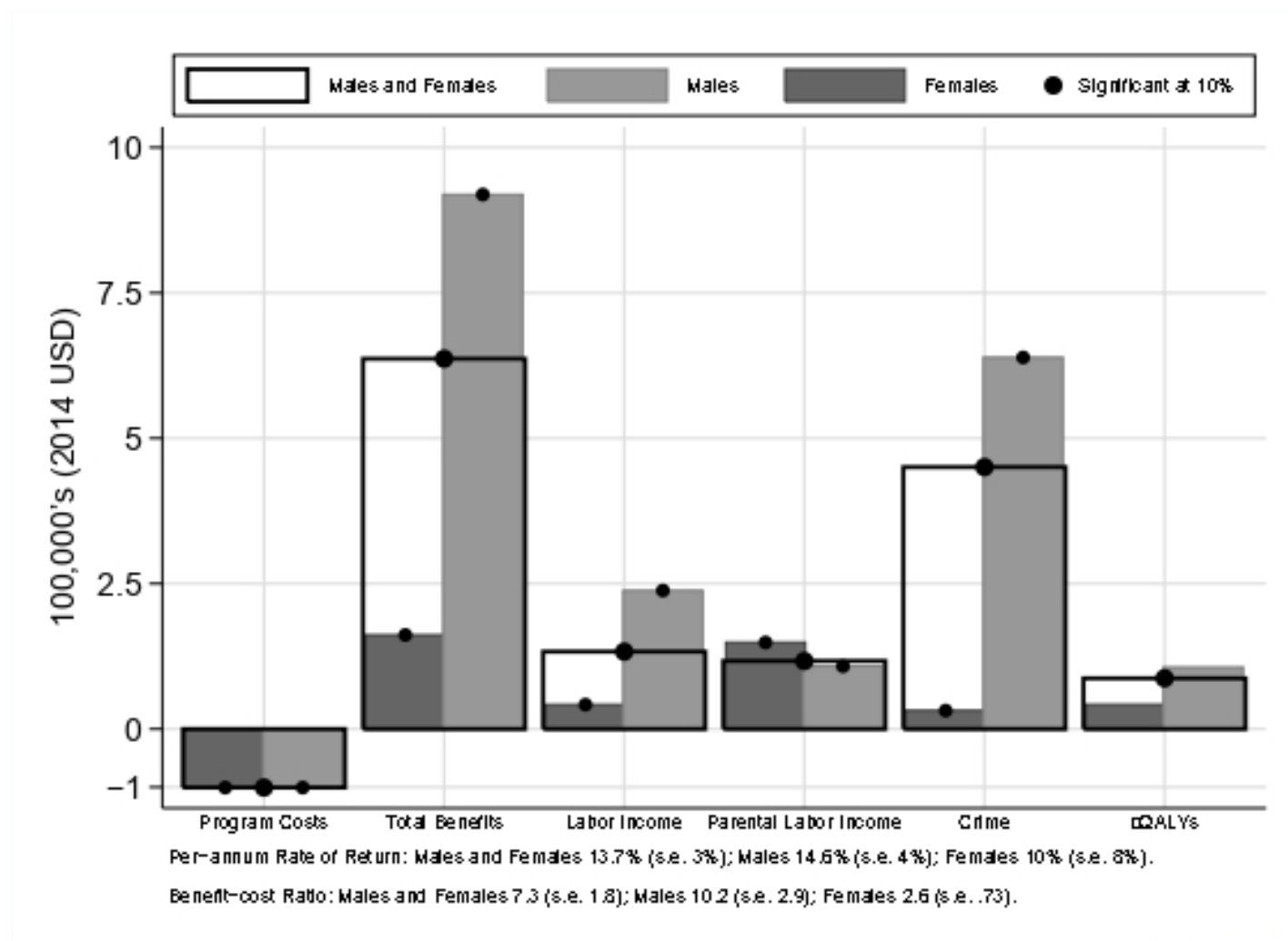
**Figure 9: Health Effects at Age 35 (Males) – From ABC Project**

	Treatment Mean	Control Mean	Treatment p-value
Systolic Blood Pressure	125.79	143.33	0.018
Diastolic Blood Pressure	78.53	92.00	0.024
Pre-Hypertension	0.68	0.78	0.235
Hypertension	0.10	0.44	0.011
HDL Cholesterol	53.21	42.00	0.067
Cholesterol/HDL-C	3.89	4.69	0.057
Abdominal Obesity	0.65	0.87	0.136
Metabolic Syndrome	0.00	0.25	0.009

Source: Campbell, Conti, Heckman, Moon, Pinto, Pungello, and Pan (2014).



Figure 10: Life-cycle Net Present Value of Main Components of the CBA



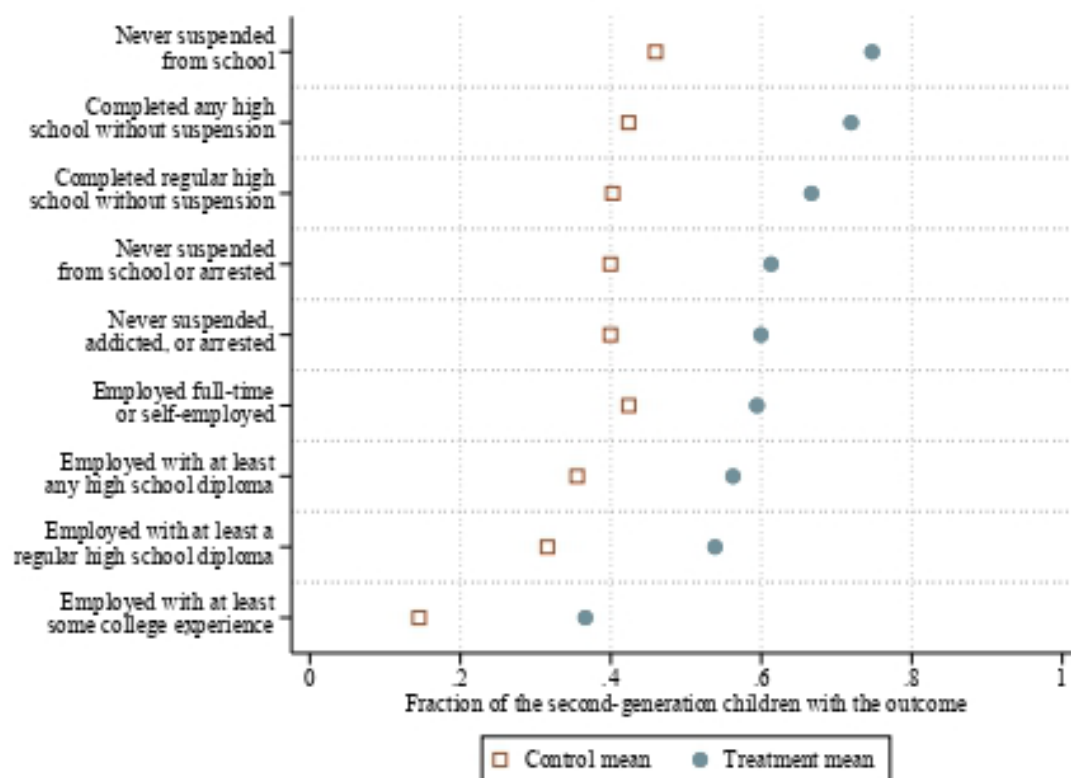
## **IV. Intergenerational Impacts**

**Effects of Perry Last Through the Next Generation**

**Recent Evidence**

# The Children of the Original Perry Participants

Figure 11: Statistically Significant Intergenerational Effects at the 10% Worst-Case Level\*



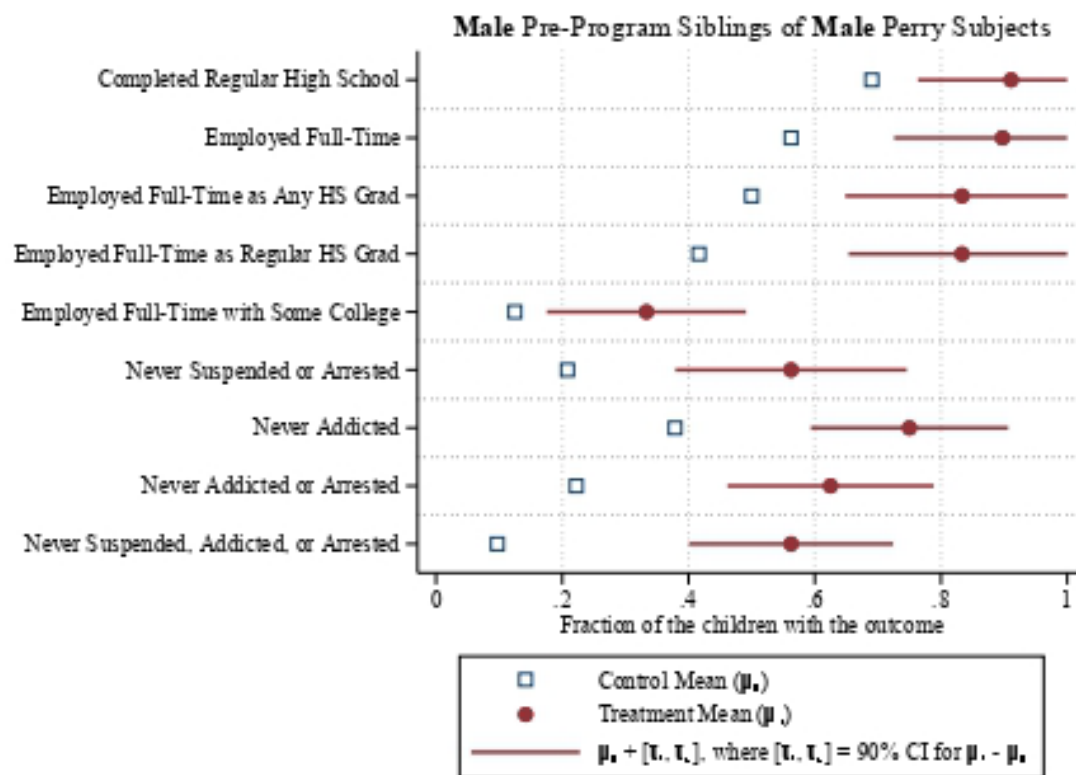
\* Note: These estimates of the intergenerational treatment effects are statistically significant at the 10% level using the conservative worst-case test procedures developed in Heckman and Karapakula (2019).

**Table 1:** Intergenerational Outcomes: Children of the Original Participants of Perry and ABC

	Male Children		
	Control Mean	Mean Difference (MD)	MD p-value
<b>Panel a. Perry</b>			
High School Graduate (Age 18 or older)	0.67	-0.01	0.582
College Graduate (Age 23 or older)	0.04	0.08	<b>0.063</b>
Employed (Age 23 or older)	0.48	0.19	<b>0.040</b>
Never Arrested (Age 18 or older)	0.37	0.14	<b>0.089</b>
In Good Health (Age 18 or older)	0.82	0.12	<b>0.006</b>
Not a Parent (Ages 14 to 22)	1.00	0.00	1.000
Never Divorced (Age 23 or older)	0.93	0.07	<b>0.028</b>
<b>Panel b. ABC</b>			
High School Graduate (Age 18 or older)	0.66	-0.06	0.718
College Graduate (Age 23 or older)	0.55	-0.08	0.683
Not Idle (Age 15 or older) <sup>†</sup>	0.91	0.06	<b>0.083</b>
In Good Health (Age 18 or older)	0.83	0.18	<b>0.000</b>
Not a Parent (Ages 14 to 22)	0.63	0.17	<b>0.069</b>

# Spillover Effects

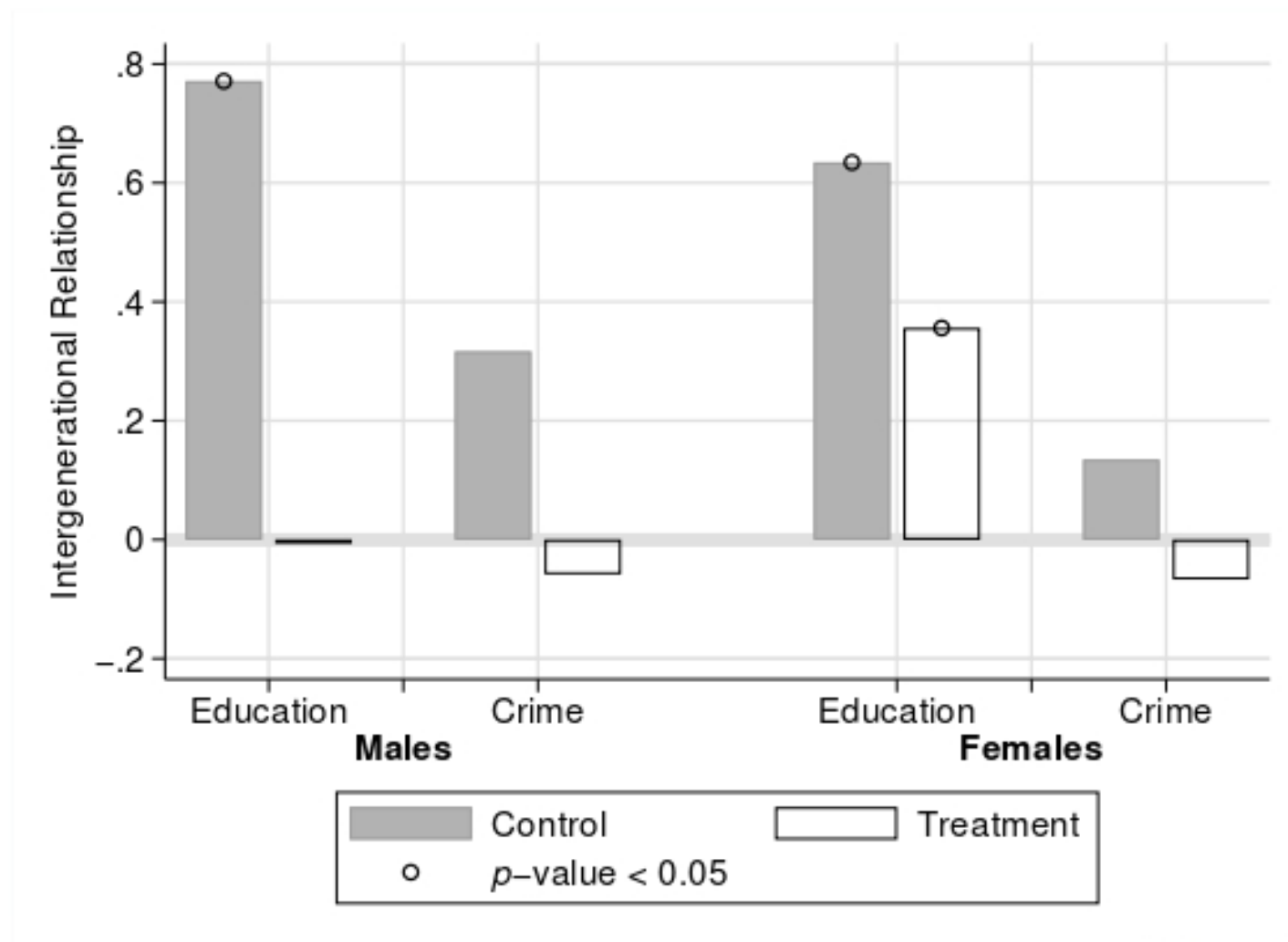
Figure 12: Treatment Effects on Male Pre-Program Siblings of Male Perry Subjects



# Benefit-Cost Ratio of the Perry Preschool Project, Summary of Estimates Including Dynastic Benefits and Spillovers

	[1]	[2]	[3]
	1st Generation	Dynasty	Extended Dynasty
	(Participants)	([1] + Child Spillovers)	([2] + Sibling Spillovers)
<b>B/C Program Benefit</b>	9.0	9.5	11.2
<b>(<i>p</i>-value)</b>	(0.03)	(0.00)	(0.00)
<b>Subtract Deadweight Loss</b>	6.0	6.3	7.5
<b>(<i>p</i>-value)</b>	(0.03)	(0.00)	(0.00)

# Intergenerational Dependence (Regression $Y_C = \alpha + \beta Y_P + \varepsilon_C$ )





# Intergenerational Transmission of IQ: Preparing for Life



- Correlation between IQ scores of mothers & children in *low* treatment group
  - Age 4:  $r=0.31$ ;  $p=0.018$
  - Age 9:  $r=0.57$ ;  $p=0.001$
- Correlation between IQ scores of mothers & children in *high* treatment group
  - Age 4:  $r=0.07$ ;  $p=0.562$
  - Age 9:  $r=0.18$ ;  $p=0.148$
- Some evidence that program reduced intergeneration transmission of low IQ

## Part 4: Cross Disciplinary Cooperation is Essential to Develop New Sources of Data and Enrich Traditional Sources

- Synthesize psychology, sociology, education, economics, statistics, AI, and neuroscience
- Beyond standardized tests—games, assessments
- Social media; internet-based interaction and monitoring
- Collaboration – not siloing

## Part 5: The Challenge and Opportunity

- Exciting challenge – unified systems of data and learning
- Families, firms, schools, other institutions

- Multiple sources of data
  - Administrative data on achievement and behaviors
  - Data from social media
  - Surveys
  - Media scraping
  - Feedback systems, reinforcement
- Multiple sources
- Develop an understanding of the interconnected, interrelated sources of learning. Schools and tests yes, but families, firms, peers, and neighborhoods

# Good Luck, Best Wishes, and Let Me Help



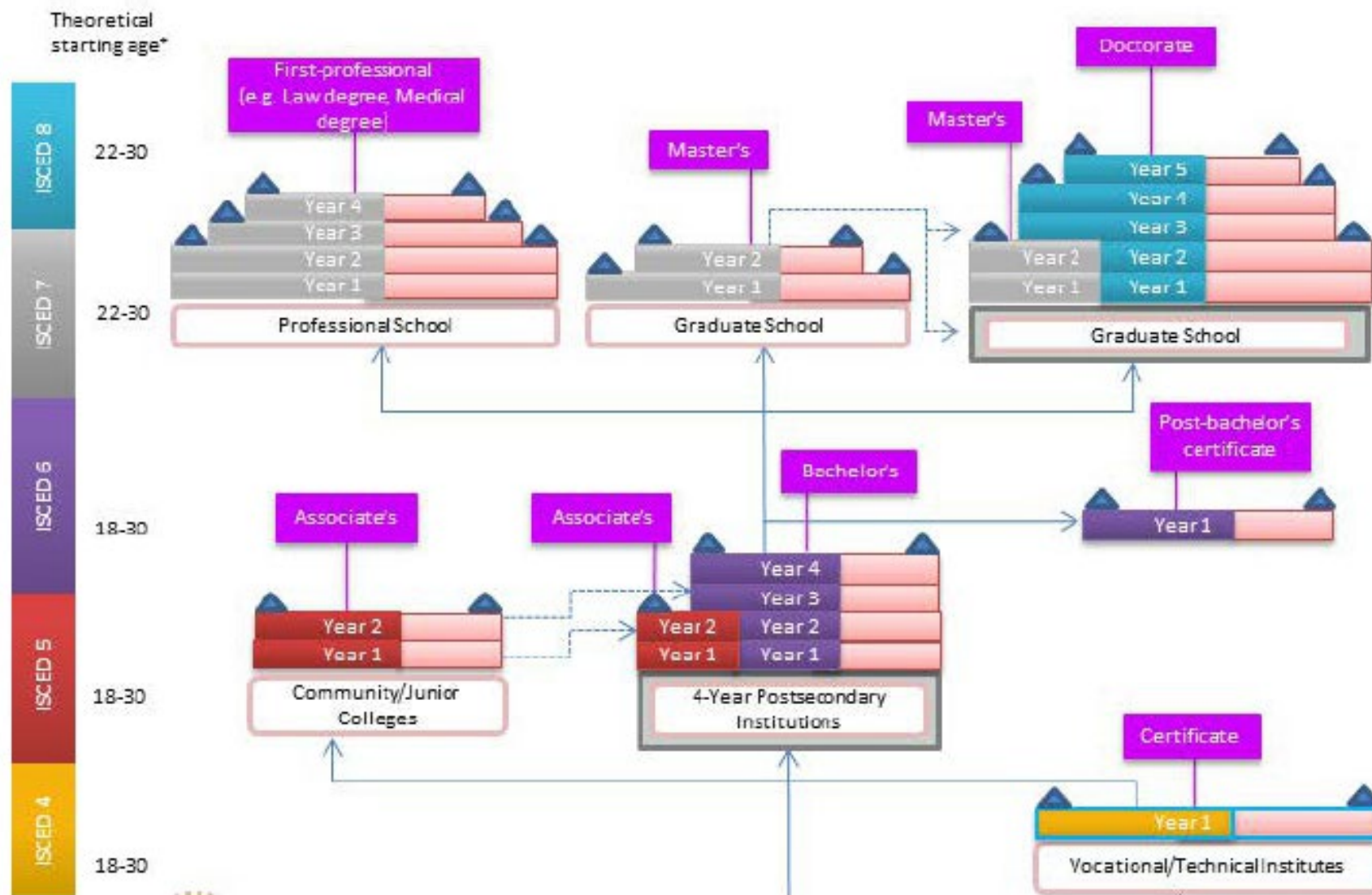
CENTER FOR THE ECONOMICS  
OF HUMAN DEVELOPMENT  
THE UNIVERSITY OF CHICAGO

[cehd.uchicago.edu](http://cehd.uchicago.edu)



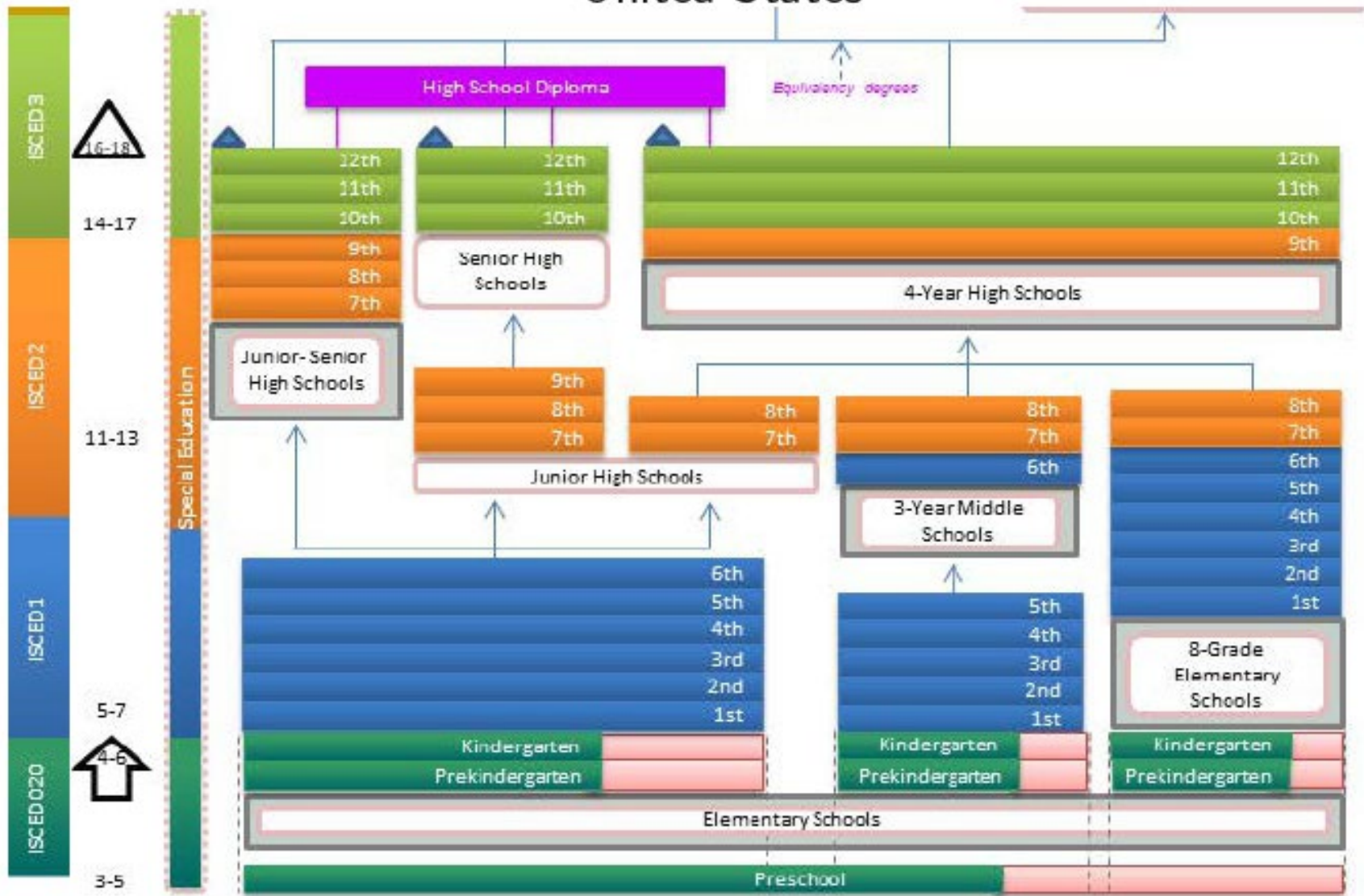
# United States

2023



Source: Education GPS (2023).

# United States

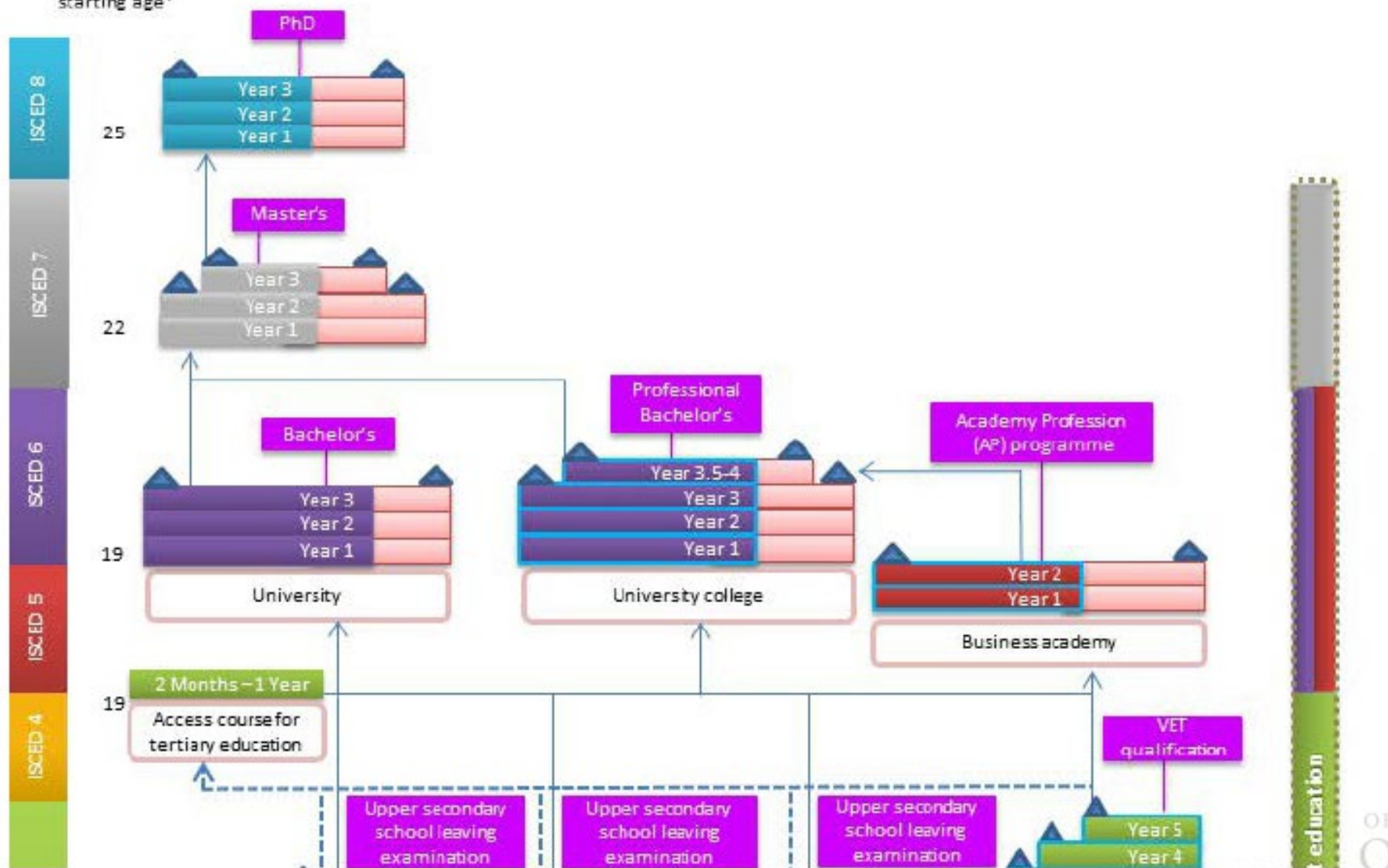


Source: Education GPS (2023).

# Denmark

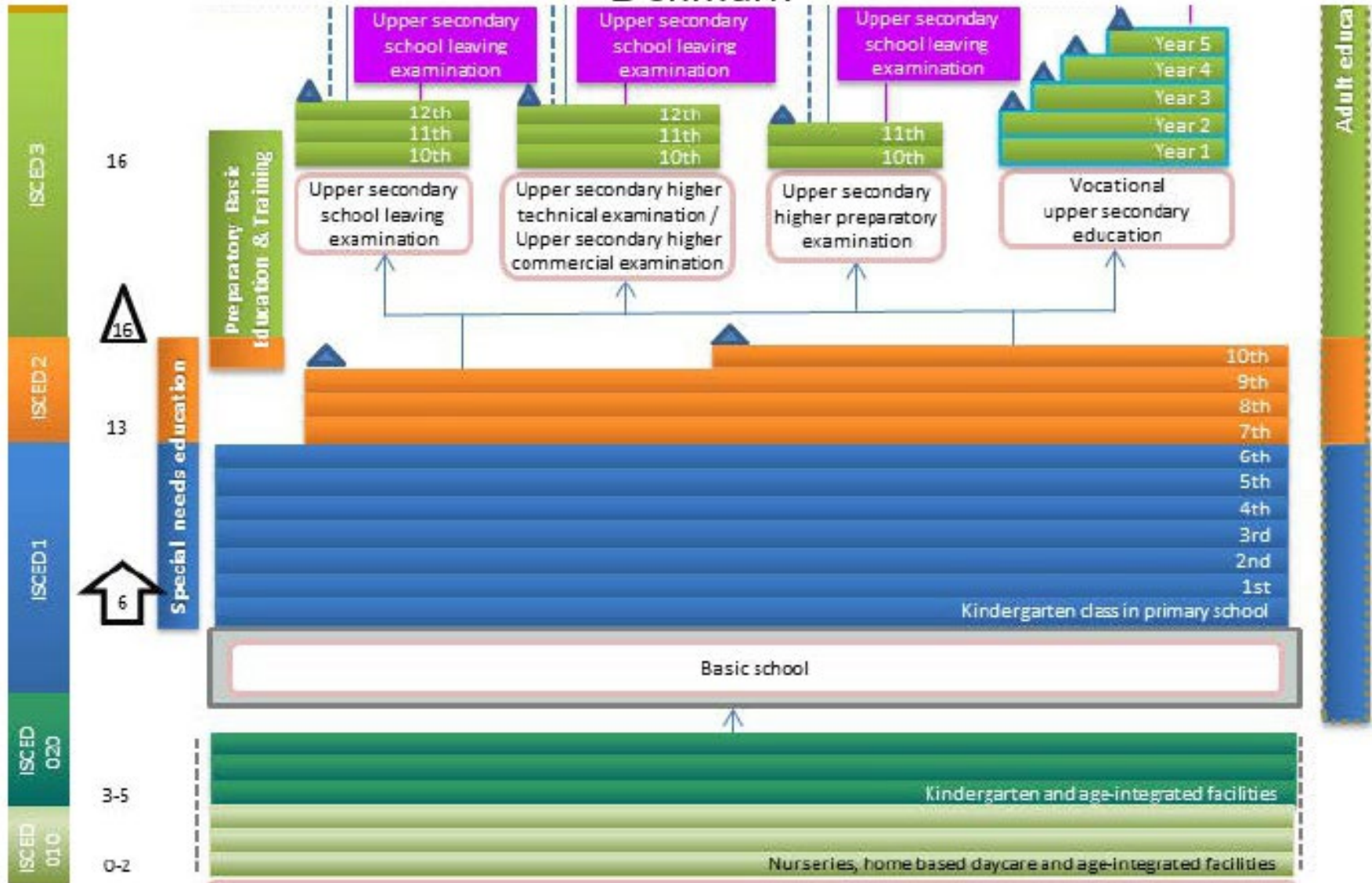
2023

Theoretical starting age\*





# Denmark



Source: Education GPS (2023).



[Return to main text](#)