MIT 14.662 Spring 2026
Lecture 6 slides – Dimensions of skill

David Autor, MIT and NBER

Feb 18, 2026 (rev 2026/02/18)

What is ‘skill’ anyway?

The canonical narrow view and its limits

- Economists typically equate ‘skill’ with fixed cognitive ability (IQ)
- But: IQ rises across cohorts (Flynn effect); cognitive endurance appears trainable

Skill is more than just static cognitive compute

- Personality traits—especially interpersonal skills—matter greatly and may be rising in importance
- *Why* do interpersonal skills matter? (See Deming 2017 *QJE*)
- Managerial skill—almost unstudied
- Lots of work on ‘non-cognitive’ skills—but I still don’t know what this means

A broader conception of skill: Efficiency vs. allocation

- Classic lit distinguishes *factor efficiency* from *allocative efficiency* (Welch 1970; Schultz 1975)
- Is work best understood as “accomplishing tasks” or as “exerting agency to yield results”—choosing among tasks according to capabilities and applicability?
- Tasks are bundled in jobs, creating interdependencies that shape expertise requirements
- What determines production of skill in equilibrium (schooling is only one dimension)?

Agenda

1. **The Flynn puzzle**
2. **The Bell Curve controversy**
3. **Malleability of cognitive human capital**
4. **Social skills: Not just "IQ"**
   - Rising supply of economically valuable personality traits in Finland
   - Rising premium to social skills in Sweden
5. **What makes a good manager?**
6. **Why do social skills matter—and more so over time?**
7. **Production of skills in equilibrium**

Standardized IQ data, 1932 – 1972

Table 1
Stanford-Binet and Wechsler Standardization Data

Tests	Acronym	Date: Duration	Date: Midpoint	*M*[a]	*SD*[a]
Stanford-Binet Form L	SB-L	1931–1933	1932	100.00	16.00[b]
Stanford-Binet Form M	SB-M	1931–1933	1932	100.00	16.00[b]
Stanford-Binet Form L-M	SB-LM	1931–1933	1932	98.00[b]	16.00
Stanford-Binet 1972 Norms	SB-72	1971–1972	1971½	102.81	15.00
Wechsler-Bellevue Form I	WB-I	1935–1938	1936½	100.00	14.83
Wechsler Intelligence Scale for Children	WISC	1947–1948	1947½	100.00	15.00
Wechsler Adult Intelligence Scale	WAIS	1953–1954	1953½	102.26	14.00
Wechsler Preschool and Primary Scale of Intelligence	WPPSI	1963–1966	1964½	102.26	14.00
Wechsler Intelligence Scale for Children—Revised	WISC–R	1971–1973	1972	102.26	14.00
Wechsler Adult Intelligence Scale—Revised	WAIS–R	1976–1980	1978	102.26	14.00

Note. Sources were Kaufman & Doppelt (1976, p. 167); Terman (1942, pp. 2–3); Terman & Merrill (1937, pp. 12–15); Terman & Merrill (1973, pp. 26–28, 64, 339, 353, 359–361); Thorndike (1975, p. 6); Seashore, Wesman, & Doppelt (1950, p. 102); Wechsler (1939, pp. 35–36, 41, 107–110); Wechsler (1949, pp. 4, 7); Wechsler (1955, pp. 3, 6, 10); Wechsler (1967, pp. 5, 13–15); Wechsler (1974, pp. iii, 17–19); Wechsler (1981, pp. 9, 16–19).
[a] Values given are for whites only.
[b] As discussed in the text, if these values are used to translate reported scores into our uniform convention, they will give only approximate results.

Flynn, *Psychological Bulletin* 1984

“The Mean IQ of Americans: Massive Gains 1932 to 1978”

Table 2
White Americans: Evidence for IQ Gains 1932 to 1978

Test combination	Dates: Test 1	Dates: Test 2	Studies	*N*	Means: Test 1	Means: Test 2	Gain	Years	Rate	Ages (years)
1. SB-L & WISC	1932	1947½	17	1,563	107.13	101.64	5.49	15½	.354	5–15
2. SB-M & WISC	1932	1947½	1	46	125.13	107.56	17.57	15½	1.134	5
3. SB-LM & WISC	1932	1947½	6	460	114.64	109.67	4.97	15½	.321	5–15
4. SB-L & WAIS	1932	1953½	3	271	113.02	105.48	7.54	21½	.351	16–32
5. SB-LM & WAIS	1932	1953½	2	79	109.08	101.75	7.33	21½	.341	16–48
6. SB-LM & WPPSI	1932	1964½	8	416	101.74	92.78	8.96	32½	.276	4–6
7. SB-LM & SB-72	1932	1971½	1	2,351	107.08	97.19	9.89	39½	.250	2–18
8. WB-I & WISC	1936½	1947½	2	110	103.51	105.54	−2.03	11	−.185	11–14
9. WB-I & WAIS	1936½	1953½	3	152	122.94	118.25	4.69	17	.276	16–39
10. WISC & WAIS	1947½	1953½	4	436	101.76	99.12	2.64	6	.440	14–17
11. WISC & WPPSI	1947½	1964½	2	108	93.56	90.86	2.70	17	.159	5–6
12. WISC & SB-72	1947½	1971½	1	30	96.40	84.42	11.98	24	.499	6–10
13. WISC & WISC–R	1947½	1972	17	1,042	97.19	88.78	8.41	24½	.343	6–15
14. WAIS & WISC–R	1953½	1972	1	40	102.94	96.29	6.65	18½	.359	16–17
15. WAIS & WAIS–R	1953½	1978	1	72	109.69	101.65	8.04	24½	.328	35–44
16. WPPSI & SB-72	1964½	1971½	1	35	93.06	88.65	4.41	7	.630	4–5
17. WPPSI & WISC–R	1964½	1972	2	140	112.84	108.58	4.26	7½	.568	5–6
18. WISC–R & WAIS–R	1972	1978	1	80	99.61	98.65	0.96	6	.161	16

Note. See Table 1 for full test names. Totals: 73 studies and 7,431 subjects. Age range in years: 2 to 48 (*Mdn* = 10.6). All means are weighted in terms of the number of

Flynn, *Psychological Bulletin* 1984

"The Mean IQ of Americans: Massive Gains 1932 to 1978"

Table 4
White Americans: IQ Gains Selected Periods

Period	Rate[a]	No. of studies	*N*	Data used[b]
1932–1948	.368	29	2,419	1–3, 4[c], 5[c]
1948–1972	.353	29	1,903	10–14, 15[c], 16, 17
1932–1948	.368	29	2,419	1–3, 4[c], 5[c]
1948–1960	.347	6	544	10, 11[c]
1960–1972	.359			Derived from 1948–1960 and 1948–1972

[a] IQ points per year.
[b] From Table 2. [c]Prorated.

Flynn, *Psychological Bulletin* 1984

Summary of Implications

Assuming that American IQ gains 1932 to 1978 are not real, but an artifact of sampling error, they have acted as a confounding variable in hundreds of studies and require altered practices from testing organizations. Assuming that these gains are semi-real, due primarily to test sophistication, they imply all of the above and also reveal the inexplicable combination of IQ gains and SAT-V losses. Assuming that these gains are real, they imply all of the above and pose a serious problem of causal explanation. Moreover, all of these implications hold even if real gains ceased in 1978 or even 1972: The period in question shows the radical malleability of IQ during a time of normal environmental change; other times and other trends cannot erase that fact.

Flynn, *Psychological Bulletin* 1984

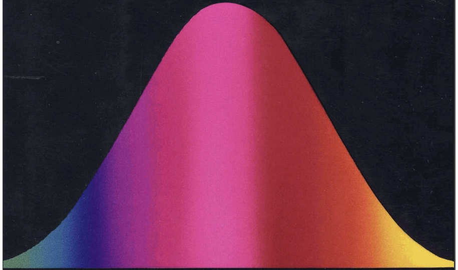
The Flynn puzzle

"If there was a prize to be offered in the field of human intelligence research, it would be for the person who can explain" the [Flynn] effect (Deary, 2001).

Agenda

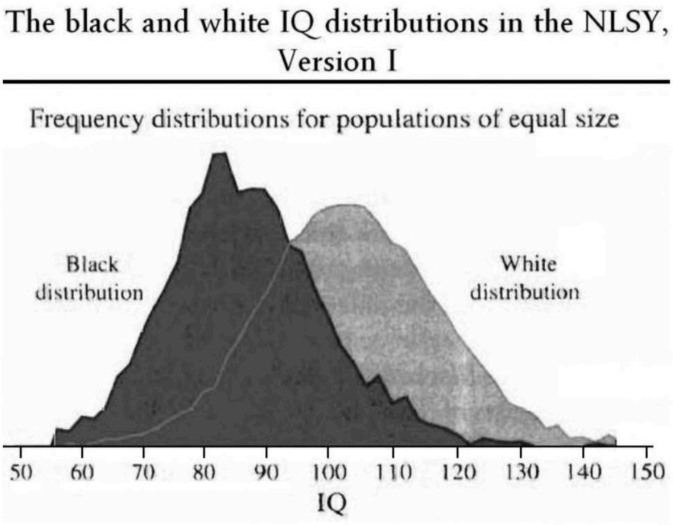
THE BELL CURVE

Intelligence and Class Structure in American Life



RICHARD J. HERRNSTEIN
CHARLES MURRAY

The Bell Curve, Herrnstein and Murray, 1994



The Bell Curve, Herrnstein and Murray, 1994: Immutable differences?

Reductions in the Black-White Difference on the National Assessment of Educational Progress

	White-Black Difference, in Standard Deviations*		
	1969–1973	**1990**	**Change**
9-year-olds			
Science	1.14	.84	-.30
Math	.70	.54	-.16
Reading	.88	.70	-.18
Average	*.91*	*.69*	*-.21*
13-year-olds			
Science	.96	.76	-.20
Math	.92	.54	-.38
Reading	.78	.40	-.38
Average	*.89*	*.57*	*-.32*
17-year-olds			
Science	1.08	.96	-.12
Math	.80	.42	-.38
Reading	1.04	.60	-.44
Average	*.97*	*.66*	*-.31*
Overall average	*.92*	*.64*	*-.28*

Source: National Center for Education Statistics, 1991b.
*The computations assume a standard deviation of 50.

Agenda

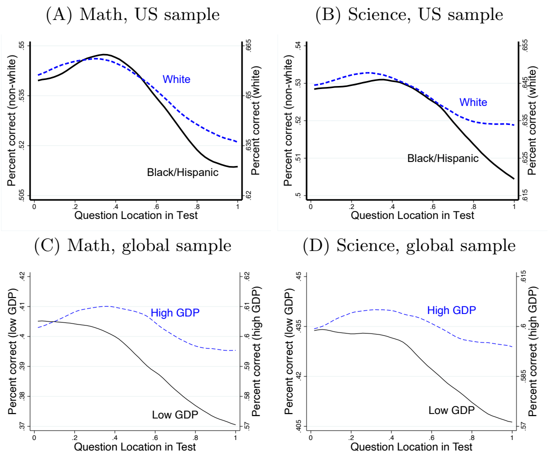
Cognitive endurance as human capital: Motivation

In this article, we focus on one specific feature of formal education: schooling engages students in effortful thinking for continuous stretches of time. From doing in-class exercises to reading a textbook, the act of learning often involves periods of sustained concentration. Using a field experiment with elementary school students, we test whether such intellectual practice can, in and of itself, expand a particular mental ability: cognitive endurance.

Cognitive endurance as human capital: Evidence from TIMSS

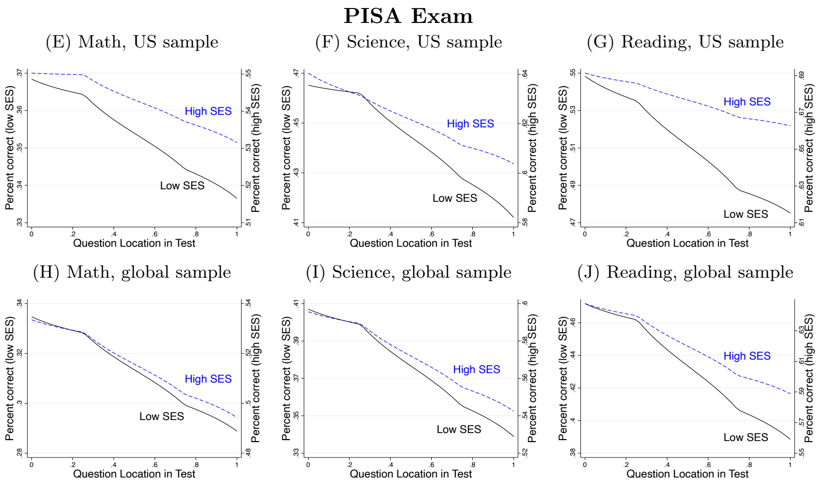
FIGURE I: Performance Declines in Achievement Tests

TIMSS Exam



Brown, Kauer, Kingdon, and Schofield, *QJE* 2024

Cognitive endurance as human capital: Evidence from PISA



Brown, Kauer, Kingdon, and Schofield, *QJE* 2024

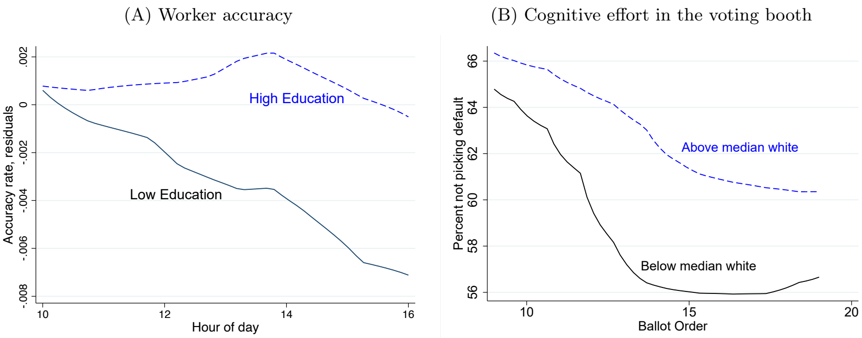
Figures notes on TIMSS and PISA results

Performance Declines in Achievement Tests

Notes: The figures show student performance over the length of the TIMSS and PISA tests, both administered in over 50 countries. Graphs plot residuals after removing question fixed effects for TIMSS and question block fixed effects for PISA (the unit at which questions are randomized). The x-axis denotes where in the exam the question item appeared normalized on a scale of zero to one (i.e., the question number in subject in TIMSS and the question block number across the exam in PISA). The y-axes plot the average score (i.e., percent answered correctly) for each question location on the test. The plots display the smoothed values of a kernel-weighted local polynomial regression, with a bandwidth of 0.15 for TIMSS and a larger bandwidth of 0.33 for PISA (due to the smaller number of randomization blocks). The intercept of the y-axis varies by group—with more (less) advantaged students on the left (right) axis. In the TIMSS U.S. sample (Panels A and B), relative advantage is proxied by race (white and nonwhite, respectively). In the TIMSS global sample, these differences are proxied by the top (bottom) quartile of GDP/capita (Panels C and D). In the PISA data (Panels E–J), high (low) socioeconomic status is proxied by the top (bottom) quartile of the ESCS measure, an index capturing parental income, occupation, and education.

Cognitive endurance as human capital: Evidence from (A) Indian data entry workers and (B) US voters

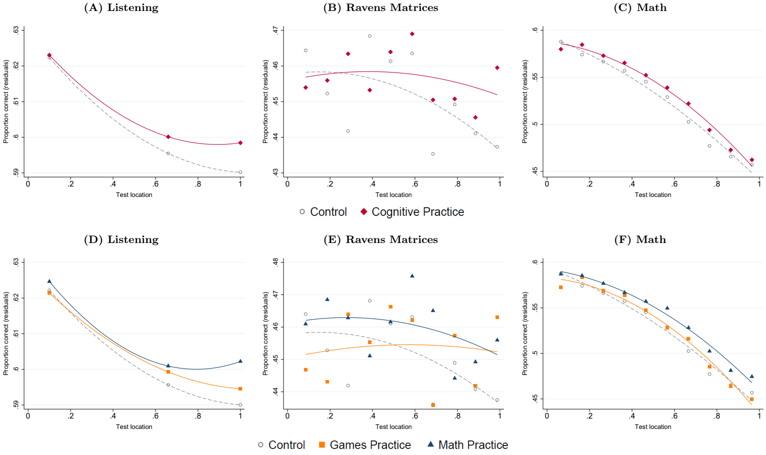
FIGURE III: Cognitive Endurance among Adults



Brown, Kauer, Kingdon, and Schofield, *QJE* 2024

Teaching cognitive endurance: Evidence from Indian schoolchildren

FIGURE II: Performance Declines on Experimental Tests by Treatment Group



Brown, Kauer, Kingdon, and Schofield, *QJE* 2024

Children's arithmetic skills do not transfer btwn applied v. academic math
Working children in Kolkata can do complex math in markets but not on tests

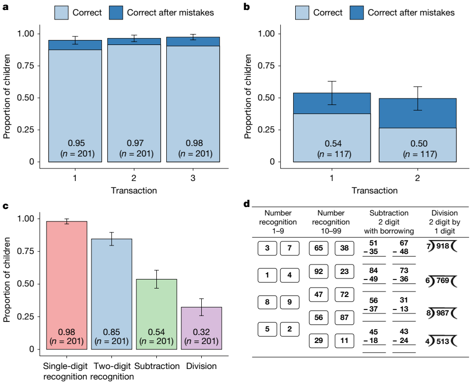


Fig. 1 | Performance of children in Kolkata working in a market (study 1). **a**, Proportion of children who correctly answered the total amount due in transactions involving two goods sold in unusual quantities. **b**, Proportion of children who correctly answered the total amount due in hypothetical transactions. **c**, Proportion of children who were credited with labelling single-digit numbers, labelling two-digit numbers, subtracting and dividing on the ASER test, a tool for assessing numeracy used across India for the ASER. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.). **d**, An example of this arithmetic assessment tool. From left to right, each column presents items that respectively task children with identifying single-digit numbers, identifying two-digit numbers, subtracting one two-digit number from another with carrying, and long division of a three-digit number by a single-digit number with a remainder. For grade 5 children, the ASER test begins with the subtraction task and proceeds to the right (division) if children succeed and to the left (identifying two-digit numbers) if they fail. Because success at each level requires mastery of the operations required by the tasks to its left, children who succeed at a given level are credited with mastery of all the levels below it.

Banerjee et al., *Nature* 2025

Working children can do market math but not school math
Non-working children can do school math but not market math

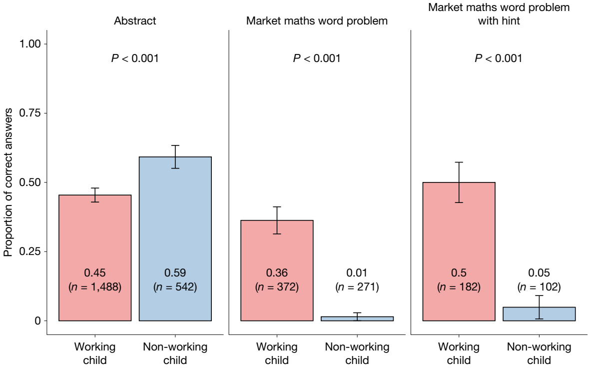


Fig. 4 | Comparison of working and non-working children in oral abstract and market maths word problems (study 3). Left, comparison of the performance of working and non-working children on the same abstract subtraction and division problems. Middle, comparison of the performance of working and non-working children on the market maths word problem after a single attempt. Right, comparison of the performance of working and non-working children on the market maths word problem after receiving a hint to break down the problem and to make use of rounding strategies (process + rounding hint). The oral abstract problems and market maths word problems were not matched in difficulty with one another. A child was coded as correct if they initially correctly answered the question or after receiving the hint. Error bars show 95% CIs around the mean (mean ± 1.96 × s.e.m.). *P* values are calculated using two-tailed *t*-tests without adjustments for multiple comparisons.

Banerjee et al., *Nature* 2025

Agenda

1. **The Flynn puzzle**
2. **The Bell Curve controversy**
3. **Malleability of cognitive human capital**
4. **Social skills: Not just “IQ”**
   - Rising supply of economically valuable personality traits in Finland
   - Rising premium to social skills in Sweden
5. **What makes a good manager?**
6. **Why do social skills matter—and more so over time?**
7. **Production of skills in equilibrium**

Agenda

Rise in economically valuable personality traits among Finnish men—relative to 1962 distribution

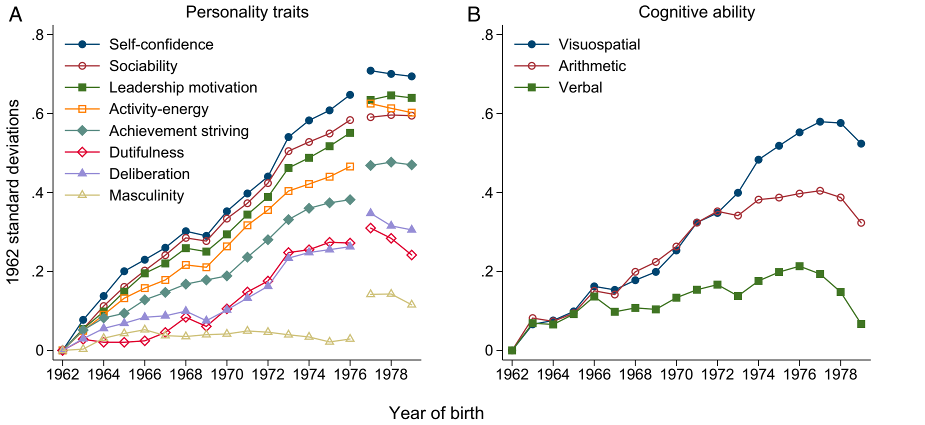


Fig. 1. Average scores for measures of (*A*) personality traits and (*B*) cognitive ability by birth year for native-born military conscripts in Finland. All scores are depicted in base year SDs, with base year means normalized at zero. The break in personality test scores reflects a change in test administration.

Jokela et al., *PNAS* 2017

Rank correlation with earnings has not changed much

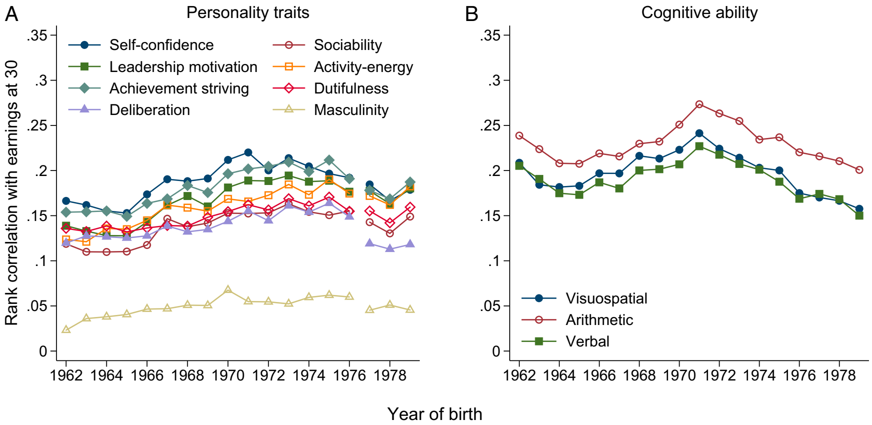


Fig. 2. The relation of earnings and (*A*) personality traits and (*B*) cognitive ability by birth cohort, measured as the within-cohort rank correlation between the test score and annual earnings at age 30. *SI Appendix*, Fig. S1 shows the same relations for average annual earnings at age 30–34, which is a better measure of lifetime earnings but not observed for the last three cohorts. The break in personality test scores reflects a change in test administration.

Jokela et al., *PNAS* 2017

Predicted earnings using time invariant $\beta' s$: 10-12% higher predicted earnings for personality and cognitive ability

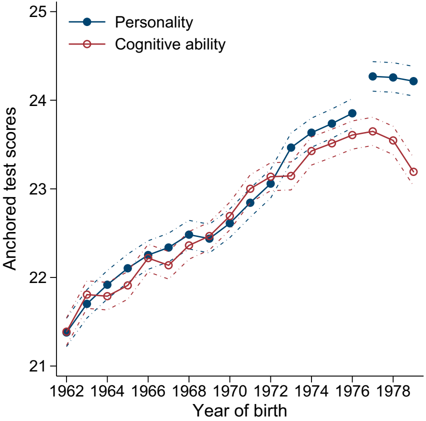


Fig. 3. Average of anchored test scores by birth cohort, with anchoring to average annual earnings at age 30–34 (in 1,000s of 2010 Euros) using the 1962–76 birth cohorts for estimating the prediction model. Dashed lines depict 95% confidence intervals. The break in personality test scores reflects a change in test administration.

Rising predicted earnings across family backgrounds

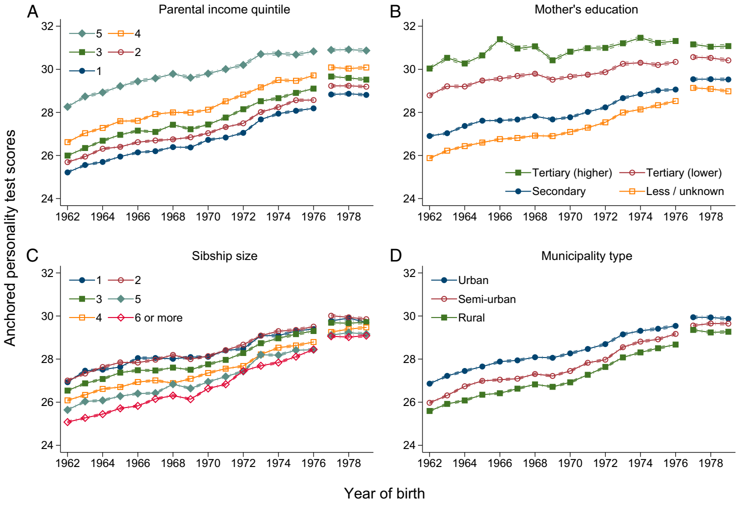


Fig. 4. Evolution of economically valuable personality traits across birth cohorts by (*A*) parental income quintile, (*B*) mother's education level, (*C*) sibship size, and (*D*) urbanization of birth place. Test scores are anchored to earnings at age 30–34. *SI Appendix*, Fig. S18 shows that the broad picture is similar when this analysis is repeated one trait at a time. Dashed lines depict 95% confidence intervals. The break in personality test scores reflects a change in test administration.

Improving family background can explain a lot of improvement...

Table 1. Cohort trends and demographic backgrounds

Variable	Change between 1962 and 1976 cohorts: Observed	Change between 1962 and 1976 cohorts: Predicted	Share predicted, %
Personality			
Self-confidence	0.65	0.16	25
Sociability	0.58	0.15	26
Leadership motivation	0.55	0.19	34
Activity-energy	0.47	0.09	20
Achievement striving	0.38	0.17	44
Dutifulness	0.27	0.11	41
Deliberation	0.26	0.04	14
Masculinity	0.03	0.00	–15
All (anchored)	0.57	0.19	33
Cognitive ability			
Visuospatial	0.55	0.25	45
Arithmetic	0.40	0.26	65
Verbal	0.21	0.25	119
All (anchored)	0.44	0.28	64

"Observed" is the actual difference in means between the birth cohorts, and "predicted" is the mean of predicted values for this difference, based on age at test, parental income, mother's and father's levels of education, sibship size, and rural/urban status, using the model estimated for the 1962 cohort. All variables were measured in 1962 SDs. Bootstrapped SEs are below 0.007 for all observed and below 0.015 for all predicted means.

Jokela et al., *PNAS* 2017

Agenda

Rising return to ‘non-cognitive’ skills in Sweden, 1992–2013

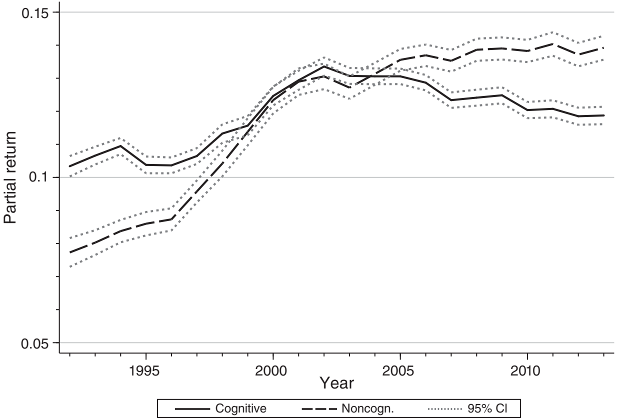


Figure 1. The Wage Return to Cognitive and Noncognitive Skill, 1992–2013

Notes: Confidence bands are based on robust standard errors. All estimates are corrected for measurement error using reliability ratios estimated by Grönqvist, Öckert, and Vlachos (2017). Online Appendix Section A1.3 outlines the procedure.

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

Selection into work: Skills, annual earnings, pr(working)

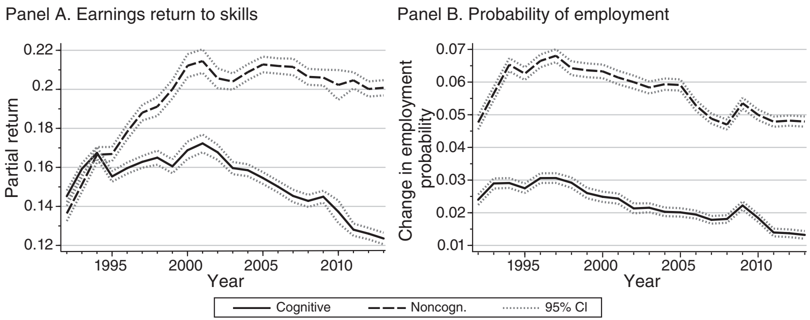


Figure 2. Annual Earnings, Employment, and Skills (All Males Aged 38–42)

Notes: Confidence bands are based on robust standard errors. All estimates are corrected for measurement error using reliability ratios estimated by Grönqvist, Öckert, and Vlachos (2017). Online Appendix Section A1.3 outlines the procedure.

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

Correct for measurement error w/ siblings (could also impute scores to sisters!)

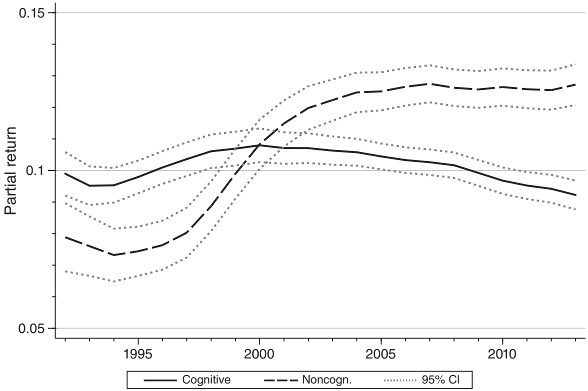


Figure 3. IV Estimates Using Brothers' Skills as Instruments for Own Skills

Notes: IV estimates of equation (1) using the skills of the brother as an instrument for own skills. To decrease the variability in the estimates, we use data for the year in question ± 2 years. Confidence bands are based on standard errors, which are clustered on individuals.

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

What about selection into – and conditioning on – education?

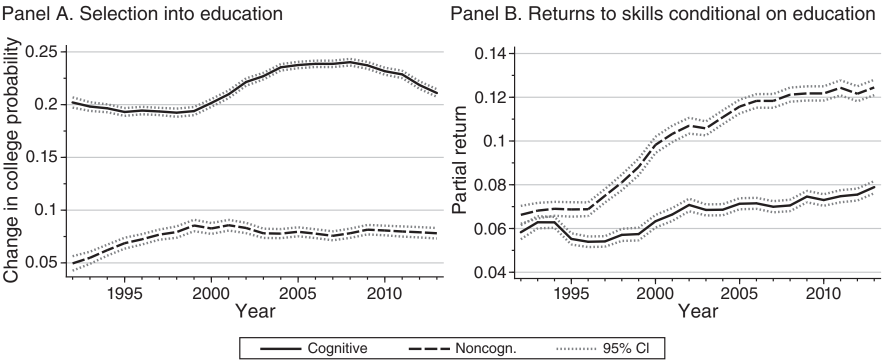


FIGURE 4. EDUCATION AND WAGE RETURNS

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

Changing wage slopes on skills

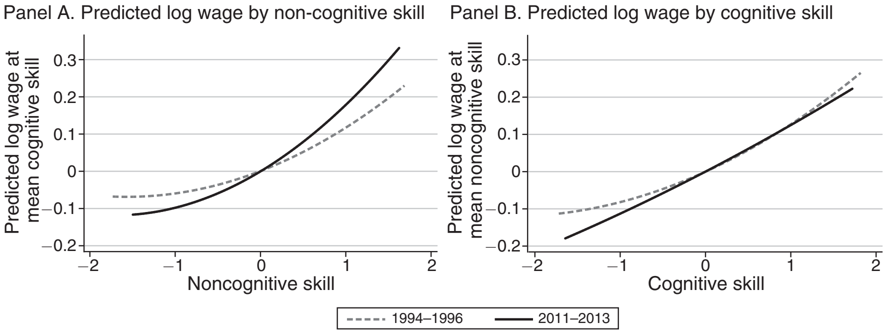


Figure 6. Predicted log Wages across the Skill Distributions

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

Increasing selection of social relative to cognitive skills into skilled occupations?

We mainly focus on changes in the *relative* intensity of noncognitive skill use and *relative* return to noncognitive skill between two points in time, 1995 and 2012.[29] We define relative skill intensity as the average difference between noncognitive and cognitive skill for individuals employed in occupation j at time t:[30]

$$(2) \qquad \left(\textit{relative skill intensity}\right)_{jt} = \left(s_{jt}^{n} - s_{jt}^{c}\right).$$

Analogously, the relative return is given by

$$(3) \qquad \left(\textit{relative return}\right)_{jt} = \left(\beta_{jt}^{n} - \beta_{jt}^{c}\right).$$

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

High paying occupations are increasingly social skill intensive

TABLE 3—CHANGES IN SKILLS AND RETURNS ACROSS OCCUPATIONS

Ranked (0/1) occ. characteristic	$\Delta\,(\textit{relative skill intensity})_j$ (1)	$\Delta\,(\textit{relative return})_j$ (2)	$\Delta\,(\textit{cog. return})_j$ (3)
Abstract	0.129	0.030	−0.025
(corr $\ln w_{j,95} = 0.85$)	(0.016)	(0.008)	(0.004)
Routine	−0.137	−0.041	0.012
(corr $\ln w_{j,95} = -0.61$)	(0.016)	(0.009)	(0.005)
Automatable	−0.072	−0.038	0.050
(corr $\ln w_{j,95} = -0.75$)	(0.017)	(0.008)	(0.004)
Offshorable	0.128	0.032	−0.001
(corr $\ln w_{j,95} = 0.46$)	(0.016)	(0.009)	(0.005)
Social	0.136	0.025	−0.000
(corr $\ln w_{j,95} = 0.72$)	(0.016)	(0.009)	(0.005)
Wage in 1995	0.147	0.040	−0.012
	(0.016)	(0.008)	(0.004)

Notes: The results come from separate regressions for each cell. The regressions are estimated at the individual level as outlined in online Appendix Section A4.3. Standard errors clustered on individuals are in parentheses. Occupational information has been matched to the O*NET database to obtain job requirements. The classification of Abstract, Routine, and Offshorable jobs follows Acemoglu and Autor (2011), the classification of occupations requiring social skills comes from Deming (2017), and the information on automatable occupations from Frey and Osborne (2017); we thank Fredrik Heyman for providing the information on automatable jobs.

Edin, Fredriksson, Nybom, and Öckert, *AEJ: Applied* 2022

Agenda

1. The Flynn puzzle
2. The Bell Curve controversy
3. Malleability of cognitive human capital
4. Social skills: Not just "IQ"
   - Rising supply of economically valuable personality traits in Finland
   - Rising premium to social skills in Sweden
5. **What makes a good manager?**
6. Why do social skills matter—and more so over time?
7. Production of skills in equilibrium

Why study added value of managers?

Management probably matters a lot

- Vast set of compelling correlations from Bloom, Sudan, and Van Reenen (World Management Survey)
- Large, persistent productivity differences between managers within and across firms
- Good managers increase productivity through: monitoring, training, motivating staff, reallocating workers to better tasks

But do we know how to identify/predict who will be a good manager?

- Heavily reliant on judgment of existing managers $\rightarrow$ well-known biases
- Often based on personality traits, cognitive ability, or past performance
- The "Peter Principle": employees promoted from things they're good at to things they're bad at (Benson, Li, Shue QJE 2019)

Central question: Can we prospectively identify people with strong managerial potential using measurable skills—and if so, what are those skills?

Identification challenges

Non-random assignment

- Managers *not* randomly assigned to teams in the field
- Difficult to separate causal managerial contributions from:
  - Worker quality (sorting/selection)
  - Match-specific effects
  - Firm/team characteristics

Selected sample

- Managers in the field are a highly non-random sample
- Benson et al. (2019): Sales managers selected based on line worker performance
- Can induce *negative* correlation between worker skills and managerial performance
- Even if the true causal effect of skills on management is positive.

$\Rightarrow$ Need experimental approach to isolate management skill from productive capacity

The ideal experiment

1. **Random assignment** of managers to teams
   - Breaks correlation between manager quality and team composition
2. **Repeated assignment** of same manager to multiple teams
   - Separates persistent manager effects from idiosyncratic match effects
3. **Control for productive capacity**
   - Isolates "management skill" from ability to do the underlying task
4. **Variation in manager selection mechanism**
   - Compare self-selected vs. randomly assigned managers

Overview of the experiment

Sample & setting

- $n = 555$ participants
- 728 groups (3 people each)
- Pre-registered at AEA RCT Registry
- Essex University Economics Lab (UK)

Four phases

1. Broad skill assessments (at home)
2. Individual production skills (in lab)
3. Group testing: 4 rounds
4. Post-experiment survey

Design

- Each manager randomly assigned to **4 different teams**
- Teams work on "Collaborative Production Task"
- Roles persist throughout experiment
- Manager preferences elicited before assignment

Treatments (session-level):

- **Lottery**: Managers chosen randomly
- **Self-promotion**: Those with strongest preference to manage become managers

Skill measures

Broad measures of individual skill

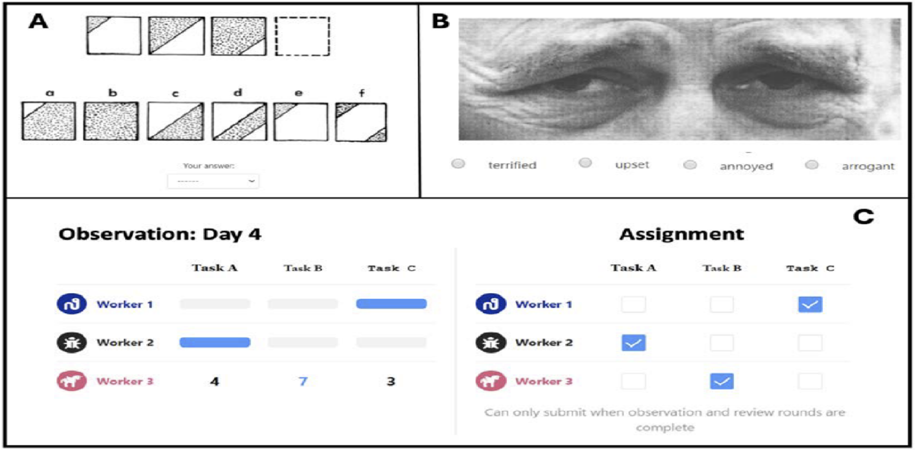
- **Economic decision-making:** Assignment Game (Caplin et al., 2024)
  - Attentionally-demanding numerical environment
  - Understanding comparative advantage
  - Avoiding biases (e.g., anchoring)
- **Fluid intelligence:** Ravens Advanced Progressive Matrices
  - Ability to solve novel problems
- **Emotional perceptiveness:** RMET (Baron-Cohen et al., 2001)
  - Identify emotions from photos of eyes
- **Production skills:**
  - Analogous to (later) team tasks: numerical,

Self-reported personality & working styles

- **Big 5 personality:** 10-item inventory (Gosling et al., 2003)
  - Positive correlations with job performance (Hurtz & Donovan, 2000)
- **Political skill inventory:** (Ferris et al., 2005)
  - Ability to understand and influence others at work
- **Risk appetite:** Single question (Dohmen et al., 2011)
  - Scale from 1 (risk averse) to 10 (risk seeking)
  - Strongly predictive of actual risky behavior

Example items from skill tests

Figure 4: Example items from broad tests of individual skill



Notes: Panel A is an example item from CFIT, a measure of fluid intelligence. Panel B is an example item from RMET, a measure of emotional perceptiveness; the correct answer is 'upset'. Panel C is a screenshot from the Assignment Game. For a full description of the game, see Caplin et al. (2024).

Who wants to be a manager? Preference to be in charge (scale 1–10)

	Full Sample (1)	Men (2)	Women (3)
Demographics			
Female	-0.46^{*}	—	—
Age, education, experience	n.s.	n.s.	n.s.
Skill Measures			
Production skills	0.27^{*}	0.43^{**}	0.03
Economic decision-making (AG)	0.21	0.60^{***}	-0.18
Fluid IQ (Ravens)	0.23	0.04	0.42^{*}
Emotional perceptiveness (RMET)	-0.13	-0.02	-0.17
Personality & Working Styles			
Extraversion	0.49^{***}	0.78^{***}	0.22
Agreeableness	-0.24	-0.50^{**}	-0.01
Risk appetite	0.42^{***}	0.43^{**}	0.39^{*}
n	509	265	244

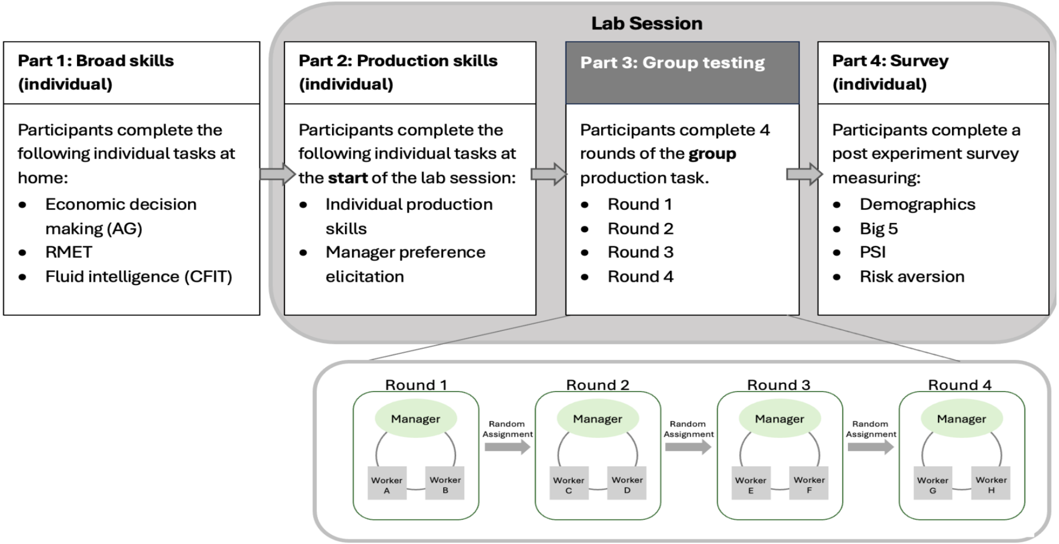
Key patterns: Women less likely to want to manage; extraversion and risk appetite predict desire to manage; skill measures show gender differences

Self-selected managers don't look worse on predictive dimensions...

Table 1: Sample and balance

	Overall sample (1)	Self-promoted arm (2)	Lottery arm (3)	p-value (4)
Demographics				
Female (%)	46.3%	49.1%	43.3%	0.17
Age mean (yrs)	25.0	25.1	24.9	0.37
Work experience mean (yrs)	2.5	2.5	2.5	0.98
Asian or Asian British (%)	53.9%	52.6%	55.3%	0.54
White (%)	18.3%	17.8%	18.8%	0.76
Black, Caribbean or African (%)	15.5%	17.0%	13.9%	0.32
Other ethnic identity[1]	12.3%	12.6%	12.0%	0.10
Graduate students (%)	67.6%	70.0%	65.0%	0.22
Skill assessments				
Task skills	0.00	-0.01	0.01	0.86
Fluid intelligence (Ravens)	0.00	-0.01	0.01	0.81
Economic Decision-making (AG)	0.00	-0.04	0.05	0.30
Emotional perceptiveness (RMET)	0.00	-0.07	0.08	0.08

Experimental flow (fig 1)



The collaborative production task

Task structure:

- Groups work on 3 question modules: *numerical, spatial, analytical*
- Each person works on their own computer solving problems
- $+1$ point for correct, -0.5 for incorrect answers

Scoring rule: 'Weakest link'

$$G_g = \min\{\text{Numerical Score}, \text{Spatial Score}, \text{Analytical Score}\}$$

- Team's score $=$ **minimum** module score
- Mimics real production: all components needed for success
- Requires coordination, not just individual effort

Manager's role

1. **Allocate:** Decide who works on which module (can reassign dynamically)
2. **Monitor:** Only manager sees overall team score; must identify bottlenecks
3. **Motivate:** Workers have no financial incentives; tasks are demanding

Identification strategy: Estimating manager effects

Model for group output

$$G_g = \gamma_M \sum_i X_i M_{ig} + \gamma_W \sum_i X_i W_{ig} + \epsilon_g$$

- G_g = group g's performance (standardized within round)
- X_i = individual i's productive capacity (measured before group task)
- M_{ig}, W_{ig} = indicators for manager/worker status in group g

Technique: Residual $\hat{\epsilon}_g$ captures performance *beyond* productive skill endowment

Estimating manager i's contribution

$$\hat{a}_i = \frac{1}{\sum_g M_{ig}} \sum_g M_{ig} \hat{\epsilon}_g$$

Average residual across manager i's teams = their **causal contribution**
(Random assignment + repeated observation identifies this)

Multilevel model and inference

Testing whether managers matter

$$\hat{\epsilon}_{gi} = \alpha_i + e_{gi}$$
$$\alpha_i \sim N(0, \sigma_\alpha^2)$$
$$e_{gi} \sim N(0, \sigma^2)$$

- σ_α = SD of manager effects = "typical manager effect"
- Impact of having a manager who is 1 SD above average

Null hypothesis: $H_0 : \sigma_\alpha = 0$ (managers have no impact after controlling for skills)

Inference

- Pre-registered: Randomization inference (permute manager assignments)
- Robustness: Wald estimator, Profile Likelihood

Profile likelihood for variance estimation

The challenge: Testing whether $\sigma_\alpha > 0$ is a *boundary problem*

- Null hypothesis ($\sigma_\alpha = 0$) lies at edge of parameter space (since variance can't be negative)
- Standard Wald tests unreliable in this setting

Profile likelihood approach:

1. Fix σ_α at a particular value
2. Maximize likelihood over all *other* parameters (worker effects, round effects, σ^2, etc.)
3. Repeat for many values of σ_α to trace out a "profile"

This yields a likelihood function depending only on σ_α, used to:

- Find the MLE of σ_α
- Construct confidence intervals (values where profile likelihood isn't too far below maximum)
- Test $H_0 : \sigma_\alpha = 0$ reliably despite the boundary problem

Treatment: Self-promotion vs. lottery (fig 2)

Figure 2: Randomization scheme

Notes: the figure describes the participant flow and randomization scheme. Z is a random variable that determines the way in which managers will be assigned. S is a self-selection mechanism, based on participants preferences for being a manager. D is a random variable that assigns 1/3rd of participants to be a manager in a lottery.

Result 1: Managers matter

	Estimate	p-value
Manager effect ($\hat{\sigma}_\alpha$)	0.22 SD	0.03
Worker effect ($\hat{\sigma}_\Omega$)	0.04 SD	0.49
Manager's production skills	0.21***	
Workers' production skills	0.27***	

Interpretation

- A 1 SD better manager $\rightarrow$ 0.22 SD improvement in team performance
- *After* controlling for productive capacity of all team members
- Total manager effect (without conditioning): 0.28 SD

Comparison: Manager quality $\approx$ as important as combined productive capacity of workers

Result 1: Managers matter (table 3 detail)

	Dependent variable: Group Performance (G)					
	(1)	(2)	(3)	(4)	(5)	(6)
Manager effect ($\hat{\sigma}_\alpha$)	0.228	0.286	0.219	0.218	0.220	0.218
[Randomization inference]	[0.04]	[<0.01]	[0.05]	[0.05]	[0.04]	[0.05]
{Wald}	{< 0.005}	{< 0.005}	{< 0.005}	{< 0.005}	{< 0.005}	{< 0.005}
(Profile likelihood)	(0.04)	(0.01)	(0.05)	(0.05)	(0.04)	(0.05)
Worker effect ($\hat{\sigma}_\Omega$)	0.041	0.207	0.078	0.066	0.051	0
[Randomization inference]	[0.48]	[0.03]	[0.39]	[0.41]	[0.46]	
{Wald}	{0.40}	{< 0.005}	{0.04}	{0.04}	{0.20}	
(Profile likelihood)	(0.49)	(0.03)	(0.40)	(0.44)	(0.47)	
Controls						
Manager's production skills[1]	0.208		0.215	0.209	0.198	
	(0.040)		(0.040)	(0.041)	(0.041)	
Workers' production skills[2]	0.261		0.263	0.263	0.238	
	(0.035)		(0.035)	(0.035)	(0.040)	
Manager familiar w/ participants[3]			x	x	x	x
Manager risk appetite[4]				x	x	x
Variance team production skills[5]					x	x
Granular production skills[6]						x

Result 2: Self-promoters perform worse

	(1)	(2)	(3)	(4)
Self-promoted vs Lottery	-0.13^{*}	-0.13^{*}	-0.12^{*}	-0.10
	(0.07)	(0.07)	(0.07)	(0.07)
Production skills		✓	✓	✓
Fluid IQ, AG, RMET			✓	✓
Demographics, Personality				✓

Key finding: Teams led by self-promoted managers perform **0.10 SD worse**

- Equivalent to having a manager with 1 SD lower fluid IQ
- Robust to extensive controls

Mechanism: overconfidence

- Self-promoters overestimate their performance ($d = 0.41$ SD, $p < 0.01$)
- Particularly overconfident about social skills (negative correlation with RMET)
- Among self-promoters: self-reported "people skills" $\rightarrow$ *worse* performance

Result 3: What predicts good management?

Correlations with manager contributions $\hat{a}_i$ (lottery arm only)

Strong predictors

- Economic Decision-Making (AG): $\rho = 0.23^{**}$
- Fluid Intelligence (Ravens): $\rho = 0.22^{**}$

Not predictive

- Emotional Perceptiveness (RMET): $\rho = -0.02$
- Extraversion: $\rho = -0.07$
- Conscientiousness: $\rho = 0.00$
- Age, Gender, Experience: n.s.

Robust to controls for:

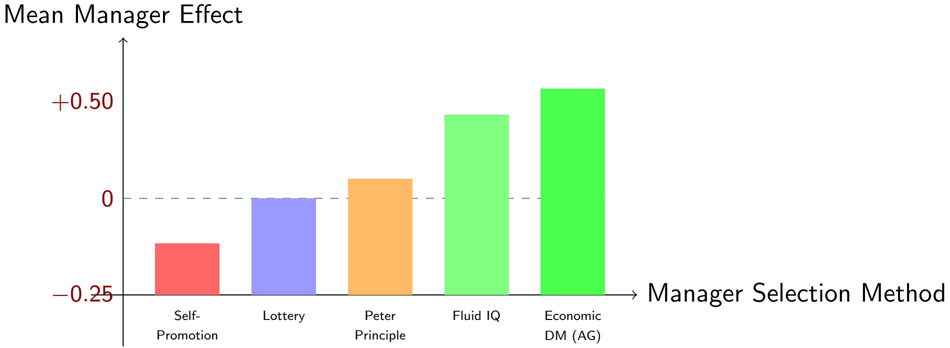
- Age, gender, education
- Work experience
- All Big 5 personality traits
- Risk preferences
- Political Skill Inventory

Among self-promoters:

- Extraversion: $\rho = -0.23^{**}$
- Political Skill: $\rho = -0.24^{**}$

("People persons" do *worse*!)

Result 4: Impact of selection mechanisms



Key finding: Selecting on economic decision-making skill improves manager quality by **0.7 SD** relative to self-promotion **Note:** Selecting on Peter Principle is based on tercile of production skills

Validation study 1: LinkedIn career outcomes

Sample: $n = 73$ experimental participants with LinkedIn job histories

Outcome: Promotions per year (base rate $= 0.25$/year)

Predictor	Effect on Promotions/Year	p-value
Lab manager performance (α_i)	$+0.16$	< 0.001
Economic decision-making	$+0.10$	0.03
Fluid intelligence	$+0.13$	0.01

Interpretation

- "Good managers" in lab (+1 SD) promoted every 2.3 years
- Others promoted every 5.3 years
- Robust to controls for IQ, demographics, personality
- Lab measure captures something *beyond* traditional predictors

Validation study 2: Retail firm in South America

Setting:

- Multi-billion dollar retail firm, 500 grocery stores
- 225 store managers completed same skill assessments as lab
- 28 months of store-level sales data (Feb 2020 – May 2022)
- Manager rotations across stores (quasi-random assignment)

Empirical strategy: Two-way fixed effects (Abowd et al., 1999)

$$\log(\text{Sales}_{smt}) = \alpha + \theta_m + \psi_s + \delta_t + \gamma \text{demos}_{st} + \epsilon_{smt}$$

Identification assumption: Manager assignments conditionally mean-independent of transitory shocks

- Supported by event study: symmetric gains/losses from manager moves
- No pre-trends before manager arrivals

Field results: The value of good managers

Event study result:

- Arrival of "above average" manager → **+13%** sales after 4 months
- 1 SD better manager → **+25%** annual sales
- $4.1 million USD per store per year

What predicts manager performance in the field?

	Coefficient	SE
Economic Decision-Making (AG)	0.19**	(0.07)
Fluid Intelligence (Ravens)	−0.05	(0.09)
Emotional Perceptiveness (RMET)	−0.15	(0.09)

Striking consistency: Same predictor (economic decision-making) works in lab *and* field!

What do good managers actually do? (table 7)

	Dep. var: $\hat{a}_i$		
	(1)	(2)	(5)
Initial allocation	0.204^{***}		0.136^{**}
	(0.072)		(0.064)
Monitoring errors		-0.335^{***}	-0.348^{***}
		(0.069)	(0.064)
Worker motivation			0.383^{***}
			(0.064)
R^2	0.042	0.113	0.292

- **Initial allocation:** Frequency manager finds optimal worker-to-task assignment
- **Monitoring errors:** Times workers expended effort on modules that couldn't improve group score (negative = good)
- **Worker motivation:** Manager's causal effect on rate workers solve puzzles (controlling for worker skill)

What do good managers actually do?

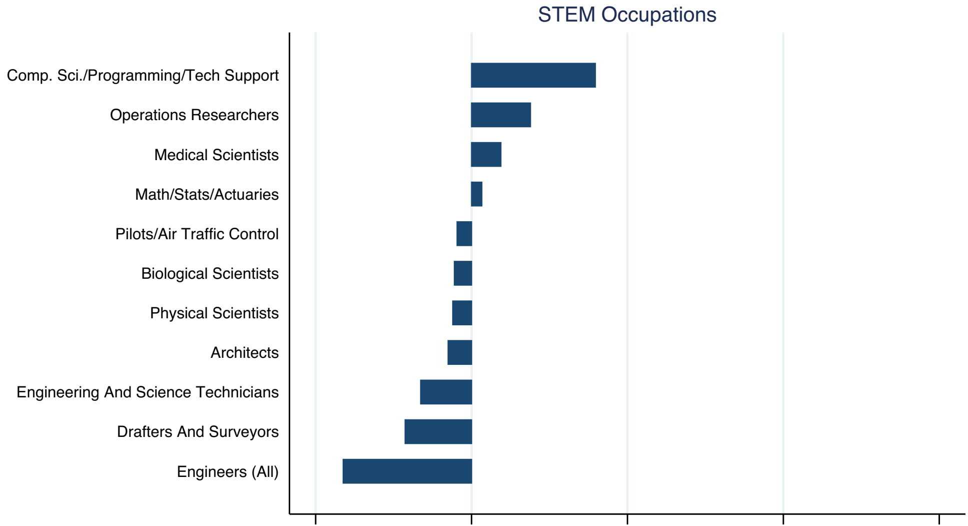
1. **Monitoring** (avoiding wasted effort)
   - Good managers (+1 SD): error rate 7.1% vs. 13.3% for average
   - Correlation with $\hat{a}_i$: $\rho = -0.34^{***}$
   - Field: Good managers $\rightarrow$ fewer stockouts (-3.7% of mean)
2. **Allocation** (matching workers to tasks)
   - Optimal initial assignment $\rightarrow$ 0.50 SD higher performance
   - Correlation with $\hat{a}_i$: $\rho = 0.18^{**}$
3. **Motivation** (increasing worker productivity)
   - Manager fixed effects on worker problem-solving rate
   - Correlation with $\hat{a}_i$: $\rho = 0.39^{***}$

Relative importance: Monitoring $\approx$ Motivation $>$ Allocation

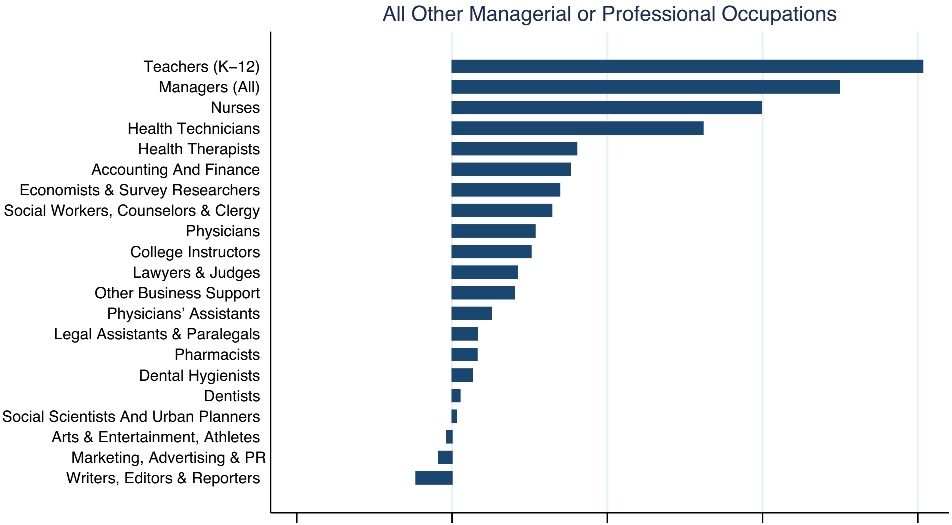
Agenda

1. **The Flynn puzzle**
2. **The Bell Curve controversy**
3. **Malleability of cognitive human capital**
4. **Social skills: Not just “IQ”**
   - Rising supply of economically valuable personality traits in Finland
   - Rising premium to social skills in Sweden
5. **What makes a good manager?**
6. **Why do social skills matter—and more so over time?**
7. **Production of skills in equilibrium**

Changes in employment shares, 2000–2012
By math and social skill requirements



Changes in employment shares, 2000–2012
By routine vs. non-routine cognitive tasks



A model of social skills

Motivation

- Economists focus on cognitive skills (IQ, schooling) as the key dimension of human capital
- But many jobs require teamwork—requires social skills

Deming's key idea: Social skills enable trade in tasks

- Workers have comparative advantage in different tasks
- Teams let workers specialize—like countries in Ricardian trade
- Social skills reduce transaction costs of trade
- Trade frictions lead to 'under-specialization'
  - When trade is costly: nontraded region of tasks that both (many) workers do for themselves

Begin with a simple task model

Task-specific production—labor is the only input

$$y_j(i) \;=\; A_j\,\alpha_j(i)\,l_j(i)$$

- $\alpha_j(i)$ = task-specific productivity; $l_j(i)$ = labor allocated to task i

Cobb-Douglas aggregator over tasks

$$Y_j \;=\; \exp\left[\int_0^1 \ln y_j(i)\,di\right]$$

Labor supply constraint

$$\int_0^1 l_j(i)\,di \;=\; L_j \;=\; 1$$

- Each worker supplies one unit of labor, allocated across tasks
- When $\alpha_j(i)$ is constant across i, eq. (2) collapses to eq. (1)—no gains from trade

Adding another worker: Tasks, and comparative advantage

Adding comparative advantage (2 workers)

- Two workers 1 and 2, each must complete a continuum of tasks $i \in [0,1]$
- Worker j's output on task i:

$$y_j(i) \; = \; A_j \cdot \alpha_j(i)$$

- $A_j =$ cognitive ability (a level shifter—scales up productivity on *all* tasks)
- $\alpha_j(i) =$ comparative advantage schedule on task i

Comparative advantage schedules

$$\alpha_1(i) = e^{-\theta i}, \qquad \alpha_2(i) = e^{-\theta(1-i)}$$

- Worker 1 has comparative advantage in low-i tasks; Worker 2 in high-i tasks
- Relative comparative advantage: $\gamma(i) \equiv \alpha_1(i)/\alpha_2(i) = e^{-\theta(2i-1)}$, strictly decreasing in i
- $\theta > 0$ governs how steeply comparative advantage varies across tasks
- Low θ: Little variation in comparative advantage across tasks (workers are similar)

Autarky: Working alone

In autarky, each worker must do all tasks $i \in [0,1]$ herself

Total output for worker j:

$$Y_j^A = A_j \int_0^1 \alpha_j(i)\, di$$

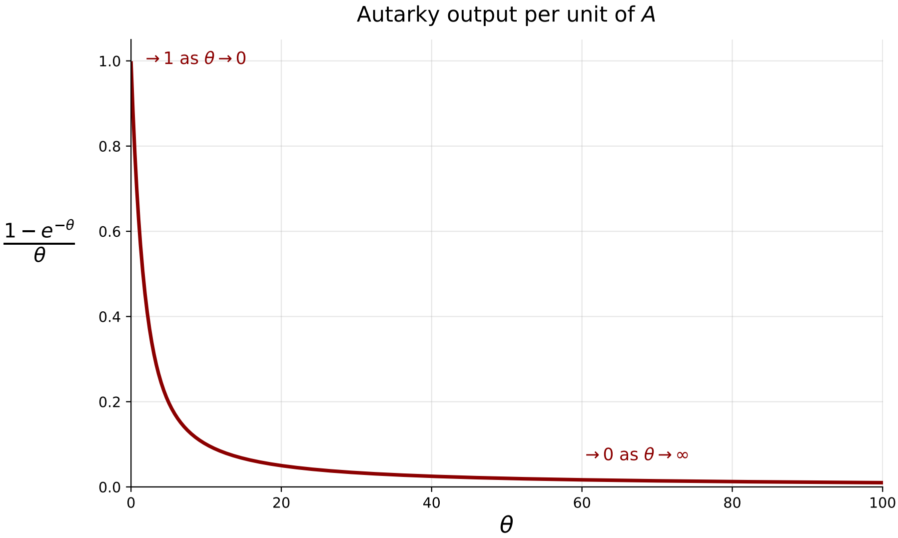
- For worker 1: $Y_1^A = A_1 \int_0^1 e^{-\theta i}\, di = A_1 \cdot \frac{1-e^{-\theta}}{\theta}$

Key features of autarky output:

- Proportional to cognitive ability A_j
- Decreasing in θ: when comparative advantage is steep, being forced to do all tasks is costly
- Worker 1 wastes effort on high-i tasks where she has no comparative advantage (and vice versa for worker 2)

$\Rightarrow$ This waste is what teamwork can remedy

Autarky output is decreasing in θ



Costless trade: Full specialization

If workers can trade tasks without friction ($S^* = 1$)

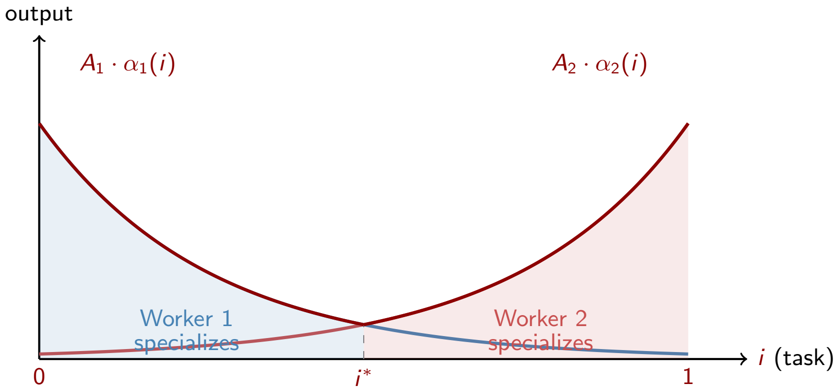
- For each task, assign it to the worker with comparative advantage
- Cutoff task i^* where both workers are equally productive:

$$A_1 \cdot \alpha_1(i^*) = \omega \cdot A_2 \cdot \alpha_2(i^*) \qquad \text{where } \omega \text{ is the relative wage}$$

Task allocation with costless trade:

- Worker 1 does tasks $i \in [0, i^*]$ (her comparative advantage region)
- Worker 2 does tasks $i \in [i^*, 1]$ (her comparative advantage region)
- Full specialization: Each task is done by exactly one worker
- Team output exceeds autarky output for both workers—gains from trade

Visualizing costless trade ($S^* = 1$)



- With costless trade, every task is done by the worker with comparative advantage
- **Upper envelope** = team output > either worker's autarky output

Social skills as the cost of trade

The friction: Teamwork is imperfect

- When worker 2 performs a task on behalf of worker 1, some output is lost
- Social skills determine how much of a teammate's output you can use

Formalizing social skills

- Each worker has social skill $S_j \in [0, 1]$
- The team's effective social skill:

$$S^* = S_1 \cdot S_2 \in [0, 1]$$

- $S^* = 1$: Frictionless trade—full specialization as on prior slide
- $S^* = 0$: No effective teamwork—autarky
- S^* is multiplicative: both partners need social skills for teamwork to work
- Traded output is scaled by S^*, so worker 1 receives $S^* \cdot y_2(i)$ for tasks done by worker 2

With trade frictions ($S^* < 1$), not all tasks are worth trading

Worker 1 trades task i to worker 2 only if

$$S^* \cdot A_2 \cdot \alpha_2(i) \; > \; A_1 \cdot \alpha_1(i)$$

- This defines a threshold i^H: Worker 1 trades tasks $i > i^H$ to worker 2

Symmetrically, worker 2 trades task i to worker 1 only if

$$S^* \cdot A_1 \cdot \alpha_1(i) \; > \; A_2 \cdot \alpha_2(i)$$

- This defines a threshold i^L: Worker 2 trades tasks $i < i^L$ to worker 1

Result: Three regions of task allocation

- $[0, i^L]$: Worker 1 specializes (does these for both workers)
- $[i^L, i^H]$: Nontraded—each worker does these tasks for herself
- $[i^H, 1]$: Worker 2 specializes (does these for both workers)

The size of the nontraded region

Width of non-traded region

$$i^H - i^L = \frac{-\ln(S^*)}{\theta}$$

Key comparative statics

- When $S^* = 1$ (no friction): $i^H - i^L = 0$—the nontraded zone vanishes, full specialization at i^*
- As $S^* \to 0$ (no social skill): $i^H - i^L \to \infty$—approaches autarky (whole task range is nontraded)
- Higher θ (steeper comparative advantage) $\Rightarrow$ smaller nontraded zone
  - When comparative advantage is steep, the gains from trade are large enough to overcome friction even for tasks near i^*
- Lower θ (flatter comparative advantage) $\Rightarrow$ larger nontraded zone
  - When workers are similar across tasks, even small frictions deter trade

How comparative advantage shape the nontraded region

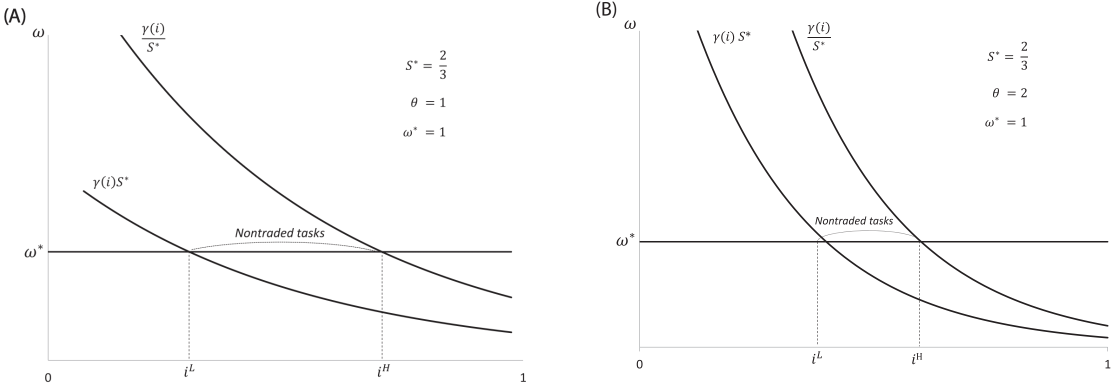


FIGURE II

Equilibrium Task Thresholds with Different Values of Theta

Low θ, weak comparative advantage

High θ, strong comparative advantage

Deming, *QJE* 2017

Equilibrium wages

Easy to see that wages rising in social skills

$$\ln w_1 = \ln P + i^H \ln A_1 + (1 - i^H)\left[\ln(S^* \cdot A_2) + \ln \omega\right] + \int_0^{i^H} \ln \alpha_1(i)\, di + \int_{i^H}^{1} \ln \alpha_2(i)\, di$$

Components

- $i^H \ln A_1$: Own cognitive ability, weighted by the share of tasks you do yourself
- $(1 - i^H) \ln(S^* \cdot A_2)$: Teammate's productivity on traded tasks, discounted by social skill friction
- $\omega = \frac{i^L}{1-i^H}$: Relative productivity in the marginal task

Three determinants of wages

1. **Cognitive ability** A_j: Scales up productivity on all tasks
2. **Social skill** S_j: Shrinks the nontraded zone, enabling more gains from trade
3. **Nonroutineness** θ: Larger potential gains from specialization

Q: *Why doesn't θ enter the wage equation directly—where did it go?*

Why cognitive skill and social skill are complements

Intuition from comparative advantage

- High-A workers are highly productive at their best tasks
- In autarky, must do all tasks
- Teamwork lets them specialize in their best tasks, to the degree that they have social skills
- Opportunity cost of the nontraded zone is *higher* for high-A workers

Formally

$$\frac{\partial^2 \ln w_j}{\partial A_j \, \partial S_j} > 0$$

- The return to social skill is *increasing* in cognitive ability
- The return to cognitive ability is *increasing* in social skill

$\Rightarrow$ Workers with both high A and high S earn a premium beyond the sum of separate returns

Why social skills matter more in nonroutine jobs

The role of θ (nonroutineness):

- High-θ jobs: Comparative advantage schedules are steep—each worker is very good at some tasks, very bad at others
- The cost of the nontraded zone is large: doing tasks where you lack comparative advantage is very wasteful
- Social skills shrink the nontraded zone by $-\ln(S^*)/\theta$—but the value of each traded task is higher when θ is large

Formally:

$$\frac{\partial^2 \ln w_j}{\partial S_j \, \partial \theta} > 0$$

- The return to social skill is *increasing* in job nonroutineness θ
- Routine jobs (θ low): Workers are near-generalists; little to gain from trade
- Nonroutine jobs (θ high): Specialists benefit enormously from teamwork

$\Rightarrow$ Social skills are most valuable in specialized, nonroutine occupations

Model yields four testable predictions

1. **Social skill has a positive return:** Workers with higher S_j earn higher wages because they shrink the nontraded zone and capture more gains from trade
2. **Cognitive and social skills are complements:** The return to social skill is increasing in A_j (and vice versa)—high-A workers waste more talent in autarky
3. **Social skill is more valuable in nonroutine jobs:** The return to S_j is increasing in θ, so high-S workers sort into nonroutine occupations
4. **Rising importance of social skills:** If jobs are becoming less routine over time (rising θ), the return to social skill is increasing

Deming presents suggestive evidence on each

Connection to the Ricardian trade analogy

Trade model	Deming model
Countries	Workers 1, 2
Goods	Tasks $i \in [0, 1]$
Technology	$A_j \cdot \alpha_j(i)$
Comparative advantage	$\gamma(i) = \alpha_1(i)/\alpha_2(i)$
Steepness of comparatve advantage schedule	θ
Transport costs / tariffs	Social skill frictions $(1 - S^*)$
Nontraded goods	Tasks in $[i^L, i^H]$
Gains from trade	Team output $>$ autarky output

Limitations and extensions

Insights

- Why social skills have labor market value (they shrink the nontraded zone, enabling specialization)
- Why cognitive and social skills are complements (high-A workers waste more talent in autarky)
- Why the return to social skills is rising (jobs are becoming less routine, i.e., θ increasing)

Important simplifications

- Team size is fixed at 2—no analysis of optimal team size
- Matching is exogenous—doesn't model who teams with whom
- No dynamics: Social skills are fixed, not investable
- Tasks are symmetric—doesn't distinguish types of social interaction (negotiation, persuasion, coordination, empathy)

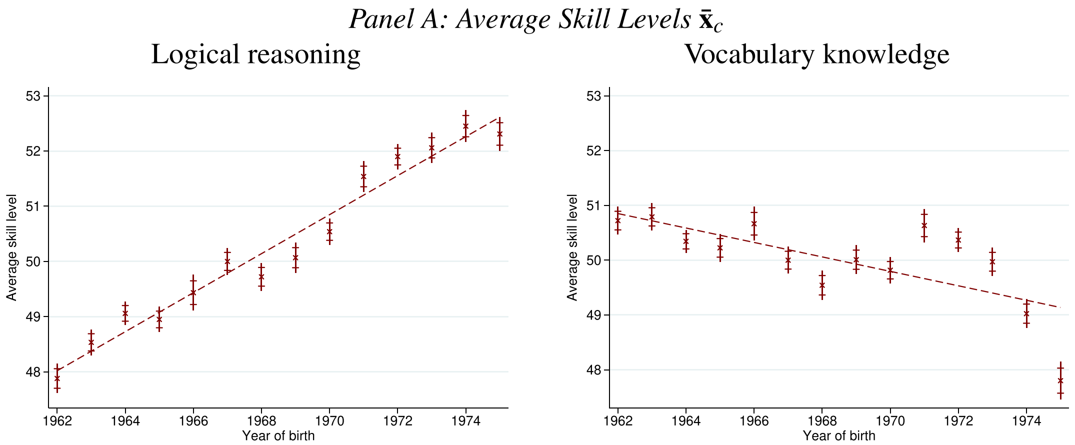
Agenda

1. The Flynn puzzle
2. The Bell Curve controversy
3. Malleability of cognitive human capital
4. Social skills: Not just "IQ"
   - Rising supply of economically valuable personality traits in Finland
   - Rising premium to social skills in Sweden
5. What makes a good manager?
6. Why do social skills matter—and more so over time?
7. **Production of skills in equilibrium**

‘Fluid’ versus ‘crystallized’ intelligence (Hermo et al. ’22)

- **Flynn effect gains are especially pronounced for fluid intelligence**
  - **Ability to solve novel reasoning problems** and is correlated with a number of important skills such as comprehension, problem-solving, and learning (Cattell 1963, 1971)
- **Less pronounced gains, or even declines, in crystallized intelligence**
  - **Ability to deduce secondary relational abstractions** by applying previously learned primary relational abstractions (Cattell 1971)
  - Hermo et al. ’22 equate this with factual knowledge such as vocab tests, though that’s arguably not the same thing
- **These are arguably complements not substitutes in prod’n of work**
  - Could be q-complements, p-substitutes
- **Measures from male military enlistment data**
  - Logical reasoning test consisted of drawing correct conclusions based on statements that are made complex by distracting negations or conditional clauses and numerical operations
  - Vocabulary knowledge test consisted of correctly identifying synonyms to a set of words

Average fluid and crystallized skills by Swedish birth cohort, YOB 1963-1975



Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Estimated lifetime skill premia: Swedish birth cohorts, 1963-1975

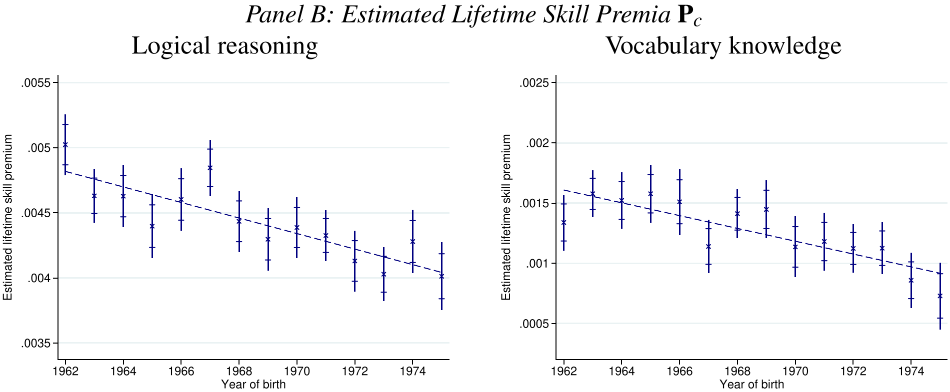


FIGURE I

Trends in Skills and Skill Premia across Birth Cohorts 1962–1975, Military Enlistment Sample

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Framework: Skills, endowments, and investment

1. An economy whose aggregate output is determined by the aggregate skills of workers
2. Skills, which can be multidimensional, are determined by an exogenous endowment (e.g., health) and an investment decision made early in life (by parents, children, and schools)
3. The investment decision is in turn influenced by the lifetime labor market returns to different skills.

Framework: Identification strategy

1. Identify the relative returns to different skills by assuming that unobserved determinants of an individual's earnings are correlated with the individual's skill endowment only through its market value
   - Under this assumption, the relative returns to different skills can be recovered from a Mincerian regression of the log of earnings on skills in a cross section of individuals
2. A single unknown parameter governs the degree to which individuals can substitute investment across skill dimensions
   - Identify this parameter by assuming that long-run average shocks to the technology for producing skills are proportional across fluid and crystallized intelligence
   - Violated if long-run improvements in skill production technology favor one skill dimension
   - Testing this assumption difficult. *Why?*
   - Imposes a restriction on changes in relative skill levels that would have occurred in the absence of changes in relative skill premia

Estimated skill quantities and prices

Panel A: Relative Skill Levels and Relative Skill Premia

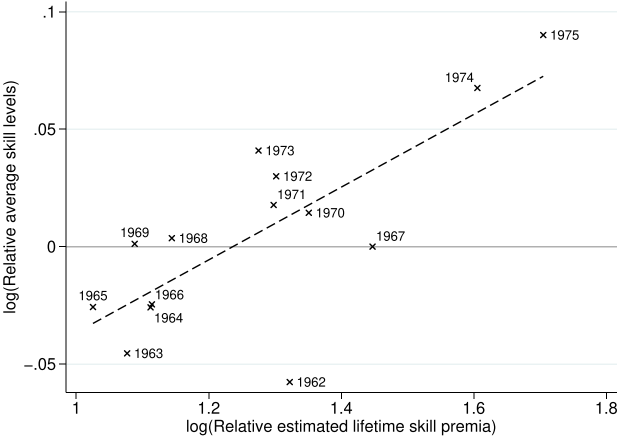


FIGURE II

Evolution of Relative Skill Levels and Relative Skill Premia, Military Enlistment Sample

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Estimating the supply function I

Under the model, the relative supply of fluid intelligence obeys

$$\text{(5)} \qquad \ln\left(\frac{\tilde{x}_{c1}}{\tilde{x}_{c2}}\right) = \frac{1}{\rho - 1}\ln\left(\frac{P_{c1}}{P_{c2}}\right) - \ln\left(\frac{K_{c1}}{K_{c2}}\right).$$

A standard difficulty in learning the elasticity of substitution $\frac{1}{\rho-1}$ is that the unobserved costs $\mathbf{K}_c$ may affect both skill investments (via the workers' incentives) and skill premia (via the labor market). We assume that, on average, there is no trend in the relative costs of the two skill dimensions.

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

A standard difficulty in learning the elasticity of substitution $\frac{1}{\rho-1}$ is that the unobserved costs $\mathbf{K}_c$ may affect both skill investments (via the workers' incentives) and skill premia (via the labor market). We assume that, on average, there is no trend in the relative costs of the two skill dimensions.

Assumption 1 (Zero average relative supply shock). We assume that

$$\frac{1}{\overline{c}-\underline{c}}\sum_{c=\underline{c}}^{\overline{c}-1}\left[\ln\left(\frac{K_{c+1,1}}{K_{c+1,2}}\right)-\ln\left(\frac{K_{c1}}{K_{c2}}\right)\right]=0.$$

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Assumption 2 is sufficient to identify the cohort-and-period-specific skill premia $\mathbf{p}_{t,t-c}$, and hence the lifetime skill premia $\mathbf{P}_c$, up to scale, from the conditional expectation function of the log of earnings.

PROPOSITION 2. Under Assumption 2, for some scalar $\alpha > 0$, a multiple $\alpha \mathbf{P}_c$ of the lifetime skill premia for each cohort c is identified from the conditional expectation function of the log of earnings,

$$\mathrm{E}(\ln(w_{it})|\mathbf{x}_i = \mathbf{x},\ \mathbf{d}_{it} = \mathbf{d},\ c(i) = c),$$

for each time period $t \in \{c+1, \ldots, c+A\}$.

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Estimated supply function

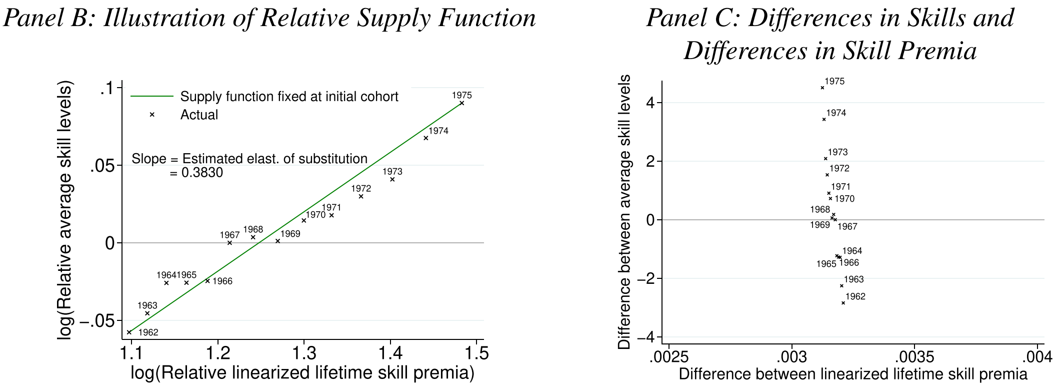


FIGURE II

Evolution of Relative Skill Levels and Relative Skill Premia, Military Enlistment Sample

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Counterfactuals: Effects of changing skill returns on skill endowments

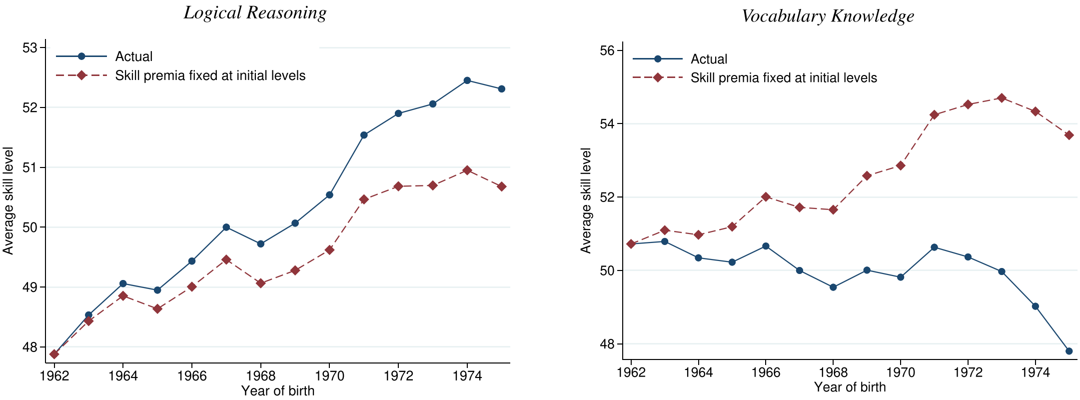


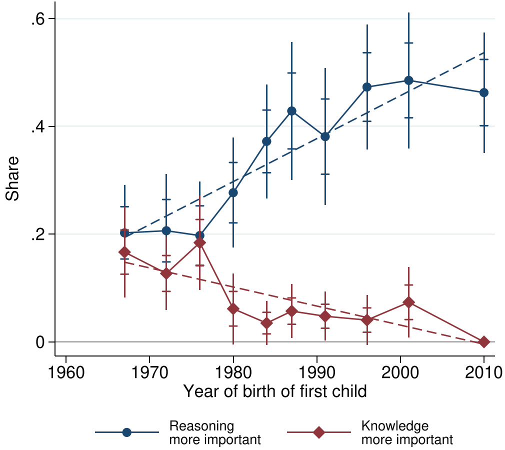
FIGURE IV

Decomposition of Change in Average Skill Level, Military Enlistment Sample

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Parental encouragement of skill formation by YOB of first child

Panel A: Which Skill Did Parents Encourage More in Their Own Children?



Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Skill demand changes implied by occupational change

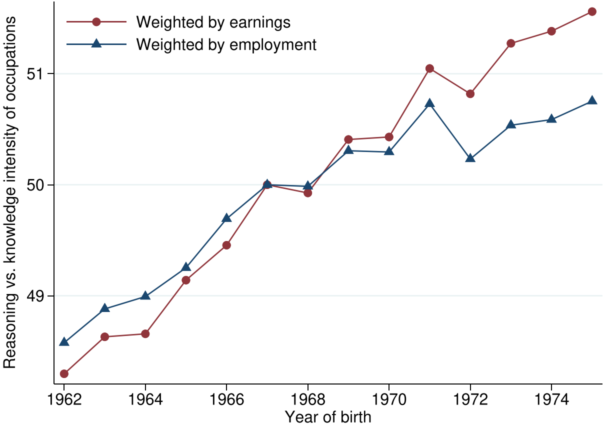


FIGURE VIII

Trends in the Reasoning versus Knowledge Intensity of Men's Occupations in Sweden

Hermo, Päällysaho, Seim, and Shapiro, *QJE* 2022

Conclusions: Hermo et al.

- **Key results**
  - A significant portion of the puzzling "Flynn effect" of rising fluid intelligence is potentially explained by investment substitution across different dimensions of skill
- **Potentially also explains decline in 'crystallized' intelligence across cohorts, if**
  - Individuals must substitute investment across these two skill dimensions
  - Technology for producing crystallized versus fluid intelligence has not become more or less biased over time towards one or the other skill

Lecture conclusions: Rethinking “skill”

The canonical narrow view and its limits

- Economists typically equate ‘skill’ with fixed cognitive ability (IQ)
- But: IQ rises across cohorts (Flynn effect); cognitive endurance appears trainable

Skill is more than just static cognitive compute

- Personality traits—especially interpersonal skills—matter greatly and may be rising in importance
- *Why* do interpersonal skills matter? Deming’s “trading tasks” theory is a nice exception
- Managerial skill appears distinct from both cognitive and social skill—and is almost unstudied
- Lots of work on ‘non-cognitive’ skills—but I still don’t know what this means

A broader conception of skill: Efficiency vs. allocation

- Classic lit distinguishes *factor efficiency* from *allocative efficiency* (Welch 1970; Schultz 1975)
- Is work best understood as “accomplishing tasks” or as “exerting agency to yield results”—choosing among tasks according to capabilities and applicability?
- Tasks are bundled in jobs, creating interdependencies that shape expertise requirements
- What determines production of skill in equilibrium (schooling is only one dimension)?