(Non)cogs in the Wheel

Non-Cognitive Traits and High School Graduation

Jonah Smith

April 2014

Overview

- What do we know?
- What do we want to know?
- How will we study it?
- What do we find?
- Conclusions

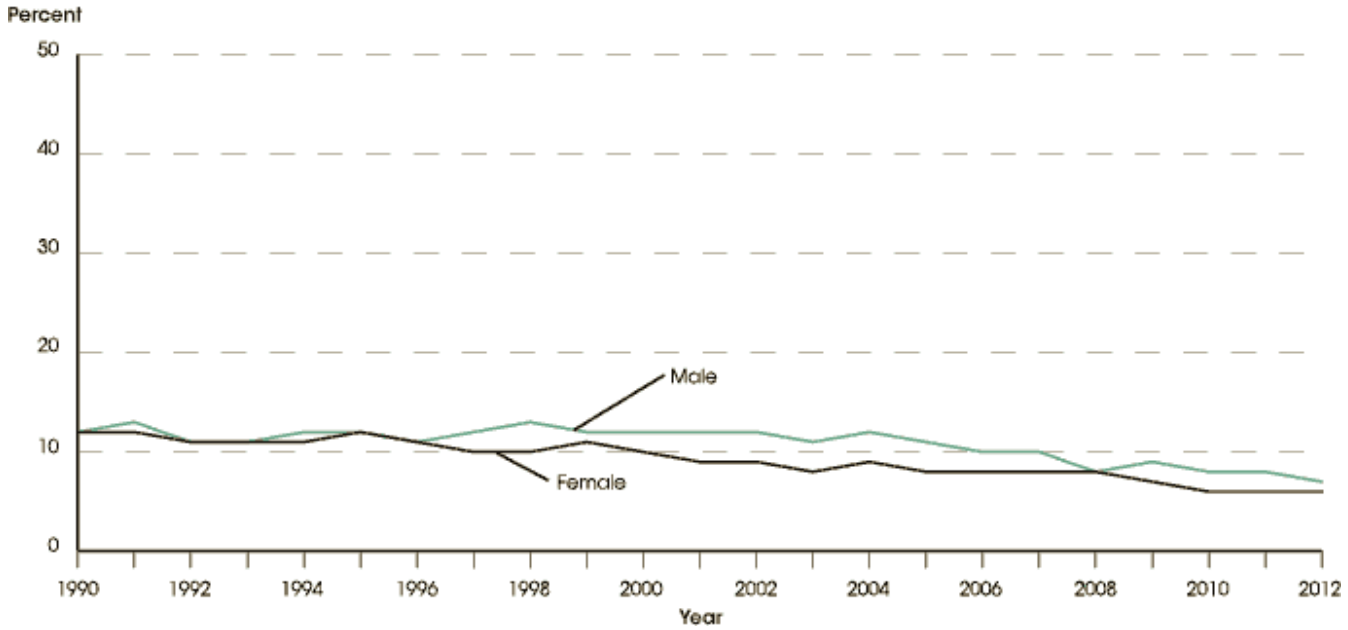
What do we know?

Does high school graduation matter?

- High school graduation is important
  - Economically
    - Secondary schooling was important for economic growth of the United States in the twentieth century (Aaronson and Sullivan, 2001; Delong, Katz, and Goldin, 2003).
    - Individual returns to completing high school (vs. not completing high school) have grown (Heckman, Lochner, and Todd 2008).
    - The wage advantage for completing high school operates through opening up opportunities for further education (Heckman, Lochner, and Todd 2008).
  - Socially
    - High school dropouts are more likely to…
      - suffer adverse health outcomes
      - participate in criminal activity
      - be civically disengaged

What do we know?

Does high school graduation matter?



Source: U.S. Department of Education, National Center for Education Statistics. (2013). The Condition of Education 2013 (NCES 2013-037), Status Dropout Rates.

- High school dropout rates have been falling steadily in the US… sort of.
  - Heckman and LaFontaine (2008) argue that dropout rate has *increased* among natives if you do not consider the growth of high school equivalency attainment

What do we know?

What do non-cognitive traits have to do with it?

- Studied in many disciplines
  - Psychology
    - "Cognitive self-regulation" (or: discipline; executive function)
    - "Emotional self-regulation"
  - Economics
    - "Non-cognitive" skills
      - Distinguish from cognitive skills emphasized in early literature
    - Personality traits
  - Sociology
    - Use terms above
    - Another notable term: "social and behavioral skills"
- **The bottom line:** factors other than "book smarts" are important predictors of many outcomes.

What do we know?

What do non-cognitive traits have to do with it?

TABLE 1 Examples of cognitive skills and noncognitive behaviors and traits whose relationship to stratification outcomes have been studied by previous researchers

Cognitive skills	**Noncognitive behaviors and traits**
English language composite	*Conscientious work habits*
Vocabulary, reading comprehension, spelling, capitalization, etc.	Effort (industriousness and perseverance), organization, discipline, attendance, participation, enthusiasm
Mathematics composite	*Other behaviors and traits*
Math concepts, problem solving, computation, etc.	Leadership, sociability (extraversion), self-confidence, social sensitivity, impulsiveness, openness to experience, emotional stability (calmness), vigor, aggressiveness, disruptiveness, high culture, locus of control, self-esteem
Aptitude and ability tests	
Abstract or mechanical reasoning, visualization, clerical checking, etc.	
Academic subject knowledge	
Measures of rote memory	

Source: Farkas (2003)

What do we know?

What do non-cognitive traits have to do with it?

- The relationship between non-cognitive traits and educational outcomes is *very* well established.
- Largely unexplored:
  - Predictive power of early measures
    - Dropout as a *process*, not an *event*
  - Single statistical construct for ‘non-cognitive traits’
    - Often divided up into ‘behaviors’ and ‘attitudes’
  - Nationally representative data
    - Usually district- or city- level

What do we want to know?

- What is the **relationship** between early levels of non-cognitive traits and later high school graduation?
- What factors **mediate** this relationship?
- How does the relationship **differ** between groups?

How will we study this?

- Panel Study of Income Dynamics (PSID)
  - Demographic and family background information
  - Child Development Supplement (grades 1-6 in 1997)
    - Teacher questionnaires evaluating student behavior (first use for evaluating noncognitive)
    - Standardized reasoning assessments (language and logical)
    - Information about behavior and performance in high school
  - Transition to Adulthood (after 18, left high school)
    - Information about graduation status
- Challenges
  - Summarizing a many-dimensional characteristic
    - Approach: factor analysis
  - Substantial missingness
    - Approach: multiple imputation (funded by Columbia)

Meet the students

Externalizing

	Student A 84th Percentile	Student B 50th Percentile	Student C 16th Percentile
Sudden mood swings	3	3	2
High strung	3	2	3
Cheats/tells lies	3	3	2
Argues too much	3	2	3
Difficulty concentrating	3	3	1
Bullies or is mean	3	3	3
Disobedient at school	3	3	2
Doesn't feel sorry	3	3	3
Trouble getting along	3	2	2
Acts without thinking	3	2	1
Restless/over-active	3	2	2
Stubborn or irritable	3	3	3
Strong temper	3	3	3
Breaks things	3	3	3
Clings to adults	2	3	2
Cries too much	3	3	3
Demands attention	3	2	2
Hangs around with trouble	3	3	2
Makes excessive demands	3	2	3
Academic underachiever	3	3	1
Goes through the motions	3	3	1
Acts up in class	3	3	2

Internalizing

	Student A 84th Percentile	Student B 50th Percentile	Student C 16th Percentile
Feels no love	3	3	3
Fearful/anxious	3	3	3
Easily confused	3	3	2
Feels inferior	3	3	2
Not liked by others	3	3	3
Has obsessions	3	3	2
Unhappy/sad	3	3	2
Withdrawn	2	3	3
Dependent on others	3	3	3
Feels paranoid	3	3	2
Secretive	3	3	3
Worries too much	3	2	3
Withdrawn from activities	3	3	3

What do we want to know?

- What is the **relationship** between early levels of non-cognitive traits and later high school graduation?

- What factors **mediate** this relationship?

- How does the relationship **differ** between groups?

Method

- Compute ‘non-cognitive’ construct
  - One-factor solution to 20 item evaluation
  - Procrustean rotation to maximally align between-imputation loading matrices

- Simple model
  - Logistic regression:

$$logit(\pi) = \beta_0 + \beta_1 X_1$$

    - where π is probability of graduation,
    - and X_1 is the variable of interest (i.e. non-cognitive traits)

- Controlled model

$$logit(\pi) = \beta_0 + \boldsymbol{\beta}\mathbf{X_c} + \beta_n X_n$$

    - where π is probability of graduation,
    - $\mathbf{X_c}$ is a vector of demographic control variables (gender, size-adjusted family income, race, and family structure),
    - and X_n is the variable of interest.

What do we find?

What is the **relationship** between early non-cognitive traits and high school graduation?

One standard deviation increase	**Not controlling for background**	**Controlling for background**
non-cognitive traits	**61%** increase in odds of graduation	**40%** increase in odds of graduation
cognitive skill	**64%** increase in odds of graduation	(Not significant)
size-adjusted household income	**338%** increase in odds of graduation	N/A

What do we find?

What is the **relationship** between early non-cognitive traits and high school graduation?

- What does this mean?
  - The baseline probability of graduation is **84%**
  - All else equal, if you are **one standard deviation above** the mean in non-cognitive levels, your probability of graduating rises to **89%**
  - All else equal, if you are **one standard deviation below** the mean, your probability of graduating falls to **79%**

What do we find?

What is the **relationship** between early non-cognitive traits and high school graduation?

- If…
  - their families all make average income
  - they are all male
  - they are all white
  - they all live with both parents

  - Student A has an **90%** chance of graduating

  - Student B has a **86%** chance of graduating

  - Student C has a **82%** chance of graduating

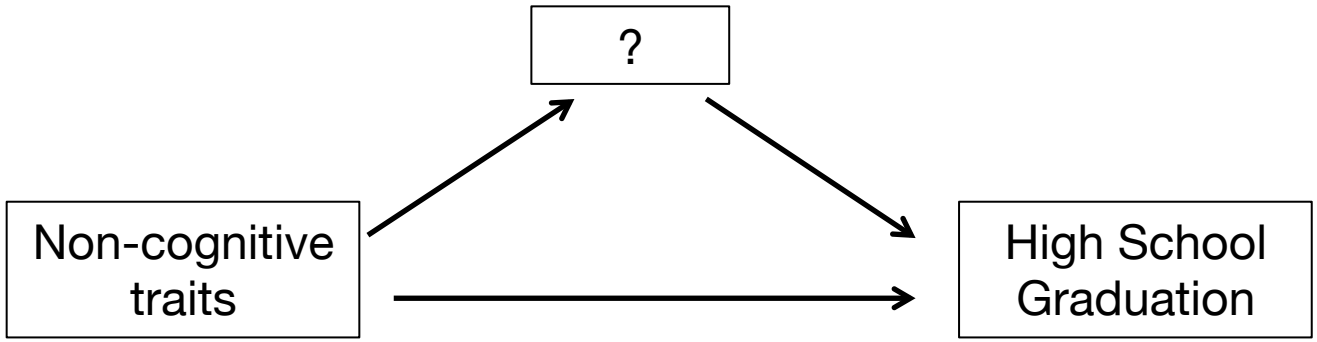
What do we want to know?

- What is the **relationship** between early levels of non-cognitive traits and later high school graduation?
- What factors **mediate** this relationship?
- How does the relationship **differ** between groups?

What do we find?

What factors **mediate** this relationship?

- We measure early non-cognitive traits and high school graduation… but what happens in-between?

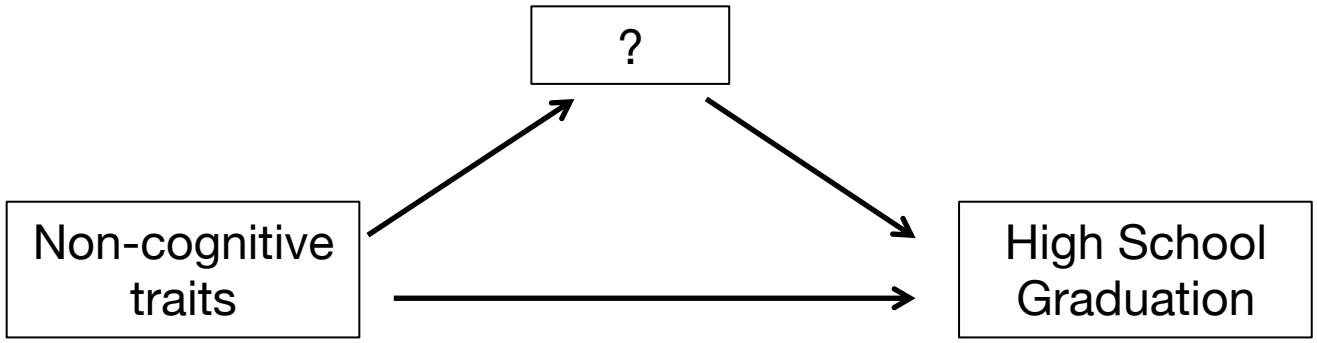


What do we find?

What factors **mediate** this relationship?

- Academic performance
  - measured by high school GPA (linear regression)

- Troublesome behavior
  - measured by counts of problem behaviors over last 6 months (Poisson regression)
    - Intentionally damaged school property
    - Skipping school (without permission)
    - Parents called into school because of misbehavior

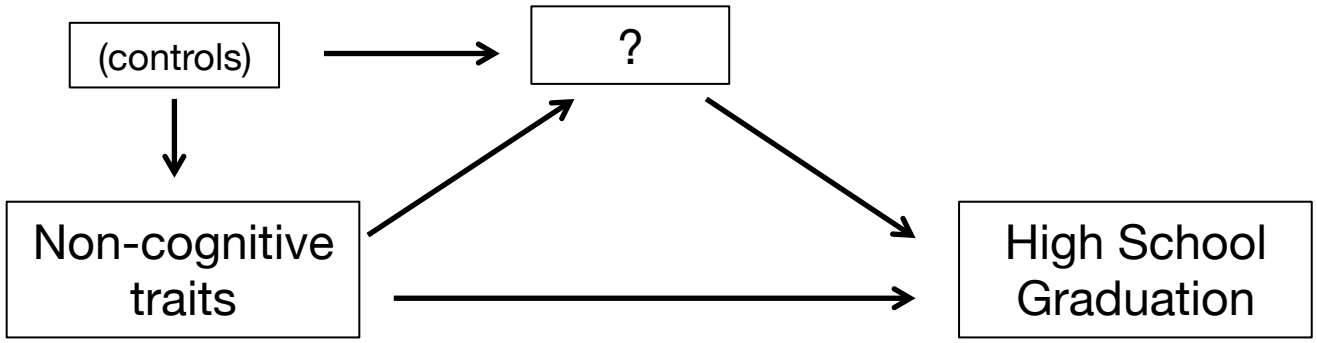


What do we find?

What factors **mediate** this relationship?

- Academic performance
  - measured by high school GPA (linear regression)

- Troublesome behavior
  - measured by counts of problem behaviors over last 6 months (Poisson regression)
    - Intentionally damaged school property
    - Skipping school (without permission)
    - Parents called into school because of misbehavior

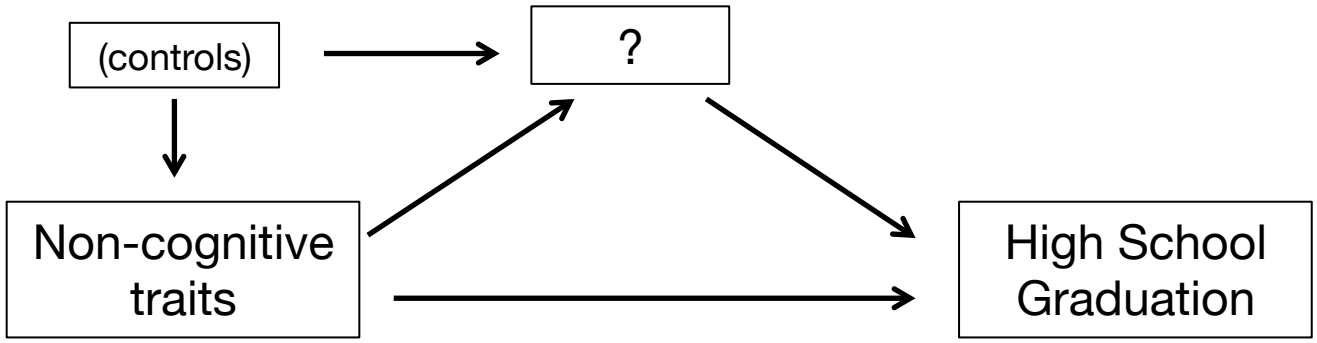


What do we find?

What factors **mediate** this relationship?

- My analysis was unable to uncover any variable that fits in the box with the below model at a statistically significant level.

- Why?
  - There might be no relationship
  - It could be the data quality
  - It could be the model



What do we find?

What factors **mediate** this relationship?

- If Student A, Student B, and Student C drop out…
  - it is *not* because their non-cognitive traits led them to get bad grades, nor because their noncognitive traits led them to act out.
  - it seems to be just the general day-to-day experience associated with non-cognitive traits…
    - their lack of attention
    - their lack of ambition
    - their short-term orientation
    - …

What do we want to know?

- What is the **relationship** between early levels of non-cognitive traits and later high school graduation?

- What factors **mediate** this relationship?

- How does the relationship **differ** between groups?

Method

- Model for interactions

$$logit(\pi) = \beta_0 + \boldsymbol{\beta}\mathbf{X_c} + \beta X_{noncog} + \beta_n(X_{noncog} \times X_n)$$

  - where π is probability of graduation,
  - $\mathbf{X_c}$ is a vector of demographic control variables (gender, size-adjusted family income, race),
  - X_{noncog} is the non-cognitive score,
  - and X_n is the variable of interest (among X_c).

(we interested in the significance and direction of X_n)

What do we find?

How does the relationship **differ** between groups?

- No evidence that the importance differs by…
  - Race
  - Gender
  - Household income

- Why?
  - There is no relationship
  - The data
  - The model

Conclusions

- Even early measures of non-cognitive traits are highly predictive of eventual high school graduation
  - More predictive than cognitive aptitude when controlling for demographic background
- Students whose non-cognitive traits lead them to drop out do not drop out because of poor grades or behavioral problems
  - The culprit is not clear, but it seems to be the ‘day-to-day experience’
- Non-cognitive traits seem to be equally predictive of outcomes for all students, regardless of background

Where do we go from here?

- Research
  - What are the pathways through which early non-cognitive traits influence outcomes?
  - What non-cognitive traits, specifically, are most predictive at this early stage? Behavioral? Attitudinal?

- Policy
  - Non-cognitive traits are important, even at an early age
  - Rhetorical and policy focus on improving test scores (cognitive skills) should be reevaluated