On the interpretation of non-cognitive skills – what is being measured and why it matters*

John Eric Humphries (University of Chicago)†

Fabian Kosse (University of Bonn)‡

January 2017

Abstract

Across academic sub-fields such as labor, education, and behavioral economics, the measurement and interpretation of non-cognitive skills varies widely. As a result, it is difficult to compare results on the importance of non-cognitive skills across literatures. Drawing from these literatures, this paper systematically relates various prototypical non-cognitive measures within one data set. Specifically, we estimate and compare several different strategies for measuring non-cognitive skills. For each strategy, we compare their relative effectiveness at predicting educational success and decompose what is being measured into underlying personality traits and economic preferences. We demonstrate that the construction of the non-cognitive factor greatly influences what is actually measured, how it relates to more standard taxonomies and what conclusions are reached about the role of non-cognitive skills in life-outcomes such as educational attainment. Furthermore, we demonstrate that, while sometimes difficult to interpret, factors extracted from self-reported behaviors can have predictive power similar to well established taxonomies, such as the Big Five.

Keywords: non-cognitive skills, personality, preferences, educational success.

JEL-Classification: J24, I20, D03, D90

*For valuable comments and discussions we thank Deborah Cobb-Clark, Thomas Dohmen, Angela Duckworth, Armin Falk, James Heckman, Robert Kaestner, Tim Kautz, Julia Nafziger, Pia Pinger, Daniel Schunk, Renata Sielecki, two anonymous reviewers and participants at various seminars.

†johneric@uchicago.edu, Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637, USA.

‡Corresponding Author, kosse@uni-bonn.de, +49 228 739222, Institute for Applied Microeconomics, University of Bonn, Adenauerallee 24 - 42, 53113 Bonn, Germany.

1 Introduction

Many traits, skills, and abilities matter for success in life.[1] Yet, the underlying dimension and classification of these traits are widely contested within economics and across the social sciences. In economics, "non-cognitive" skills have recently become very popular in applied research, but there is little agreement on what these types of skills represent. In labor economics, non-cognitive skills are usually seen as a broadly defined second dimension of individual heterogeneity (next to cognitive skills). In the education and early childhood intervention literatures, non-cognitive skills are broadly categorized as skills not captured by standardized tests and are commonly measured by observing behavior. In economic psychology and behavioral economics, non-cognitive skills are seen as a superordinate concept summarizing various specific concepts (i.e. economic preferences such as time and risk preferences) as well as personality measures (as the Big Five). Overall, across sub-fields, and across papers within sub-fields, the measurement and interpretation of non-cognitive skills varies widely due to the different motives and available data sources.

The aim of this paper is to compare several different strategies for measuring non-cognitive skills and to decompose and interpret their relative effectiveness in predicting educational success. Using data from the youth survey of the German Socio-Economic Panel (GSOEP), we construct four stylized factors based on measures from previous literature and relate the different estimates of non-cognitive skills to each other and to established taxonomies. To the best of our knowledge, this is the first paper that systematically relates the various non-cognitive measures used in the literature within one data set. We also contribute to the literature on skills in the education context and evaluate how the different non-cognitive skill constructs relatively perform in predicting educational success.[2] We shed light on what previous papers have measured when using different methods to generate proxies of non-cognitive skills

[1] The literature uses the expressions "traits", "skills" and "abilities" to describe unobserved individual heterogeneities. Some papers use "traits" to describe immutable characteristics of individuals while using "skills" when referring to malleable characteristics. For the most part, including this paper, these terms are used interchangeably.

[2] For papers that focuses on the relation between psychological personality measures and economic preference measures, see Becker et al. (2012) and Rustichini et al. (2016).

and show that the construction of the non-cognitive factor greatly influences what is actually measured, how it relates to more standard taxonomies and what conclusions are reached about the role of non-cognitive skills.

Measuring non-cognitive skills: Different sub-fields tend to have different intentions when constructing non-cognitive factors, which may explain the lack of standardization. Researchers sometimes rely on formal models to determine the nature and dimension of the skills to be estimated. At other times, researchers may only be concerned with fully capturing and controlling for pre-existing differences. In the latter case, the researchers may not care about interpretation of the extracted skills, as their only aim may be to span as much of the underlying multidimensional heterogeneity as possible. Below we review three literatures that study non-cognitive skills, highlighting differences in their methodologies for measuring non-cognitive skills and their motivations for including them in their analyses.

In labor economics, a one dimensional skill or ability has been used to differentiate workers (Becker, 1964; Herrnstein and Murray, 1994; Neal and Johnson, 1996; Carneiro, 2002). More recent research has used a two-factor framework which usually consists of a "cognitive" component and a "non-cognitive," "socio-emotional," or "personality" component that is an aggregate of skills or traits other than cognition that are important determinants of educational and labor market outcomes (Heckman and Rubinstein, 2001; Heckman et al., 2006). For historical reasons, the "non-cognitive" component has been constructed from measures included in social surveys such as the National Longitudinal Survey of Youth 1979 (NLSY79). Commonly, these are one dimensional constructs combining self-reported measures of self-esteem, locus of control, or other similar measures. When labor-economists extract multi-dimensional heterogeneity, they commonly aim to simply control for pre-existing differences, and they do not focus on the interpretation of the additional traits.[3]

Non-cognitive skills have also become important components of the education economics and

[3]See, for example, Keane and Wolpin (1994) and the related literature on latent types. This literature uses types to span a potentially multi-dimensional unobservable component of individuals but generally do not focus on what skills, traits, or differences the types are capturing.

early childhood intervention literature. Some early interventions were found to have no lasting effect on test scores, yet still improved later-life performance of participants. Similarly, research on the GED (general education development) high school equivalency exam found that GED certificate recipients performed similarly to high school graduates on achievement tests, yet performed worse later in life. The research on early interventions and the research on the GED raised questions about what skills other than cognition were being formed by education. Looking beyond test scores, researchers in education economics have turned to behaviors measurable in their data, such as behavioral issues, absences, engagement, and teacher reports (Heckman, Humphries and Kautz, 2013a).

More recent research in economic psychology and behavioral economics have taken a somewhat different approach. This literature uses the term "non-cognitive" (or soft) skills as a generic term for a whole set of constructs to differentiate individuals (e.g. Borghans et al., 2008b; Koch et al., 2015).[4] The most frequently used constructs are either key economic variables, such as preferences for risk and time,[5] or, are borrowed from psychology as, e.g., the Big Five personality inventory (Costa and McCrae, 1992).[6] This work has focused on precisely measuring and describing multiple facets of personality and preferences, using incentivized laboratory experiments or extensive questionnaire batteries. However, these precise measures are often not connected to information about later-life outcomes. Exceptions that connect (incentivized) preference measures and real-life outcomes include Burks et al. (2015, 2012), Golsteyn et al. (2014), Sutter et al. (2013), Rustichini et al. (2016) and Chabris et al. (2008). Burks et al. (2015) is most closely related to this paper and considers the relation between education outcomes, personality measures, and economic preferences.

All three strands of literature concur that skills other than cognition are important in explaining heterogeneity in behaviors and outcomes between people.[7] Research across these fields shows

[4]For an overview about the concepts and for a discussion on using the term "non-cognitive" in this context, see Borghans et al. (2008a).

[5]For an overview and discussion, see Becker et al. (2012)

[6]The five personality dimensions are labeled as conscientiousness, agreeableness, neuroticism, openness, and extraversion.

[7]There are a number of overviews of the literature on non-cognitive skills: Koch et al. (2015) on behavioral economics of education, Thiel and Thomsen (2013) on models and measurement, Gutman and Schoon (2013)

the importance of non-cognitive skills, but little consensus exists on what is being measured and what matters. Due to data availability, the labor and education literature have focused on survey questions and reported early-life behaviors to construct a single measure of non-cognitive skills. Heckman et al. (2006) state in this context "we choose these measures because of their availability in the NLSY79. Ideally, it would be better to use a wider array of psychological measurements and ... to connect them with more conventional measures of preference parameters in economics" (p. 429).

While cognitive ability clearly maps into the concept of intelligence, the interpretation of the non-cognitive factor depends on the various measures and behaviors used in its construction. Non-cognitive factors are generally difficult to interpret and do not easily map into pre-existing taxonomies such as economic preferences or personality traits. As different papers use different constructions of non-cognitive skills, it is difficult to reach consilience on the importance of non-cognitive skills on educational and labor market outcomes.

We provide a first step in reconciling the evidence on non-cognitive skills by directly comparing different measures of non-cognitive skills. Specifically, we use new data from the Youth Survey of the German Socio-Economic Panel (SOEP, 2013; Wagner et al., 2007) to compare several prototypical non-cognitive constructs drawn from the different approaches taken in the different fields. The GSOEP Youth Survey is particularly suited for this task as it contains measures of IQ, schooling-related behavior, social behavior, personality measures, and measures of economic preferences. Using these measures, this paper constructs four stylized models of cognitive and non-cognitive skills in order to intrepret how different constructions of non-cognitive skill may result in different conclusions about the relative importance of non-cognitive skills in an important educational outcome. These models are compared to a baseline model that includes Big Five personality traits, economic preferences, and IQ. We use Big Five personality traits, economic preferences, and IQ as our baseline model as these measure are established taxonomies with deep roots in the personality psychology, cognitive psychology, and economic

on effects on various outcomes and Brunello and Schlotter (2011) on effects on educational and labor market outcomes. For an early overview focusing on labor market returns see Groves (2005). Almlund et al. (2011) refer to cognitive and non-cognitive skills and summarize empirical results.

literature.[8] As shown by Becker et al. (2012), the Big Five personality traits and economic preference parameters permit a wide range of individual heterogeneities and complement each other in predicting life outcomes.

We find that the various stylized non-cognitive factors are not strongly related to each other, and ,in fact, some have negative correlations. The differences between the non-cognitive factors can be explained by different correlations with the traits of our baseline model. For example, all four of our non-cognitive factors are positively correlated with conscientiousness and patience, and some are positively correlated with willingness to take risk and extraversion, while others are negatively correlated. In this paper we also consider the role of skills in determining educational success (measured by high school GPA) and show that different non-cognitive factors have different explanatory power on educational success. As a second analysis, we also demonstrate that observed behaviors can be used to extract latent factors with approximately the same amount of predictive power to traditional survey-based taxonomies such as the Big Five or economic preferences, though the interpretation of the factors becomes much more difficult. Different policy conclusions can be reached depending on which measurement system of non-cognitive skills is used.

2 Data and measures

Our analysis uses data from the German Socio-Economic Panel (SOEP, 2013), a large panel data set started in 1984 that is representative of the German population. Today it consists of more than 20,000 individuals in more than 10,000 households. For details, see Wagner et al. (2007). Since 2001, all young people in the sample have received a special "Youth Questionnaire" at age 17. Starting in the year 2006, the youth questionnaire measures preferences and personality and includes an IQ test. The answers and measures collected in the youth questionnaire are our main source of information. We complement this information

[8] Admittedly, there is no proof that the Big Five personality traits, economic preferences, and IQ span the full set of latent traits. Instead, we focus on these measures as they are widely studied and may be easier to directly interpret.

with data about consumption behavior from the main questionnaire of the GSOEP.[9] Our main analysis uses data on more than 1,300 adolescents interviewed at age 17 in the years 2006 to 2012.

Using the GSOEP, we are able to construct useful measures of personality traits, risk and time preference, and IQ. The Big Five are measured using a validated 15-item questionnaire (Gerlitz and Schupp, 2005) that is commonly used in empirical personality research (see, e.g., Becker et al., 2012). Risk preference is measured by the question "How do you see yourself: Are you generally a person who is fully prepared to take risks, or do you try to avoid taking risks?" Answers were given on an 11-point scale, where zero corresponds to "unwilling to take risks" and 10 corresponds to "fully prepared to take risks." This question has been studied in various papers and is highly correlated with incentivized experimental measures and revealed behavior (see, for example, Dohmen et al., 2011). To measure time preference, the participants rated how strongly they agree with the two statements "I abstain from things today to be able to afford more tomorrow" and "I prefer to have fun today and don't think about tomorrow" (reversed) on a 7-point Likert scale. We construct our measure of time preference by summing the standardized responses. The resulting score is well correlated with incentivized experimental measures of time preference.[10] To measure cognitive skills, the participants took part in a validated short version of the well-established "I-S-T 2000 R" (Amthauer et al., 2001), covering all three subsets which are verbal, numerical, and figural abilities (for details see Solga et al., 2005).

Below, we estimate four stylized two-factor models. Each model constructs a non-cognitive factor using different reported behaviors or questionnaire responses. The different non-cognitive factors are chosen to be similar to measurement systems used in applied labor economics and the economics of education literature. For all four models, the cognitive factor is uniformly constructed from three IQ sub-tests.[11] Each factor model assumes there are two factors, and

[9] This information was usually collected one or two years after the youth questionnaire. For details see below and Wagner et al. (2007).

[10] Using data of Vischer et al. (2013) indicates a highly significant correlation ($p<0.001$, $N=965$).

[11] Some papers use achievement tests to measure cognition rather than IQ. As shown in Borghans et al. (2011), achievement tests may capture traits other than IQ. As the GSOEP does not include achievement test scores, we cannot consider variation on this dimension.

each model is estimated by maximum likelihood. As in standard factor analysis, we assume that the covariance between a set of measures can be fully explained by a low-dimensional vector of unobservables. In this case, we assume the unobservables are distributed as a correlated bi-variate normal distribution. In the web-appendix, we relax the strong assumption of normality and estimate factor scores using the minimum distance estimator. As expected from a less efficient estimator, standard errors are somewhat larger when using minimum distance estimation, but the point estimates are very similar to those obtained using maximum likelihood. In order to identify and interpret the factors, we assume that the IQ measures only load on the cognitive factors, while the non-cognitive measures load only on the non-cognitive factor. No restrictions are placed on the correlations between the factors. We normalize the variance of both factors to unity, pinning down the scale of each unobservable.[12] From our estimated factor model, we use maximum likelihood to predict individual factor scores (for an overview see Heckman et al., 2013b), which are further standardized to have a mean of zero and a standard deviation of one.

For each two-factor system, a different set of behaviors or responses is used to identify the non-cognitive factor. For an overview see Table 1. Model 1 (NC-LOCUS) uses responses on a 10-item Locus of Control questionnaire (Rotter, 1966) and a single item self-esteem questionnaire (Robins et al., 2001), as has been done in work using the National Longitudinal Survey of Youth 1979, such as in Heckman et al. (2006). Model 2 (NC-ENGAGEMENT) uses participation in extra-curricular activities, time used on productive leisure time engagement, and number of close friends.[13,14] Model 3 (NC-RELATIONS) uses the quality of relations to parents and friends covering topics such as arguments and fights as well as constructive problem solving, love and bonding,[15,16] and the number of close friends. Model 4 (NC-BEHAVIORS)

[12]Given we have more than five measures for our correlated two-factor model, and at least three dedicated measures in each of our models, our factor models are non-parametrically identified, as laid out in Williams (2013).

[13]Information is based on time-use questionnaires. Individuals were asked about their general engagement in school-based extra-curricular activities and indicated their degree of involvement (5-point Likert scale) in a list of (leisure time) activities.

[14]For example, Borghans et al. (2008c) use similar engagement measures as a measure of non-cognitive skills.

[15]Individuals rated the frequency of respective situations on a 5-point Likert scale.

[16]For example, see Heckman et al. (2013b) for a paper that also uses reported relations as a measure of non-cognitive skills.

uses reported behavior on drinking, smoking and eating habits,[17] and how often the individual argues with their parents or best friend.[18,19] BASE is the label of the baseline model composed of the Big Five and economic preferences.

Table 1 Measurement System of Different Non-cognitive Constructs

Model	**Measurement System**
NC-LOCUS (NC-L)	Rotter's Locus of Control, Self-esteem.
NC-ENGAGEMENT (NC-E)	Frequency of engagegment (volunteering, sport, technical work, reading), number of close friends.
NC-RELATIONS (NC-R)	Relation to parents and friends (bonding, love, argues or fights, problems solving), number of close friends.
NC-BEHAVIORS (NC-B)	Consumption behavior of alcohol and tabacco, eating behavior, argues or fights with family or friends.
Baseline (BASE)	Big-5 (conscientiousness, agreeableness, neuroticism, openness, extraversion), economic preferences (risk and time).

Across the four models, we believe that we have embodied many of the different definitions of non-cognitive skills found in literatures from different disciplines. Tables 2 and 3 give an overview of which measures are used to capture non-cognitive skills in a number of top papers.[20] The last column relates the constructions used in these various papers to our stylized models.

[17]Individuals rated on 4-point Likert scales how much they follow a healthy diet and how often they consume beer, wine, spirits, and long drinks. They also indicated if they smoke or not.

[18]There are only around 750 observations for this measure as it uses variables collected in later questionnaires that have not yet been answered by all those who have answered the age 17 questionnaire.

[19]Papers using self-reported behaviors to construct non-cognitive factors are common, for example Heckman et al. (2017).

[20]Some papers that study cognitive and non-cognitive skills do not explicitly use a two-factor structure (for example, Cobb-Clark and Tan (2011); Farkas (2003); Rustichini et al. (2016); Lleras (2008)). Rather, these papers include a number of measures which they believe serve as proxies for non-cognitive skills and discuss how these measures affect the outcome of interest.

Table 2 Summary of Effects of Non-cognitive Skills in the Education Literature

				Cognitive		Non-Cognitive (validated)								Non-Cognitive (revealed)						
Paper	Data Set	Outcome	Results	IQ	Achievement Test	Big-5	Self-Percep.	Engagement	Persistence	Self-control	Social Skills	Risk Pref.	Discount	GPA	Behav. Probs.	Smoke/Drink	Risk. Behav.	Participation	Other School Meas.	Related NC Factors
Almås et al. (2016)	Private	Education	■	◍	○	●	○	○	○	○	○	●	●	○	○	○	○	○	○	BASE
Andrisani and Nestel (1976)	NLS72	Labor Market	◧	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	◍	L
Antecol and Cobb-Clark (2010)	AddHealth	Labor Market	◧	○	○	○	●	○	●	●	○	○	○	○	○	○	○	○	○	L
Barron et al. (2000)	NLSY, NLS-72	Labor Market, Education	▨	○	●	○	○	○	○	○	○	○	○	○	○	○	○	●	○	E
Becker et al. (2012)	GSOEP	Multiple	◧	○	○	●	●	○	○	○	●	●	●	○	○	○	○	○	○	L,R,BASE
Belzil and Leonardi (2007)	SHIW	Education	◧	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	BASE
Borghans et al. (2008c)	BSS,IAB,BCS	Labor Market	■	○	◍	○	○	●	○	○	●	○	○	○	○	○	○	●	○	B,E
Burks et al. (2015)	Private	Education	■	◍	○	●	○	○	○	○	○	●	●	○	○	○	○	○	○	BASE
Chiteji (2010)	PSID	Health	■	○	○	○	●	○	○	○	○	○	◍	○	○	○	○	○	○	L
Cobb-Clark and Tan (2011)	HILDA	Occupation	◧	○	○	●	●	○	○	○	○	○	○	○	○	○	○	○	○	L
Delaney et al. (2013)	IUS	Education	◧	○	○	●	○	○	○	○	○	●	●	○	○	○	○	◍	○	BASE
Duckworth et al. (2007)	Multiple	Education, GPA	■	●	●	●	○	○	●	●	○	○	○	○	○	○	○	○	○	BASE
Duncan and Morgan (1981)	PSID	Education	◧	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	BASE
Fortin (2008)	NLS72, NELS	Labor Market	■	○	○	○	●	○	○	○	◍	○	○	○	○	○	○	○	○	L,R
Gensowski (2014)	Terman	Labor Market	■	●	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	P
Groves (2005)	NLSYW, NCDS	Labor Market	■	●	●	○	●	◍	○	◍	◍	○	○	○	○	○	○	○	○	L

Notes: Table shows papers that demonstrate the role of cognitive and non-cognitive skills on outcomes. The "Cognitive" columns show measures of cognition, the "non-cognitive (validated)" columns show survey-based measures of various non-cognitive skills. The "non-cognitive (revealed)" column shows reported behaviors or early life outcomes. The "Related NC Factors" column lists the models from our paper that are most similar to the measures being used in the cited paper. ○ – Not used. ◐ – Considered. ◍ – Somewhat used. ● - Used.. ■ – Both Cog and Non-cog effects. ◧ – Non-cog Effects. ▨ – Mixed Effects. □ – No Effects. NLSY79– The National Longitudinal Survey of Youth 1979, NLSY97– The National Longitudinal Survey of Youth 1997, NLSW– National Longitudinal Survey of Mature and Young Women, IUS– Irish University Study, NLS-72– National Longitudinal Study of the High School Class, SHIW– Bank of Italy Survey of Income and Wealth, BSS– British Skills Survey, IAB data– Data of the Institute for Employment Research, BCS– British Cohort Study, GSOEP– German Socio-Economic Panel, NELS– National Education Longitudinal Study of 1988, CLS– Chicago Longitudinal Study, AddHealth– The National Longitudinal Study of Adolescent to Adult Health, Perry– Perry Preschool Project, HILDA– Household Income and Labour Dynamics in Australia, PSID– Panel Survey of Income Dynamics, Terman– Survey by Lewis Terman started at Stanford in 1921/22. of 1972

Table 3 Summary of Effects of Non-cognitive Skills in the Education Literature (Continued)

				Cognitive		Non-Cognitive (validated)								Non-Cognitive (revealed)						
Paper	Data Set	Outcome	Results	IQ	Achievement Test	Big-5	Self-Percep.	Engagement	Persistence	Self-control	Social Skills	Risk Pref.	Discount	GPA	Behav. Probs.	Smoke/Drink	Risk. Behav.	Participation	Other School Meas.	Related NC Factors
Heckman et al. (2017)	NLSY79	Labor Market, Health, Education	■	○	●	○	○	○	○	○	○	○	○	●	○	●	●	○	○	B
Heckman et al. (2013b)	Perry	Labor Market, Social	■	●	○	○	○	○	●	○	○	○	○	○	●	○	○	○	●	B,R
Heckman et al. (2006)	NLSY79	Labor Market, Education	■	○	●	○	●	○	○	○	○	○	○	○	○	◍	◍	○	○	L,B
Heineck and Anger (2010)	GSOEP	Labor Market	▨	●	○	●	●	○	○	○	○	○	○	○	○	○	○	○	○	L,BASE
Jacob (2002)	NELS	Education	◧	○	●	○	●	○	○	○	○	○	○	●	●	○	●	○	●	E,B
Kaestner and Callison (2011)	NLSY79	Health	▨	○	●	○	●	○	○	○	○	○	○	○	○	●	●	●	○	L,B,E
Lleras (2008)	NELS	Labor Market, Education	■	○	●	○	○	●	○	○	◍	○	○	○	○	○	○	●	●	E
Murnane et al. (2001)	NLSY	Labor Market	■	○	●	○	●	○	○	○	○	○	○	○	○	○	○	○	○	L
Piatek and Pinger (2015)	GSOEP	Education, Labor Market	▨	○	◍	○	●	○	○	○	○	○	○	○	○	○	○	○	○	L
Reynolds et al. (2010)	CLS	Education, Crime	■	○	●	○	●	●	◍	◍	●	○	○	○	●	○	●	○	●	E,B
?	Private	Health, Perf., Cred. Score	■	●	○	●	○	○	○	○	○	●	●	○	○	○	○	○	○	BASE
Stevenson (2010)	NLSY79	Labor Market, Education	◧	○	◍	○	○	○	○	○	○	○	○	○	○	○	○	●	○	E
Waddell (2006)	NLS72	Education, Labor Market	◧	○	○	○	●	●	○	○	○	○	○	○	○	○	○	○	○	L

Notes: Table shows papers that demonstrate the role of cognitive and non-cognitive skills on outcomes. The "Cognitive" columns show measures of cognition, the "non-cognitive (validated)" columns show survey-based measures of various non-cognitive skills. The "non-cognitive (revealed)" column shows reported behaviors or early life outcomes. The "Related NC Factors" column lists the models from our paper that are most similar to the measures being used in the cited paper. ○ – Not used. ◐ – Considered. ◍ – Somewhat used. ● - Used.. ■ – Both Cog and Non-cog effects. ◧ – Non-cog Effects. ▨ – Mixed Effects. □ – No Effects. NLSY79– The National Longitudinal Survey of Youth 1979, NLSY97– The National Longitudinal Survey of Youth 1997, NLSW– National Longitudinal Survey of Mature and Young Women, IUS– Irish University Study, NLS-72– National Longitudinal Study of the High School Class, SHIW– Bank of Italy Survey of Income and Wealth, BSS– British Skills Survey, IAB data– Data of the Institute for Employment Research, BCS– British Cohort Study, GSOEP– German Socio-Economic Panel, NELS– National Education Longitudinal Study of 1988, CLS– Chicago Longitudinal Study, AddHealth– The National Longitudinal Study of Adolescent to Adult Health, Perry– Perry Preschool Project, HILDA– Household Income and Labour Dynamics in Australia, PSID– Panel Survey of Income Dynamics, Terman– Survey by Lewis Terman started at Stanford in 1921/22. of 1972

3 Results

In order to structure and guide the following analysis, we present a small descriptive model. Let y_i represent the outcome of interest of individual i and assume the following additive-separable and linear-in-parameter model:

$$y_i = \sum_{j=1}^{J} \beta_j NC_i^j + \beta_C C_i + u_i \tag{1}$$

As described in the previous section, we will always use the same cognitive factor C and combine it with various single or multiple non-cognitive factors NC^j. u is the idiosyncratic error.

When we refer to the stylized two-factor models, there is only one non-cognitive factor ($J = 1$) which is NC-LOCUS (NC-L), NC-ENGAGEMENT (NC-E), NC-RELATIONS (NC-R) or NC-BEHAVIORS (NC-B). When we refer to the baseline model (BASE), th e term "non-cognitive" is used as a superordinate concept summarizing various specific concepts recently used in behavioral economics such as economic preferences as well as personality measures. Therefore, when using the baseline model, there are seven non-cognitive factors ($J = 7$) which are conscientiousness, agreeableness, neuroticism, openness and extraversion, as well as time and risk preference.[21]

The empirical analysis is structured in three steps. First, to analyze what is measured by the different stylized non-cognitive factors, we relate them to each other and to the baseline model. We show that the extracted non-cognitive factors have only moderate (sometimes negative) correlations across models and have plausible but varying correlations with measures of the Big Five personality traits and economic preferences. We decompose each non-cognitive factor by regressing each non-cognitive factor on the covariates from the baseline model (IQ, Big Five

[21] See Almlund et al. (2011) for a more detailed model of agent behavior. In that paper, they highlight that the distinction between test scores and behaviors is somewhat arbitrary in that both are observed responses to a particular action. In their model they highlight that the response will depend on underlying latent skills but may also depend on incentives for the task and the environment in which the tasks in administered.

personality traits, time preference, and risk preference). We find that different personality traits and economic preferences play important roles in the different constructions of the non-cognitive factors.

Second, we use the descriptive model presented in Equation (1) to evaluate how much the predictive power of non-cognitive skills changes when using different constructions of non-cognitive skills. Specifically, we compare regressions of GPA on each of our two-factor models and on our baseline model. We find that the different non-cognitive factors vary widely in their ability to explain educational performance; therefore, different constructions of the non-cognitive factor lead to different conclusions about their relative importance. The baseline model outperforms every two-factor model, and the stylized non-cognitive factors from the two-factor models add little or no additional explanatory power to the baseline model.

Third, we evaluate the overall ability of behavioral measures to predict GPA in comparison to the baseline model consisting of questionnaire-based taxonomies. We use all behavioral measures (input to NC-E, NC-R, NC-B) to perform an exploratory factor analysis and to predict new factors. We use these new factors to predict educational success and show, while difficult to interpret, they have similar explanatory power to the baseline model.

3.1 Relations of the stylized factors and the baseline model

As presented in Table 4, the correlations between the four non-cognitive factors are quite low. The correlation between the different non-cognitive factors ranges between -0.12 and 0.21. Interestingly, the correlation between NC-LOCUS and NC-BEHAVIORS is negative while the correlations between NC-RELATIONS and NC-LOCUS and between NC-RELATIONS and NC-BEHAVIORS are positive. This suggests that each factor may be capturing different aspects of a vector of unobservable non-cognitive characteristics. Looking at the lower half of Table 4 indicates that all stylized non-cognitive factors are significantly positively correlated with conscientiousness and patience. Personality psychology would interpret this as "construct validation" for our non-cognitive factors. Furthermore, different non-cognitive factors are correlated with different aspects of the Big Five, economic preference parameters, and IQ.

Moreover, NC-LOCUS is also positively correlated with IQ and has a relatively large negative correlation with neuroticism. NC-ENGAGEMENT has lower correlations than the previously discussed factors but is correlated most with risk, openness, and extraversion. NC-RELATIONS is relatively highly correlated with many of the other factors. It is positively correlated with agreeableness and openness. NC-BEHAVIORS is positively correlated with both agreeableness and neuroticism but negatively correlated with willingness to take risk, extraversion, and IQ.

Table 4 Correlations (Pearson) Between Different Noncog. and Cog. Constructs

	NC-L	NC-E	NC-R	NC-B
NC-L	1			
NC-E	0.113***	1		
NC-R	0.214***	0.0968***	1	
NC-B	-0.116***	-0.0367	0.0844**	1
Cons.	0.204***	0.0959***	0.186***	0.134***
Agree.	0.134***	0.0217	0.218***	0.0668*
Neuro.	-0.302***	-0.0125	-0.0741***	0.110***
Open.	0.128***	0.187***	0.182***	-0.0546
Extra.	0.173***	0.125***	0.151***	-0.144***
Time	0.0954***	0.0741***	0.123***	0.0974***
Risk	0.0911***	0.0952***	-0.0320	-0.186***
IQ	0.227***	0.116***	0.137***	-0.131***

Notes: Number of observations between 760 and 1416, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7 regresses the cognitive factor and each of the (standardized) non-cognitive factors on the Big Five personality traits and economic preferences.[22] This table provides similar evidence to Table 2, but now considers partial correlations between parameters in the baseline model and the factors. For each factor, we regress the factor on Big Five personality traits, IQ, time preference, and risk preference. Overall, we find that the partial correlations confirm the relationships found in Table 2. As found in the psychology literature, IQ is most strongly correlated with openness. Extraversion and neuroticism are negatively correlated with IQ,

[22]Each model has a uniquely estimated cognitive factor as the factors in each model are estimated jointly. Yet, the cognitive factor is estimated using the same measures across models and have correlations of nearly unity.

while time preference is positively correlated.

All four of the non-cognitive constructs have positive partial correlations with conscientiousness. These relationships are statistically significant, except in the specifications regarding NC-ENGAGEMENT and NC-RELATIONS, where conscientiousness is not statistically significant once we control for time preference.[23] Evaluating the non-cognitive factors according to the R-squared by personality and preference measures reveals notable differences. While Big Five personality traits and economic preference measures can explain more than 13% of the variation in NC-LOCUS, they explain only about 5% of the variation in NC-Engagement.[24]

The analysis in Table 8 and 9 focus on time and risk preferences. We regress time and risk preferences on the stylized 2-factor models, the Big Five personality measures and IQ, as well as combinations of the 2-factor models and the Big Five. The results of this multivariate regression analysis largely confirm the results of the raw correlation analysis presented in Table 4. Time preference (patience) is partially positively correlated with all non-cognitive factors, as well as with cognitive skills and conscientiousness. For risk preference (willingness to take risk) we find positive significant partial correlations with NC-LOCUS and NC-ENGAGEMENT, no significant relation to NC-RELATIONS and negative significant partial correlations with NC-BEHAVIORS.

3.2 Educational outcomes: stylized factors and the baseline model

Previous literature provides evidence for the importance of school performance and educational decisions for later life outcomes (see, e.g., Heckman et al., 2017). While there is a broad consensus that non-cognitive traits play a critical role in the determination of these outcomes and decisions (see discussion in Almlund et al., 2011), little is known about the actual drivers behind the abstract concept of non-cognitive traits.

In this section, we compare the four stylized two-factor models and the baseline model regarding

[23] Time preference and conscientiousness are closely related concepts and have a correlation of 0.31 in our sample.

[24] Tables report R-squared. Given the low number of parameters, the R-squared and adjusted R-squared agree to three or more decimal places.

predictive power. Tables 5 considers how the different two-factor models predict high school GPA. We regress GPA on each of the four stylized two-factor models as well as on our baseline model of personality traits and economic preferences. In the last four columns we combine each two-factor model without the baseline model to check if the extracted non-cognitive factor has predictive power beyond the baseline model. The dependent variable in Table 5, GPA, can range from zero to four and is coded such that higher values indicate better performance.[25] GPA is measured at age 17 and calculated as the average of grades in mathematics, German, and first foreign language. The distribution is displayed Figure 1. In the German grading system, the best and the worst grade are relatively rarely used and so we observe little to no censoring at the bottom or the top of the GPA distribution. The mean (standard deviation) in our sample is 2.08 (0.71).

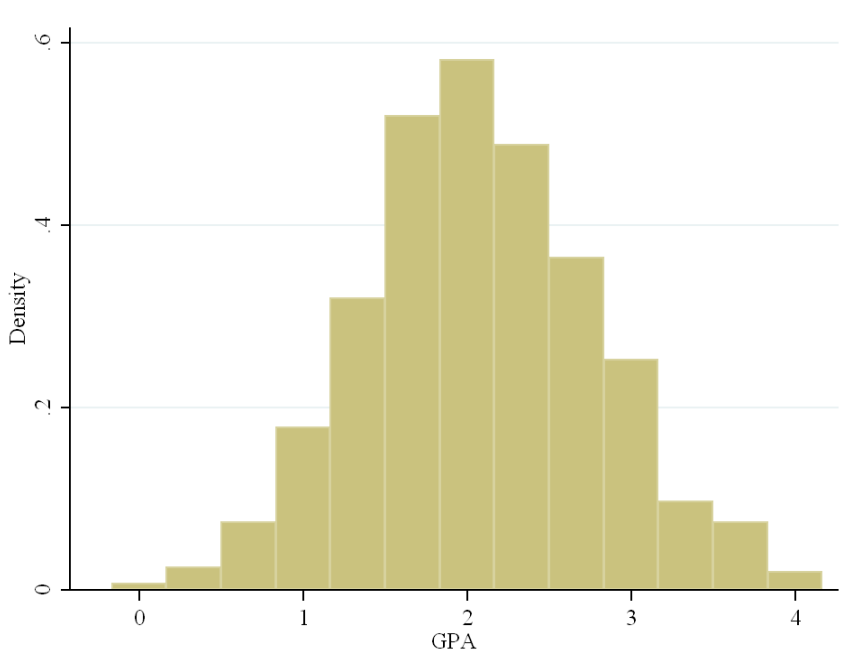

Figure 1 Distribution of GPA. Higher values indicate better performance, $N = 1333$.

OLS regressions indicate that cognitive ability predicts GPA, but so do all of the non-cognitive factors (though the loadings vary in size and significance). The loadings on the non-cognitive

[25]In the German grading system, theoretically, one is the best grade and six is the worst grade. Both, 5 and 6, indicate failing. The grade 6 is hardly ever used. The worst average grade (in the German system) that we observe in the data is 5. The grades are recoded in a way that 1 translates to 4, 2 to 3, 3 to 2, 4 to 1 and 5 to 0.

factor vary, but even the largest is less than one third the size of the loading on cognition.[26] Comparing the predictive power of the two-factor models to a one-factor model which only consists of the cognitive factor (Cog-only) indicates that the non-cognitive factors of the stylized two-factor models add little explanatory power in terms of R squared. In the baseline model, cognition and conscientiousness are positively correlated with GPA, while risk preference and agreeableness are negatively correlated. The coefficient on conscientiousness is approximately three-fourths of the size of the coefficient on cognition. The baseline model explains six percent more of the variance than the two-factor models.[27] No substantial additional explanatory power over the baseline model alone is observed when we include the different non-cognitive factors in the baseline model. In general, we find similar results regarding size and significance if we control for gender, urban status, residence in Eastern Germany and the student's educational tier[28] or if we use the minimum distance estimator to estimate the non-cognitive factors (see Table 10 and 11).

In sum, this horse race analysis shows that conclusions regarding the importance of the non-cognitive factor strongly depend on the measurement system. If we just consider the non-cognitive factors from the stylized two-factor models, we would, e.g., conclude that the cognitive factor is much more important than the non-cognitive factor in predicting GPA. In contrast, if we consider the baseline model, including validated measures on conscientiousness, we would conclude that cognitive and non-cognitive skills are of similar importance in predicting GPA.

[26] All skills are standardized to have a mean of zero and a standard deviation of one.

[27] Note that personality and economic preference have substantial predictive power and explain 10.6% of the variance in GPA without the inclusion of IQ.

[28] The secondary education system in Germany consists essentially of three tracks (low, medium, high) which are supplemented by comprehensive schools and vocational school.

Table 5 Model Comparison: GPA

	Cog-only	NC-L	NC-E	NC-R	NC-B	BASE	Comb-L	Comb-E	Comb-R	Comb-B
Cog	0.253***	0.242***	0.248***	0.244***	0.258***	0.233***	0.227***	0.229***	0.226***	0.234***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.026)	(0.019)	(0.020)	(0.019)	(0.019)	(0.026)
Noncog		0.040**	0.031*	0.049***	0.085***		0.011	0.018	0.020	0.051**
		(0.019)	(0.018)	(0.019)	(0.025)		(0.019)	(0.018)	(0.019)	(0.025)
Cons.						0.178***	0.175***	0.176***	0.176***	0.187***
						(0.022)	(0.022)	(0.022)	(0.022)	(0.029)
Agree.						-0.050**	-0.058***	-0.054**	-0.060***	-0.078***
						(0.021)	(0.021)	(0.021)	(0.021)	(0.028)
Neuro.						-0.021	-0.025	-0.032	-0.031	-0.028
						(0.021)	(0.021)	(0.020)	(0.020)	(0.028)
Open.						0.026	0.050	0.041	0.048	0.037
						(0.034)	(0.034)	(0.034)	(0.034)	(0.046)
Extra.						-0.014	-0.033	-0.029	-0.033	0.008
						(0.033)	(0.033)	(0.033)	(0.033)	(0.045)
Risk						-0.067***	-0.068***	-0.070***	-0.065***	-0.070***
						(0.019)	(0.019)	(0.019)	(0.019)	(0.027)
Time						0.024	0.022	0.021	0.022	0.033
						(0.019)	(0.019)	(0.019)	(0.019)	(0.026)
Observations	1333	1333	1333	1333	736	1333	1333	1333	1333	736
R^2	0.123	0.126	0.124	0.126	0.125	0.195	0.196	0.197	0.196	0.209

Notes: The table shows OLS estimates of GPA on one of the four constructed 2-factor models, our baseline model, or combined models; Standard errors in parentheses. NC-L is based on the Rotter's Locus of control. NC-E is based on engagement behavior, NC-R is based on self-reported relationships, and NC-B is based on self reported risky behaviors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.3 Educational outcomes: ad hoc factors extracted from behaviors and the baseline model

In the previous section the main intention was to use interpretable stylized factors to predict school performance. In the last step of the analysis, we give up the structure concerning the non-cognitive factors. We focus on revealed behavior and use the engagement, relations and behavior related input data of NC-E, NC-R and NC-B to perform an exploratory factor analysis and to predict new ad hoc non-cognitive factors.[29] These ad hoc factors are harder to interpret but capture the heterogeneities concerning the input measures optimally. We use these new factors to predict educational success and show that that they have similar explanatory power as the baseline model.

Performing an explanatory factor analysis on the standardized behavior input measures of NC-E, NC-R and NC-B yields six factors according to Horn (1965). The labeling concerning the factors resulting from an exploratory factor analysis is always arbitrary; therefore, we just label them as non-cognitive factor 1 to 6. The structure of the factor loadings suggests that ad hoc factor Noncog1 (more or less) corresponds to NC-B, Noncog2 corresponds to NC-R and Noncog3 to NC-E.

Due to the use of exploratory factor analysis, the sample is restricted to observations for which all variables are available. In Table 6 we compare the predictive power of the six ad hoc NC factors resulting from the exploratory factor analysis and the baseline model. In column 1 GPA is regressed on the six ad hoc NC factors. For comparison, in column 2, GPA is regressed on the Big Five personality traits and economic preferences (the baseline model). While the R-squared is slightly higher for the baseline model that uses the established taxonomies, the Bayesian information criteria (BIC) that accounts for the number of used coefficients indicate a slightly better fit for the ad hoc model using information on behavior in column 1.

Column 3 and 4 repeat the estimations of column 1 and 2 but use cognitive skills as additional

[29] NC-L consists of items of validated Locus of Control and Self-esteem questionnaires (see section 2), and therefore, we are intentionally not using the items to extract ad hoc non-cognitive factor from the revealed behavior data.

Table 6 Predictive Power of ad hoc Non-cognitive Factors and the Baseline Model.

	Dependent Variable: GPA			
Noncog 1	0.054**		0.075***	
	(0.027)		(0.026)	
Noncog 2	0.060**		0.037	
	(0.027)		(0.026)	
Noncog 3	0.113***		0.093***	
	(0.027)		(0.026)	
Noncog 4	0.170***		0.103***	
	(0.027)		(0.027)	
Noncog 5	0.045*		0.019	
	(0.027)		(0.026)	
Noncog 6	-0.092***		-0.101***	
	(0.027)		(0.026)	
Cons		0.160***		0.186***
		(0.032)		(0.030)
Agree		-0.116***		-0.086***
		(0.030)		(0.029)
Neuro		-0.089***		-0.023
		(0.029)		(0.029)
Open		0.143***		0.015
		(0.048)		(0.049)
Extrav		-0.073		0.023
		(0.047)		(0.047)
Risk		-0.107***		-0.079***
		(0.029)		(0.028)
Time		0.064**		0.045*
		(0.028)		(0.027)
IQ			0.201***	0.218***
			(0.029)	(0.028)
N	657	657	657	657
R^2	0.113	0.120	0.175	0.195
BIC	1387.911	1388.719	1346.233	1337.006

Notes: The table shows OLS estimates of GPA on ad hoc non-cognitive factors and our baseline model. Non-cognitive factors are constructed using exploratory factor analysis. No imputation is done in the estimation and therefore factors are only predicted for individuals with no missing measures, explaining differences in observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

predictor. As expected, the R-squared in both models increases remarkably. The increase is weaker in the ad hoc model, which indicates that the behavioral measures also includes information on cognitive skills, leading the overall R-squared to be somewhat smaller than that of our preferred model.

4 Conclusion

This aim of this paper is to compare several different strategies for measuring non-cognitive skills and to decompose and interpret their relative effectiveness in predicting educational success. Using data from the youth survey of the GSOEP,[30] we construct four stylized factors based on measures from the previous literature and relate the different estimates of non-cognitive skills to each other and to established taxonomies. The correlation between the different non-cognitive factors varies greatly, ranging between -0.12 and 0.21. Given the heterogeneity among the different non-cognitive factors, we compare them with a baseline model that contains personality traits and economic preference parameters. While all stylized factors are positively correlated with conscientiousness and patience, their relationships to other dimensions vary substantially.

In the context of educational performance, we find that the baseline model outperforms all of the 2-factor models in terms of predictive power. Moreover, there is heterogeneity in the relevance of the different non-cognitive factors. In general, the higher the correlation with traits from our baseline model, such as conscientiousness and time preference, the higher their predictive power concerning high school GPA. When we perform an exploratory factor analysis where we remove the structure concerning the non-cognitive factors and capture the information of the behavioral input measures optimally, we find similar explanatory power for baseline model and the non-cognitive factors. This suggests that, if a researcher simply cares about proxying for unobserved heterogeneity, rich sets of behaviors may perform as well as

[30] While this study focuses on the late adolescence, in principle, we expect the general pattern to also hold for adults. Cognitive skills have been shown to have relatively stable rank orderings by the middle of childhood. For many non-cognitive skills, rank-order stability peaks after age 50. For a detailed discussion see Borghans et al. (2008a).

questionnaires and taxonomies developed specifically for the task.

This paper demonstrates that building a single non-cognitive factor from behavioral or survey measures to complement measures of cognitive skills provides useful insights. Because the relationship of such factors to more standard measures of individual differences varies, the value of those insights is increased by the road map provided in the present paper, in which personality traits and economic preferences are connected to the sorts of two-factor models used in different parts of the existing literature. This decomposition can serve as a first step towards integrating findings from studies that use different measurement systems for non-cognitive factors. In summary, this paper suggests that many two-factor models may be under-predicting the importance of non-cognitive skills compared to models which use richer sets of (validated) measures. Moreover, papers that reach different conclusions about the importance of non-cognitive skills may do so because of differences in how their non-cognitive factors were constructed.

By understanding what different non-cognitive constructs measure, we can better understand which skills matter for success in life. For example, knowledge about specific traits that drive educational success could lead to more targeted interventions as e.g. Schunk et al. (2016) or Alan et al. (2015). Ultimately, understanding what various constructs measure allows us to create consilience across the various studies, creating the potential for improved policy making.

References

Alan, S., Boneva, T., Ertac, S., 2015. Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. HCEO Discussion Paper 2015-009.

Almås, I., Cappelen, A.W., Salvanes, K.G., Tungodden, B., et al., 2016. What explains the gender gap in college track dropout? experimental and administrative evidence. The American Economic Review 106, 296–302.

Almlund, M., Duckworth, A., Heckman, J.J., Kautz, T., 2011. Personality psychology and economics, in: Hanushek, E.A., Machin, S., Wößmann, L. (Eds.), Handbook of the Economics of Education. Elsevier, Amsterdam. volume 4, pp. 1–181.

Amthauer, R., Brocke, B., Liepmann, D., Beauducel, A., 2001. Intelligenz-Struktur-Test 2000R: IST 2000. Hogrefe & Huber.

Andrisani, P.J., Nestel, G., 1976. Internal-external control as contributor to and outcome of work experience. Journal of Applied Psychology 61(2), 156–165.

Antecol, H., Cobb-Clark, D.A., 2010. Do non-cognitive skills help explain the occupational segregation of young people? Claremont McKenna College Robert Day School of Economics and Finance Research Paper .

Barron, J.M., Ewing, B.T., Waddell, G.R., 2000. The effects of high school athletic participation on education and labor market outcomes. Review of Economics and Statistics 82, 409–421.

Becker, A., Deckers, T., Dohmen, T., Falk, A., Kosse, F., 2012. The relationship between economic preferences and psychological personality measures. Annual Review of Economics 4, 453–478.

Becker, G.S., 1964. Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education. National Bureau of Economic Research, New York.

Belzil, C., Leonardi, M., 2007. Can risk aversion explain schooling attainments? Evidence from Italy. Labour Economics 14, 957–970.

Borghans, L., Duckworth, A.L., Heckman, J.J., Ter Weel, B., 2008a. The economics and psychology of personality traits. Journal of Human Resources 43, 972–1059.

Borghans, L., Golsteyn, B.H., Heckman, J., Humphries, J.E., 2011. Identification problems in personality psychology. Personality and Individual Differences 51, 315–320.

Borghans, L., Meijers, H., Ter Weel, B., 2008b. The role of noncognitive skills in explaining cognitive test scores. Economic Inquiry 46, 2–12.

Borghans, L., Ter Weel, B., Weinberg, B.A., 2008c. Interpersonal styles and labor market outcomes. Journal of Human Resources 43, 815–858.

Brunello, G., Schlotter, M., 2011. Non cognitive skills and personality traits: Labour market relevance and their development in education & training systems. IZA Discussion Paper Series 5743.

Burks, S., Carpenter, J., Götte, L., Rustichini, A., 2012. Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment. Journal of Economic Behavior & Organization 84, 308–320.

Burks, S., Lewis, C., Kivi, P., Wiener, A., Anderson, J., Goette, L., DeYoung, C.G., Rustichini, A., 2015. Cognitive skills, personality, and economic preferences in collegiate success. Journal of Economic Behavior & Organization 115, 30–44.

Carneiro, P., 2002. Heterogeneity in the Returns to Schooling: Implications for Policy Evaluation. Ph.D. thesis. University of Chicago.

Chabris, C.F., Laibson, D., Morris, C.L., Schuldt, J.P., Taubinsky, D., 2008. Individual laboratory-measured discount rates predict field behavior. Journal of Risk and Uncertainty 37, 237–269.

Chiteji, N., 2010. Time-preference, non-cognitive skills and well-being across the life course: Do non-cognitive skills encourage healthy behavior? American Economic Review 100, 200.

Cobb-Clark, D.A., Tan, M., 2011. Noncognitive skills, occupational attainment, and relative wages. Labour Economics 18, 1–13.

Costa, P.T., McCrae, R.R., 1992. Revised neo personality inventory (neo pi-r) and neo five-factor inventory (neo-ffi). Psychological Assessment Resources Odessa, FL.

Delaney, L., Harmon, C., Ryan, M., 2013. The role of noncognitive traits in undergraduate study behaviours. Economics of Education Review 32, 181–195.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association 9, 522–550.

Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: Perseverance and passion for long-term goals. Journal of Personality and Social Psychology 92, 1087–1101.

Duncan, G.J., Morgan, J.N., 1981. Sense of efficacy and subsequent change in earnings-a replication. Journal of Human Resources , 649–657.

Farkas, G., 2003. Cognitive skills and noncognitive traits and behaviors in stratification processes. Annual Review of Sociology 29, 541–562.

Fortin, N.M., 2008. The gender wage gap among young adults in the united states the importance of money versus people. Journal of Human Resources 43, 884–918.

Gensowski, M., 2014. Personality, IQ, and lifetime earnings. IZA Discussion Paper Series 8235.

Gerlitz, J.Y., Schupp, J., 2005. Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. DIW Research Notes 4, 2005.

Golsteyn, B.H., Grönqvist, H., Lindahl, L., 2014. Adolescent time preferences predict lifetime outcomes. The Economic Journal 124, F739–F761.

Groves, M.O., 2005. How important is your personality? Labor market returns to personality for women in the US and UK. Journal of Economic Psychology 26, 827–841.

Gutman, L.M., Schoon, I., 2013. The impact of non-cognitive skills on outcomes for young people. Technical Report. Institute of Education, University of London.

Heckman, J., Humphries, J.E., Veramendi, G., 2017. The returns to education: The causal effects of education on earnings, health, and smoking. Journal of Political Economy, forthcoming .

Heckman, J.J., Humphries, J.E., Kautz, T., 2013a. The Myth of Achievement Tests: The GED and the Role of Character in American Life. University of Chicago Press, Chicago, IL.

Heckman, J.J., Pinto, R., Savelyev, P.A., 2013b. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. American Economic Review 103(6).

Heckman, J.J., Rubinstein, Y., 2001. The importance of noncognitive skills: Lessons from the ged testing program. The American Economic Review 91, 145–149.

Heckman, J.J., Stixrud, J., Urzúa, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics 24, 411–482.

Heineck, G., Anger, S., 2010. The returns to cognitive abilities and personality traits in germany. Labour Economics 17, 535–546.

Herrnstein, R.J., Murray, C.A., 1994. The Bell Curve: Intelligence and Class Structure in American Life. Free Press, New York.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30, 179–185.

Jacob, B.A., 2002. Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. Economics of Education Review 21, 589–598.

Kaestner, R., Callison, K., 2011. Adolescent cognitive and noncognitive correlates of adult health. Journal of Human Capital 5, 29–69.

Keane, M.P., Wolpin, K.I., 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. The Review of Economics and Statistics , 648–672.

Koch, A., Nafziger, J., Nielsen, H.S., 2015. Behavioral economics of education. Journal of Economic Behavior & Organization 115, 3–17.

Lleras, C., 2008. Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. Social Science Research 37, 888–902.

Murnane, R.J., Willett, J.B., Braatz, M.J., Duhaldeborde, Y., 2001. Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. Economics of Education Review 20, 311–320.

Neal, D.A., Johnson, W.R., 1996. The role of premarket factors in black-white wage differences. Journal of Political Economy 104, 869–895.

Piatek, R., Pinger, P., 2015. Maintaining (locus of) control? Data combination for the identification and inference of factor structure models. Journal of Applied Econometrics , forthcoming.

Reynolds, A.J., Temple, J.A., Ou, S.R., 2010. Preschool education, educational attainment, and crime prevention: Contributions of cognitive and non-cognitive skills. Children and Youth Services Review 32, 1054–1063.

Robins, R.W., Hendin, H.M., Trzesniewski, K.H., 2001. Measuring global self-esteem: Construct validation of a single-item measure and the rosenberg self-esteem scale. Personality and Social Psychology Bulletin 27, 151–161.

Rotter, J.B., 1966. Generalized expectancies for internal versus external control of reinforcement. Psychological monographs: General and applied 80, 1.

Rustichini, A., DeYoung, C.G., Anderson, J.E., Burks, S.V., 2016. Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. Journal of Behavioral and Experimental Economics 64, 122–137.

Schunk, D., Berger, E., Fehr, E., Mueller, H., Winkel, K., 2016. Does working memory training at primary school improve self-control skills and school performance? evidence from a randomized field experiment. Mimeo .

SOEP, 2013. Socio-economic panel, data for years 1984-2012, version 29, doi:10.5684/soep.v29.

Solga, H., Stern, E., von Rosenbladt, B., Schupp, J., Wagner, G.G., 2005. The measurement and importance of general reasoning potentials in schools and labor markets. DIW Research Notes 10.

Stevenson, B., 2010. Beyond the classroom: Using title ix to measure the return to high school sports. Review of Economics and Statistics 92, 284–301.

Sutter, M., Kocher, M.G., Glätzle-Rützler, D., Trautmann, S.T., 2013. Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. American Economic Review 103, 510–531.

Thiel, H., Thomsen, S.L., 2013. Noncognitive skills in economics: Models, measurement, and empirical evidence. Research in Economics 67, 189–214.

Vischer, T., Dohmen, T., Falk, A., Huffman, D., Schupp, J., Sunde, U., Wagner, G.G., 2013. Validating an ultra-short survey measure of patience. Economics Letters 120, 142–145.

Waddell, G.R., 2006. Labor-market consequences of poor attitude and low self-esteem in youth. Economic Inquiry 44, 69–97.

Wagner, G., Frick, J., Schupp, J., 2007. The German Socio-Economic Panel study (SOEP)-evolution, scope and enhancements. Journal of Applied Social Science (Schmollers Jahrbuch) 127, 139–169.

Williams, B., 2013. Identification of factor models. Unpublished Working Paper .

Online Appendix

The following material is for online publication only.

A Appendix Tables

Table 7 Relations of Traits to Cog. and Noncog. Factors

	Cog		NC-L		NC-E		NC-R		NC-B	
Cons	-0.037	-0.064**	0.105***	0.088***	0.062**	0.047	0.067**	0.037	0.167***	0.150***
	(0.030)	(0.031)	(0.030)	(0.031)	(0.031)	(0.033)	(0.031)	(0.032)	(0.041)	(0.043)
Agree	-0.098***	-0.103***	0.063**	0.068**	-0.091***	-0.085***	0.155***	0.147***	0.012	-0.009
	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)	(0.032)	(0.031)	(0.031)	(0.041)	(0.041)
Neuro	-0.251***	-0.256***	-0.279***	-0.276***	-0.020	-0.016	-0.057*	-0.064**	0.085**	0.070*
	(0.028)	(0.028)	(0.029)	(0.029)	(0.030)	(0.030)	(0.029)	(0.029)	(0.039)	(0.039)
Open	0.549***	0.549***	0.045	0.036	0.255***	0.245***	0.084*	0.087*	0.013	0.038
	(0.046)	(0.046)	(0.046)	(0.046)	(0.048)	(0.048)	(0.047)	(0.047)	(0.064)	(0.064)
Extrav	-0.441***	-0.422***	0.019	0.013	-0.071	-0.080*	0.020	0.048	-0.167***	-0.144**
	(0.045)	(0.045)	(0.045)	(0.045)	(0.047)	(0.047)	(0.046)	(0.047)	(0.063)	(0.063)
Time		0.078***		0.069**		0.066**		0.077***		0.014
		(0.027)		(0.027)		(0.028)		(0.028)		(0.037)
Risk		-0.031		0.056**		0.068**		-0.065**		-0.130***
		(0.027)		(0.028)		(0.029)		(0.028)		(0.038)
N	1382	1382	1382	1382	1382	1382	1382	1382	758	758
R^2	0.113	0.120	0.123	0.129	0.043	0.050	0.070	0.079	0.055	0.070

Notes: The table shows OLS regressions of cognition and different non-cognitive constructs on the Big-5 Personality traits, discount rate, and risk preference. NC-L is based on the Rotter's Locus of control. NC-E is based on engagement behavior, NC-R is based on self-reported relationships, and NC-B is based on self reported risky behaviors. Standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8 Relations of Time preference to 2-factor models and the Big 5 model. Dep. var.: Time preference (patience)

	NC-L	NC-E	NC-R	NC-B	Big5 & Cog	Comb-L	Comb-E	Comb-R	Comb-B
Cog	0.061**	0.073***	0.063**	0.081**	0.071**	0.063**	0.069**	0.063**	0.082**
	(0.029)	(0.028)	(0.028)	(0.038)	(0.028)	(0.028)	(0.028)	(0.028)	(0.038)
Noncog	0.072**	0.065**	0.107***	0.105***		0.037	0.047*	0.060**	0.041
	(0.028)	(0.027)	(0.028)	(0.038)		(0.028)	(0.026)	(0.027)	(0.037)
Cons.					0.325***	0.321***	0.322***	0.321***	0.348***
					(0.030)	(0.031)	(0.030)	(0.030)	(0.041)
Agree.					0.037	0.032	0.040	0.025	0.026
					(0.030)	(0.030)	(0.030)	(0.031)	(0.040)
Neuro.					0.051*	0.058*	0.048	0.050*	0.075*
					(0.030)	(0.030)	(0.029)	(0.029)	(0.040)
Open.					0.001	0.010	-0.005	0.006	-0.045
					(0.050)	(0.049)	(0.049)	(0.049)	(0.067)
Extra.					-0.100**	-0.109**	-0.102**	-0.109**	-0.028
					(0.047)	(0.047)	(0.047)	(0.047)	(0.066)
Observations	1333	1333	1333	736	1333	1333	1333	1333	736
R^2	0.012	0.011	0.018	0.014	0.121	0.122	0.123	0.124	0.130

Notes: The table shows OLS regressions of time preference (patience) on the four stylized 2-factor models, Big-5 Personality traits and cognitive skills, and combined models. NC-L is based on the Rotter's Locus of control. NC-E is based on engagement behavior, NC-R is based on self-reported relationships, and NC-B is based on self reported risky behaviors. Standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9 Relations of Risk preference to 2-factor models and the Big 5 model. Dep. var.: Risk preference (willingness to take risk)

	NC-L	NC-E	NC-R	NC-B	Big 5 & Cog	Comb-L	Comb-E	Comb-R	Comb-B
Cog	-0.048*	-0.034	-0.023	-0.079**	-0.058**	-0.067**	-0.058**	-0.050*	-0.122***
	(0.029)	(0.028)	(0.028)	(0.037)	(0.028)	(0.028)	(0.027)	(0.027)	(0.036)
Noncog	0.092***	0.106***	-0.033	-0.186***		0.047*	0.065**	-0.064**	-0.127***
	(0.028)	(0.027)	(0.028)	(0.036)		(0.028)	(0.026)	(0.027)	(0.035)
Cons.					-0.091***	-0.096***	-0.095***	-0.086***	-0.098**
					(0.030)	(0.030)	(0.030)	(0.030)	(0.039)
Agree.					-0.102***	-0.104***	-0.094***	-0.090***	-0.158***
					(0.030)	(0.030)	(0.030)	(0.030)	(0.038)
Neuro.					-0.096***	-0.083***	-0.092***	-0.095***	-0.133***
					(0.029)	(0.030)	(0.029)	(0.029)	(0.038)
Open.					0.144***	0.141***	0.122**	0.140***	0.255***
					(0.049)	(0.048)	(0.048)	(0.048)	(0.064)
Extra.					0.236***	0.236***	0.245***	0.245***	0.098
					(0.047)	(0.046)	(0.046)	(0.046)	(0.063)
Observations	1333	1333	1333	736	1333	1333	1333	1333	736
R^2	0.008	0.012	0.002	0.037	0.137	0.138	0.140	0.141	0.162

Notes: The table shows OLS regressions of risk preference (willingness to take risk) on the four stylized 2-factor models, Big-5 Personality traits and cognitive skills, and combined models. NC-L is based on the Rotter's Locus of control. NC-E is based on engagement behavior, NC-R is based on self-reported relationships, and NC-B is based on self reported risky behaviors. Standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10 Model Comparison: GPA, including controls

	NC-L	NC-E	NC-R	NC-B	BASE	Comb-L	Comb-E	Comb-R	Comb-B
Cog	0.198***	0.204***	0.201***	0.217***	0.185***	0.182***	0.184***	0.183***	0.194***
	(0.023)	(0.023)	(0.023)	(0.031)	(0.023)	(0.023)	(0.023)	(0.023)	(0.031)
Noncog	0.053***	0.027	0.056***	0.082***		0.021	0.016	0.026	0.053**
	(0.019)	(0.019)	(0.019)	(0.026)		(0.020)	(0.019)	(0.019)	(0.026)
Cons.					0.180***	0.177***	0.179***	0.178***	0.179***
					(0.023)	(0.023)	(0.023)	(0.023)	(0.030)
Agree.					-0.059***	-0.061***	-0.058***	-0.064***	-0.074***
					(0.021)	(0.022)	(0.022)	(0.022)	(0.028)
Neuro.					-0.028	-0.023	-0.028	-0.027	-0.027
					(0.022)	(0.022)	(0.022)	(0.022)	(0.028)
Open.					0.022	0.025	0.019	0.021	0.008
					(0.036)	(0.036)	(0.036)	(0.036)	(0.047)
Extra.					-0.030	-0.033	-0.029	-0.032	0.008
					(0.034)	(0.034)	(0.034)	(0.034)	(0.046)
Risk					-0.054***	-0.055***	-0.055***	-0.052***	-0.051*
					(0.020)	(0.020)	(0.020)	(0.020)	(0.027)
Time					0.031	0.031	0.030	0.030	0.044*
					(0.019)	(0.019)	(0.020)	(0.020)	(0.026)
Observations	1217	1217	1217	715	1217	1217	1217	1217	715
R^2	0.175	0.171	0.175	0.172	0.238	0.239	0.238	0.239	0.242

Notes: The table shows OLS estimates of GPA on one of the four constructed 2-factor models, our baseline model, or combined models. NC-L is based on the Rotter's Locus of control. NC-E is based on engagement behavior, NC-R is based on self-reported relationships, and NC-B is based on self reported risky behaviors. All estimated models include the following controls: parent's education, gender, urban status, and residence in Eastern Germany. Standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11 Model Comparison: GPA (factors estimated using minimum distance distance estimator)

	NC-L	NC-E	NC-R	NC-B	BASE	Comb-L	Comb-E	Comb-R	Comb-B
Cog	0.252***	0.236***	0.227***	0.245***	0.254***	0.260***	0.240***	0.236***	0.248***
	(0.022)	(0.024)	(0.025)	(0.026)	(0.019)	(0.022)	(0.024)	(0.025)	(0.026)
Noncog	0.034	0.048**	0.035	0.075***		0.003	0.037	-0.008	0.037
	(0.022)	(0.024)	(0.025)	(0.025)		(0.022)	(0.023)	(0.025)	(0.026)
Cons.					0.189***	0.211***	0.188***	0.182***	0.180***
					(0.023)	(0.025)	(0.028)	(0.028)	(0.030)
Agree.					-0.040*	-0.039	-0.039	-0.048*	-0.048*
					(0.021)	(0.024)	(0.027)	(0.027)	(0.029)
Neuro.					-0.015	-0.008	-0.017	-0.028	-0.031
					(0.021)	(0.024)	(0.026)	(0.026)	(0.028)
Open.					-0.006	0.036	0.001	0.007	-0.016
					(0.026)	(0.029)	(0.033)	(0.034)	(0.035)
Extra.					0.001	-0.039	0.011	-0.002	0.020
					(0.026)	(0.029)	(0.033)	(0.033)	(0.036)
Risk					-0.062***	-0.065***	-0.079***	-0.076***	-0.059**
					(0.020)	(0.022)	(0.026)	(0.026)	(0.028)
Time					0.018	0.008	0.028	0.029	0.031
					(0.019)	(0.022)	(0.025)	(0.025)	(0.026)
Observations	1058	813	797	716	1292	1037	792	779	697
R^2	0.130	0.111	0.110	0.116	0.198	0.223	0.203	0.195	0.200

Notes: The table shows OLS estimates of GPA on one of the four constructed two-factor models, our baseline model, or combined models. NC-L is based on the Rotter's Locus of control. NC-E is based on engagement behavior, NC-R is based on self-reported relationships, and NC-B is based on self-reported risky behaviors. The factors are estimated using the minimum distance estimator. No imputation is done in the estimation and therefore factors are only predicted for individuals with no missing measures, explaining differences in observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.