Journal of Pension Economics and Finance (2021), **20**, 532–546
doi:10.1017/S1474747219000064

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Alternative measures of non-cognitive skills and their effect on retirement preparation and financial capability

Gema Zamarro*

University of Arkansas & University of Southern California, Fayetteville, USA
*Corresponding author. Email: gzamarro@uark.edu

(Received 11 September 2018; revised 9 January 2019; accepted 24 January 2019; first published online 28 February 2019)

Abstract

Individuals are increasingly asked to take responsibility for preparing for retirement and available financial products to do so are growing in sophistication. A better understanding of how non-cognitive skills influence financial capability and retirement preparation could help effective policy design. This area of research has been hampered by the struggle to find reliable measures of these skills. I argue that questionnaires themselves can be seen as performance tasks, such that measures of survey effort could lead to meaningful measures of non-cognitive skills. I exploit the fact that I observe respondents taking multiple survey modules covering different topics in different moments of time to build survey effort measures in a nationally representative internet panel. I use survey effort measures along with self-reports to study the role of non-cognitive skills on retirement preparation and financial capability. My results show that non-cognitive skills can have a significant role, beyond the role of cognitive ability.

Key words: Financial capability; non-cognitive skills; retirement preparation

JEL Codes: C83; C91; D14

1. Introduction

Non-cognitive skills and personality traits, such as conscientiousness, have been found to play a prominent role in shaping important long-term outcomes, such as educational attainment and labor outcomes, even after controlling for cognitive ability (Almlund *et al.*, 2011). However, we still lack a good understanding of how they might affect preparation for retirement and financial capability.

A limited amount of recent research has highlighted the potential role that non-cognitive skills and personality traits could have for retirement planning and savings. Hershey and Mowen (2000), using a small sample of Arkansas households, studied the link between personality characteristics, financial knowledge, and financial preparedness. They found that both self-reported personality characteristics such as conscientiousness and neuroticism, as well as financial knowledge, were significantly correlated with retirement planning. Hurd *et al.* (2012) also highlight the role of self-reported conscientiousness for retirement preparation. Using data from the Health and Retirement Study the authors find self-reported conscientiousness to be associated with a higher accumulation of resources for retirement both through an increased level of reported earnings but also through higher levels of reported savings. Finally, in a recent paper, Parise and Peijnenburg (forthcoming) study the relationship between self-reported conscientiousness and emotional stability (reverse of neuroticism) and financial choices among a panel of Dutch adults. They find that both personality traits are negatively associated with several measures of financial distress. Also, these self-reported personality traits were associated with higher levels of reported retirement planning and saving and negatively associated with impulse buying and unsecured borrowing. In this paper, I build on this research and study the effect of different alternative measures of non-cognitive skills and personality traits to explain individuals' reported preparation for retirement and financial capability.

A key challenge in this area of research is finding robust measures of non-cognitive skills. The limited research available has only used self-reports but these are prone to potential important biases due to reference group bias and social desirability bias (Krosnick *et al.*, 1996; Dobbie and Fryer, 2015; West *et al.*, 2016). Also, some respondents expend low effort on surveys. The problem this creates for non-cognitive skills research is that effort on surveys is likely related to the very skills that researchers are attempting to measure. For example, respondents who lack conscientiousness are unlikely to report that they lack those skills. This indicates that measurement error on surveys is potentially related to the underlying skills we seek to measure, which then could lead to invalid research findings. In particular, one could expect that in studies that use self-reported measures of non-cognitive skills and self-reported outcomes, correlated measurement error may lead to upward-biased estimates of the estimated relationships. This would be the case if, for example, respondents who overstate their personality or non-cognitive skills are also more likely to exaggerate their financial wellbeing and retirement preparation.

Because of these limitations, Duckworth and Yeager (2015) have urged the research community to exercise caution when using existing self-reported measures of non-cognitive skills. The authors highlight the importance of developing novel measures by capitalizing on advances in theory and technology. This is precisely what I do in this paper. Here, I argue that questionnaires themselves can be seen as performance tasks, such that measures of survey effort can lead to meaningful measures of non-cognitive skills. Surveys take the effort to complete, resemble paperwork and clerical tasks in daily life, and respondents reveal something about their non-cognitive skills through the effort they exhibit on them.

I study survey effort measures, as proxy measures of non-cognitive skills, among adults in the Understanding America Study (UAS), a nationally representative internet panel of households. The UAS provides me with a unique framework to test for the validity of these measures as respondents are observed taking multiple survey modules on different topics and different moments of time, receive a payment to do so, independently of survey effort, and are not made aware that survey effort could be tracked. Using UAS metadata, including detailed information on whether a respondent skipped a question he or she should have answered, I am able to assess respondents' survey taking effort, independently of survey topics and time. In addition, I also study an adaptation, for the adult population, of a standard performance task, the Academic Diligence Task (ADT) developed by Galla *et al.* (2014).

My results show the difficulty of adapting the ADT to a different context and population and the promise of survey effort measures to proxy for relevant non-cognitive skills. In particular, measures of careless answering in surveys, show great promise for being good proxy measures of conscientiousness and neuroticism. When related with measures of financial capability and retirement preparation, I find that both self-reported measures of conscientiousness, neuroticism, grit, and, more importantly, measures of careless-answering are significant determinants of the level of financial capability and retirement preparation of UAS respondents, beyond the role of cognitive ability.

These findings have important policy implications. As an increasing share of the responsibility for a good financial plan for the future is given to individuals and financial products are growing on their sophistication, there are growing concerns about Americans' low readiness for retirement financial needs. As a result, a better understanding of the personal factors driving to sound financial decisions is crucial for the design of targeted policies and interventions that could help promote financial capability and retirement preparation. This paper's results highlight the importance of considering psychological factors. For instance, traditional financial educational interventions might not be fully successful among those individuals who lack the level of conscientiousness to act on their provided information. In this respect, Lusardi *et al.* (2017) show theoretically substantial levels of heterogeneity on returns to financial education. They argue that, in order to change behavior, financial education programs should be targeted to specific groups of the population. My results highlight the potential importance of considering an individual's non-cognitive skills and personality traits in creating such target groups.

2. Data

This project uses the UAS, an ongoing internet panel of American households run by the University of Southern California, comprising a nationally-representative sample of approximately 6,000 respondents.[1]

2.1 Self-reported measures of non-cognitive skills

This study includes self-reported measures of the Big Five personality traits and grit. The Big Five is a taxonomy of five major personality traits including conscientiousness, agreeableness, neuroticism, extraversion, and openness. Overall, the Big Five model is one of the most widely used schemas in personality research and practice. More recently, economists have also used measures of the Big Five personality traits and found them to be related to life outcomes in a variety of ways (e.g., Almlund *et al*., 2011).

My measures of the Big Five personality traits were collected in the very first survey the UAS respondents take after joining the panel (UAS1[2]) and it is based on a 44-item scale developed by John *et al*. (1991). Based on the answers to this scale, each respondent receives a continuous score from one to five on each of the five personality dimensions described above.

This paper builds on the work by Zamarro *et al*. (2018) for which a wave of data that included self-reported grit (Duckworth and Quinn, 2009) was collected through survey modules UAS15 and UAS37. Grit is defined as 'perseverance and passion for long-term goals' (Duckworth *et al*., 2007: 1087). The grit scale has eight items where respondents are asked to evaluate themselves on a five-point scale (very much like me; mostly like me; somewhat like me; not much like me; not like me at all) on a series of statements including, among others, 'I am a hard worker', 'I am diligent', or 'Setbacks don't discourage me'. A grit score is then computed for each respondent to the survey by averaging the scores from responses to each of the items in the scale.

Tables 1 and 2 present descriptive statistics for these self-reported measures of personality traits and grit. Similarly to results in Duckworth and Quinn (2009), self-reported grit measures in my sample exhibit strong significant positive correlations with self-reported measures of conscientiousness, a moderate significant negative correlation with self-reported neuroticism, a moderate significant positive correlation with agreeableness, and a weak significant positive correlation with extraversion and openness to experience. Observed correlations, however, are generally of smaller size, with the exception of openness to experience, than those observed by Duckworth and Quinn (2009). However, one could expect these correlations not to be fully comparable, due to differences in samples, as I use a nationally representative sample and their work used a convenience sample of adults who volunteered to participate in the study. I also observe some intercorrelations among the Big Five personality traits measures in this table. For instance, conscientiousness exhibits positive moderate correlation with agreeableness and negative moderate correlation with neuroticism. These intercorrelations are to be expected as certain behaviors, used for their measure, may reflect multiple traits, and are not unusual in the personality literature (Costa and McCrae, 1992).

2.2 Performance task measures of non-cognitive skills

2.2.1 Measures of survey effort

Survey effort can be measured by analyzing response patterns within surveys. Recent evidence has highlighted the potential of studying response patterns to questionnaires as a way of quantifying non-cognitive skills (see Hitt, 2015; Hitt, Trivitt and Cheng, 2016; Zamarro *et al*., 2018). For example, the rate at which students skip questions on surveys is predictive of later educational attainment and labor-market outcomes (Hitt, Trivitt and Cheng, 2016). Similarly, measures of 'careless answering' on

[1]For more information about the UAS, please see the Introduction section to this special issue.

[2]UAS1 refers to the first survey module respondents take. UAS data used in this paper is publicly available and can be accessed here: https://uasdata.usc.edu/index.php

Table 1. Summary statistics for measures of non-cognitive skills

Measure	Mean	St. Dev.	Min.	Max.	*N*. Obs
1. Grit	3.58	0.60	1.37	5.00	4,906
2. Conscientiousness	4.05	0.62	1.00	5.00	5,224
3. Agreeableness	4.02	0.62	1.00	5.00	5,223
4. Neuroticism	2.64	0.82	1.00	5.00	5,222
5. Extraversion	3.35	0.79	1.00	5.00	5,218
6. Openness	3.61	0.63	1.00	5.00	5,218
7. Item non-response	0.08	0.02	0.01	0.48	5,021
8. Careless answers	0.01	1.01	−1.96	4.43	5,075
9. Correct answers	0.97	0.06	0	1	901
10. Time on task	92.90	17.15	0.34	100	904

Note: Summary statistics presented using population weights.

Table 2. Correlation matrix of non-cognitive traits measures

	1	2	3	4	5	6	7	8
1. Grit	–							
2. Conscientiousness	0.50**	–						
3. Agreeableness	0.24**	0.42**	–					
4. Neuroticism	−0.33**	−0.39**	−0.42**	–				
5. Extroversion	0.18**	0.24**	0.21**	−0.28**	–			
6. Openness	0.14**	0.23**	0.23**	−0.21**	0.32**	–		
7. Item nonresponse	−0.05**	−0.04**	−0.04**	0.02†	−0.01	−0.05**	–	
8. Careless answers	−0.16**	−0.18**	−0.10**	0.27**	−0.10**	−0.03*	0.05**	–
9. Correct answers	0.06	0.04	0.03	−0.01	0.02	0.01	−0.10**	−0.10**
10. Time on task	−0.006	0.01	−0.0003	−0.01	0.01	0.005	0.001	−0.03
11. Cognitive ability	−0.04	−0.008	−0.14**	−0.07**	−0.06*	0.14**	0.004	−0.23**

Note: †p < 0.1; *p < 0.05; **p < 0.01.

surveys by both teenage students and adults are found to be predictive of educational and labor-market outcomes in adulthood (see Hitt, 2015; Zamarro *et al*., 2018).

By quantifying the extent to which individuals put forward effort in surveys I am able to obtain information about respondents who otherwise may provide unreliable self-reported information. In addition, these performance task-based measures are not prone to reference group bias as respondents simply reveal personal attributes by their behavior. Also, respondents are typically unaware that they are being assessed on survey effort, which avoids issues such as social desirability bias or experimenter bias. An added cost-effective benefit of survey-based effort measures is that these measures often will not require new data to be collected. Therefore, one could obtain measures of non-cognitive skills from existing surveys to complement the already-collected information, expanding the opportunity for researchers to answer new questions with existing data. In this study, I follow the work by Zamarro *et al*. (2018) and study the potential of measures of item nonresponse and careless answering in the UAS to proxy for relevant non-cognitive skills.

Item nonresponse rates are defined as the percentage of items that respondents skipped out of the total number of items they were required to complete in a given survey. Upon request, I obtained metadata information on whether a respondent skipped questions he or she should have answered. With this information, I computed the item nonresponse rates for surveys in ten different modules in the UAS[3]. These were survey modules all UAS respondents were asked to participate in and were particularly long and so, presented more potential for observing patterns of item nonresponse. These survey modules were fielded at different points in time and varied in topics including demographic and family background information, health status and knowledge, housing, income, employment and labor market,

[3]The UAS survey modules included in this measure were the following: UAS16, UAS18, UAS20, UAS21, UAS22, UAS23, UAS24, UAS25, UAS26, and UAS38.

retirement, pensions, social networks and opinion on economics and politics. Altogether, respondents were asked an average of 93.3 questions in each of these ten survey modules. I then take the average item nonresponse rate across survey modules and within each respondent. By averaging nonresponse rates along multiple survey modules covering different topics, I aim to identify a behavioral pattern independent of a specific survey topic and less affected by random fluctuations.[4]

Tables 1 and 2 present summary statistics for item nonresponse measures as well as correlations with other measures of non-cognitive skills and cognitive ability. On average, UAS respondents exhibited item nonresponse rates of about 8%. Item nonresponse rates, however, did not present much construct validity in my sample as they showed very weak correlations with self-reported measures of grit and personality traits. Although weak, correlations were, however, significant and in the expected direction. Weak negative and significant correlations were observed with self-reported grit, conscientiousness, agreeableness and openness to experience. A marginally significant positive weak correlation was also observed with self-reported neuroticism. No significant correlation was found between item nonresponse and the cognitive ability measure, described in detail below. These small correlations could be due to the fact that item nonresponse is discouraged in the UAS. If respondents leave an answer blank, this triggers a screen that reminds them of the importance of their answers and asks them to return and provide a response.[5] Additionally, respondents might avoid leaving questions blank in order to build a good reputation as panel members. As a consequence, I have doubts that item nonresponse is a good proxy for non-cognitive skills in this case. However, in the results section below, I still present results of correlations between item nonresponse rates and financial capability and retirement preparation.

Careless Answering Measures. Instead of skipping items, some respondents may provide thoughtless and incoherent answers. For instance, some respondents may report the same answer to every question (i.e., straight-lining) in order to complete the survey with minimal effort and quickly (O'Conner *et al*., 1982). Others may simply provide random answers. My second measure of survey effort identifies these patterns.

I follow Hitt (2015) to build a measure of careless answering by generalizing diagnostic techniques that psychologists have used to analyze data quality (Johnson, 2005; Huang *et al*., 2012; Meade and Craig, 2012). First, I identify reliable self-reported scales that respondents had to answer. I study answer patterns in several survey modules, fielded at different points in time and covering an array of topics. I restrict the analysis to survey modules different from the modules that contain other data for the analysis to eliminate confounding variation. I chose the following three scales to build the careless answering measure: A life satisfaction scale, a well-being scale, and a depression scale.[6] All these scales in the data had high-reliability coefficients, ranging from 0.7 to 0.9 Cronbach's α scores.

Within each of the selected scales, I regress responses from each item on the average score of the rest of items. Answers among items on a reliable scale should be well correlated with each other. However, an individual who is careless in responding to a scale will submit answers that are more weakly correlated with each other. Residuals from each of the regressions will capture the response inconsistencies between each item and the remaining items, based upon the responses that the individual and others in the analytic sample provided on those remaining items.

I standardize the absolute values of these residuals to account for any differences across the items, within the same scale, and then average these standardized residuals within scales. Finally, after standardizing each of these averages to take into account differences across scales (e.g., the different total number of items, or answer options), I create a composite careless answering score by averaging these standardized averages of residuals at the individual level[7].

[4]Research has found that respondents tend to skip items that are sensitive in nature (Tourangeau and Yan, 2007). By averaging item nonresponse over a set of survey modules covering a range of topics, I mitigate the possibility that the measure is driven by one survey containing several sensitive questions.

[5]Obviously, respondents can choose to ignore the alert and continue answering subsequent items, hence the nonzero item nonresponse rates.

[6]The life satisfaction and the well-being scales are from UAS2. The depression scale is in UAS20.

[7]See Hitt (2015) for additional technical details and explanation on this measure of careless answering.

Tables 1 and 2 show summary statistics for careless answering measures as well as correlations with self-reported measures of non-cognitive skills and cognitive ability. Careless answering is a standardized measure and so the mean and standard deviation are not so informative. However, I observe a significant range in values of careless answering behavior with some respondents giving well-predicted answers (negative values) and others presenting higher unexpected responses (positive values). As it was also the case in results presented in Zamarro *et al.* (2018), I find that careless answering is most correlated with self-reported measures of neuroticism (positively correlated) and with self-reported measures of conscientiousness and grit (negatively correlated). This result speaks to the construct validity of this measure. Finally, I also observe a negative and significant correlation between careless answering and the cognitive ability variable of about −0.2, indicating that this behavior is more common among respondents with lower cognitive ability levels. In the next section, I explore the relationship between careless answering, financial capability and retirement preparation along with the self-reported measures of grit and personality traits, after controlling for cognitive ability differences across respondents.

2.2.2 The academic diligence task

I also collected data through a standard performance task, using an adaptation of the Academic Diligence Task (ADT) (Galla *et al.*, 2014). In the original task, a convenience sample of high school students was given the option to perform simple math problems, which they were told to be beneficial, or play computer games. Galla *et al.* (2014) found that the number of questions answered correctly and the time spent on task were weakly but significantly correlated with students' self-reported conscientiousness and grit. Additionally, the number of correct answers and the time on task were also significantly correlated with high school GPA, academic achievement, on-time high school graduation, and college enrollment.

I adapted the ADT and collected data on a subsample of UAS respondents. First, respondents were prompted about the importance of simple mental exercises and their potential role in preventing mental diseases. Secondly, they were asked to choose five web pages, from a list of 23, that would be available during the task, the distractors. Finally, respondents were asked to perform as many verbal and math problems as possible in 10 min but allowed to take breaks to surf the web through their selected five web pages.

Tables 1 and 2 present descriptive statistics of the percentage of correct responses answered as well as the percentage of total time they were on task. The big majority of respondents did not seem to be tempted by the distractors, they took the task very seriously and devoted all or almost all their time to perform the task. This lead to very high percentages of correct answers. As a result, there was a lack of variation across respondents on their performance in the task leading to very small correlations with self-reported non-cognitive skills. Given the low construct validity of the ADT in my sample, I do not think this would be a meaningful measure of relevant non-cognitive skills in this case. As it turns out I also found no correlation between ADT performance and financial capability or retirement preparation in my sample[8].

2.3 Retirement preparation and financial capability measures

My analysis uses three sets of outcome measures with the aim to capture different dimensions of respondents' financial capability, consumer financial wellbeing, and retirement preparation. As part of my measures of financial capability, I include respondent's financial literacy scores based on respondent's responses to 20 questions developed to measure their financial knowledge (Knoll and Houts, 2012). Respondents then get scored on a scale of 0–20 representing the number of questions they answered correctly. In that same survey module, respondents were also asked to self-report how many questions they think they have answered correctly, from 0 to 20. This measure constitutes the perceived financial literacy scale[9]. Finally, I include information about the respondent's total value

[8]Results available from the authors upon request.

[9]The financial literacy score and perceived financial literacy measures were collected as part of UAS6.

of assets, excluding the value of their secondary residence[10], measured in 10,000's of dollars, as another measure of financial capability.

Consumer financial wellbeing, defined by the Consumer Financial Protection Bureau (CFPB) as the level to which a person can fully meet current and ongoing financial obligations, can feel secure in their financial future, and is able to make choices that allow them to enjoy life, is captured through the CFPB financial well-being scale[11]. The scale is based on a set of 10 questions and a specific scoring system by which a financial well-being score on a scale of 0–100 is provided, with higher scores representing higher levels of financial well-being[12]. In addition, I also use information on respondents' reported credit scores and generated an indicator variable for the respondent reporting have a good or excellent level of credit score (credit score above 700)[13].

The final set of outcome variables aims to capture respondents' reported levels of preparation for retirement. In particular, I developed two indicator variables that capture if the respondent reported being very well or somewhat prepared financially for retirement and whether the respondent has thought and developed a plan for retirement through answering yes to both of the following questions: 'In the past, have you ever tried to figure out how much your household should save for retirement?' and 'Have you ever tried to develop a plan for your retirement?'.[14]

Table 3 presents summary statistics for the outcome variables. Out of 20 financial literacy questions, on average, UAS respondents responded correctly to almost 14 questions while they perceived they had responded correctly to 13 of such questions. On average, respondents report about $287,000 in their total value of assets. On a scale from 0 to 100, the average of the consumer wellbeing index in my sample is about 54 points. Forty-nine percent of respondents report having good or excellent credit scores, 22% report being financially prepared for retirement while only 13% report having thought about and tried to develop a retirement plan.

2.4 Cognitive ability and other relevant information

There are multiple sources of information on cognitive ability in the UAS that I use in this analysis. These include the Lipkus Numeracy Scale (Lipkus *et al*., 2001), responses to a Cognitive Reflection Test (Frederick, 2005; Toplak *et al*., 2014), and a quantitative reasoning, picture vocabulary, and verbal analogies battery from the Woodcock-Johnson Tests of Cognitive Abilities (Mather and Jaffe, 2016). I combined information on all these scales to form a unique cognitive ability index using factor analysis of the total number of correct responses in each of these tests. All scales loaded onto a unique factor with relative equal size weights.[15,16]

Other relevant demographic controls included in the analysis include information about respondent's age, gender, ethnicity, whether born in the USA, region of residence (West, Midwest, Northeast or South), whether the respondent is currently working, whether the respondent is currently retired, education level (college degree, high school degree), and whether the respondent is currently married or living together with a partner.

Table 3 presents summary statistics for demographic variables and the cognitive ability measure included in the analysis. On average, respondents are about 47 years old, a majority are working (61%), have a high school degree (50%) or a college degree (40%) and are born in the USA (91%).

[10]This variable is obtained from the UAS-HRS public use dataset.

[11]This variable is obtained from UAS38.

[12]For more information see: https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/

[13]Information on credit scores is obtained from UAS48.

[14]Information to build these two indicator variables was obtained from UAS16 and UAS26.

[15]Information on the Lipkus Numeracy Scale and Cognitive Reflection Test were collected during the very first survey of the UAS (UAS1), while the quantitative reasoning, picture vocabulary, and verbal analogies battery from the Woodcock-Johnson Tests of Cognitive Abilities where collected during later survey modules in UAS42, UAS43, and UAS44.

[16]Results are available from the authors upon request.

Table 3. Summary statistics for outcome variables, demographic variables, and cognitive ability

Measure	Mean	Standard deviation	Minimum	Maximum
Financial capability				
Financial literacy	13.84	3.08	0	20
Perceived fin. liter.	13.20	4.35	0	20
Tot. val. assets (10,000s)	28.75	107.31	−687.51	3607
Consumer financial wellbeing				
Consumer fin. well.	53.99	12.88	14	95
Good or excellent credit.	0.49	0.50	0	1
Retirement preparation				
Prepared retirement	0.22	0.42	0	1
Thought of retirement	0.13	0.34	0	1
Age	47.33	16.78	18	98
Female	0.52	0.50	0	1
Black	0.12	0.32	0	1
Hispanic	0.001	0.025	0	1
Other race	0.24	0.43	0	1
Born in USA	0.91	0.29	0	1
West	0.18	0.39	0	1
Midwest	0.08	0.28	0	1
Northeast	0.11	0.32	0	1
South	0.27	0.45	0	1
Working	0.61	0.41	0	1
Retired	0.19	0.31	0	1
High school degree	0.50	0.50	0	1
College	0.40	0.49	0	1
Married/living togeth	0.57	0.49	0	1
Cognitive ability-factor	−0.10	1.00	−3.02	2.64

Note: Sample sizes range from 3,104 to 5,949. Summary statistics use population weights.

3. Studying the effect of non-cognitive skills on retirement preparation and financial capability

Next, I study the role of both self-reported measures of non-cognitive skills and measures of survey effort, through item nonresponse and careless answering, on explaining financial capability and retirement preparation. I estimate slight variations of the following linear regression model:

$$Y = \beta_0 + \beta_1 X_i + \beta_2 \text{ Non\_cognitive skills}_i + \gamma_i^S + \varepsilon_i \tag{1}$$

where Y is an outcome measure, as described above. β_2 is the coefficient of interest representing the association between respondents' non-cognitive skills and preparedness for retirement. My regressions include the following alternative measures of non-cognitive skills: self-reported Big five personality traits, self-reported grit measures, item non-response rates and measures of careless answering. Four sets of separate regressions are obtained including each of these four alternative measures of different non-cognitive skills. X_i includes relevant socio-economic background information, education level, work status, marital status, and, importantly, cognitive ability. Finally, I also control for regional dummies collected in γ_i^S as a means of controlling for any unobserved differences across regions in the USA.

4. Results

Table 4 presents regression coefficients for the effect of non-cognitive skills on financial capability measures. Columns one, three and five of Table 4, panel A, present the estimated effect of each of the self-reported Big five personality traits. As can be seen, I fail to find a statistically significant effect of self-reported conscientiousness on financial literacy scores or total value of assets. From the Big five personality traits only openness to experience shows a small but statistically significant effect on financial literacy scores. A one point increase in openness is associated with about a 0.2 point increase in the financial literacy score. Interestingly, all self-reported personality traits, except for openness to

Table 4. Financial capability and self-reported and survey effort measures of non-cognitive skills (OLS estimates)

Panel A Variables	(1) Fin. Lit	(2) Fin. Lit	(3) Perc. Fin. lit	(4) Perc. Fin. lit	(5) Total assets	(6) Total assets
Cognitive ability	1.456***	1.452***	1.189***	1.202***	6.801***	6.954***
	(0.053)	(0.056)	(0.097)	(0.095)	(2.478)	(2.678)
Conscientiousness	−0.069		0.335**		4.230	
	(0.080)		(0.149)		(4.603)	
Agreeableness	0.131		−0.343**		−1.249	
	(0.085)		(0.169)		(4.013)	
Neuroticism	0.037		−0.349***		−4.497	
	(0.059)		(0.115)		(3.474)	
Extraversion	−0.065		0.248**		2.454	
	(0.058)		(0.107)		(2.510)	
Openness	0.198***		0.116		2.732	
	(0.069)		(0.138)		(2.751)	
Grit		−0.024		0.434***		4.904
		(0.072)		(0.135)		(3.993)
Observations	4,381	4,048	4,037	3,741	2,846	2,799
Adjusted R^2	0.469	0.459	0.271	0.274	0.0446	0.0432
Panel B	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Cognitive ability	1.459***	1.463***	1.175***	1.152***	7.120***	6.073**
	(0.053)	(0.053)	(0.094)	(0.094)	(2.602)	(2.513)
Item non-response	−6.671***		−7.705		164.874***	
	(2.341)		(7.449)		(54.377)	
Careless answering		−0.094**		−0.308***		−3.628**
		(0.045)		(0.093)		(1.429)
Observations	4,395	4,395	4,046	4,046	2,856	2,856
Adjusted R^2	0.469	0.469	0.265	0.265	0.049	0.044

Note: Demographic variables, educational attainment levels, and employment, and marital status included as controls.
Standard errors in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1.

experience, are significantly associated with perceived levels of financial literacy. Conscientiousness and extraversion are positively associated with perceived financial literacy levels while agreeableness and neuroticism are negatively related. Columns two, four and six of Table 4, panel A, present the results when self-reported grit is included as an explanatory variable in the analysis, instead. Similarly to the results I observed for self-reported conscientiousness, a personality trait found to be related to grit, I observe that self-reported grit does not show any statistically significant association with financial literacy scores or the total value of asset measures. However, an increase of one point in self-reported grit is associated with a 0.4 point increase in perceived financial literacy scores. The fact that I fail to find a relationship between reported conscientiousness or grit and financial literacy scores but find a relationship with perceived performance in the literacy test, goes in line with the notion that correlated measurement errors could lead to overstated estimated relationships between self-reported variables.

Table 4, panel B, presents the results for the survey effort measures. Columns one, three and five present the results for measures of item non-response. In this case, I observe that a 1% increase in item nonresponse is associated with a 0.07 statistically significant decrease in financial literacy scores. A similar effect is found for perceived financial literacy levels but this effect is not statistically significant. Surprisingly, I find that higher item non-response is associated with the higher value of total assets. A 1% increase is associated with about a $16,000 increase in assets. This result is contrary to what I expected if item non-response were to be a good proxy for non-cognitive skills related to conscientiousness and generates doubts about this measure being a good proxy for relevant non-cognitive skills in this data. Finally, columns two, four and six of Table 4, panel B, present the results when careless answering measures are used. Interestingly, in this case, I do observe small but statistically significant effects of careless answering behavior not only on self-reported financial literacy but also on actual financial literacy scores and the total value of assets. Note that careless answering, as it is a performance-task measure, will not be affected by the correlated measurement errors problem described above. A one standard deviation increase in careless answering is associated with a 0.09 decrease in financial literacy scores, a 0.3 decrease in perceived financial literacy and a $36,000 decrease in total value of assets. It should be stressed that these estimates are obtained after controlling for cognitive ability, educational levels, and other relevant socio-demographic information. In that respect, my estimates are conservative effects beyond the effect of cognitive ability and demographic information. Overall, in all regressions, cognitive ability seems to be a significant driver of financial capability and retirement preparation measures.

Table 5 presents results when financial wellbeing variables are used as dependent variables. Looking at columns one and three of Table 5, panel A, I observe that self-reported conscientiousness is significantly associated with financial wellbeing levels as well as the probability of reporting good or excellent credit scores. An increase of one point in conscientiousness is associated with a 2.6 increase on the CFPB financial wellbeing scale and a 6 percentage point increase in the probability of reporting good or excellent credit scores. Neuroticism and extraversion are also found to be significantly associated with CFPB financial wellbeing levels. A one point increase in reported neuroticism and extraversion is associated with a 2.4 decrease and a 0.6 increase in the financial wellbeing scale, respectively. Agreeableness and openness to experience, on the other hand, are found to be significantly correlated with the probability of reporting having a good or excellent credit score. A one point increase in agreeableness or openness is associated with a three and a 4 percentage points decrease in the probability of having a good or excellent credit score, respectively. Columns two and four of Table 5, panel A, show the results for self-reported grit. In this case, I find that self-reported grit only shows a significant effect on the CFPB financial well-being index but not on the probability of reporting a good or excellent credit score. A one point increase in self-reported grit is associated with an almost four-point increase in the financial well-being index. In contrast, as presented in columns two and four of Table 5, panel B, careless answering is found to be correlated with both the financial well-being index and reporting having good or excellent credit scores. A standard deviation increase in careless answering is associated with an almost three-point decrease in financial wellbeing and a 5 percentage point decrease in the

Table 5. Consumer financial wellbeing and self-reported and survey effort measures of non-cognitive skills (OLS estimates)

Panel A Variables	(1) Fin. Well.	(2) Fin. Well.	(3) Good/excell. credit	(4) Good/excell. credit
Cognitive ability	1.803***	1.814***	0.076***	0.074***
	(0.263)	(0.274)	(0.012)	(0.011)
Conscientiousness	2.636***		0.066***	
	(0.402)		(0.018)	
Agreeableness	−0.543		−0.030*	
	(0.387)		(0.018)	
Neuroticism	−2.391***		−0.011	
	(0.318)		(0.013)	
Extraversion	0.622**		0.010	
	(0.295)		(0.013)	
Openness	−0.593		−0.040**	
	(0.371)		(0.016)	
Grit		3.834***		0.021
		(0.394)		(0.017)
Observations	4,324	4,021	3,467	3,415
Adjusted R^2	0.268	0.258	0.232	0.226
Panel B	(1)	(2)	(3)	(4)
Cognitive ability	1.576***	1.376***	0.074***	0.067***
	(0.256)	(0.260)	(0.011)	(0.011)
Item non-response	−71.306***		0.315	
	(18.808)		(0.534)	
Careless answering		−2.850***		−0.051***
		(0.253)		(0.010)
Observations	4,338	4,330	3,482	3,482
Adjusted R^2	0.229	0.264	0.227	0.236

Note: Demographic variables, educational attainment levels, and employment, and marital status included as controls.
Standard errors in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1.

probability of reporting good credit. Item non-response rates were only weakly correlated with the financial well-being index, as can be seen in columns one and three of this table.

Results for regressions of self-reported retirement preparation are presented in Table 6. Looking at columns one and three of Table 6, panel A, I observe that, among the Big five personality traits, both conscientiousness and extraversion are statistically significantly related to the probability of reporting having prepared for retirement and having developed a retirement plan. A one-point increase in the conscientiousness level is associated with an almost 6 percentage point increase and a 3 percentage point increase in the probability of reporting having prepared for retirement and having developed a plan, respectively. The effect of extraversion is somewhat smaller. A one-point increase in the extraversion scale is associated with an almost 3 percentage point increase in the probability of having prepared for retirement and a 1.5 percentage point increase in the probability of having developed a plan. Agreeableness and neuroticism are also found to be correlated with reported retirement preparation but their effect is negative. A one point increase in agreeableness and neuroticism is associated with a 3 and a 2 percentage point decrease in the probability of having prepared for retirement, respectively. Looking at columns two and four of Table 6, panel A, I observe that self-reported grit is also significantly related to reported retirement preparation. A one-point increase in the grit scale is associated with a 3 percentage point increase in the probability of both having prepared for retirement and having developed a retirement plan. Careless answering is a behavior that is also found to be correlated with these outcomes as it can be seen in columns two and four of Table 6, panel B. However, the correlation is found to be bigger for the probability of reporting having prepared for retirement than for the probability of actually having developed a plan. A one standard deviation increase in careless answering behavior is associated with a 4.5 percentage point decrease in the probability of being prepared for retirement but only a 1.4 percentage point decrease in the probability of

Table 6. Retirement preparation and self-reported and survey effort measures of non-cognitive skills (OLS estimates)

Panel A Variables	(1) Prep. Retire.	(2) Prep. Retire.	(3) Thought Ret.	(4) Thought Ret.
Cognitive ability	0.020** (0.009)	0.021** (0.010)	0.037*** (0.007)	0.042*** (0.008)
Conscientiousness	0.058*** (0.013)		0.034*** (0.011)	
Agreeableness	−0.032** (0.013)		0.005 (0.011)	
Neuroticism	−0.021** (0.010)		0.011 (0.009)	
Extraversion	0.026** (0.011)		0.015* (0.008)	
Openness	−0.012 (0.013)		0.008 (0.011)	
Grit		0.032** (0.014)		0.033*** (0.011)
Observations	4,566	4,062	4,566	4,062
Adjusted R^2	0.195	0.183	0.143	0.142
Panel B	**(1)**	**(2)**	**(3)**	**(4)**
Cognitive ability	0.019** (0.009)	0.011 (0.009)	0.035*** (0.007)	0.034*** (0.007)
Item non-response	0.691 (0.476)		−0.180 (0.240)	
Careless answering		−0.045*** (0.008)		−0.014** (0.006)
Observations	4,579	4,512	4,579	4,512
Adjusted R^2	0.185	0.193	0.139	0.140

Note: Demographic variables, educational attainment levels, employment, and marital status included as controls.
Standard errors in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1.

having thought of a retirement plan. Item non-response presented no association with retirement preparation variables.

5. Conclusions

As the population ages and increasing responsibility is in the hands of individuals on making sound financial decisions for their future, understanding the factors that contribute to financial capability and retirement preparation becomes increasingly important. In this paper, I explore the potential role that non-cognitive skills could have on promoting financial wellbeing and retirement preparation. Some prior research has highlighted their potential role but it has done so only using self-reported measures, which could be affected by reference group bias and social desirability bias. In particular, one would expect that in studies that use self-reported measures of non-cognitive skills and self-reported outcomes, like the ones described above, correlated measurement error may lead to upward-biased estimates of the estimated relationships. Using data from the UAS, I further study the role of non-cognitive skills on financial capability and retirement preparation, not only using self-reports but also exploring innovative alternative measures of non-cognitive skills based on survey effort.

I argue that questionnaires themselves can be seen as performance tasks, such that measures of survey effort can lead to meaningful measures of non-cognitive skills. As respondents are typically unaware that they are being assessed on survey effort, these measures are not affected by reference group bias and social desirability bias that affect self-reports. In particular, I studied measures based on item nonresponse rates and careless answering behaviors. My results for item nonresponse rates show how the construct validity of these measures could be affected by survey design decisions. Item nonresponse is discouraged in the UAS. If respondents leave an answer blank, this triggers a screen that reminds them of the importance of their answers and asks them to return and provide a response. Since respondents know that, they may be tempted to provide a less than thoughtful answer rather than leaving a question unanswered. I believe this could have contributed to the finding that item nonresponse does not appear to be a good proxy for relevant non-cognitive skills in the UAS. In contrast, measures of careless answering showed promise to be good proxy measures of non-cognitive skills related to conscientiousness and neuroticism.

For comparison, I also collected data on a standard performance task, with an adaptation, for the adult population, of the ADT. My results, however, showed the difficulty of adapting the ADT to a different context and population. Future research is needed to better design standard performance task measures that could work in an internet panel like the UAS. Alternatively, researchers could further exploit the context of internet panels and explore the use of other metadata information as potential proxies for relevant non-cognitive skills. For example, Soland (2018) shows how rates of rapid guessing in achievement tests were related to socio-emotional outcomes. Rapid guessing could also be studied in the UAS. Cheng *et al.* (2018) studied the potential of paradata at recruitment, e.g., number of reminders respondents needed before enrolling in the panel, as proxy measures of personality traits. However, more research is needed to explore the validity of these alternative proxy measures.

Finally, I explore the relationship between self-reported measures of non-cognitive skills, survey effort measures and measures of financial capability, financial well-being, and retirement preparation. My results show that both self-reported measures of non-cognitive skills, as well as careless-answering behaviors, are important determinants of the level of financial capability and retirement preparation among UAS respondents, even after controlling for cognitive ability and relevant demographic information. Cognitive ability was found to be an important and significant predictor of all measures of financial capability and retirement preparation. These results highlight the importance of considering cognitive ability but also psychological factors when designing targeted policies that aim to improve the level of financial capability and retirement preparation in the population.

I acknowledge several limitations. The validity of my results lies on the assumption that low survey effort is not translating necessarily in systematically different reports of the outcome and personality measures. Personality measures were collected in the very first survey respondents take (UAS1) in which respondents appear to put forward more effort, for instance, item response rates were very high at 98%. Respondents likely exhibited more diligence on the first survey module, assuaging concerns that self-reported measures of non-cognitive traits are distorted by low survey effort or other sources of bias. This feature helps in to strengthen the validation exercise. The rest of outcome measures, however, come from different survey modules and so I still have to assume measurement error is not systematic to affect the results. It would be good if in future work researchers are able to obtain outside measures of outcome variables to assess this assumption. Finally, my survey effort measures are based on observing respondent's behavior over multiple survey modules. I do so because I aim to identify a behavioral pattern independent of a specific survey topic and less affected by random fluctuations. However, personality traits could also be related to the probability of completing more survey modules in the panel and be part of my sample. Cheng *et al.* (2018) found that both conscientiousness and openness to experience predicted the incidence of unit nonresponse in subsequent survey modules. In this sense, it is possible that my sample of respondents is more conscientious and open to experience than the average population. However, I believe this would only make it harder for me to find the significant effects I do find.

Acknowledgements. The project described in this paper relies on data from surveys administered by the Understanding America Study, which is maintained by the Center for Economic and Social Research (CESR) at the University of Southern California. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of USC or UAS. The research reported herein was performed pursuant to a grant from the United States Social Security Administration (SSA) funded as part of the Retirement Research Consortium through the University of Michigan Retirement Research Center Award RRC08098401. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof. Special thanks to Angela Duckworth and Arie Kapteyn for their feedback and support for this project. I also thank conference participants at the 2017 University of Michigan Retirement Research Center (MRRC) Research Workshop and JPEF workshop for all their comments.

References

Almlund M, Duckworth A, Heckman J and Kautz T (2011) Personality psychology and economics. In Hanushek EA, Machin S and Wossmann L (eds), *Handbook of the Economics of Education*. Amsterdam: Elsevier, 1–181.

Cheng A, Zamarro G and Orriens B (2018) Personality as a predictor of unit nonresponse in an internet panel. *Sociological Methods and Research* 1–27. https://journals.sagepub.com/doi/abs/10.1177/0049124117747305.

Costa Jr PT and McCrae RR (1992) *Revised NEO Personality Inventory* (*NEO-PI-R*) *and NEO Five-Factor Inventory* (*NEO-FFI*) *Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Dobbie W and Fryer RG (2015) The medium-term impacts of high-achieving charter schools. *Journal of Political Economy* **123**(5), 985–1037.

Duckworth AL and Quinn PD (2009) Development and validation of the short grit scale (GRIT–S). *Journal of Personality Assessment* **91**(2), 166–174.

Duckworth AL and Yeager DS (2015) Measurement matters assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher* **44**(4), 237–251.

Duckworth AL, Peterson C, Matthews MD and Kelly DR (2007) Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology* **92**(6), 1087–1101.

Frederick S (2005) Cognitive reflection and decision making. *The Journal of Economic Perspectives* **19**(4), 25–42.

Galla BM, Plummer BD, White RE, Meketon D, D'Mello SK and Duckworth AL (2014) The Academic Diligence Task (ADT): assessing individual differences in effort on tedious but important schoolwork. *Contemporary Educational Psychology* **39**(4), 314–325.

Hershey DA and Mowen JC (2000) Psychological determinants of financial preparedness for retirement. *The Gerontologist* **40**(6), 687–697.

Hitt CE (2015) Just Filling In the Bubbles: Using Careless Answer Patterns on Surveys as a Proxy Measure of Noncognitive Skills. UARK-EDRE WP 2015-06. http://www.uaedreform.org/downloads/2015/07/edre-working-paper-2015-06.pdf.

Hitt CE, Trivitt JR and Cheng A (2016) When you say nothing at all: the predictive power of student effort on surveys. *Economics of Education Review* **52**, 105–119.

Huang JL, Curran PG, Keeney J, Poposki EM and DeShon RP (2012) Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology* **27**(1), 99–114.

Hurd MD, Duckworth A, Rohwedder S and Weir DR (2012) Personality Traits and Economic Preparation for Retirement. MRRC WP 2012-09. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.391.2886&rep=rep1&type=pdf.

John OP, Donahue EM and Kentle RL (1991) *The Big Five Inventory—Versions* 4a *and* 54. Berkeley, CA: Institute of Personality and Social Research, University of California.

Johnson JA (2005) Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality* **39**(1), 103–129.

Knoll MAZ and Houts CR (2012) The financial knowledge scale: an application of item response theory to the assessment of financial literacy. *The Journal of Consumer Affairs* **3**, 381–410.

Krosnick JA, Narayan S and Smith WR (1996) Satisficing in surveys: initial evidence. *New Directions for Evaluation*, Special Issue: Advances in Survey Research, **70**, 29–44.

Lipkus IM, Samsa G and Rimer BK (2001) General performance on a numeracy scale among highly educated samples. *Medical Decision Making* **21**(1), 37–44.

Lusardi A, Michaud P and Mitchell O (2017) Optimal financial knowledge and wealth inequality. *Journal of Political Economy* **125**(2), 431–477.

Mather N and Jaffe LE (2016) *Woodcock-Johnson IV: Reports, Recommendations, and Strategies*. Hoboken, NJ: John Wiley & Sons.

Meade AW and Craig SB (2012) Identifying careless responses in survey data. *Psychological Methods* **17**(3), 437–455.

O'Conner PJ, Sullivan GL and Jones WH (1982) An evaluation of the characteristics of response quality induced by follow-up survey methods. *Advances in Consumer Research* **9**, 257–259.

Parise G and Peijnenburg K (forthcoming) Noncognitive abilities and financial distress: evidence from a representative household panel. *The Review of Financial Studies*.

Soland J (2018) Can the amount of time students spend on test items help illustrate social-emotional learning needs? Initial evidence from an international achievement test. Paper presented at the American Education Finance and Policy 43rd Annual Conference, Portland, Oregon.

Toplak ME, West RF and Stanovich KE (2014) Assessing miserly information processing: an expansion of the cognitive reflection test. *Thinking & Reasoning* **20**, 147–168.

Tourangeau R and Yan T (2007) Sensitive questions in surveys. *Psychological Bulletin* **133**(5), 859–883.

West MR, Kraft MA, Finn AS, Martin R, Duckworth AL, Gabrieli CFO and Gabrieli JDE (2016) Promise and paradox: measuring students' Non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis* **38**(1), 148–170.

Zamarro G, Cheng A, Shakeel M and Hitt C (2018) Comparing and validating measures of non-cognitive traits: performance task measures and self-reports from a nationally representative internet panel. *Journal of Behavioral and Experimental Economics* **72**, 51–60.

Cite this article: Zamarro G (2021). Alternative measures of non-cognitive skills and their effect on retirement preparation and financial capability. *Journal of Pension Economics and Finance* **20**, 532–546. https://doi.org/10.1017/S1474747219000064