Online Appendix

The Development of Students' Engagement in School, Community, and Democracy

[*Not intended for publication in printed versions*]

Survey Information

Starting in the 2014-15 school year (Wave I), and continued through the 2015-2016 school year, WCPSS implemented a new, annual student survey that included 49 questions drawn from three main sources: The 33-item Student Engagement Instrument (Appleton, Christenson, Kim, and Reschly 2006), the 8-item Short Grit scale (Duckworth and Quinn, 2009) and 6 civic engagement items drawn from the political science literature. Two additional questions about academic rigor were developed within the district. In both years, the survey was fielded in the Spring term.[1] The survey was administered using the open-source LimeSurvey (v. 2.05) software situated behind the district's firewall. This feature enabled the district to capture personally-identifiable information in order to facilitate matching to administrative data. To begin the survey, students in grades 5, 8, and 9 entered their unique student identifiers.[2] The survey was available in either English or Spanish and the survey readability registered at grade 4.9 on the Flesch–Kincaid scale (Kincaid, Fishburne Jr, Rogers, and Chissom, 1975) and as such, was appropriate for grade 5 and higher.[3]

All schools in the district were required to participate, but some schools did not administer the survey to fidelity. There were a few potential reasons for this. First, some schools simply chose not to administer the survey. We suspect this was because the stakes were low for students in grades 5 and 8, as their results would not inform student engagement in the same school for the subsequent year. Second, the survey was initially socialized among principals through a short note in a weekly newsletter. Follow-up only occurred late in the survey window when some schools reported low response rates. Third, the end of the survey window conflicts with various assessments across grade levels. Schools intending to administer the survey toward the end of the window may have opted to focus on test administration at the cost of survey administration. Fourth, the first wave, 2014-15, was the first year of bring-your-own-device (BYOD) administration, during which schools participating accessed the internet through a unique wireless network that did not support the LimeSurvey link. This may have happened in any number of the district's 13 BYOD schools. Students in these schools attempting to access the survey would have been blocked from doing so since the survey was only accessible through the official district wireless network. While the district's IT administrators ultimately resolved the issue, the initial restricted access may have led to an ultimate decline in survey-takers. Finally, the student survey in previous years (which did not have the grit or civic engagement items) was administered anonymously through SurveyMonkey and, prior to that, through bubble sheets. Some school leaders and teachers unfamiliar with the new survey format may have been unprepared to proceed with administration. Excluding the 18 schools that did not fully participate (as evidenced by no students responding or just a couple (<5%) of students responding), the survey achieved a response rate of 72.9% of students in the three grades. Including those schools who were eligible to participate but did not, the unadjusted response rate was 63.1%. Survey

[1] For example, the 2015 survey was fielded March 18 to May 15 of that year (with a two week extension for low-response schools and those with a slightly different technology environment explained further below)

[2] Most students were familiar with this number because it is also their lunch number, but students who did not know their number would typically ask their classroom teacher, who could locate it in the student information system

[3] To preserver comparability, we restrict our analyses to the English version of the survey.

non-response was less of a problem in the second wave in 2015-2016. Only 2 schools did not fully participate and many more students participated in schools overall, with the survey achieving a response rate of 82.5% of students in the three grades. Including the schools who were eligible to participate but did not, the unadjusted response rate was 81.3%. Combining both years, the unadjusted response rate was 72.8% and the adjusted response rate was 78.3%.

While we include the most commonly used student-level characteristics, we cannot use the district's variable for free- and reduced-price lunch (FRL) status. This is because the North Carolina Department of Public Instruction (NCDPI) restricts access to and use of this measure based on its interpretation of National School Lunch Act disclosure provisions. NCDPI's annual memorandum to this effect summarizes entities and/or programs permitted to access and/or use FRL data, including state and federal child nutrition programs, federal education programs (e.g., Title I), and state education programs administered by a local education agency (LEA). Research and evaluations of LEA programs are not permitted to use FRL data. The annual memorandum, entitled "Disclosure of Student's Eligibility Status for Free and Reduced Price Meals: Memorandum of Agreement" (dated August 24, 2016) is available on NCDPI's website (NCDPI, 2016).[4]

[4] Cite: NCDPI. (2016). Disclosure of Student's Eligibility Status for Free and Reduced Price Meals: Memorandum of Agreement.

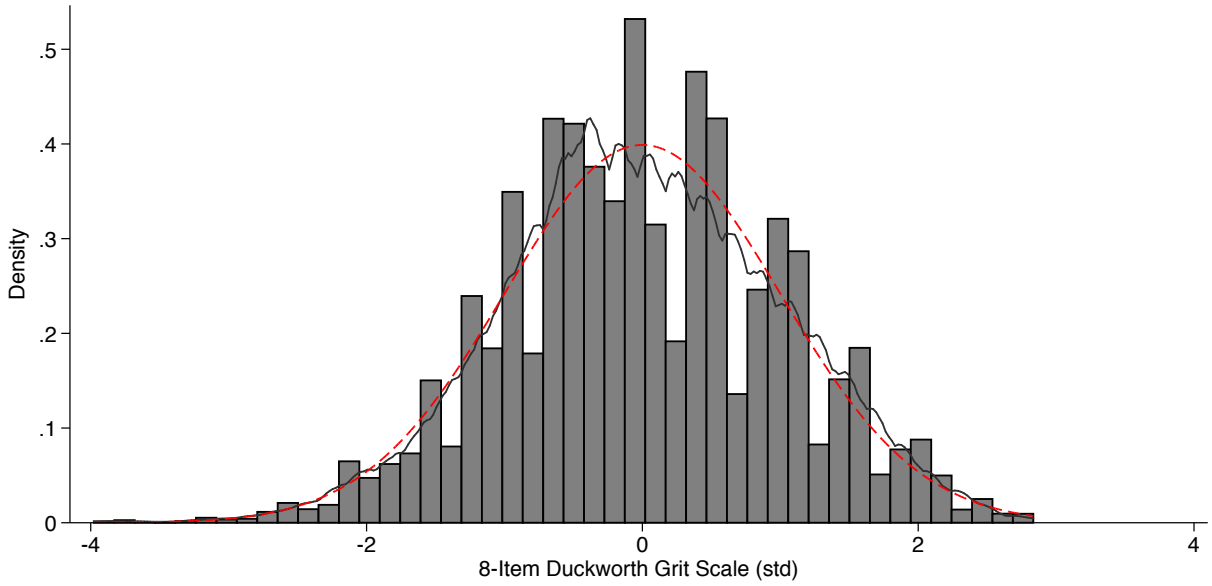
Table A.1: WCPSS Summary Statistics

	Variable	N	μ	σ	Min.	Max.
IV	Duckworth Grit (std.)	50,094	0.00	1.00	-3.98	2.83
DVs	Math (std.)	51,283	0.00	1.00	-2.84	4.28
	Reading (std.)	46,945	0.00	1.00	-3.43	2.43
	Number of Unexcused Absences	74,006	2.4	4.2	0	146
	Number of Unexcused Tardies	95,414	3.5	9.4	0	244
	School Engagement Scale (std.)	52,513	0.00	1.00	-4.68	1.93
	Vote When Older	49,934	0.84	0.37	0	1
	Volunteer	49,679	0.63	0.48	0	1
	Political Efficacy	49,893	3.20	0.72	1	4
	Civic Engagement Scale (std.)	48,136	0.00	1.00	-3.72	1.86
Other	Age (student)	95,412	13.65	2.13	8	20
	Prop. Academically Gifted (student)	95,414	0.11	0.32	0	1
	Female (student)	95,412	0.48	0.50	0	1
	Limited English Proficiency (student)	95,414	0.05	0.22	0	1
	Magnet School (school)	91,270	0.26	0.44	0	1
	Title I School (school)	91,220	0.13	0.34	0	1
	% Free-Reduced Price Lunch (school)	90,993	34.6	18.5	3.3	95.7
	Sibling in Survey Sample (student)	23,596	0.31	0.46	0	1
	Twin in Survey Sample (student)	99,139	0.02	0.13	0	1
	Proportion Answers Strongly Agree (student)	51,032	0.33	0.22	0	1
	Proportion of Questions Missing (student)	51,350	0.06	0.18	0	1
	Straight-line on Grit Scale (student)	51,350	0.01	0.12	0	1
	Conflict Response Political Interest (student)	51,350	0.07	0.25	0	1
	Minutes to Finish Survey (student)	47,855	7.39	3.54	0.47	101.5

Table A.1 provides summary statistics for some of the variables we have from our WCPSS data. Variables labeled "std." are standardized within grade to have a mean of zero and standard deviation of one. The first panel shows grit—our primary independent variable of interest. The second panel shows our primary outcomes of interest for achievement, school engagement, and civic engagement. The columns show the number of observations measuring the various variables, while the rest show the mean, standard deviation, minimum, and maximum values.

Figure A.1 shows the distribution of our grit scale—plotting a simple histogram, with both a smoothed kernel density of our grit scale's distribution and a standard normal distribution overlaid. As can be seen, the distribution of grit in the WCPSS is roughly normal, perhaps slightly right skewed.

Figure A.1: Duckworth Grit Scale Distribution in WCPSS (Waves 1 and 2)



Notes: Distribution of the 8-item Duckworth Child Grit Scale (Duckworth et al. 2007; Duckworth and Quinn 2009). Overlaid is a smoothed kernel-density (solid black line) and a normal distribution (dashed red line). The grit scale is standardized within grade.

Table A.2 shows systematic differences in levels of grit across student characteristics. This provides evidence that grit varies systematically across individual student attributes such as socio-economic status, grade, race/ethnicity, and gender and school characteristics. These descriptive patterns show that there is a gender gap—with female students being more gritty than males, a racial gap—with Asian and White students scoring substantially higher on the Duckworth grit scale than Black, Hispanic, and Other students, and a gap in academic performance—with those students being labeled academically gifted scoring noticeably higher than those students without this label. There also appears to be a socio-economic gap at the school level, with schools having the highest third of free-reduced price lunch students scoring 0.11 standard deviations below low free-reduced prince lunch schools. The gap in magnet schools is largely explained by the fact that WCPSS magnet schools largely draw low SES students in the urban center of the district.

Table A.2: Overview of Grit's (std.) Variation

Variable	Mean	*p*
Female (student)	0.060	0.00
Male (student)	-0.061	
White (student)	0.048	0.00
Black (student)	-0.070	
Hispanic (student)	-0.138	
Asian (student)	0.227	
Other (student)	-0.086	
5th Grade (student)	0.203	0.00
8th Grade (student)	-0.115	
9th Grade (student)	-0.131	
Academically Gifted (student)	0.196	0.00
Not Academically Gifted (student)	-0.031	
Limited English Proficiency (student)	-0.236	0.00
Not Limited English Proficiency (student)	0.011	
High Free/Reduced Lunch (school)	-0.054	0.00
Medium Free/Reduced Lunch (school)	-0.009	
Low Free/Reduced Lunch (school)	0.055	
Magnet School (school)	-0.080	0.00
Non-Magnet School (school)	0.023	

Table I shows differences in students' levels of grit across several student- and school-level dimensions. The second column shows mean levels of grit standardized, but not within grade (to allow for differences across grades). The third column shows the p-value from an analysis of variance test across the group levels.

Heterogeneities

Some may be curious which student subgroups benefit the most from grit development. Figure A.2 shows the heterogeneities in our grit estimates. We focus on three heterogeneities of interest: race, cognitive ability, and political motivation.[5]

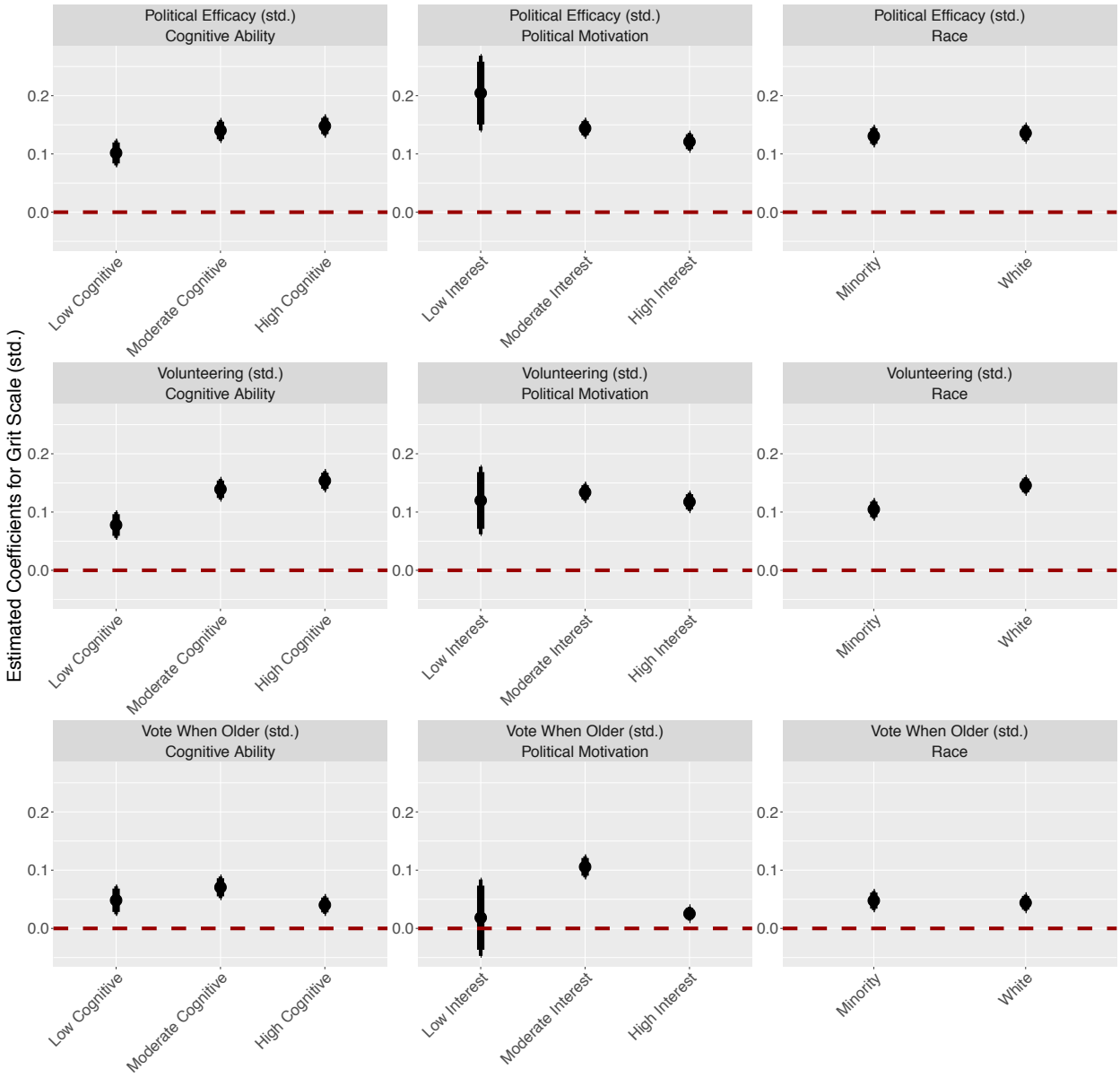
We first consider whether grit is a complement or substitute to cognitive ability. As can be seen, there appears to be some evidence that students with moderate and high levels of cognitive ability benefit most from developing grit in terms of their levels of political efficacy and volunteering. However, for voting when older, the benefits seem to mostly accrue for students in the middle ability category.

By political motivation, it seems that the lowest interest individuals benefit most from developing grit when it comes to their political efficacy development. For volunteering, the benefits seem to be similar by levels of motivation. When it comes to voting when older, the benefits are mostly accrued by students with moderate levels of political interest. The effects are almost 5 times as large than for individuals with minimal or high levels of political interest. This jives with our expectation of grit's role for voting. As we have hypothesized, it may take some minimal orientation towards participation before grit's effects kick in.

There do not seem to be large differences by race. Our grit estimates are similar for white and minority students when we consider political efficacy and voting when older. White students do tend to benefit a little bit more when it comes to volunteering. While this difference is statistically significant ($p < 0.001$), it is not larger substantively.

[5] Unfortunately, WCPSS cannot legally release student-level free-reduced price lunch data to researchers (internal to or external to the district). As is the case in most school districts, race is strongly correlated with socioeconomic status in the Wake County Schools.

Figure A.2 Grit Estimate Heterogeneities



Notes: Figure A.2 shows the heterogeneities in grit estimates for political efficacy (first row), volunteering (second row), and voting when older (third row). These are broken by cognitive ability (first column), political motivation (second column), and race (third column). Cognitive ability is proxied by lag reading scores, which are broken into terciles. Political motivation is proxied by the "care who is president" item, which is broken by the top and bottom categories representing low and high levels of interest and the two middle categories constituting moderate interest. Race comes from the school administrative records. Estimates correspond to a one standard deviation increase in grit. Outcomes are also standardized to make visual comparisons as easy as possible. Model controls the same as listed in Table 1 in the paper (excepting the corresponding heterogeneity variables).

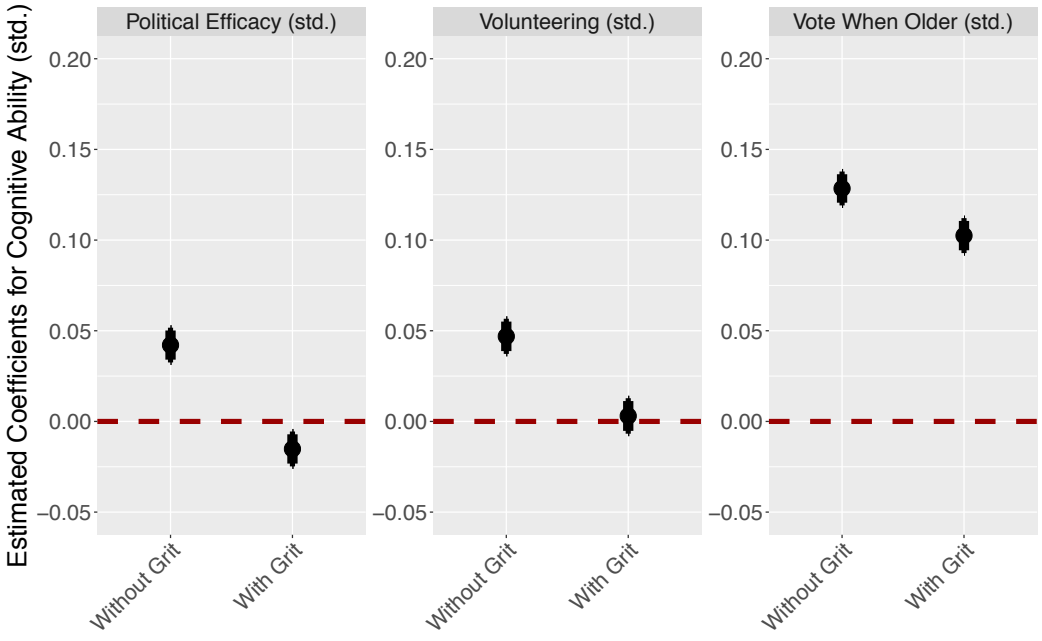
Cognitive Ability

As we argue in the paper, psychosocial or noncognitive skills play an underappreciated role in the development of youth civic attitudes and behaviors. Another way to see this is to see what happens to the estimates for cognitive ability—the common foil and predominant focus of the political science literature—once we begin to account for psychosocial ability.

Figure A.3 plots our unadjusted cognitive ability estimates. These are broken by whether the statistical models do (right estimates) or do not (left estimates) control for grit. As can be seen, when we do not take into consideration psychosocial or noncognitive ability, cognitive ability's role seems to be much larger than when we do. In the case of the development of volunteering, cognitive ability's estimate goes from being positive and statistically significant to being indistinguishable from 0. In terms of substantive size, the estimates are reduced by 94% when noncognitive skills are accounted for. A similar pattern comes in the political efficacy models (136% reduction) and vote when older (20% reduction) outcomes.

This illustrates that the role of cognitive ability in the development of civic attitudes and behaviors may be overstated. A good portion of the relationship identified between cognitive aptitude and civic engagement may actually be rightfully ascribed to the vitally important set of psychosocial or noncognitive skills individuals possess.

Figure A.3: The Effect of Controlling for Grit on Cognitive Ability Estimates



Notes: Figure A.3 displays the relationship between cognitive ability (factor scale constructed from lagged reading and math scores, standardized with grade) and our civic engagement outcomes (also standardized). Estimates are broken out by whether grit (also standardized by grade) is included.

Online Appendix References

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, *44*(5), 427-445.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of personality assessment*, *91*(2), 166-174.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel* (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.

Tucker-Drob, E.M., Briley, D.A., Engelhardt, L.E., Mann, F.D., & Harden, K.P. (Forthcoming). Genetically-mediated associations between measures of childhood character and academic achievement. *Journal of Personality and Social Psychology*.