Economics of Education Review 75 (2020) 101969

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Jack-of-all-subjects? The association between individual grade variance and educational attainment

Astrid Marie Jorde Sandsør

Nordic Institute for Studies in Innovation, Research and Education and Department of Economics, University of Oslo, Norway

ARTICLE INFO

Keywords:
Grades
Cognitive skills
Non-cognitive skills
Human capital

JEL classification:
I21
J24

ABSTRACT

This paper uses detailed register information on students in lower secondary school in Norway to study the importance of the second moment of individual grade distribution: grade variance. Students receive discrete-value grades from 1 to 6 in the same 13 subjects, and the grade point average (GPA) is used to determine entrance into upper secondary school. This leads to a limited number of possible GPA values and the within-GPA-value variation in grades is used to investigate the association between grade variance and educational attainment. Grade variance is found to be negatively associated with educational attainment across the grade distribution and for both genders. US data confirm this finding. Results suggests that being a generalist with similar skills across subjects predicts educational attainment and that educational institutions may benefit from considering more than just grade point average when making admission decisions.

1. Introduction

Cognitive skills are an important predictor for future outcomes for the individual, including education and labor market outcomes (Heckman, 1995; Herrnstein & Murray, 2010; Murnane, Willett, & Levy, 1995), and aggregate measures of cognitive skills are important for economic growth and development (Hanushek & Kimko, 2000; Hanushek & Woessmann, 2008). One measure of cognitive skills is student grades received in school, commonly measured as the grade point average. Grades are strongly correlated with short-term and long-term outcomes such as educational attainment and income and often have direct consequences for students, for example by contributing to the college admission decision and determining their post-education job qualifications. The grade point average captures the first moment of the individual grade distribution: the mean. This paper is the first to investigate the importance of the second moment of individual grade distribution: grade variance. For a given grade point average, which student might be expected to have the higher educational attainment: the student with high or low grade variance? Is it beneficial to be a specialist, particularly good at some subjects, or rather to be a generalist, a jack-of-all-subjects? Or does grade variance really measure something else, such as non-cognitive skills?

In Norway, students at the end of lower secondary education receive grades in the same 13 subjects ranging from 1 (lowest) to 6 (highest), and the grade point average is used to determine acceptance into upper secondary schools and programs. Students with the same grade point average have the same educational opportunities, but may have different grade variance. For instance, receiving two grades of value four results in the same grade point average as receiving two grades of values three and five or two and six. Thus the system is ideal for investigating the association between grade variance and educational outcomes in that it allows for the inclusion of grade point average fixed effects in the analysis. Also, grading practices are monitored by the central authorities, reducing potential measurement error.

This paper uses detailed Norwegian register data to investigate the association between grade variance and educational attainment. The data cover the entire population of students graduating from lower secondary education in Norway from 2002 to 2004 and include transcript data, educational attainment and socioeconomic characteristics. Grade variance is found to be negatively associated with graduating from upper secondary school and continuing on to higher education. Estimates are negative across the grading distribution and results are unaffected by including socioeconomic characteristics and school-by-cohort fixed effects. Heterogeneity analyses reveal that the negative association is stronger for girls than for boys. The association between grade variance and educational attainment is further investigated using US data from the National Longitudinal Survey of Youth, 1979 (NLSY79) where the same negative association is found. Data from the NLSY79 and Character Development in Adolescence Project (CDAP) allow for the inclusion of non-cognitive skills, but neither data source finds an association between grade variance and non-cognitive skills.

This paper suggests that it is beneficial to be a generalist, a jack-of-

E-mail address: astrid.sandsor@nifu.no.

https://doi.org/10.1016/j.econedurev.2020.101969
Received 19 February 2019; Received in revised form 30 January 2020; Accepted 30 January 2020

all-subjects, as students with lower grade variance have higher educational attainment. If institutions are interested in students with high ability and effort, but only use the grade point average in the admission decision, as is the case in Norway, they may not be accepting the best students. Students with low grade variance who are just below the grade point average cutoff are likely to outperform students with high grade variance just above the cutoff. My findings suggest that institutions might benefit from taking more than the grade point average into account when making admission decisions.

The paper is structured as follows : Section 2 presents a discussion of why grade variance might differ, Section 3 presents institutions and data while Section 4 presents the empirical strategy and results. Section 5 presents results using the US NLSY79 data, Section 6 introduces measures of cognitive and non-cognitive skills to investigate potential mechanisms and Section 7 presents conclusions.

2. Why might grade variance differ across individuals, and why might it be associated with educational attainment?

Standardized tests, such as the PISA and SAT tests,[1] are designed to determine a student's skills in a specific subject relative to all other students. School grades are a much more subjective measure, however. Grades are usually decided by the teacher of the subject, are not standardized across classes and schools and can be absolute measures or measured relative to classmates. They often measure a combination of knowledge in the subject (cognitive skills), and other skills such as attending and participating in class (non-cognitive skills).[2] In addition, the degree to which cognitive or non-cognitive abilities matter will depend on the subject.

One reason why we might observe differences in grade variance, is that students have different innate subject-specific skills. For a given grade point average, high grade variance students have both good and bad subject-specific skills while low grade variance students have more similar skills across subjects. We can refer to the first group as specialists and the second group as generalists. On the one hand, as higher education allows students to specialize in their preferred field, high variance students might be expected to have higher educational attainment. This might be especially true for students at the upper end of the grade distribution, as these students are more likely to go on to higher education. On the other hand, it might be beneficial to be a generalist for some studies or occupations. Lazear (2004) finds that individuals with balanced skills (jacks-of-all-trades) are more likely to become entrepreneurs. Rather than having a comparative advantage in a specific skill, entrepreneurs have a comparative advantage in having a range of skills, which is necessary to be successful as an entrepreneur. Being a jack-of-all-trades, or jack-of-all-subjects in this case, might be beneficial for the educational outcomes studied in this paper, as higher education is often based on general knowledge. This could particularly be true in the United States where there is a long tradition for a liberal arts education in four-year colleges and a specialist might see the benefit of a short and specialized education rather than a long and general education. Hanushek, Woessmann, and Zhang (2011) propose an alternative explanation, that it might be beneficial to be a generalist in the long run because generalists have greater adaptability. They study the impact of vocational versus general education and find that although individuals with vocational education have an early labor-market advantage due, for instance, to their higher employability, these gains are often offset by reduced adaptability later in life.

Another reason why we might observe differences in grade variance, is that students have different cognitive and non-cognitive skills that are important across subjects. Cognitive skills such as reading and writing or logical reasoning are likely important for many subjects, and having these skills could therefore reduce grade variability. Non-cognitive skills, such as perseverance and self-esteem, can be used to compensate for limited subject-specific skills. The degree to which these cognitive and non-cognitive skills matter will likely vary by subject. Falch, Nyhus, and Strøm (2014), for instance, use math and science grades in school as a proxy for cognitive skills, and grades in physical education, home economics, arts and crafts and music as a proxy for non-cognitive skills. A non-cognitive skill that might be particularly relevant to grade variance is loss aversion, i.e. "the impact of a difference on a dimension is generally greater when that difference is evaluated as a loss than when the same difference is evaluated as a gain" (Tversky and Kahneman, 1991, p. 1040). In this context, if a student views the "loss" from having a grade below average as greater than the "gain" from having a grade above average, then even though improving each grade would have the same impact on their grade point average, they would invest more effort in the subject with the below average grade, and therefore decrease their grade variance. Both cognitive and non-cognitive skills have been shown to be meaningful predictors of educational, labor market and behavioral outcomes (Borghans, Duckworth, Heckman, & Ter Weel, 2008; Carneiro, Crawford, & Goodman, 2007; Falch et al., 2014; Heckman, Stixrud, & Urzua, 2006; Kautz, Heckman, Diris, ter Weel, & Borghans, 2014), especially for the lower part of the distribution for non-cognitive skills (Lindqvist & Vestman, 2011). If higher cognitive or non-cognitive skills are associated with lower grade variance, then we would expect a negative association between grade variance and educational attainment.

Another potential mechanism is related to the big fish in a small pond effect (Marsh, 1987), where students wrongly believe they have high absolute ability when their ability is high relative to their classmates (high rank), and therefore invest more in human capital (see e.g. Elsner & Isphording, 2017; Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007; Murphy & Weinhardt, 2018). Elsner and Isphording (2017), for instance, study the impact of a student's rank in high school by comparing students with similar ability but different rank, and find that students with a higher rank are more likely to finish high school and attend college, and also have higher expectations about their future careers and higher perceived intelligence. Students with high grade variance are more likely to be big fish in some subjects (high grades and therefore high rank), but are also more likely to be small fish in other subjects (low grades and low rank), whereas students with low grade variance have a more stable fish size. It is not clear how varying rank across subjects relates to educational attainment.

It might also be the case that for a given grade point average, the association between grade variance and educational attainment differs across subgroups. A common finding, for instance, is that while average skill differences between boys and girls tend to be small, skills variance is higher for boys than for girls.[3] This is known as the greater male

[1] The Programme for International Student Assessment (PISA) is a standardized test carried out every three years among a representative sample of 15-year-olds, and measures their ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges. Around 540,000 students in a total of 72 countries participated in PISA in 2015 (OECD, 2019). The SAT is a standardized test developed to test students' academic readiness for college. The SAT, along with the ACT, forms a large part of the admission decision for many colleges (ACT, 2015; SAT, 2015).

[2] Non-cognitive skills are referred to as soft skills, personality traits, non-cognitive skills, non-cognitive abilities or character and socio-emotional skills, among others. Heckman and Kautz (2013) refer to them as character skills, rather than traits, as they are constant at any age but may change over time. Character skills include "conscientiousness, perseverance (grit), self-control, trust, attentiveness, self-esteem, self-efficacy, resilience to adversity, openness to experience, empathy, humility, tolerance of diverse opinions and the ability to engage productively in society" (Heckman and Kautz, 2013, p. 6).

[3] Hedges and Nowell (1995) study six representative large scale surveys of data on mental abilities and find that although average gender differences are generally small, males consistently have a larger variance in their test scores. Similar results were found in a recent meta-study comparing gender differences

variability hypothesis (Johnson, Carothers, & Deary, 2008; Shields, 1982). Although variance across individuals is higher among boys than girls, there is no reason to believe that individual variance is higher for boys than for girls. And even if individual grade variance is higher for one gender, it does not necessarily mean that the association between grade variance and education attainment, conditional on grade point average, varies by gender. If such differences exist, for gender or for other subgroups, they might be explained by differences in the potential mechanisms described above.

As there are explanations for both positive and negative associations between grade variance and educational attainment, the question becomes an empirical one. Is there an association between grade variance and educational attainment, and if so, in what direction and how large? And do further analyses rule out or confirm potential mechanisms?

3. Institutions and data

3.1. The educational system

In Norway, compulsory education consists of primary education (grades 1–7) and lower secondary education (grades 8–10). Entry into primary and lower secondary education is determined by catchment areas and there is no possibility of failing a class, implying that all students finish compulsory education the year they turn 16.[4] There is no tracking, a common national curriculum for all students, and very few private schools.[5]

In lower secondary education, students receive grades for the same 13 subjects: oral and written Norwegian,[6] oral and written English, mathematics, natural science, social science, religion, home economics, music, physical education and arts and crafts. Teachers teach a subset of subjects and in lower secondary school students are commonly exposed to different teachers in different subjects. Students receive grades with integer values of between 1 (lowest) and 6 (highest) from their teachers every semester, primarily based on their performance in the subject. These grades have no consequences for the students prior to grade 10. The teacher-assessed grades received in the last semester of grade 10, along with 2–3 externally graded oral or written exams, are used to determine acceptance for upper secondary education.

When applying for upper secondary education, students rank their preferred study programs and the schools offering that study program. All students have been guaranteed admission to upper secondary education since 1994, but whereas acceptance to one of their three ranked choices is guaranteed, the grade point average determines which school and study program the student is accepted for. How important grades are for entering the school or study program of their choice will vary from county to county, as counties are free to determine how acceptance into upper secondary education is organized (Haraldsvik, 2003). In upper secondary education, academic programs have a duration of 3 years and qualify students for higher education, while vocational programs typically last for four years, including two years of apprenticeship training. Subject requirements differ depending on the study program and there are both mandatory and elective subjects. If students from vocational programs wish to continue on to higher education, they can attend a year of supplementary studies to obtain an academic program degree that qualifies them for higher education. The application system to higher education is centralized for the entire country and is based almost entirely on grades from upper secondary school.

3.2. Data

This paper uses register data provided by Statistics Norway for all individuals leaving lower secondary education in the period 2002–2004. The data make it possible to combine detailed information on an individual's background and education, including grades, measures of educational attainment and socioeconomic characteristics.

Grade point average (GPA) is measured as the unweighted mean of all 13 teacher-assessed grades received on leaving lower secondary education. Grade variance is measured as the standard deviation of an individual's grades (GSD), using the same grades as were used to calculate the individual's grade point average. Descriptive statistics are presented in panel A of Table 1. The sample is restricted to students graduating from lower secondary education at the age of 16 with information on the lower secondary school they attended. The sample is restricted to students with 13 valid teacher-assessed grades, and is thereby reduced by 12%. The data reduction is presented in Table A1 in the online appendix.

Fig. 1 a and b display GPA and GSD distribution, with the dashed and dotted lines displaying kernel densities with a bandwidth of 0.15 for girls and boys, respectively. The average GPA is higher for girls (4.23) than for boys (3.85) while the spread is slightly higher for boys (the standard deviation of the GPA is 0.74 for girls and 0.78 for boys). These are both common findings in the literature (Herrnstein & Murray, 2010). The average GSD is higher for boys (0.68) than for girls (0.64), while the spread in GSD is the same (standard deviation of GSD is 0.19 for girls and 0.18 for boys).

Fig. 2 displays the distribution of grades for the whole sample. The most common grade is 4 (34%), while the least common grade is 1 (0.6%).

Outcome variables are measures of educational attainment: (1) Started academic program (Started ACA), (2) Graduated upper secondary (Grad UPE) (3) GPA upper secondary (GPA UPE) and (4) started higher education (Started HE). *Started academic program* is an indicator variable equal to one if the student started an academic study program in the first year of upper secondary education. 98% of students in the sample go on to upper secondary education in the fall after completing lower secondary education, with 48% starting an academic program and 50% starting a vocational program. *Graduated upper secondary* is an indicator variable equal to one if the student graduates from upper secondary school within five years. Students have a legal right to five years of upper secondary education and this is the standard measure of upper secondary education completion used by the authorities. 73% of students in the sample graduate from upper secondary education within five years. *GPA upper secondary* is measured as the unweighted mean of all teacher-assessed grades on the upper secondary education transcript. The measure only covers students who complete an academic program or who transfer from a vocational to an academic program. GPA upper secondary has a mean of 4.15 and a standard deviation of 0.68. The last measure, *started higher education*, is an indicator variable equal to one if a student started, but did not necessarily complete, a higher education program before 2012. In the complete sample, 55% started higher education. In the sample with students that graduated from the academic program, 88% started higher education. Descriptive statistics for outcome variables are presented in panel B of Table 1.

Socioeconomic characteristics consist of gender, birth month, immigration status, parental employment status and parental education. Immigration status is divided into two categories, the first indicating that a student is a first generation immigrant born abroad with parents born abroad and the second that the student is a second-generation immigrant, born in Norway but with both parents born abroad. Parental

(*footnote continued*)
in the academic grades of over 1.6 million students (O'Dea, Lagisz, Jennions, & Nakagawa, 2018).

[4] In a very few cases, students do not start primary education at the expected age. If a child is not considered to be mature enough, the parents together with the school and psychologists can postpone enrollment by one year. In addition, some older students return to improve their grades, and immigrants are often over-aged at graduation.

[5] Only 3.5% of students attended a private elementary or lower secondary school in 2015. (The Norwegian Directorate for Education & Training, 2015).

[6] Norwegian has two written languages and students therefore have two grades in written Norwegian.

Table 1
Descriptive statistics.

	Total		Boys		Girls	
	mean	(sd)	mean	(sd)	mean	(sd)
A. Transcript data						
Grade Point Average (GPA)	4.04	(0.79)	3.85	(0.78)	4.23	(0.74)
Grade Standard Deviation (GSD)	0.66	(0.19)	0.68	(0.18)	0.64	(0.19)
B. Outcome Variables						
Started academic program (ACA)	0.48		0.44		0.52	
Graduated upper secondary (Grad UPE)	0.73		0.69		0.77	
GPA upper secondary (GPA UPE)	4.16		4.08		4.22	
Started higher education						
- complete sample (Started HE)	0.55		0.45		0.65	
- academic program (Started HE (2))	0.88		0.88		0.88	
C. Socioeconomic characteristics						
Girl	0.50					
Birth month	6.40	(3.36)	6.37	(3.36)	6.43	(3.37)
First generation immigrant	0.012		0.011		0.013	
Second generation immigrant	0.0073		0.0068		0.0078	
Parental education: less than upper secondary	0.092		0.089		0.094	
Parental education: upper secondary	0.48		0.47		0.48	
Parental education: bachelor	0.30		0.30		0.30	
Parental education: master +	0.11		0.11		0.11	
Parental education: unknown	0.026		0.025		0.027	
Only mother working	0.12		0.12		0.13	
Only father working	0.12		0.12		0.12	
Both parents working	0.71		0.71		0.71	
No parent working	0.049		0.045		0.052	

Note: N = 142,257 with 71,010 boys and 71,247 girls. For vocational program graduate and academic program graduate, N = 70,954 (38,281 boys and 32,673 girls) and N = 67,751 (31,009 boys and 36,742 girls), respectively. For GPA upper secondary education and started higher education - academic program, N = 78,909 (31,077 boys and 47,832 girls).

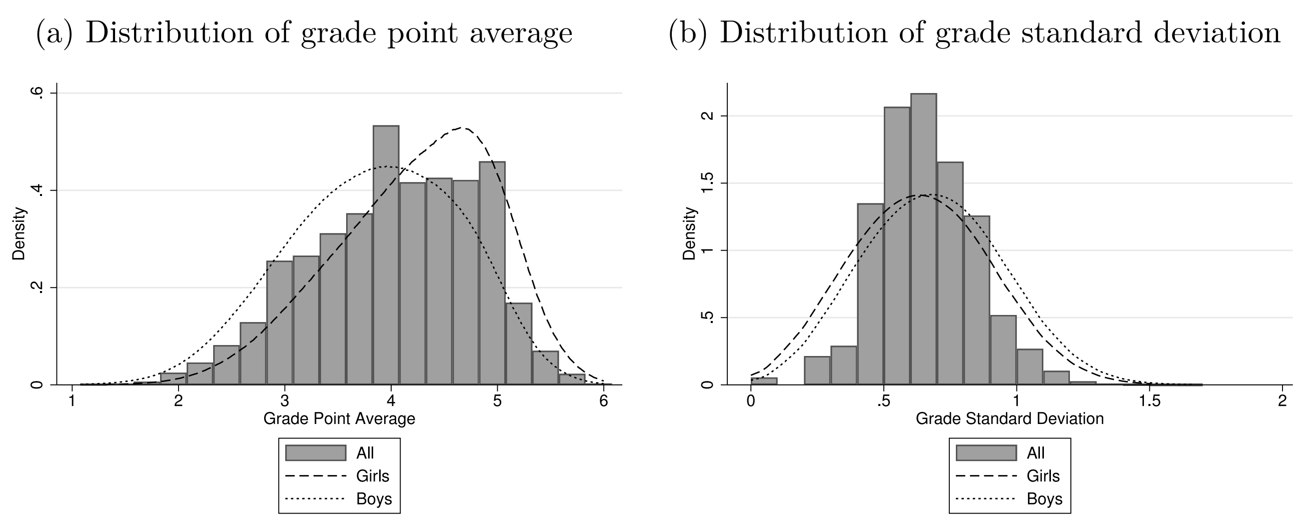

Fig. 1. Distribution of grade point average and grade standard deviation Note: For grade point average, each bin has a width of 0.25, while each bin has a width of 0.1 for grade standard deviation. Lines display kernel densities with bandwidth 0.20 for each variable for girls (dashed) and boys (dotted).

education is measured as the highest completed education of one of the parents: less than upper secondary, upper secondary, a Bachelor's degree, a Master's degree or higher, and having an unknown education. Parental employment status is an indicator of whether only the mother, only the father, both parents or no parents are working. Variables are measured the year the student turns 16. Descriptive statistics for socioeconomic characteristics are presented in panel C of Table 1. The last columns of Table 1 present descriptive statistics for girls and boys separately. Boys are less likely to start an academic program, have lower GPA and higher GSD in upper secondary education, are less likely to complete upper secondary education and less likely to start higher education.

4. Empirical strategy and results

4.1. Empirical strategy

A unique feature of the Norwegian grading system allows us to include grade point average fixed effects. As grades in Norway can only take on integer values of from 1 to 6, and all students receive grades in the same 13 subjects, students are bunched at certain values of GPA. When calculating the grade point average, receiving two grades of 4 result in the same grade point average as receiving grades of 3 and 5 or 2 and 6. This means that even if students have exactly the same GPA, they can have different values of GSD. This feature not only makes it possible to investigate whether there are heterogeneous results across

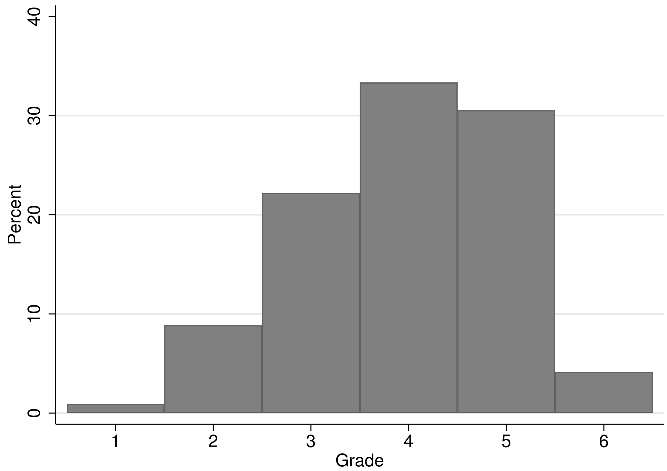

Fig. 2. Distribution of grade values Note: 1,849,341 grades ranging from 1 (lowest) to 6 (highest) for 142,257 students leaving lower secondary education 2002–04.

the grade distribution; it also makes it possible to eliminate any concern that the coefficient for grade standard deviation is the result of a mechanical correlation. Our main specification includes all values of GPA in the sample, 65 indicator variables, as controls. Fig. 3a shows the spread in GSD for separate values of GPA, which is the variation used in the estimation.

The association between grade standard deviation (GSD) and educational attainment is estimated using the following model:

$$y_{istr} = \gamma GSD_{istr} + X_{istr}^{'}\beta + \theta_s \times \delta_t + \alpha_r + \epsilon_{istr} \qquad (1)$$

where y_{ist} is the outcome for student i from school s in year t with grade point average r. GSD_{istr} is grade standard deviation from lower secondary education and is standardized with mean 0 and standard deviation 1. $\theta_s \times \delta_t$ is school by cohort fixed effects and α_r is GPA fixed effects. X_{istr} is a vector of socioeconomic characteristics consisting of gender, immigrant status, parental education, parental employment status and birth month (socioeconomic characteristics are listed in Table 1). The error term ϵ_{ist} is clustered at the school level. The coefficient of interest, γ, can be interpreted as the association between grade standard deviation and educational attainment, conditional on socioeconomic characteristics, school by cohort fixed effects and GPA fixed effects.

4.2. Results

4.2.1. Main results

Table 2 reports the main results, where the outcome is whether the student has started higher education (columns (1)-(4)). Across all specifications, grade standard deviation is negatively associated with starting higher education. Column (1) includes GPA and cohort fixed effects and the estimate between GSD and starting higher education is negative. Column (2) adds socioeconomic characteristics, and the estimate does not change much, suggesting that background characteristics cannot explain the negative association. Column (3) adds school fixed effects while column (4) adds school by cohort fixed effects. The estimate is slightly reduced for both specifications, indicating that the results are not explained by school characteristics. Column (4) is the preferred specification and corresponds to Eq. (1). A one standard deviation increase in GSD decreases the likelihood of a student starting higher education by 1.3 percentage points. This is equivalent to 0.026 of a standard deviation decrease in the likelihood of starting higher education.

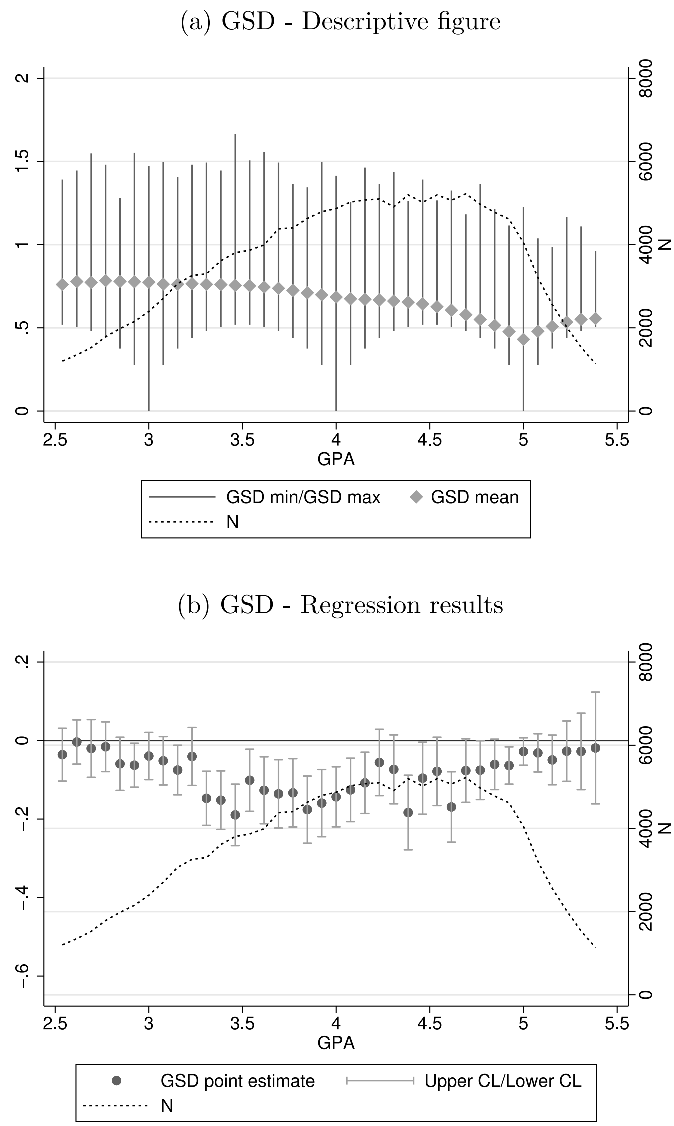

Fig. 3. Started higher education - separate regressions for each GPA value Note: GSD is standardized for the entire sample with mean 0 and standard deviation 1. GPA corresponds to the 39 GPA values where there are at least 1000 observations. a: Dots indicate mean values while bars indicate the minimum and maximum GSD values for each GPA value. Estimates in b are based on the following regression: $y_{ist} = \gamma GSD_{ist} + X_{ist}^{'}\beta + \delta_t + \epsilon_{ist}$, with one regression for each GPA value. Regressions include socioeconomic characteristics and cohort fixed effects and standard errors are clustered at the school level. Socioeconomic characteristics include gender, birth month, immigration status, parental employment status and parental education. Dots indicate the coefficient for each regression while the bars indicate the 95% confidence interval. The dotted line in both figures indicates the number of students at each GPA value. The outcome variable for regressions in b is started higher education, which is an indicator variable equal to 1 if the student started higher education before 2012.

4.2.2. Other outcomes

Our rich data allow us to investigate other outcome measures related to educational attainment. Columns (5)-(8) in Table 2 display the results of estimating Eq. (1) for all educational attainment measures described in Section 3.2. Column (8) is equivalent to column (4) of Table 2, except that only students graduating from academic programs are included. Estimates show that GSD is negatively associated with graduating from upper secondary and starting higher education for the subsample of academic program graduates. A one standard deviation increase in GSD decreases the likelihood of graduation from upper

Table 2
Main results.

	(1) Started HE	(2) Started HE	(3) Started HE	(4) Started HE	(5) Started ACA	(6) Grad USE	(7) GPA USE	(8) Started HE(2)
GSD	−0.019***	−0.017***	−0.013***	−0.013***	−0.001	−0.015***	0.005	−0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Soc. char	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	No	No	No	No	No
School FE	No	No	Yes	No	No	No	No	No
Schoolxcohort FE	No	No	No	Yes	Yes	Yes	Yes	Yes
GPA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.392	0.416	0.404	0.407	0.306	0.294	0.567	0.119
N	142,257	142,257	142,257	142,257	142,257	142,257	78,909	78,909
Number of groups			1207	3345	3345	3345	3166	3166
Mean outcome	0.55	0.55	0.55	0.55	0.48	0.73	4.16	0.88

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Note: *Started higher education* (columns (1)-(4)) is an indicator variable equal to 1 if the student started higher education before 2012. Estimations are based on Eq. (1). The first column includes cohort fixed effects, the second column adds socioeconomic characteristics and the third column adds school fixed effects. Socioeconomic characteristics include gender, birth month, immigration status, parental employment status and parental education. The remaining columns include socioeconomic characteristics and school by cohort fixed effects. *Started ACA* (column (5)) is an indicator variable for whether the student goes on to an academic program in upper secondary school. *Grad USE* (column (6)) is an indicator variable for graduating from upper secondary within 5 years. *GPA USE*(column (7)) is GPA from upper secondary for students who have graduated from an academic program in upper secondary. This includes students who have transferred from a vocational program during upper secondary school. *Started HE (2)* (column (8)) is an indicator variable for whether the student started higher education before 2012 and is the same sample as GPA USE. Standard errors are clustered at the school level.

secondary by 1.5 percentage points and decreases the likelihood of starting higher education for the subsample of academic program graduates by 0.7 percentage point. However, grade variance is not associated with starting an academic program or GPA in upper secondary. Students with higher grade variance are less likely to graduate from upper secondary school and start higher education, but the association is not explained by a decreased likelihood of starting an academic program or reduced grade point average in upper secondary school.

While GSD negatively predicts most of the outcomes in Table 2 conditional on GPA, the magnitudes of the relationships are smaller than the corresponding relationships with GPA, which is unsurprising. Table A2 shows this using a model that includes flexible controls for GPA instead of GPA fixed effects. In column (3), the GSD coefficient is about 1/10 of the GPA coefficient. The coefficients imply that a 1 standard deviation increase in the GPA increases the likelihood of starting higher education by 26 percentage points, whereas a 1 standard deviation increase in the GSD decreases the likelihood of starting education by 2.2 percentage points.

Tables A3 and A4 in the online appendix further investigate upper secondary education outcomes related to specialization. The academic track consists of three different specialization tracks, general (82%), music, dance and drama (7%) and physical education (11%). Table A3 investigates whether grade variance is associated with starting a general or a non-general specialization, as well as the association between educational attainment and GSD for each subsample. If high grade variance is associated with being a specialist, we would expect the GSD to be positively associated with starting the non-general specialization. This is indeed found to be the case, as the GSD is positively associated with starting a non-general specialization (0.012). When investigating the association between educational attainment and GSD for each subsample, a negative coefficient is found for both (−0.018 and −0.029 for general and non-general, respectively). There is still a negative association between the GSD and educational attainment even when choice of specialization is taken into account.

Table A4 investigates upper secondary STEM subject choice. The general specialization track consists of many elective subjects, and 49% of the sample take a voluntary math or science course. Finding a positive association between GSD and STEM choice could be due to STEM specialists (high grade variance and high grades in STEM subjects from lower secondary education) choosing STEM subjects. However, there is a negative relationship between GSD and the choice of math or science (columns 1 and 2), also when one controls for STEM grades from lower education (columns 3 and 4). Taking this choice into account by including controls for math/science courses in upper secondary reduces the association between GSD and starting higher education (columns 6 and 7). Students with a higher GSD are less likely to choose STEM subjects. STEM is strongly correlated with going on to higher education, and this does explain part (though not all) of the conditional correlation between GSD and higher education. The interpretation of the results is not straightforward, however. The centralized intake system to higher education places extra weight on STEM subject grades in order to incentivize students to take them and the subjects are required for certain studies such as engineering and medicine. Thus, any student considering higher education is more likely to take these subjects. It is unclear whether low variance students choose these subjects because they prefer them to other subjects, or because they know they want to go on to higher education.

4.2.3. Grade distribution

The limited number of GPA values makes it possible to study the association between grade standard deviation and educational attainment across the grade distribution by running separate regressions for each GPA value. We restrict the analysis to GPA values where at least 1000 students have the same GPA, resulting in 39 unique values. Fig. 3a displays the mean, minimum and maximum GSD value for each of these 39 values. The dotted line shows the number of students in each regression. There are more than 4000 students for each GPA value between 3.5 and 5. There is a GSD spread for each GPA value, which is the variation used to identify how the GSD is associated with educational attainment.

A separate regression is run for each of these values, with starting higher education as the outcome measure, and results are reported in Fig. 3b. Each regression includes socioeconomic characteristics and cohort fixed effects. The dots are the point estimates, the bars indicate the 95% confidence band and the dotted line shows the number of students in each regression. The point estimates are always negative and are significantly below zero at the middle of the grade distribution, while they are typically not significantly different from zero at the lower and higher ends of the grade distribution. This finding strongly supports the main finding that grade standard deviation is negatively

Table 3
Started higher education - interaction analysis of background characteristics.

Outcome Interacted background char.	(1) Started HE Gender	(2) Started HE Immigrant status	(3) Started HE Parental employment	(4) Started HE Parental education
GSD	−0.008*** (0.002)	−0.015*** (0.001)	−0.015*** (0.001)	−0.015*** (0.002)
GSD*girl	−0.012*** (0.002)			
GSD*immigrant		0.002 (0.008)		
GSD*(only mother working)			0.006 (0.004)	
GSD*(only father working)			−0.001 (0.004)	
GSD*(no parent working)			0.008 (0.005)	
GSD*(par. ed. < upper secondary)				0.014*** (0.004)
GSD*(par. ed. bachelor)				−0.000 (0.003)
GSD*(par. ed. master +)				0.002 (0.004)
GSD*(par. ed. unknown)				0.007 (0.007)
SchoolxCohort FE	Yes	Yes	Yes	Yes
GPA FE	Yes	Yes	Yes	Yes
R-squared	0.394	0.392	0.392	0.405
N	142,257	142,257	142,257	142,257
Number of groups	3345	3345	3345	3345
Mean outcome	0.55	0.55	0.55	0.55

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Note: The outcome variable is *started higher education*- an indicator variable equal to 1 if the student started higher education before 2012. All columns report estimates based on the following regression: $y_{istr} = \gamma_1 GSD_{itr} + \gamma_2 GSD_{itr} \times C_i + \gamma_3 C_i + \delta_t + \delta_t \times C_i + \alpha_r + \alpha_r \times C_i + \theta_s \times \delta_t + \epsilon_{itr}$. Regressions include GSD, the background characteristic, C_i and the background characteristic interacted with GSD and school by cohort fixed effects. In addition, the regressions include cohort fixed effects and GPA fixed effects, and both interacted with the background characteristic, making the results comparable to running a regression with cohort and GPA fixed effects separately on each subsample (see Table A6 in the online appendix). Socioeconomic characteristics are not included as control variables. Standard errors are clustered at school level. In column (2), *immigrant* consists of first and second generation immigrants. In column (3), the excluded category is both parents working and in column (4) the excluded category is parental education upper secondary.

associated with educational outcomes and that this is not solely due to a mechanical correlation between the two variables. We also find no evidence of the association differing across the grade distribution, as point estimates are consistently negative.[7]

The point estimates are larger, and more often significant, in the middle of the distribution than in the tails. If we believe that student performance in the middle of the distribution largely depends on effort, whereas students at the lower and upper ends of the distribution have lower grade variance, then this could be consistent with low variance students allocating effort more evenly across subjects, possibly motivated by loss aversion, as discussed in Section 2. However, negative estimates are found when N is high, and this finding is not as clear when we see the results from Table A5 in the online appendix, where the sample is divided into quartiles, and the number of observations is therefore similar for each sample. Estimates are fairly stable and always significantly below zero across specifications. The result that high and low ability students have low grade variance does not seem to be the clear takeaway from Table A2 in the online appendix, which includes flexible controls for GPA instead of GPA fixed effects.

4.2.4. *Heterogeneity*

The next step is to investigate whether the association between GSD and educational attainment is heterogeneous for different subgroups of the population. Column 1 of Table 3 displays the results when the GSD variable is interacted with gender. The negative association between GSD and educational attainment is significantly larger in absolute value for girls (−0.020) than for boys (−0.008). Girls have a 1.2 percentage point lower likelihood of starting higher education than boys when their GSD increases by one standard deviation. Table 3 also displays results when the GSD variable is interacted with immigrant status, parental employment status and parental education. There is a significant difference between students whose parents have less than upper secondary education and parents who have upper secondary education, with no negative association between GSD and starting higher education for the former group. The remaining interaction coefficients are not significantly different from the excluded groups, suggesting that the negative association is stable across subgroups (see Table A6 in the online appendix for separate regressions for each subsample).

4.2.5. *Robustness*

One possible concern is that the results may be driven by certain subjects. If students with high grade variance are more likely to have a low grade in mathematics, for instance, we might be picking up that mathematics, rather than grade variance, predicts starting higher education.[8] To investigate this further, Table 4 shows the results from

[7] Table A5 in the online appendix shows a more aggregated version of this exercise where separate regressions are run for observations below and above the median grade point average (columns (1) and (2)), and for each quartile of the grade point average (columns (3)-(6)). The coefficient is negative and strongly significant across all specifications. The strongest relationship between GSD and educational attainment is for the second quartile of the grade distribution, although the estimates are fairly stable for all quartiles.

[8] Table A7 in the online appendix shows descriptive statistics for each subject and the results of running a regression of the following model, $y_{ist} = \gamma + \sum_{k=1}^{13} \alpha^k S_{kist} + X_{ist}^{'}\beta + \theta_s \times \delta_t + \epsilon_{ist}$, where S is the grade for each

Table 4
Started higher education - removing subjects.

Removed subject(s)	γ	Removed subject(s)	γ	Removed subject(s)	γ
Oral English	−0.014*** (0.001)	Music	−0.011*** (0.001)	Physical education	−0.012*** (0.001)
Written English	−0.015*** (0.001)	Arts and crafts	−0.009*** (0.001)	STEM	−0.011*** (0.001)
Written Norwegian 1	−0.014*** (0.001)	Home economics	−0.009*** (0.001)	Non-academic	−0.004*** (0.001)
Oral Norwegian	−0.013*** (0.001)	Religion	−0.013*** (0.001)	Norwegian	−0.017*** (0.001)
Written Norwegian 2	−0.016*** (0.001)	Mathematics	−0.011*** (0.001)	Other	−0.022*** (0.001)
Social science	−0.015*** (0.001)	Natural science	−0.014*** (0.001)		
Socioeconomic characteristics			Yes		
Schoolxcohort FE			Yes		
GPA FE			Yes		
R-squared			0.393–0.408		
N			142,257		
Number of groups			3345		
Mean outcome			0.55		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Note: The outcome variable is *started higher education*- an indicator variable equal to 1 if the student started higher education before 2012. For each regression, the grade standard deviation and grade point average are calculated using a subsample of grades, removing either one subject at a time or groups of subjects. Each regression corresponds to Eq. (1) using the subsample from which a subject or group of subjects has been removed. STEM removes mathematics and natural science, Non-academic removes music, arts and crafts, home economics and physical education, Norwegian removes the three Norwegian grades and Other removes the remaining subjects: English, social science and religion. All specifications include socioeconomic characteristics, school by cohort fixed effects and GPA fixed effects. Socioeconomic characteristics include gender, birth month, immigration status, parental employment status and parental education. Standard errors are clustered at school level.

running the main specification, column (3) of Table 2, with subjects removed one by one when calculating GSD. Each new GSD is standardized with mean 0 and standard deviation 1. All coefficients remain significant and point estimates are between −0.009 and −0.016, and there is no evidence that certain subjects are driving the results. Similarly, the last four specifications remove groups of subjects. STEM removes mathematics and natural science, Non-academic removes arts and crafts, home economics, music and physical education, Norwegian removes the three grades in Norwegian (two written and one oral) and Other removes the remaining subjects; English (oral and written), social science and religion. Point estimates again remain significant although it is worth noting that the estimate is smaller when "non-academic" subjects are excluded.[9] Non-academic subjects are the subjects with the highest average grade (see Table A7 in the online appendix), and removing them from the sample reduces the GPA from 4.04 to 3.92 and reduces the GSD from 0.66 to 0.58. This suggests that the results could be driven by students who have high variance because they are good at non-academic subjects as compared to academic subjects. To investigate this further, an indicator variable is created for whether average grades in non-academic subjects are higher than average grades in all other subjects (non-academic students).[10] Including the non-academic student variable in the main analysis (or excluding non-academic students from the main analysis) reduces the estimate from −0.012 to −0.008. Results are partly driven by non-academic students, but the estimate is still negative and significant when this is taken into account.

4.2.6. Summary of results

The results lead to the conclusion that grade variance is negatively associated with educational attainment. Grade point average fixed effects rule out that the possibility that the association is due to mechanical correlation. We can also rule out the possibility that the negative association is due to individual subjects or groups of subjects. When a regression is run for each GPA value, the negative association is found across the grading distribution. Students with both high and low GPAs are less likely to start higher education if they have higher grade variance. Investigating heterogeneity reveals little or no difference related to immigrant background, parental employment or parental education. However, there is a significantly larger negative association for girls. Girls have a 1.2 percentage point lower likelihood of starting higher education than boys when their GSD increases by one standard deviation.

The results support the hypothesis that it is beneficial to be a generalist, rather than a specialist, with similar skills across subjects. The results also support the hypothesis that higher grade variance is associated with lower non-cognitive skills, as we know that non-cognitive skills are important for educational attainment (Lindqvist & Vestman, 2011). Falch et al. (2014) show that the association between non-cognitive skills and educational attainment is higher for girls than for boys, which could explain why there is a stronger negative association between GSD and educational attainment for girls. To investigate the second hypothesis further, we now turn to the US data, which include non-cognitive measures along with measures of GPA and GSD, as good measures of non-cognitive skills are not available in the Norwegian data.

5. Grade variance using US data

This section uses US data from the National Longitudinal Survey of Youth, 1979 (NLSY79) to investigate whether the association between grade variance and educational attainment is similar in a different educational context. The NLSY79, unlike Norwegian register data, also includes measures of cognitive and non-cognitive skills and can therefore be used to investigate further potential mechanisms (Section 6). There are, however, a few differences that need to be taken into account

(*footnote continued*)
subject *k*. Coefficients for social science, religion, mathematics and natural science (ranging from 0.036 to 0.050) are higher than coefficients for the other subjects (ranging from −0.011 to 0.027) indicating that there is variation across subjects in predicting starting higher education.

[9] These estimates are not directly comparable as they are standardized for the subsample of grades included in the analysis.

[10] The difference between the average GPA for non-academic subjects and the average GPA for all other subjects is a variable with a mean of 0.4, standard deviation Of 0.55, minimum value of −2.5 and maximum value of 3.3. The indicator variable created is equal to 1 if the difference is greater than 1, which is the case for14% of students.

Table 5
NLSY79 - descriptive statistics.

	Total		Boys		Girls	
	mean	(sd)	mean	(sd)	mean	(sd)
A. Transcript data						
Grade point average (GPA)	2.49	(0.81)	2.34	(0.81)	2.63	(0.78)
Grade standard deviation (GSD)	0.84	(0.25)	0.88	(0.25)	0.81	(0.24)
Number of grades	25.8	(6.28)	25.7	(6.41)	26.0	(6.16)
B. Outcome variable						
Years of education	13.6	(2.23)	13.6	(2.34)	13.6	(2.12)
C. Socioeconomic characteristics						
Girl	0.51		0		1	
Black	0.11		0.10		0.11	
Hispanic	0.060		0.061		0.060	
Living in South	0.30		0.29		0.32	
Living in urban area	0.76		0.76		0.76	
Broken home	0.21		0.21		0.22	
Number of siblings	3.19	(2.14)	3.13	(2.14)	3.24	(2.14)
Month of birth	6.47	(3.37)	6.52	(3.41)	6.42	(3.34)
Family income 1979 (thousands)	17.2	(15.1)	17.8	(15.2)	16.6	(14.9)
Mother: Years of education	11.4	(3.45)	11.3	(3.60)	11.4	(3.30)
Father: Years of education	11.2	(4.61)	11.3	(4.68)	11.1	(4.55)
D. Cognitive skills						
Arithmetic reasoning (ASVAB 1)	18.3	(7.19)	19.3	(7.33)	17.3	(6.91)
Word knowledge (ASVAB 2)	26.5	(7.07)	26.4	(7.33)	26.6	(6.81)
Paragraph comprehension (ASVAB 3)	11.2	(3.15)	10.9	(3.33)	11.5	(2.94)
Mathematical knowledge (ASVAB 4)	46.7	(15.2)	42.9	(14.8)	50.4	(14.6)
Coding speed (ASVAB 5)	14.1	(6.31)	14.5	(6.50)	13.9	(6.10)
Cognitive	0	(1.00)	−0.044	(1.05)	0.042	(0.94)
E. Non-cognitive skills						
Rotter locus of control scale	7.57	(2.38)	7.62	(2.37)	7.53	(2.39)
Rosenberg self-esteem scale	22.7	(4.05)	22.9	(3.95)	22.5	(4.13)
Non-cognitive	0	(1.00)	0.043	(0.98)	-0.042	(1.02)

Note: The analysis is restricted to the $N = 4{,}136$ students from the main sample with non-missing transcript data, non-missing educational outcome, 10 or more valid grades and non-missing cognitive and non-cognitive skill measures. The sample consists of 2110 girls and 2026 boys.

when interpreting the results. The Norwegian register data covers the entire population while the NLSY79 is a longitudinal survey of a nationally representative sample of young Americans with fewer observations. In the Norwegian setting, all grades are received in the last semester of lower secondary school and are based on the past semester's performance, whereas the US data include grades received at various points during upper secondary education, grades 9 to 12. Whereas all Norwegian students take the same course bundle in the data analyzed, which allows for the inclusion of GPA fixed effects, students can select which courses they take in the US data. Some students might opt out of difficult courses and have higher and more even grades as a result, a choice that is likely tied to educational and career aspirations. Since students receive grades at different times, students might be affected by disruptive shocks at household level (e.g. divorce, parental) or at individual level (student sickness) during their education which could affect their performance in one or more subjects and likely increase grade variance. If this is the case, then low grade variance in the US data might proxy for stable health or a stable family environment, which we know to be important for educational outcomes. Taken together, this suggests that the results from the NLSY79 need to be interpreted with more caution.

5.1. Data

The NLSY79 includes high school transcript data, educational attainment and socioeconomic characteristics (see Section B of the online appendix for a more detailed description of the data). The grade point average (GPA) is measured as the unweighted mean of all grades received in all years of high school (grades 9–12), and is restricted to students with at least 10 valid grades. Grade variance is measured as the standard deviation of an individual's grades (GSD), using the same grades as were used to calculate the individual's grade point average. Both measures are standardized with mean 0 and standard deviation 1 to facilitate interpretation. Socioeconomic characteristics consists of number of siblings, father's highest completed grade, mother's highest completed grade and family income in 1979 as well as a dummy for a broken home at age 14, a dummy for living in the South at age 14 and a dummy for living in an urban area at age 14, and race and ethnicity dummies. The measures of socioeconomic characteristics correspond to

Table 6
NLSY79: main results.

Sample	(1) Total	(2) Total	(3) Q1 + Q2	(4) Q3 + Q4	(5) Q1	(6) Q2	(7) Q3	(8) Q4	(9) Boys	(10) Girls
GSD	−0.166***	−0.162***	−0.115**	−0.240**	−0.152**	−0.076	−0.223*	−0.242	−0.139**	−0.157**
	(0.039)	(0.037)	(0.041)	(0.078)	(0.052)	(0.067)	(0.091)	(0.159)	(0.049)	(0.055)
GPA	1.060***	0.956***	1.225***	0.543	−1.255	1.055	−4.871	−4.295	1.146***	0.770***
	(0.055)	(0.054)	(0.318)	(0.938)	(1.476)	(1.680)	(4.274)	(12.917)	(0.077)	(0.073)
GPA^2	0.154***	0.141***	0.313	0.760	−1.261	−1.421	14.968	3.913	0.193***	0.159***
	(0.034)	(0.031)	(0.291)	(1.143)	(0.922)	(6.420)	(11.663)	(10.173)	(0.047)	(0.043)
GPA^3	0.051**	0.044**	0.075	−0.236	−0.218	−1.802	−11.111	−0.908	0.054*	0.046*
	(0.017)	(0.016)	(0.074)	(0.404)	(0.179)	(6.616)	(9.414)	(2.611)	(0.025)	(0.021)
Soc. char	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.319	0.406	0.232	0.290	0.204	0.131	0.146	0.259	0.466	0.349
N	4136	4136	2071	2065	1040	1031	1031	1034	2026	2110
Mean outcome	13.6	13.6	12.6	14.6	12.1	13.1	13.8	15.3	13.6	13.6

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Note: The outcome variable is years of education at age 30. Estimations are based on the following regression: $y_{it} = \gamma GSD_{it} + \alpha GPA_{it} + \mu(GPA_{it})^2 + \eta(GPA_{it})^3 + X_{it}^{'}\beta + \delta_t + \epsilon_{it}$. Columns (1) and (2) use the total sample with socioeconomic characteristics added to column (2). Socioeconomic characteristics consist of number of siblings, father's highest completed grade, mother's highest completed grade and family income in 1979 as well as a dummy for broken home at age 14, a dummy for living in the South at age 14 and a dummy for living in an urban area at age 14, and race and ethnicity dummies. The measures of socioeconomic characteristics correspond to those in Heckman et al. (2006). Columns (3)-(4) and (5)-(8) report results for median and quartile subsamples of GPA, respectively. Columns (9) and (10) report the results for boys and girls separately. Regressions include cohort fixed effects. Robust standard errors in parentheses.

those in Heckman et al. (2006). The outcome of interest is educational attainment and is measured as years of education at age 30, measured from 1 in 1st grade to 20 in the 8th year of college. Descriptive statistics for transcript data, the outcome variable and socioeconomic characteristics are reported in panels A, B and C of Table 5, with the last columns presenting descriptive statistics for boys and girls separately. The average number of years of education is 13.5, with a standard deviation of 2.22, and is similar across genders. As with the Norwegian data, the grade point average is higher among girls than boys while the grade standard deviation is higher among boys.

5.2. Results

The association between years of education and GSD is estimated using an OLS model, controlling for socioeconomic characteristics and including cohort fixed effects, where cohort corresponds to birth year. Grade point average is added as linear, quadratic and cubic terms in order to take possible mechanical correlation between GPA and GSD into account. The specification is comparable to column (5) of Table A2 in the online appendix for the Norwegian data, which showed a similar point estimate to the specification using GPA fixed effects (column (4) of Table 2).

Results are presented in Table 6. The GSD coefficient in columns (1) and (2) tells us how GSD predicts educational attainment when flexible controls for GPA are included.[11] The coefficient is -0.166 in column (1) and is only slightly reduced, to -0.162 when socioeconomic characteristics are included, indicating that student background does not explain the negative association. The results show that for a given grade point average, students with higher variance complete fewer years of education than students with low grade variance. If GSD increases by one standard deviation, educational attainment is reduced by approximately 2 months. This corresponds to 0.07 of a standard deviation decrease in years of education. The US data allow us to investigate other long-term outcomes as well. Using the same specification as column (2) of Table 6, Table B6 in the online appendix shows that increasing GSD by one standard deviation decreases the probability of the student being married by 3 percentage points and decreases the student's net family income by 5000 dollars in 2012, when respondents are aged 48–55. This is in line with the findings for years of education, as schooling leads to both pecuniary and non-pecuniary benefits (Oreopoulos & Salvanes, 2011).

Columns (3)-(10) in Table 6 present the results of regressions run separately for each median and quartile of the grade point average distribution and for boys and girls separately (comparable to columns (1) and (2) in Table A6 for the Norwegian data). Columns (3) and (4) show the results for observations below and above the median grade point average. Both coefficients are negative and significant, but the coefficient is more negative for the sample above the median (-0.24) than below the median (-0.12). The same pattern emerges when the regression is run for each quartile (columns (5)-(8)). In the final two columns, the main estimation is run separately for boys and girls. Estimates are both negative but not statistically different between the genders. The results for all subsamples are less significant, as the standard errors increase due to fewer observations.

The results are very similar to the Norwegian results. The results from Norway and the United States both show a negative association between grade variance and educational attainment when GPA is controlled for. Estimates are negative across the grade distribution and negative for both genders. The coefficient is larger for girls than for boys, as with the Norwegian data, but the difference is not significant.

6. Cognitive and non-cognitive skills

The NLSY79 allows for the inclusion of measures of cognitive and non-cognitive skills.[12] Cognitive and non-cognitive skills are added to the analysis for two reasons. First, including cognitive skills might improve the analysis by reducing potential bias, as GPA likely only proxies for cognitive skills.[13] Second, adding cognitive and non-cognitive skills allows for an investigation of the association between these skills and grade variance. If cognitive skills (e.g. logical reasoning or reading and writing) or non-cognitive skills (e.g. perseverance and conscientiousness) are important across subjects, these skills could reduce grade variance. This could potentially explain the main results, as both cognitive and non-cognitive skills are meaningful predictors of educational outcomes.

Measures of cognitive and non-cognitive skills correspond to those previously used by Heckman et al. (2006). The measure of cognitive skills is a standardized measure based on the Armed Services Vocational Aptitude Battery (ASVAB).[14] The cognitive measure represents the standardized average over the ASVAB scores for arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge and coding speed. The measure for non-cognitive skills is a standardized measure based on a combination of the Rotter Locus of Control Scale (Rotter, 1966), and the Rosenberg Self-Esteem Scale (Rosenberg, 1965). The Rotter Locus of Control Scale is designed to measure the extent to which individuals believe they have control over their lives through self-motivation or self-determination (internal control) as opposed to the extent that the environment (chance, fate, luck) controls their lives (see Table B3 in the online appendix) while the Rosenberg Self-Esteem Scale describes one's degree of approval or disapproval of oneself (see Table B4 in the online appendix). Descriptive statistics are reported in panels D and E of Table 5. Both measures are standardized, with mean 0 and standard deviation 1.

Table 7 displays the conditional correlation between GPA and GSD in the NLSY79 data when cognitive and non-cognitive measures are included. Column (1) is the conditional correlation between GPA and GSD when school fixed effects, socioeconomic characteristics and quadratic and cubic GPA terms are included, column (2) adds the measure of non-cognitive skills, column (3) adds the measure of cognitive skills and column (4) adds both measures. The measure of non-cognitive skills is not significant and does not change the conditional correlation between GPA and GSD. The measure of cognitive skills is negatively associated with GSD.

In columns (5)-(8) of Table 7, cognitive and non-cognitive measures are added to the main analysis. Column (5) is equivalent to the main result in column (2) of Table 6. The estimate of non-cognitive skills, as shown in column (6) is significant and positive, as expected, with a one standard deviation increase in non-cognitive skills predicting an increase in educational attainment of 0.26 of a year. However, the GSD estimate is unchanged, suggesting that the measure of non-cognitive skills does not explain why GSD is negatively associated with educational attainment. The measure for cognitive skills, as shown in column (7), is also significant and positive. A one standard deviation increase in cognitive skills predicts an increase in educational attainment of 0.8 of a year. Also, the GSD estimate is reduced, from -0.16 to -0.11. Column (8) includes both measures, with GPA and GSD estimates remaining

[11] See Table B5 in the online appendix for results when various flexible controls for GPA are added.

[12] Adding improved measures of non-cognitive skills to the Norwegian register data would have allowed us to investigate the relationship between grade variance and non-cognitive skills further, but such data are unfortunately not available.

[13] Roth et al. (2015), for instance, investigate the relationship between standardized intelligence tests and school grades employing a psychometric meta-analysis and find a population correlation of $\rho = 0.54$.

[14] The Armed Services Vocational Aptitude Battery (ASVAB) is a battery of tests administered to applicants to the United States military to determine their qualifications and job assignment (ASVAB, 2015).

Table 7
NLSY79: cognitive and non-cognitive skills.

Outcome	(1) GSD	(2)	(3)	(4)	(5) Years of education	(6)	(7)	(8)
GSD					−0.162***	−0.166***	−0.110**	−0.115***
					(0.037)	(0.036)	(0.034)	(0.034)
GPA	−0.664***	−0.667***	−0.627***	−0.629***	0.956***	0.887***	0.550***	0.535***
	(0.024)	(0.024)	(0.026)	(0.026)	(0.054)	(0.054)	(0.053)	(0.053)
GPA^2	−0.396***	−0.397***	−0.394***	−0.394***	0.141***	0.133***	0.129***	0.125***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.031)	(0.031)	(0.030)	(0.030)
GPA^3	0.006	0.006	0.005	0.005	0.044**	0.046**	0.053***	0.054***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.016)	(0.016)	(0.015)	(0.015)
Non-cognitive		0.009		0.021		0.262***		0.141***
		(0.012)		(0.012)		(0.029)		(0.028)
Cognitive			−0.066***	−0.073***			0.791***	0.747***
			(0.017)	(0.017)			(0.036)	(0.037)
Soc. char	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.544	0.544	0.546	0.546	0.406	0.418	0.462	0.466
N	4136	4136	4136	4136	4136	4136	4136	4136
Mean outcome	13.6	13.6	13.6	13.6	13.6	13.6	13.6	13.6

Note: GSD is the outcome variable in columns (1)-(4). Estimations are based on the following regression: $GSD_{it} = \alpha GPA_{it} + \mu(GPA_{it})^2 + \eta(GPA_{it})^3 + \varphi_1 NonCog + \varphi_2 Cog + X_{it}^{'}\beta + \delta_t + \epsilon_{it}$. Years of education at age 30 is the outcome variable in columns (5)-(8). Estimations are based on the following regression: $y_{it} = \gamma GSD_{it} + \alpha GPA_{it} + \mu(GPA_{it})^2 + \eta(GPA_{it})^3 + \varphi_1 NonCog + \varphi_2 Cog + X_{it}^{'}\beta + \delta_t + \epsilon_{it}$. Column (5) is equivalent to column (2) in Table 6. Columns (2) and (6) include a measure of non-cognitive skills, columns (3) and (7) include a measure of cognitive skills and columns (4) and (8) include both. Regressions include socioeconomic characteristics and cohort fixed effects. Socioeconomic characteristics consist of number of siblings, father's highest grade completed, mother's highest grade completed and family income in 1979 as well as a dummy for broken home at age 14, a dummy for living in the South at age 14 and a dummy for living in an urban area at age 14, and race and ethnicity dummies. The measures of socioeconomic characteristics correspond to those in Heckman et al. (2006). Robust standard errors in parentheses.

Table 8
CDAP: non-cognitive skills, conditional correlations.

Samples	(1) GSD All	(2) GSD	(3) GSD	(4) GSD	(5) GSD Boys	(6) GSD	(7) GSD Girls	(8) GSD
GPA	−0.362***	−0.363***	−0.396***	−0.396***	−0.294***	−0.316***	−0.429***	−0.476***
	(0.050)	(0.051)	(0.054)	(0.054)	(0.070)	(0.078)	(0.076)	(0.083)
GPA^2	−0.301***	−0.301***	−0.303***	−0.303***	−0.338***	−0.340***	−0.263***	−0.270***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.035)	(0.035)	(0.039)	(0.040)
GPA^3	−0.014	−0.014	−0.013	−0.013	−0.033	−0.034	−0.001	0.001
	(0.018)	(0.018)	(0.018)	(0.018)	(0.024)	(0.024)	(0.033)	(0.034)
Non-cognitive: SR		0.004		−0.005		−0.066		0.066
		(0.026)		(0.027)		(0.036)		(0.039)
Non − cognitive: TR			0.053	0.055		0.071		0.030
			(0.039)	(0.040)		(0.060)		(0.052)
Soc. char	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.361	0.360	0.361	0.361	0.297	0.300	0.430	0.432
Observations	1015	1015	1015	1015	514	514	501	501

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Note: GSD is the outcome variable. SR denotes self-reported. TR denotes teacher-reported. Estimates are based on the following regression: $GSD_{it} = \alpha GPA_{it} + \mu(GPA_{it})^2 + \eta(GPA_{it})^3 + \varphi_1 NonCogSR + \varphi_2 NonCogTR + X_{it}^{'}\beta + \delta_t + \epsilon_{it}$. Regressions include socioeconomic characteristics and school fixed effects. Socioeconomic characteristics comrpise gender, ethnicity (dummy variables for Hispanic, Asian, African American, multiethnic or other), birth date, being an English language learner, receiving reduced/free lunch and receiving special education, as well as dummy variables or missing observations. Columns (1)-(4) are based on the whole sample. Column (2) adds self-reported non-cognitive skills, column (3) adds teacher-reported non-cognitive skills, and column (4) adds both. Columns (5)-(6) and columns (7)-(8) report results for boys and girls, respectively, where columns (6) and (8) add self-reported and teacher-reported non-cognitive skills. Robust standard errors in parentheses.

stable from column (7) to column (8). The results correspond to those found for the conditional correlations. The main inference from these estimates is that there is no evidence that the relationship between GPA and GSD or between GSD and educational attainment can be explained by non-cognitive skills in the NLSY79 data. Adding cognitive skills reduces the GSD estimate, however, suggesting that cognitive skills explain part of the association, either by improving the measure of cognitive skills or by indicating that similar cognitive skills are important across subjects. The GSD estimate remains negative and statistically significant, indicating that even if cognitive skills explain some of the association, they do not explain everything.

The results from the NLSY79 data show that grade variance is not associated with non-cognitive skills. However, the measure of non-cognitive skills is quite simple (self-esteem and locus of control) and does not necessarily include the non-cognitive skills we might expect to be associated with low grade variance. In order to explore non-cognitive skills using a richer set of skill measures, data from the Character Development in Adolescence Project (CDAP), provided by Angela Duckworth were analyzed.

The Development in Adolescence Project (CDAP) is a longitudinal survey of 1559 middle school students and their teachers from 8 different schools. It includes measures of students' self-reported non-cognitive skills and teacher-reported non-cognitive skills, along with grades from math, science, English and social studies (see Section C in the online appendix for a more detailed description of the data). Grade point average (GPA) is calculated as the average of all grades received during the fall and spring of 8th grade. Grade standard deviation (GSD), used as a measure of grade variance, is calculated as the standard deviation of the same grades as are used to calculate grade point average. GPA and GSD are then standardized for the whole sample. Socioeconomic characteristics consist of gender, ethnicity (dummy variables for Hispanic, Asian, African American, multiethnic or other) birth date, being an English language learner, receiving reduced price/ free lunch and receiving special education. Dummy variables for missing socioeconomic characteristics are included in the regressions. Descriptive statistics for GPA, GSD and socioeconomic characteristics are listed in the online appendix, Table C1.

Students' self-reported measure (Non-cognitive: SR) is a joint measure of the non-cognitive skills (1) delay discounting, (2) perseverance, (3) self-control: work, (4) self-control: interpersonal, (5) gratitude, (6) actively open-minded thinking, (7) pro-social purpose and (8) internal locus of control. The teacher-reported measure (Non-cognitive: TR) is a joint measure of the non-cognitive skills (1) perseverance, (2) self-control: work, (3) self-control: interpersonal, (4) gratitude, (5) actively open-minded thinking and (6) pro-social purpose. The joint measures are created by standardizing each measure with mean 0 and standard deviation 1 before standardizing the sum of these measures with mean 0 and standard deviation 1. The analysis was performed on all students with non-missing information on both non-cognitive skill measures, leaving a sample of 1015 students. Descriptive statistics for these measures are also reported in the online appendix, Table C1.

Table 8 presents results based on the CDAP data. Column (1) displays the conditional correlation between GPA and GSD, which is negative and significant. Column (2) adds the self-reported non-cognitive measure, column (3) adds the teacher-reported non-cognitive measure and column (4) adds both measures. The GSD estimate remains stable and the measures for non-cognitive skills are not statistically significant across specifications. The results hold when regressions are run for each student- and teacher-reported non-cognitive skill separately (not reported here). Columns (5)-(6) and (7)-(8) report results separately for boys and girls, showing that the conditional correlation between GPA and GSD is negative for both genders and that the estimate of GSD remains stable when measures of non-cognitive skills are added.

In both the NLSY79 data and the CDAP data, non-cognitive skills do not change the size or direction of the GSD estimate in the conditional correlation tables. The association between grade variance and grade point average can therefore not be explained by non-cognitive skills. Adding cognitive skills to the NLSY79 analysis reduces the estimate for GSD, suggesting that part of the association is explained by measures of cognitive skills.

7. Conclusion

Grades are a commonly used measure of cognitive skills and the grade point average, the first moment of the individual grade distribution, is often used to determine admission into further education. This paper is the first to investigate the importance of the second moment of individual grade distribution: grade variance. Grade variance is found to be negatively associated with educational attainment in Norway and the US. Estimates are negative across the grading distribution and are not driven by background characteristics. Results suggest that it is beneficial to be a generalist with low grade variance for a given grade point average - a jack-of-all-subjects.

What does it mean to be a generalist? One explanation is that the student has similar skills across subjects i.e. that the student is good at math, English, science, music etc. and that these skills matter for educational attainment. Lazear (2004) suggests that it might be beneficial to have such a range of skills for certain studies or occupations. This might also be the case for higher education, which is often based on general knowledge, particularly in the US where there is a long tradition of a liberal arts education at four-year colleges. Having a range of skills could also increase your adaptability, which might be useful in the educational system (Hanushek et al., 2011).

Another explanation is that grades measure general skills needed across subjects. These could either be non-cognitive skills, such as perseverance and conscientiousness, or cognitive skills, such as reading and writing or logical reasoning. This explanation is investigated in the US data by including measures of cognitive and non-cognitive skills. Including measures of non-cognitive skills does not change the association between grade variance and educational attainment. Non-cognitive skills therefore do not seem to explain the main finding, although we cannot rule out that improved measures in the US data or available measures in the Norwegian data could have produced other results. When measures of cognitive skills are added, however, the estimate is reduced by 1/3 while remaining negative and significant. This suggests that cognitive skills are part of the explanation.

The application system in Norway for both upper secondary education and higher education relies almost entirely on the grade point average. As grade variance predicts educational attainment conditional on the grade point average, grade variance can potentially be used as an additional skill measure. If educational institutions are interested in students with high ability and effort, but only use the grade point average in the admission decision, they may not be accepting the best students. Students with low grade variance who are just below the grade point average cutoff are likely to outperform students with high grade variance just above the cutoff. This lends support to institutions taking more than the grade point average into account when making admission decisions.

CRediT authorship contribution statement

Astrid Marie Jorde Sandsør: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Project administration, Writing - original draft, Writing - review & editing.

Acknowledgments

Thank you to Angela Duckworth and the Character Development in Adolescence Project as well as Torberg Falch and the project Governance, Management and Performance in the Norwegian Educational System financed by the Research Council of Norway (grant no. 197760) for allowing me to use their data. Thank you to Torberg Falch, Kalle Moene, Katrine Løken, Tuomas Pekkarinen, Bernt Bratsberg and Alexander Koch as well seminar participants at the Norwegian University of Science and Technology, the Workshop on Economics of Education in Mainz, and EALE in Uppsala for helpful comments and suggestions. The paper is part of the research activities at the ESOP center at the Department of Economics, University of Oslo. ESOP is supported by the Research Council of Norway (S/179552).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.econedurev.2020.101969 .

References

ACT (2015). *What is the act?* http://www.actstudent.org/faq/what.html, Reading date: 07.08.2015

ASVAB (2015). *Asvab fact sheet* http://official-asvab.com/docs/asvab_fact_sheet.pdf, Reading date: 07.08.2015

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources, 43*(4), 972–1059.

Carneiro, P., Crawford, C., & Goodman, A. (2007). *The impact of early cognitive and non-cognitive skills on later outcomesTechnical Report.* CEE DP 92.

Elsner, B., & Isphording, I. E. (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics, 35*(3), 787–828.

Falch, T., Nyhus, O. H., & Strøm, B. (2014). Performance of young adults: The importance of different skills. *CESifo Economic Studies*. https://doi.org/10.1093/cesifo/ifu005.
Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review,* 1184–1208.
Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature,* 607–668.
Hanushek, E. A., Woessmann, L., & Zhang, L. (2011). *General education, vocational education, and labor-market outcomes over the life-cycleTechnical Report*. National Bureau of Economic Research.
Haraldsvik, M. (2003). Inntaksprosedyrer for den videregæende skole: Grad av valgfrihet. Institutt for samfunnsøkonomi, NTNU.
Heckman, J. J. (1995). Lessons from the bell curve. *Journal of Political Economy,* 1091–1120.
Heckman, J. J., & Kautz, T. (2013). *Fostering and measuring skills: Interventions that improve character and cognitionTechnical Report*. National Bureau of Economic Research.
Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411–482.
Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*(5220), 41–45.
Herrnstein, R. J., & Murray, C. (2010). *Bell curve: Intelligence and class structure in American life.* Simon and Schuster.
Johnson, W., Carothers, A., & Deary, I. J. (2008). Sex differences in variability in general intelligence: A new look at the old question. *Perspectives on Psychological Science, 3*(6), 518–531.
Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime successOECD Education Working Papers*. OECD Publishing.
Lazear, E. P. (2004). Balanced skills and entrepreneurship. *American Economic Review,* 208–211.
Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics, 3*(1), 101–128.
Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of educational psychology, 79*(3), 280.
Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal, 44*(3), 631–669.
Murnane, R. J., Willett, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics, 77*(2), 251–266.
Murphy, R., & Weinhardt, F. (2018). *Top of the class: The importance of ordinal rankTechnical Report*. National Bureau of Economic Research.
O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for stem. *Nature Communications, 9*(1), 1–8.
OECD (2019). *About pisa* http://www.oecd.org/education/singapore-tops-latest-oecd-pisa-global-education-survey.htm, Reading date: 17.12.2019
Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives, 25*(1), 159–184.
Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.
Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118–137. https://doi.org/10.1016/j.intell.2015.09.002.
Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), 1.
SAT (2015). *About the sat* https://sat.collegeboard.org/about-tests/sat, Reading date: 07.08.2015
Shields, S. A. (1982). The variability hypothesis: The history of a biological model of sex differences in intelligence. *Signs: Journal of Women in Culture and Society, 7*(4), 769–797.
The Norwegian Directorate for Education and Training (2015). *Grunnskolens informasjonssystem* https://gsi.udir.no/application/main.jsp?languageId = 1, Reading date: 29.12.2015
Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106*(4), 1039–1061.