Science Advances
AAAS

Supplementary Materials for

The stereotype that girls lack talent: A worldwide investigation

Clotilde Napp* and Thomas Breda*

*Corresponding author. Email: clotilde.napp@dauphine.psl.eu (C.N.); thomas.breda@ens.fr (T.B.)

Published 9 March 2022, *Sci. Adv.* **8**, eabm3689 (2022)
DOI: 10.1126/sciadv.abm3689

This PDF file includes:

Supplementary Text

Appendix A: Data

General Description of PISA 2018

The Programme for International Student Assessment (PISA) is an every-three-year international survey of 15-year-old students aimed at determining their knowledge and skills in different domains. Students' abilities are assessed in the three curricular domains: mathematics, reading, and science. Students also answer a background questionnaire, seeking information about the students themselves, their homes, and their school and learning experiences.

The PISA target population is made up of all students in any educational institution between the ages of 15 years and 3 months and 16 years and 2 months at the time of the assessment. This specific age has been chosen because it is close to the end of compulsory education in most countries. Efforts have been made to ensure the absence of cultural or national biases in the test items and in the evaluation of performance.

We analyse data from the PISA 2018 survey. The student data set contains around 600,000 observations, representing a population of roughly 32 million 15-year-olds attending seventh grade or above in 79 countries, 37 of which belong to the Organisation for Economic Co-operation and Development (OECD) in 2018, covering in total students from 80% of the world economy.

PISA surveys systematically assess students' performance and knowledge in three core subjects: mathematics, reading and science. However, one of the three core subjects is chosen to be covered in greater depth in each survey. In 2018, reading literacy is the major subject area, as it was in 2000 and 2009. This allows us to get more in-depth information on the students' reading skills.

Test items were a mixture of multiple-choice questions and questions requiring students to construct their own responses. On top of taking tests about their math, science and reading literacy, students also answered a background questionnaire, which took about 35 minutes to complete. The questionnaire sought information about the students themselves, their attitudes, dispositions and beliefs, their homes and their school and learning experiences. It contains in particular questions about gender, home possessions, parental occupation and education. It also contains questions that

are of particular interest to us about their talent, as evaluated by the students, as well as questions about their competitiveness, their self-efficacy, their expected careers (see details below).

Variables of interest in PISA

Main variables of interest

- **Academic performance.** We use individual-level PISA2018 scores in math, reading and science to measure *students' performance.* These scores are on a 0-1000 scale. They have been scaled during the first PISA survey in 2000 to have a mean of 500 and a standard deviation of 100. We will also consider general performance, measured as an equally weighted average of the three scores in math, reading and science. Reading scores, hence general scores, are not available for Spain.
- **Attribution of failure to lack of talent.** Students in PISA2018 were asked to report the extent to which they agree with the following statement about themselves: "When I am failing, I am afraid that I might not have enough talent" (item ST183q02). The sample of students for which this item is available is restricted to 546,037 observations. Answers are given on a 4-point Likert scale from 1 ("strongly disagree") to 4 ("strongly agree") and are used in our analysis to construct the measure of Gender-Talent Stereotypes (GTS), as described below in Appendix B.
- **Competitiveness.** One item in PISA2018 deals with students' competitiveness. More precisely, item st181q02 asks students how much they agree with the following statement about themselves: 'I enjoy working in situations involving competition with others'. We rely on students' answers to this item, given on a 4-point Likert scale to construct our measure of competitiveness, and of gender gaps in competitiveness, as described in Appendix B.
- **Self-efficacy.** In questions ST188 of PISA2018, students are asked to report the extent to which they agree ("strongly disagree", "disagree, "agree", "strongly agree") with the following statements about themselves: "I usually manage one way or another"; "I feel proud that I have accomplished things"; "I feel that I can handle many things at a time"; "My belief in myself gets me through hard times"; and "When I'm in a difficult situation, I can usually find my way out of it". In our analysis, we focus on the last two statements

that involve self-efficacy in the face of difficulty and consider the index of self-confidence in face of difficulty consisting of the equally weighted average of these two items.

- **Choice of ICT fields**. PISA 2018 asked students what occupation they expect to be working in when they are 30 years old. Students could enter any job title or description in an open-entry field; their answers were classified according to the International Standard Classification of Occupations (ISCO-08) and provided in PISA data by the item *ocod3*. One may thus identify "information and communication technology (ICT) professionals" from amongst the occupations they cite. More precisely, we define the choice of ICT field by a dummy equal to one when *ocod3* is between 2500 and 2600.

Variables used for robustness checks

- **Socioeconomic background.** In PISA, a student's socio-economic background is estimated by the PISA index of *economic, social and cultural status* (ESCS), which is based on information about parental education, highest parental occupation, and home possessions including books in the home.
- **Individual controls** include grade repetition, the level of education of the student's parents, measured both in years and kind of diploma obtained, a measure of home educational resources, and a measure of attitude towards school, namely how much students think that they belong to school.
- **Attitude toward the importance of trying hard at school**. PISA2018 asks students, thinking about their school, how much they agree (from strongly agree to strongly disagree) with the following statement: Item st036q08 'Trying hard at school is important'.
- **Truancy.** PISA2018 includes questions about students' truancy, and in particular about skipping classes, or days of school or arriving late at school in the two weeks prior to the PISA test. On average across OECD countries, 21 % of students had skipped a day of school and 48 % of students had arrived late for school in the two weeks prior to the PISA test.

Variables from PISA2018 and previous PISA surveys that are used to better qualify what our main measure of Gender-Talent Stereotypes captures

PISA2012 focuses more specifically on math. Students' questionnaire includes items about math-related behaviors, about their self-efficacy in math-related fields, and about their attitudes towards math.

- **Finding the actual distance between two places on a map.** PISA2012 includes questions about students' self-efficacy in math-related fields and in particular, item st37q33 asks students how confident they feel about having to find the actual distance between two places on a map with a 1:10,000 scale. Answers are given from 1 (very confident) to 4 (not at all confident)
- **Self-responsibility for failure in math.** The index of perceived self-responsibility for failing in mathematics (FAILMAT) was constructed in PISA2012 using student responses when examining the following scenario defined in (ST44): "suppose that you are a student in the following situation: each week, your mathematics teacher gives a short quiz. Recently you have done badly on these quizzes. Today you are trying to figure out why. Are you very likely, likely, slightly likely or not at all likely to have the following thoughts or feelings in this situation? I'm not very good at solving mathematics problems; my teacher did not explain the concepts well this week; this week I made bad guesses on the quiz; sometimes the course material is too hard; the teacher did not get students interested in the material; sometimes I am just unlucky. Higher levels in this index reflects higher tendency to attribute failure in math to external factors rather than to oneself.
- **Quickness to understand things**. In PISA2012, students are asked about their openness to problem solving, and in particular item st94q05 asks them how well the following statement 'I am quick to understand things' describes them (from 1. Very much like me to 5. Not at all like me)
- **Ability to succeed**. In PISA2012, students are asked about their perceived control of academic success. In particular, item st91q01 asks them if they think that they are able to succeed with enough efforts (from 1. strongly agree to 4. strongly disagree).
- **Ambition measured by student's expected occupational status.** As in previous cycles of PISA, students were asked to report their expected occupation at age 30 and a description of this job. The responses were coded to four-digit ISCO codes and then mapped to the ISEI index (Ganzeboom et al., 2003). Recoding of ISCO codes into ISEI index results in

scores for the students' expected occupational status (BSMJ), where higher scores of ISEI indicate higher levels of expected occupational status

- **Self-confidence in reading skills**. Questions ST161 in PISA2018 asks students about the self-confidence in their reading skills and in particular, if they i) strongly agree, ii) slightly agree, iii) slightly disagree, iv) strongly disagree with the following statements: ST161q01: 'I am a good reader', and ST161q02: 'I am able to understand difficult texts'.

Other variables of interest

- **Sense of belonging.** The index of *sense of belonging* (BELONG) was constructed in PISA2018 using students' responses to a trend question about their sense of belonging to school. Students were asked whether they agree ("strongly disagree", "disagree", "agree", "strongly agree") with the following school-related statements (ST034): "I feel like an outsider (or left out of things) at school"; "I make friends easily at school"; "I feel like I belong at school"; "I feel awkward and out of place in my school"; "Other students seem to like me"; and "I feel lonely at school". Positive values on this scale mean that students reported a greater sense of belonging at school than did the average student across OECD countries.
- **Doing homework.** PISA 2012 asked students to report how much time per week they spend doing homework or other study set by teachers. Boys are overwhelmingly less likely than girls to spend time doing homework. On average across OECD countries, girls spend 5.5 hours per week doing homework while boys spend a little less than 4.5 hours. We consider the gender gap in hours, provided by country in Table 2.10a, OECD, PISA 2012 Database.
- **Sustaining performance.** Balart and Oosterveen (2019) analyzes the gender gap in sustaining performance during PISA tests. For both sexes, questions have a lower probability of being answered correctly as the position that they occupy move towards the end of the test, a pattern described as the performance decline. Balart and Oosterveen (2019) shows that the performance decline is weaker for female students than for male students. The gender gap in students' ability to sustain performance is measured by the difference between female and male students in the declines. We retrieve Balart and

Oosterveen (2019)'s data about the gender gap in sustaining performance for PISA2009 and PISA2012.

The PISA methodology, plausible values and statistical inference

Details about the PISA methodology can be found in the PISA Technical reports (see for 2018), but the following lines give the general idea. PISA adopts the Item Response Theory models and does not provide for each student actual scores in math, reading and science but plausible values for performance. These plausible values (10 for PISA2018) are random numbers drawn from the distribution of scores that could be reasonably assigned to each individual, given his or her answers - that is, the marginal posterior distribution. Any estimation procedure in PISA involving students' measured ability in math, reading or science requires the calculation of the targeted statistic for each plausible value (appropriately weighting with the reported student weights) and the final estimate is the arithmetic average of the ten estimates obtained.

Standard errors are calculated with a replication method that takes into account the stratified two-stage sample design for selection of schools and of students within schools. Sources of uncertainty in PISA are actually twofold. First, as explained above, there is some uncertainty on the ability measure of each student and PISA provides ten plausible values drawn from a posterior distribution of ability. Second, there is standard sampling error at country-level. To deal with sampling error, PISA provides 80 alternative sets of individual weights and detailed guideline to use those weights. The computation of corrected standard errors relies on bootstrap techniques: one needs to run the regression of interest for each of the ten plausible values, weighting it first by the "true" set of individual weights and then by the 80 alternative sets of weights. The correct point estimate is the average of the 10 regressions ran with the "true" set of weights, while the standard error is computed according to a formula that sums both measurement errors described above (in practice, this is done with the Stata software using the command "repest").

Other data sources

Our main data source is the 2018 Programme for International Student Assessment (PISA2018). As seen above, we also consider data from previous PISA surveys. We complete these data with several country-level measures of socioeconomic development, extent of (gender) equality, and

values or gender norms or stereotypes as well as data on country-level measures of gender gaps in competitiveness.

Data on country-level measures of development, gender equality and values

Gross Domestic Product (GDP)

GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2017 U.S. dollars.

Source: GDP2018 PPP cstt 2017, worldbank
https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD?end=2019&most_recent_year_desc=true&start=2018

Human Development Index (HDI)

The Human Development Index (HDI) is a composite statistic of life expectancy, education, and per capita income indicators. A country scores a higher HDI when life expectancy, education level and per capita income is higher. Values have been taken for year 2018.

Source: HDI2018, Human development report
http://hdr.undp.org/sites/default/files/hdr2019.pdf, p300-303

Gender Gap Index (GGI)

The Gender Gap Index, from the World Economic Forum, synthesizes the position of women in any given country by taking into account economic opportunities, economic participation, educational attainment, political achievements, and health and well-being. Larger values point to a better position of women in society.

Source: GGI2018, World economic forum
http://www3.weforum.org/docs/WEF_GGGR_2018.pdf

Individualism

"Individualism" is based on Hofstede's cultural dimensions. It captures the degree to which a society is individualistic (as opposed to collectivist), *i.e.* the extent to which individuals are integrated into groups, and how loose are social links.

Source: https://www.hofstede-insights.com/product/compare-countries/

Data on country-level measures of gender gaps in competitiveness.

The World Value Surveys (WVS2010-2014, WVS2017-2020) include an item about competition, asking individuals whether they find competition good or harmful. We measure the country-level gender gap in competition value by the ratio between the percentage of boys and the percentage of girls appreciating competition. We use this measure as a robustness check (Table S8).

Appendix B: Country-level variables computed from PISA items

PISA provides weights to make surveyed students representative of the 15-year-old students of the surveyed countries. We use these weights in all our analyses, so that the results we provide are not subject to sample selection and are representative statistics.

When the estimation of a statistic involves the use of performance (as a control variable, see below), we use the procedure described in the previous section to deal with plausible values. Unless otherwise specified (in the case of dummy variables for instance), all variables computed from the PISA survey are standardized to have a mean of 0 and a standard deviation of 1 for each country separately. This transformation is used to obtain for each country gender gaps in the variables of interest that are directly expressed as a fraction of the variable standard deviation in the country. As such, gender gaps are expressed in a similar metric and directly comparable across countries.

Our measures of Gender-Talent Stereotype (GTS)

Our country-level measure of interiorized Gender-Talent Stereotypes (GTS) is based on national differences between girls and boys in their perceived lack of talent. More precisely, we rely on students' answers to item st183q02 in PISA2018, about the degree to which they agree with the following assertion about themselves: 'When I am failing, I am afraid that I might not have enough talent'. Possible answers are coded as 1: Strongly disagree, 2: Disagree, 3: Agree and 4: Strongly agree. This implies that higher values of the variable correspond to higher perceived lack of talent.

To construct our measure of interiorized gender talent stereotypes GTS, we proceed as follows. First, we standardize the variable about the attribution of failure to lack of talent, as well as the plausible values for math, reading and science performance at the country level, so that their weighted mean (using students' weights provided in PISA to make the country samples representative) is equal to zero and their weighted standard deviation is equal to one. Second, we regress separately in each country or each region the standardized variable about lack of talent on a dummy variable equal to one for female students, controlling for students' ability in math, reading and science. Our measure is simply the estimated effect of being a female student in this regression, which captures the difference between female and male students with respect to interiorized gender-talent stereotypes for each country or region separately. We call this measure GTS. The regression is estimated by weighted least squares using students' weights.

In robustness checks, we also provide the values of the gender differences in the standardized variable about the perceived lack of talent, without controlling for performance (Table S1). As alternative measures of gender-talent stereotypes, we also provide in Table S1 the gap between the percentage of girls and boys agreeing with the assertion about their lack of talent, controlling or not for performance. To this aim, we introduce the binary variable equal to 1 when the code for student's answer to item st183q02 is 3 or 4 (i.e., the student agrees or strongly agrees with the assertion). We then measure the difference in the percentage of boys and girls attributing failure to their lack of talent by the coefficient of the dummy in the regression of the binary variable (not standardized) on a dummy for females, with or without controlling for performance in math, reading and science.

We also consider (in Table S10 as well as in Tables S13 and S14) measures of Gender Talent stereotypes that are not at the country level but at the level of groups of students of similar ability. To construct this measure of GTS at the level of groups of students of similar ability, we proceed as follows. We consider general performance, i.e., the unweighted mean of performance in math, reading and science. We standardize it to have a weighted mean equal to zero and a weighted standard deviation equal to one in each country in the sample. We split the sample in deciles. Adopting the same approach as above, we then measure Gender Talent Stereotypes among one of these performance deciles by the difference between girls' and boys' attribution of failure to lack of talent (or perceived lack of talent), standardized at the country level.

This measure of gender-talent stereotypes and its interaction with students' gender are introduced as controls in Tables S10, S13 and S14.

Other country-level or region-level gender differences

For most country-level (or region-level) gender gaps considered in our work, we proceed as above. We first standardize the variable of interest at the country level (see the description of the considered variables in Appendix A). We then regress the standardized variable on a dummy equal to 1 for females, and controlling for performance in math, reading and science. Performance is also standardized at the country-level. We then measure the country-level gender gap by the coefficient of the female dummy in the regression.

For instance, national gender differences in competitiveness are obtained by regressing for each country and region separately students' answers to item st181q02, standardized by country, on a female dummy and controlling for performance in math, reading and science.

We proceed in the same way for gender gaps in perceived ability to find the distance on a map (item in PISA2012), in self-responsibility for failure in math (index Failmat in PISA2012), in the perceived quickness to understand things (item st94q06 in PISA2012), in the perceived ability to succeed at school (item st91q01 in PISA2012), in ambition (item bsmj). For gender gaps in self-efficacy facing difficulty, we proceed in the same way with our index of self-efficacy (weighted average of items st188q06 and st188q07, standardized by country).

For gender gaps in the belief to be a good reader (item st161q01), we only control for reading performance. For gender gaps in the perceived ability to understand difficult texts (item st161q02), we propose two measures, the first controlling only for reading performance, and the second controlling for both reading performance and the belief to be a good reader.

To construct the gender gap in expectations to work in an ICT-related occupation, we perform a logistic regression of the indicator variable equal to one when the student expects to work in an ICT-related field (see Appendix A) on a female dummy, and control variables for performance in math, reading and science, standardized at the country level.

Appendix C: Methods

Country-level analyses

Some of our analyses are conducted at the country level using the variables described above either constructed from PISA or retrieved from other sources. We perform simple pairwise correlations or use non-weighted linear regression models.

In the first part of our work, we use regression models of the type:

$$GTS_c = \alpha_1 M_c + \varepsilon_c \quad (1)$$

where GTS_c is the measure of Gender-Talent Stereotypes in country c and M_c is either a measure of development or gender equality of country c or a measure of given gender differences in country c, like gender differences in playing chess, in perceived reading ability, in ambition, etc. These models permit to analyze how gender talent stereotypes are related with countries level of development. They also permit to analyze how gender talent stereotypes relate with other measures of national gender differences.

In the second part of our work, we use regression models of the type:

$$N_c = \alpha_2 GTS_c + \beta_2 D_c + \varepsilon_c \quad (2)$$

where N_c denotes a measure of given gender differences in country c (for instance, gender differences in competitiveness), GTS_c is the measure of Gender-Talent Stereotypes in country c and D_c is a measure of development or equality in country c.

Variants of (2) are estimated using only one of the two regressors GTS_c and D_c and the evolutions of α_2 and β_2 across specifications are examined.

Before any regression, we standardize all variables entering the model on the regression sample. This allows us to compare the magnitude of the coefficients across specifications as they are expressed in a similar metric. More specifically, α_2 measures by how many standard deviations N_c varies when Gender-Talent Stereotypes vary by one standard deviation. Similarly, β_2 measures by how many standard deviations N_c varies when measures of development or gender equality vary by one standard deviation.

Individual-level analyses

We provide estimates from individual-level regressions on a sample of about 540,000 students in the 73 countries included in PISA2018 survey, for which we have data on perceived talent.

In the first part of our work, we use individual level regression when analyzing the relation between the gender talent stereotype and general academic performance. We also use individual level regressions as the student-level counterparts of the macro-level analyses of the previous section, to control for unobserved individual-level heterogeneity (at the cost of reducing the sample size) and to correct estimates and standard errors for measurement errors in some of the country-level variables constructed with PISA (see details below).

The closest micro-level counterpart to our cross-country regression model (1) between GTS and country-level measures of development and gender equality is as follows:

$$(Perceived\ lack\ of\ talent)_{ic} = \delta_3 Girl_{ic} + \alpha_3 (D_c * Girl_{ic}) + \mu X_{ic} + \gamma_c + \varepsilon_{ic} \qquad (3)$$

where $(Perceived\ lack\ of\ talent)_{ic}$ is the response of student i in country c to the question st183q02, standardized at the country level, $Girl_{ic}$ is a dummy variable equal to 1 if student i in country c is a girl, D_c is a measure of development or equality in country c, X_{ic} is a vector of control variables whose content varies according to the different specifications (including academic ability, the level of education of the student's parents, measured both in years and kind of diploma obtained, grade repetition, an index of economic, social and cultural status of the household, a measure of home educational resources, and a measure of attitude towards school, namely how much students think that trying hard at school is important), and γ_c a vector of country fixed effects.

Recalling that GTS_c is the gender gap in $(Perceived\ lack\ of\ talent)_{ic}$ in country c, we see that α_3 in model (3) captures how gender-talent stereotypes vary with countries development or equality. Equation (3) is estimated by weighted least squares using students' weights normalized to sum to one in each country. Such "senate" weights ensure that each country has the same weight

in the analysis instead of contributing according to its total population. Standard errors are clustered at the country level, as it is the relevant level of analysis.

To analyze the relation between the gender talent stereotype and academic ability, we perform regression models of the following type:

$$(Perceived\ lack\ of\ talent)_{ic} = \delta_4 Girl_{ic} + \alpha_4(Ability_{ic} * Girl_{ic}) + \beta_4(Ability_{ic}) + \mu X_{ic} + \gamma_c + \varepsilon_{ic} \quad (4)$$

where $Ability_{ic}$ denotes the general performance of student *i* in country *c* (normalized to have a mean of 0 and a standard deviation of 1 in each country), and the other variables are defined as in (3).

In the second part of our work, we analyze the micro level counterparts of Equation (2) by performing regression models of the type

$$(Competitiveness)_{ic} = \delta_5 Girl_{ic} + \alpha_5(GTS_c * Girl_{ic}) + \beta_5(D_c * Girl_{ic}) + \mu X_{ic} + \gamma_c + \varepsilon_{ic} \quad (5)$$

where $(Competitiveness)_{ic}$ denotes the response of student *i* in country *c* to the question st181q02, standardized at the country level and the other variables are defined as in the previous equations. Senate weights are also used and standard errors are again clustered at the country level. Variants of (5) are estimated using only one of the two regressors GTS_c and D_c and the evolutions of α_5 and β_5 across specifications are examined.

Appendix D: Additional Tables: Tables S1-S15

Table S1: Gender Talent Stereotypes (GTS) by country and region

	GTS (1)	Gender gap in attribution of failure to lack of talent (2)	Gender gap in the percentage agreeing being afraid of lacking	Gender gap as in (3) controlling for performance in math, reading and science (4)

				talent when failing (3)	
All PISA2018 Countries		0.239***	0.254***	0.110***	0.105***
OECD countries		0.320***	0.334***	0.143***	0.140***
Non-OECD countries		0.167***	0.182***	0.082***	0.074***
Australia	AUS	0.372***	0.408***	0.167***	0.148***
Austria	AUT	0.256***	0.306***	0.144***	0.110***
Belgium	BEL	0.428***	0.450***	0.216***	0.200***
Canada	CAN	0.444***	0.477***	0.194***	0.178***
Switzerland	CHE	0.324***	0.357***	0.168***	0.149***
Chile	CHL	0.172***	0.185***	0.080***	0.075***
Colombia	COL	0.219***	0.207***	0.094***	0.102***
Czech Republic	CZE	0.339***	0.385***	0.187***	0.164***
Germany	DEU	0.394***	0.415***	0.190***	0.179***
Denmark	DNK	0.575***	0.620***	0.299***	0.283***
Estonia	EST	0.404***	0.465***	0.211***	0.184***
Finland	FIN	0.476***	0.558***	0.269***	0.235***
France	FRA	0.439***	0.466***	0.221***	0.205***
United Kingdom	GBR	0.495***	0.510***	0.227***	0.223***
Greece	GRC	0.293***	0.333***	0.162***	0.149***
Hungary	HUN	0.452***	0.468***	0.221***	0.212***
Ireland	IRL	0.395***	0.450***	0.179***	0.160***
Iceland	ISL	0.532***	0.553***	0.231***	0.217***
Italy	ITA	0.401***	0.428***	0.195***	0.183***
Japan	JPN	0.268***	0.287***	0.118***	0.115***
Korea	KOR	0.346***	0.362***	0.148***	0.142***
Lithuania	LTU	0.353***	0.398***	0.190***	0.166***
Luxembourg	LUX	0.378***	0.406***	0.196***	0.184***
Latvia	LVA	0.391***	0.427***	0.189***	0.180***
Mexico	MEX	0.077**	0.090***	0.039***	0.031*
Netherlands	NLD	0.509***	0.532***	0.234***	0.221***
New Zealand	NZL	0.397***	0.455***	0.195***	0.172***
Poland	POL	0.386***	0.414***	0.174***	0.159***
Portugal	PRT	0.378***	0.415***	0.181***	0.166***
Slovak Republic	SVK	0.273***	0.310***	0.153***	0.138***
Slovenia	SVN	0.468***	0.496***	0.228***	0.223***
Sweden	SWE	0.529***	0.539***	0.247***	0.244***
Turkey	TUR	0.144***	0.154***	0.075***	0.070***
United States	USA	0.359***	0.355***	0.135***	0.144***

Albania	ALB	0.108***	0.075**	0.040***	0.058***
United Arab Emirates	ARE	0.090***	0.124***	0.060***	0.046***
Argentina	ARG	0.179***	0.202***	0.095***	0.080***
Bulgaria	BGR	0.229***	0.214***	0.108***	0.119***
Bosnia and Herzegovina	BIH	0.157***	0.166***	0.081***	0.080***
Belarus	BLR	0.390***	0.415***	0.198***	0.185***
Brazil	BRA	0.238***	0.270***	0.122***	0.110***
Brunei Darussalam	BRN	0.255***	0.296***	0.118***	0.112***
Costa Rica	CRI	0.213***	0.199***	0.105***	0.109***
Dominican Republic	DOM	0.087**	0.094***	0.056***	0.056***
Georgia	GEO	0.180***	0.174***	0.102***	0.111***
Hong Kong	HKG	0.292***	0.319***	0.122***	0.117***
Croatia	HRV	0.266***	0.320***	0.146***	0.128***
Indonesia	IDN	0.101***	0.104***	0.031**	0.031*
Jordan	JOR	0.065*	0.030	0.020	0.034**
Kazakhstan	KAZ	0.179***	0.181***	0.088***	0.085***
Kosovo	KSV	0.197***	0.160***	0.072***	0.093***
Lebanon	LBN	0.039	0.047	0.036**	0.028*
Macao	MAC	0.289***	0.340***	0.117***	0.098***
Morocco	MAR	0.163***	0.176***	0.091***	0.077***
Moldova	MDA	0.337***	0.322***	0.149***	0.157***
FYROM	MKD	0.188***	0.134***	0.057***	0.083***
Malta	MLT	0.296***	0.367***	0.174***	0.140***
Montenegro	MNE	0.054*	0.040	0.023	0.025
Malaysia	MYS	0.222***	0.259***	0.088***	0.083***
Panama	PAN	0.010	0.012	0.014	0.015
Peru	PER	0.087***	0.084***	0.031**	0.034**
Philippines	PHL	0.133***	0.171***	0.058***	0.048***
Qatar	QAT	0.022	0.062***	0.032***	0.014
Romania	ROU	0.278***	0.280***	0.127***	0.128***
Russia	RUS	0.279***	0.288***	0.137***	0.133***
Saudi Arabia	SAU	-0.189***	-0.179***	-0.070***	-0.080***
Singapore	SGP	0.323***	0.361***	0.137***	0.125***
Serbia	SRB	0.145***	0.175***	0.082***	0.069***
Chinese Tapeï	TAP	0.281***	0.305***	0.115***	0.110***
Thailand	THA	0.114***	0.195***	0.103***	0.073***
Ukraine	UKR	0.290***	0.322***	0.158***	0.142***
Uruguay	URY	0.103***	0.117***	0.037**	0.034*

Notes: The table presents Gender Talent Stereotypes by country and region when they are measured 1. by the gender gap in the answers to PISA question about their attribution of failure to lack of talent ('when I am failing, I am afraid

that I might lack talent'), controlling for performance, 2. as in (1) without controlling for performance, 3. by the difference between the percentage of boys and girls who strongly agree or agree being afraid of lacking talent when failing, 4. as in (3) but controlling for performance in math, reading and science. Performance is standardized at the country level, as is the dependent variable. The variables are defined in more detail in Appendix B. *** *p<0.01*, ** *p<0.05*, * *p<0.1*

Table S2: Gender Talent Stereotypes (GTS) by country and region, with controls for truancy and beliefs that trying hard at school is important

	GTS	Gender gap in perceived talent controlling for truancy and attitudes towards school
All PISA2018 Countries	0.244***	0.246***
OECD countries	0.330***	0.325***
Non-OECD countries	0.167***	0.176***
Australia	0.366***	0.364***
Austria	0.267***	0.261***
Belgium	0.434***	0.405***
Canada	0.448***	0.436***
Switzerland	0.353***	0.346***
Chile	0.202***	0.204***
Colombia	0.234***	0.235***
Czech Republic	0.347***	0.335***
Germany	0.385***	0.379***
Denmark	0.583***	0.584***
Estonia	0.416***	0.416***
Finland	0.475***	0.470***
France	0.437***	0.405***
United Kingdom	0.490***	0.485***
Greece	0.302***	0.304***
Hungary	0.464***	0.462***
Ireland	0.411***	0.410***
Iceland	0.542***	0.545***
Italy	0.403***	0.395***
Japan	0.267***	0.256***
Korea	0.345***	0.346***
Lithuania	0.361***	0.367***

Luxembourg	0.385***	0.377***
Latvia	0.387***	0.381***
Mexico	0.108***	0.118***
Netherlands	0.524***	0.532***
New Zealand	0.416***	0.412***
Poland	0.373***	0.366***
Portugal	0.371***	0.374***
Slovak Republic	0.291***	0.274***
Slovenia	0.442***	0.436***
Sweden	0.537***	0.538***
Turkey	0.147***	0.147***
United States	0.362***	0.354***
Albania	0.108***	0.132***
United Arab Emirates	0.087***	0.105***
Argentina	0.147***	0.146***
Bulgaria	0.261***	0.265***
Bosnia and Herzegovina	0.162***	0.157***
Belarus	0.395***	0.391***
Brazil	0.239***	0.243***
Brunei Darussalam	0.286***	0.286***
Costa Rica	0.218***	0.224***
Dominican Republic	0.236***	0.257***
Georgia	0.179***	0.182***
Hong Kong	0.297***	0.292***
Croatia	0.270***	0.264***
Indonesia	0.096***	0.115***
Jordan	0.061*	0.068*
Kazakhstan	0.192***	0.211***
Kosovo	0.196***	0.191***
Macao	0.291***	0.284***
Morocco	0.184***	0.204***
Moldova	0.334***	0.322***
Malta	0.293***	0.277***
Montenegro	0.064**	0.062**
Malaysia	0.223***	0.232***
Panama	0.046	0.060
Peru	0.105**	0.103**
Philippines	0.127***	0.136***
Qatar	0.025	0.031

Romania	0.282***	0.284***
Russia	0.287***	0.290***
Saudi Arabia	-0.200***	-0.199***
Singapore	0.324***	0.301***
Serbia	0.161***	0.164***
Chinese Tapeï	0.281***	0.281***
Thailand	0.120***	0.130***
Ukraine	0.287***	0.290***
Uruguay	0.093**	0.100**

Notes: Analyses restricted to the 435,508 students for whom data about truancy, importance of trying hard at school and performance scores in math, reading and science are available. The table presents in the first column Gender talent Stereotypes measured, as in Table S1, by the difference between the answers of boys and girls to PISA question about their attribution of failure to lack of talent ('when I am failing, I am afraid that I might lack talent'), controlling for performance. Performance is standardized at the country level, as is the dependent variable. The second column adds as controls truancy items (how often skip school day, some classes, arrive late at school), as well the student's belief about the importance of trying hard at school. The variables are defined in more detail in Appendix A. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S3: Relationship between Gender-Talent Stereotypes and countries' development or gender equality. Estimates from individual-level linear regressions

	Dependent variable is perceived lack of talent			
a) Linking the Gender-Talent stereotype to GDP				
Girl	0.294***	0.290***	0.269***	0.295***
(s.e)	(0.0166)	(0.0166)	(0.0162)	(0.0167)
Girl*GDP	0.0806***	0.0807***	0.0800***	0.0806***
(s.e)	(0.0167)	(0.0169)	(0.0169)	(0.0168)
Number of observations	470,543	470,543	470,543	447,608
b) Linking the Gender-Talent stereotype to Human Development (HDI)				
Girl	0.294***	0.290***	0.270***	0.296***
(s.e.)	(0.0144)	(0.0144)	(0.0142)	(0.0143)
Girl*HDI	0.113***	0.114***	0.113***	0.115***
(s.e)	(0.0133)	(0.0133)	(0.0133)	(0.0133)
Number of observations	466,776	466,776	466,776	443,856
c) Linking the Gender-Talent stereotype to Gender equality (Gender Gap Index)				
Girl	0.300***	0.296***	0.275***	0.298***
(s.e.)	(0.0150)	(0.0149)	(0.0141)	(0.0151)
Girl*GGI	0.106***	0.107***	0.106***	0.108***
(s.e)	(0.0141)	(0.0144)	(0.0143)	(0.0140)
Number of observations	455,834	455,834	455,834	433,329
Country fixed effects	Yes	Yes	Yes	Yes
Individual control for general ability	No	Yes	Yes	Yes
Individual control for math, reading and science ability	No	No	Yes	Yes
Other individual controls	No	No	No	Yes

Notes: The Table is the student-level counterpart of Table 2. Other individual controls include the level of education of the student's parents, measured both in years and kind of diploma obtained, grade repetition, an index of economic, social and cultural status of the household, a measure of home educational resources, and a measure of attitude towards school, namely how much students think that trying hard at school is important. See the data section in this SM for details about the sources of country-level measures of development or equality. Standard errors have been clustered at the country level. Regressions are weighted by "senate" weights which sum to one in each country. See the method section of this SM for details on the empirical models. **** p<0.01, ** p<0.05, * p<0.1*.

Table S4: Impact of performance on perceived lack of talent for boys and girls on the whole sample

Dependent variable	*Perceived lack of talent*			
	Girls	Boys	Girls	Boys
	(1)	(2)	(3)	(4)
General performance (standardized by country)	0.109***	0.029***		
	(0.005)	(0.006)		
General performance (standardized worldwide)			0.121***	0.001
			(0.006)	(0.006)
Constant	0.117***	-0.129***	0.118***	-0.128***
	(0.005)	(0.006)	(0.005)	(0.006)
Observations	256,192	250,742	256,192	250,742
R-squared	0.012***	0.0009**	0.014***	0.000

Notes: Analyses based on a sample of 73 countries. The Table shows that perceived lack of talent (or analogously, attribution of failure to lack of talent) increases more with performance for girls than for boys (higher slopes and R-squared), whether general performance is standardized by country or not. Perceived lack of talent is based on students' answers to PISA item about their attribution of failure to lack of talent and is standardized at the country level. General performance is the unweighted mean of performance in math, reading and science. We consider two different standardizations of performance. For (1) and (2), performance is standardized by country i.e., standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country (and on the full sample of countries). For (3) and (4), performance is standardized worldwide i.e., such that the weighted mean is equal to 0 and the weighted standard deviation is equal to 1 on the full sample of countries. All estimates and standard errors are based on plausible values for math, reading and science ability and account for measurement error in these abilities on top of standard sampling error. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S5: Impact of general performance on perceived lack of talent for boys and for girls separately, and on Gender Talent Stereotypes: Breakdown by country and region

		Impact of performance on ...	
	Intercept of equation (1) (α)	boys' attribution of failure to lack of talent (β)	GTS or gender gap in attribution of failure to lack of talent (γ)
All PISA2018 Countries	0.245***	0.029***	0.081***

OECD countries	0.329***	0.053***	0.067***
Non-OECD countries	0.170***	0.008	0.094***
Australia	0.400***	0.071***	0.063***
Austria	0.301***	-0.033*	0.073***
Belgium	0.437***	0.024	0.083**
Canada	0.468***	0.058***	0.049**
Switzerland	0.347***	0.011	0.118***
Chile	0.179***	-0.004	0.140***
Colombia	0.203***	-0.033	0.047*
Czech Republic	0.378***	-0.016	0.091***
Germany	0.389***	-0.015	0.154***
Denmark	0.618***	-0.016	0.022
Estonia	0.458***	-0.024	0.130***
Finland	0.527***	0.048***	0.103***
France	0.460***	0.021	0.082***
United Kingdom	0.508***	0.055***	0.062**
Greece	0.324***	-0.004	0.076***
Hungary	0.463***	-0.044*	0.131***
Ireland	0.444***	0.046**	0.034
Iceland	0.541***	0.066***	0.011
Italy	0.426***	0.020	0.044
Japan	0.285***	0.107***	-0.012
Korea	0.360***	0.015	0.038
Lithuania	0.393***	-0.028	0.105***
Luxembourg	0.397***	0.010	0.095***
Latvia	0.417***	0.007	0.103***
Mexico	0.088***	-0.006	0.024
Netherlands	0.508***	0.081***	0.071*
New Zealand	0.449***	0.058***	0.073***
Poland	0.404***	0.029	0.098***
Portugal	0.411***	-0.044**	0.122***
Slovak Republic	0.287***	0.071***	0.111***
Slovenia	0.484***	-0.020	0.120***
Sweden	0.522***	0.006	0.130***
Turkey	0.153***	-0.003	0.031
United States	0.348***	0.123***	0.103***
Albania	0.098***	-0.154***	0.096***
United Arab Emirates	0.107***	-0.007	0.106***
Argentina	0.199***	-0.033*	0.029
Bulgaria	0.200***	-0.038*	0.118***
Bosnia and Herzegovina	0.165***	-0.119***	0.103***

Belarus	0.412***	-0.002	0.062**
Brazil	0.253***	0.073***	0.120***
Brunei Darussalam	0.288***	0.004	0.071***
Costa Rica	0.195***	-0.035	0.092***
Dominican Republic	0.082**	-0.076***	0.075**
Georgia	0.194***	-0.115***	-0.009
Hong Kong	0.312***	0.025	0.024
Croatia	0.308***	0.014	0.155***
Indonesia	0.088***	0.025	0.088***
Jordan	0.055*	-0.054**	-0.043
Kazakhstan	0.176***	-0.053***	0.112***
Kosovo	0.171***	-0.112***	0.015
Lebanon	0.042	-0.005	0.058*
Macao	0.332***	0.077***	0.024
Morocco	0.179***	-0.019	0.001
Moldova	0.339***	-0.129***	0.053
FYROM	0.162***	-0.102***	0.013
Malta	0.342***	0.059**	0.085**
Montenegro	0.035	-0.082***	0.093***
Malaysia	0.240***	0.079***	0.074***
Panama	0.014	-0.014	-0.009
Peru	0.066**	-0.061**	0.084**
Philippines	0.144***	0.061***	0.133***
Qatar	0.022	0.008	0.146***
Romania	0.272***	-0.062***	0.167***
Russia	0.282***	-0.067***	0.149***
Saudi Arabia	-0.174***	-0.017	0.008
Singapore	0.357***	0.004	0.105***
Serbia	0.159***	-0.053***	0.149***
Chinese Tapeï	0.301***	0.106***	-0.019
Thailand	0.163***	0.060***	0.078***
Ukraine	0.318***	0.038*	0.025
Uruguay	0.105***	-0.045*	0.093***

Notes: The Table shows the results of the regressions by country and regions of a variable measuring perceived lack of talent on a dummy for female (first column), general performance (second column) and their interaction (third column). Perceived lack of talent is based on students' answers to PISA item about their attribution of failure to lack of talent and is standardized at the country level. General performance is the unweighted mean of performance in math, reading and science, standardized by country i.e., standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country (and on the full sample of countries). All estimates and standard errors are based on plausible values for math ability and account for measurement error in these abilities on top of standard sampling error. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S6: Gender gap in competitiveness, self-confidence in face of difficulty and choice of IT fields by country and region

	Gender Gap in competitiveness (B-G)	Gender Gap in self-confidence (B-G)	Gender Gap in choice of IT fields (ratio B/G)
	(1)	(2)	(3)
All PISA2018 Countries	0.181***	0.100***	1.883
OECD countries	0.248***	0.160***	2.087
Non-OECD countries	0.120***	0.048***	1.686
Australia	0.272***	0.176***	2.879
Austria	0.394***	0.127***	2.278
Belgium	0.263***	0.381***	3.681
Canada	0.289***	0.223***	2.262
Switzerland	0.367***	0.230***	2.513
Chile	0.253***	0.200***	3.700
Colombia	0.155***	0.016	2.112
Czech Republic	0.427***	0.172***	2.748
Germany	0.463***	0.251***	2.116
Denmark	0.323***	0.345***	3.429
Estonia	0.321***	0.135***	2.538
Finland	0.374***	0.198***	2.537
France	0.380***	0.424***	3.715
United Kingdom	0.479***	0.368***	2.582
Greece	0.323***	0.149***	2.504
Hungary	0.307***	0.190***	2.417
Ireland	0.295***	0.273***	2.259
Iceland	0.273***	0.307***	2.421
Israel	0.242***	-0.076***	1.472
Italy	0.218***	0.238***	3.366
Japan	0.026	0.045	2.448
Korea	0.361***	0.285***	1.890
Lithuania	0.220***	0.046	2.695
Luxembourg	0.317***	0.183***	2.785
Latvia	0.163***	0.112***	2.558
Mexico	0.165***	0.108***	1.402
Netherlands	0.355***	0.343***	2.787
Norway	0.222***		2.791
New Zealand	0.239***	0.180***	1.920
Poland	0.219***	0.248***	2.988
Portugal	0.472***	0.278***	2.571
Slovak Republic	0.202***	0.151***	1.722

Slovenia	0.469***	0.264***	2.973
Sweden	0.359***	0.271***	2.655
Turkey	0.062**	-0.066**	1.712
United States	0.254***	0.133***	1.809
Albania	-0.010	-0.086***	2.680
United Arab Emirates	-0.026	0.000	1.009
Argentina	0.176***	0.157***	2.204
Bulgaria	0.107***	-0.037	2.083
Bosnia and Herzegovina	0.196***	0.056*	1.753
Belarus	0.205***	0.043	1.903
Brazil	0.249***	0.128***	2.348
Brunei Darussalam	0.129***	0.053*	2.285
Costa Rica	0.208***	0.198***	2.014
Dominican Republic	0.084**	0.023	3.791
Georgia	-0.115***	-0.092***	2.743
Hong Kong	0.109***	0.215***	1.782
Croatia	0.367***	0.131***	2.447
Indonesia	0.037	-0.012	0.832
Jordan	-0.135***	-0.149***	1.795
Kazakhstan	-0.052*	-0.089***	1.868
Kosovo	0.074**	-0.034	2.184
Lebanon	-0.014	-0.035	1.708
Macao	0.164***	0.290***	2.195
Morocco	-0.092***	0.013	1.656
Moldova	0.154***	0.000	2.744
FYROM	0.035	-0.064*	1.540
Malta	0.161***	0.149***	1.518
Montenegro	0.212***	0.037	2.054
Malaysia	0.142***	0.055*	0.959
Panama	0.125***	0.014	1.164
Peru	0.035	0.065*	2.194
Philippines	0.166***	-0.012	1.154
Qatar	-0.023	0.066***	1.504
Romania	0.105***	0.052*	2.069
Russia	0.185***	0.123***	1.895
Saudi Arabia	-0.018	-0.138***	1.612
Singapore	0.336***	0.138***	1.525
Serbia	0.220***	-0.011	2.036
Chinese Tapeï	0.174***	0.175***	1.325
Thailand	0.287***	-0.029	1.282
Ukraine	0.310***	0.040	2.109
Uruguay	0.296***	0.241***	2.445

Notes: Column 1 shows the results of the regressions by country and regions of a variable measuring competitiveness on a dummy for female, controlling for performance in math, reading and science. Competitiveness is based on students' answers to PISA item about their enjoyment 'working in situations involving competition with others' and is standardized at the country level. Performance is also standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. The second column shows the results of the linear regressions by country and regions of an index measuring self-confidence in face of difficulty on a dummy for female, controlling for performance in math, reading and science. The index of self-confidence in face of difficulty is based on students' agreement with the assertions “When I’m in a difficult situation, I can usually find my way out of it” as well as “My belief in myself gets me through hard times” and is standardized at the country level. The third column shows the results of the logistic regression by country and regions of a dummy variable representing the expectation to work in a IT field on a dummy for female, controlling for performance in math, reading and science. Performance is standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. All estimates and standard errors are based on plausible values for math ability and account for measurement error in these abilities on top of standard sampling error. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S7: Relationship between the gender gap in competitiveness, countries' measures of development and gender equality (Gender equality paradox for competitiveness) and GTS. Estimates from individual-level linear regressions

	Dependent variable is competitiveness		
a) Linking the gender gap in competitiveness to wealth: GDP			
Girl	-0.202***	-0.211***	-0.223***
(s.e)	(0.0181)	(0.0174)	(0.0173)
Girl*GDP	-0.0682***	-0.0696***	-0.0701***
(s.e)	(0.0203)	(0.0207)	(0.0209)
Number of observations	530,490	491,018	461,651
b) Linking the gender gap in competitiveness to development: HDI			
Girl	-0.202***	-0.211***	-0.224***
(s.e)	(0.0165)	(0.0160)	(0.0156)
Girl*HDI	-0.0981***	-0.0998***	-0.104***
(s.e)	(0.0156)	(0.0160)	(0.0157)
Number of observations	526,720	487,248	457,896
c) Linking gender gap in competitiveness to Gender equality (Gender Gap Index)			
Girl	-0.208***	-0.219***	-0.230***
(s.e)	(0.0164)	(0.0158)	(0.0164)
Girl*GGI	-0.0995***	-0.0962***	-0.0970***
(s.e)	(0.0184)	(0.0174)	(0.0180)
Number of observations	515,753	476,281	447,363
d) Linking gender gap in competitiveness to Gender Talent Stereotypes (GTS)			
Girl	-0.192***	-0.204***	-0.213***
(s.e)	(0.0135)	(0.0128)	(0.0130)
Girl*GTS	-0.123***	-0.120***	-0.115***
(s.e)	(0.0133)	(0.0128)	(0.0131)
Number of observations	516,472	516,472	492,378
Country fixed effects	Yes	Yes	Yes
Individual control for general ability	No	Yes	Yes
Other individual controls	No	No	Yes

Notes: Definitions and Data sources for Gross Domestic Product, Human Development Index and Gender Gap Index are more detailed in Appendix A. The variable GTS denotes a country-level measure of gender-talent stereotypes as described in Appendix B and shown in Table S1. Other individual controls include the level of education of the student's parents, measured both in years and kind of diploma obtained, grade repetition, an index of economic, social and cultural status of the household, a measure of home educational resources, and a measure of attitude towards school, namely how much students think that trying hard at school is important. See the data section in this SM for details about the sources of country-level measures of development or equality. Standard errors have been clustered at the country level. Regressions are weighted by "senate" weights which sum to one in each country. See the method section of this SM for details on the empirical models. **** p<0.01, ** p<0.05, * p<0.1.*

Table S8: Relationship between the gender gap in competitiveness, countries' measures of development and gender equality and Gender Talent Stereotypes: macrolevel analysis with World Value Survey

	Dependent variable is Gender Gap in alternative measure of competitiveness, retrieved from World Value Survey						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GGI	0.259*				-0.206		
	(0.138)				(0.168)		
log GDP		0.382***				-0.0316	
		(0.126)				(0.185)	
HDI			0.456***				0.0413
			(0.123)				(0.219)
GTS				0.537***	0.656***	0.558***	0.506**
				(0.121)	(0.160)	(0.177)	(0.207)
Constant	0.0240	-0.00359	0.00659	-0.0167	-0.00110	-0.0206	-0.0104
	(0.137)	(0.125)	(0.122)	(0.120)	(0.129)	(0.124)	(0.126)

Notes: One country (Korea) excluded due to an abnormally low number of respondents considering competition good, preventing to compute a meaningful gender gap. The Table shows the results of the regressions at the country-level of an alternative measure of gender gap in competitiveness on a measure of Gender Talent Stereotypes (GTS) and measures of development (Gross Domestic Product, Human Development Index) and gender equality (Gender Gap Index). The alternative measure of the gender gap in competitiveness relies on the difference between the percentage of males and of females who report finding competition good in the World Value Survey 2017-2020 (see Appendix A for more details). GTS denotes a country level measure of talent stereotypes as defined and shown in Table S1. The regressions whose results are reported in the first four columns involve only one explanatory variable whereas those in the last three columns involve both a measure of gender talent stereotype and a variable measuring development or gender equality. Definitions and Data sources for GDP, GGI and HDI are more detailed in Appendix A. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S9: Impact of general performance on boys' and girls' competitiveness separately, and on their gender gap: Breakdown by country and region

	Intercept of equation (2) (B-G)	Impact of performance on ...	
		Boys' competitiveness	**Gender gap in competitiveness (B-G)**
All PISA2018 Countries	0.191***	0.133***	0.054***
OECD countries	0.259***	0.119***	0.047***
Non-OECD countries	0.130***	0.144***	0.061***
Australia	0.297***	0.087***	0.037*
Austria	0.381***	0.130***	0.079***
Belgium	0.281***	0.084***	0.128***
Canada	0.300***	0.071***	-0.002
Switzerland	0.390***	0.074***	0.046
Chile	0.278***	0.086***	0.052*
Colombia	0.150***	0.232***	0.077***
Czech Republic	0.433***	0.111***	0.054*
Germany	0.467***	0.089***	-0.014
Denmark	0.341***	0.099***	0.094***
Estonia	0.386***	0.175***	0.061*
Finland	0.422***	0.129***	0.020
France	0.392***	0.104***	0.072**
United Kingdom	0.505***	0.070***	-0.002
Greece	0.328***	0.157***	0.078**
Hungary	0.346***	0.064**	0.039
Ireland	0.337***	0.096***	-0.002
Iceland	0.269***	0.231***	0.123***
Israel	0.223***	0.156***	0.168***
Italy	0.222***	0.104***	0.086***
Japan	0.085***	0.174***	-0.009
Korea	0.375***	-0.038**	-0.047**
Lithuania	0.233***	0.128***	0.101***
Luxembourg	0.325***	0.120***	0.080**
Latvia	0.182***	0.164***	0.076**
Mexico	0.139***	0.196***	0.081***
Netherlands	0.332***	0.120***	0.135***
Norway	0.233***	0.140***	0.055*
New Zealand	0.254***	0.046**	0.036
Poland	0.227***	0.097***	0.134***
Portugal	0.472***	0.063***	0.018
Slovak Republic	0.247***	0.147***	0.055*
Slovenia	0.479***	0.075***	0.079**

Sweden	0.372***	0.110***	0.064**
Turkey	0.021	0.136***	0.124***
United States	0.275***	0.110***	0.035
Albania	-0.058**	0.223***	0.076***
United Arab Emirates	-0.064***	0.253***	0.242***
Argentina	0.169***	0.112***	0.092***
Bulgaria	0.092***	0.275***	0.157***
Bosnia and Herzegovina	0.191***	0.190***	0.143***
Belarus	0.226***	0.102***	0.043
Brazil	0.259***	0.184***	0.103***
Brunei Darussalam	0.110***	0.130***	0.207***
Costa Rica	0.223***	0.214***	0.105***
Dominican Republic	0.054	0.264***	0.163***
Georgia	-0.118***	0.174***	0.084**
Hong Kong	0.124***	0.048**	0.005
Croatia	0.363***	0.106***	0.109***
Indonesia	0.055**	0.130***	-0.008
Jordan	-0.213***	0.279***	0.158***
Kazakhstan	-0.034	-0.025*	-0.061***
Kosovo	0.041	0.206***	-0.003
Lebanon	-0.027	0.307***	0.123***
Macao	0.170***	0.032	-0.024
Morocco	-0.123***	0.227***	0.088***
Moldova	0.169***	0.140***	0.057*
FYROM	-0.011	0.274***	0.126***
Malta	0.188***	0.213***	0.069*
Montenegro	0.181***	0.198***	0.208***
Malaysia	0.148***	0.142***	0.093***
Panama	0.083**	0.191***	0.133***
Peru	-0.033	0.250***	0.089***
Philippines	0.173***	0.199***	0.140***
Qatar	-0.062***	0.310***	0.216***
Romania	0.117***	0.129***	0.078**
Russia	0.179***	0.127***	0.085***
Saudi Arabia	-0.078**	0.303***	0.182***
Singapore	0.352***	-0.016	0.021
Serbia	0.206***	0.222***	0.153***
Chinese Tapeï	0.194***	0.054***	-0.037
Thailand	0.295***	0.013	0.097***
Ukraine	0.304***	0.079***	0.050
Uruguay	0.294***	0.217***	0.122***

Notes: The Table shows the results of the regressions by country and regions of a variable measuring competitiveness on a dummy for female (first column), general performance (second column) and gender interacted with performance (third column). Competitiveness is based on students' answers to PISA item about their enjoyment 'working in situations involving competition with others' and is standardized at the country level. General performance is the unweighted mean of performance in math, reading and science and is also standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. All estimates and standard errors are based on plausible values for math ability and account for measurement error in these abilities on top of standard sampling error. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S10: Relationship between the gender gap in competitiveness and countries' measures of development and gender equality with the mediation of GTS. Estimates from individual-level linear regressions

	Dependent variable is competitiveness		
a) Linking the gender gap in competitiveness to GDP with the mediation of GTS			
Girl	-0.196***	-0.208***	-0.218***
(s.e)	(0.0145)	(0.0137)	(0.0139)
Girl*GDP	-0.0170	-0.0186	-0.0207
(s.e)	(0.0168)	(0.0159)	(0.0161)
Girl*GTS	-0.114***	-0.110***	-0.104***
(s.e)	(0.0141)	(0.0132)	(0.0138)
Number of observations	479,526	479,526	455,755
b) Linking the gender gap in competitiveness to HDI with the mediation of GTS			
Girl	-0.198***	-0.210***	-0.220***
(s.e)	(0.0144)	(0.0136)	(0.0137)
Girl*HDI	-0.0356*	-0.0365**	-0.0427**
(s.e)	(0.0192)	(0.0182)	(0.0180)
Girl*GTS	-0.0993***	-0.0956***	-0.0868***
(s.e)	(0.0167)	(0.0153)	(0.0158)
Number of observations	475,756	475,756	452,000
c) Linking gender gap in competitiveness to GGI with the mediation of GTS			
Girl	-0.204***	-0.215***	-0.224***
(s.e)	(0.0143)	(0.0136)	(0.0139)
Girl*GGI	-0.0404*	-0.0389*	-0.0334
(s.e)	(0.0206)	(0.0197)	(0.0203)
Girl*GTS	-0.0944***	-0.0924***	-0.0913***
(s.e)	(0.0162)	(0.0157)	(0.0159)
Number of observations	464,789	464,789	441,467
Country fixed effects	Yes	Yes	Yes
Individual control for general ability	No	Yes	Yes
Other individual controls	No	No	Yes

Notes: Definitions and Data sources for Gross Domestic Product, Human Development Index and Gender Gap Index are more detailed in Appendix A. The variable GTS denotes a country-level measure of gender-talent stereotypes as described in Appendix B and shown in Table S1. Other individual controls include the level of education of the student's parents, measured both in years and kind of diploma obtained, grade repetition, an index of economic, social and cultural status of the household, a measure of home educational resources, and a measure of attitude towards school, namely how much students think that trying hard at school is important. See the data section in this SM for details about the sources of country-level measures of development or equality. Standard errors have been clustered at the country level. Regressions are weighted by "senate" weights which sum to one in each country. See the method section of this SM for details on the empirical models. *** p<0.01, ** p<0.05, * p<0.1.

Table S11: Relationship between the gender gap in self-confidence in face of difficulty, countries' measures of development and gender equality and Gender Talent Stereotypes: macrolevel analysis

	Dependent variable is Gender Gap in self-confidence (B-G)						
GGI	0.565***				0.114		
	(0.105)				(0.106)		
log GDP		0.585***				0.305***	
		(0.102)				(0.084)	
HDI			0.687***				0.356***
			(0.093)				(0.099)
GTS				0.746***	0.666***	0.611***	0.518***
				(0.079)	(0.102)	(0.082)	(0.096)
Constant	-0.010	-0.000	-0.019	0.025	-0.004	0.010	-0.003
	(0.104)	(0.102)	(0.092)	(0.078)	(0.081)	(0.075)	(0.074)
Observations	68	71	70	76	67	70	69
R-squared	0.304	0.322	0.477	0.547	0.591	0.639	0.647

Notes: The Table shows the results of the regressions at the country-level of the gender gap in self-confidence in face of difficulty on a measure of Gender Talent Stereotypes (GTS) and measures of development (Gross Domestic Product, Human Development Index) and gender equality (Gender Gap Index). GTS denotes a country level measure of talent stereotypes as defined and shown in Table S1. The regressions whose results are reported in the first four columns involve only one explanatory variable whereas those in the last three columns involve both a measure of gender talent stereotype and a variable measuring development or gender equality. The index of self-confidence in face of difficulty is based on students' agreement with the assertions "When I'm in a difficult situation, I can usually find my way out of it" as well as "My belief in myself gets me through hard times" and is standardized at the country level. Gender gaps in self-confidence are defined in more details in Appendix A. Definitions and Data sources for GDP and GGI are more detailed in Appendix A. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S12: Relationship between the gender gap in choice of IT fields, countries' measures of development and gender equality and Gender Talent Stereotypes: macrolevel analysis

	Dependent variable is Gender Gap in Choice of IT fields (B/G)						
GGI	0.376***				0.0378		
	(0.118)				(0.149)		
log GDP		0.202*				-0.0578	
		(0.120)				(0.121)	
HDI			0.365***				0.0381
			(0.116)				(0.147)
GTS				0.508***	0.483***	0.536***	0.486***
				(0.100)	(0.139)	(0.117)	(0.140)
Constant	0.0582	0.0323	0.0329	0.00292	0.0523	0.0238	0.0290
	(0.117)	(0.119)	(0.115)	(0.0994)	(0.110)	(0.107)	(0.108)
Observations	69	72	71	76	67	70	69
R-squared	0.132	0.039	0.125	0.259	0.265	0.266	0.265

Notes: The Table shows the results of the regressions at the country-level of the gender gap in choice of IT fields (measured as described in the Notes of Table S6) on a measure of Gender Talent Stereotypes (GTS) and measures of development (Gross Domestic Product) and gender equality (Gender Gap Index). GTS denotes a country level measure of talent stereotypes as defined and shown in Table S1. The regressions whose results are reported in the first four columns involve only one explanatory variable whereas those in the last three columns involve both a measure of gender talent stereotype and a variable measuring development or gender equality. Competitiveness is based on students' answers to PISA item about their enjoyment 'working in situations involving competition with others' and gender gaps in competitiveness are defined in more details in Appendix A. Definitions and Data sources for GDP and GGI are more detailed in Appendix A. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S13: Impact of performance on boys' and girls' self-confidence in face of difficulty on the whole sample, with and without the mediation of gender talent stereotypes

	Dependent variable: Self-Confidence in face of difficulty	
	(1)	(2)
General performance	0.087***	0.062***
	(0.006)	(0.007)
Girl*General performance	-0.051***	0.002
	(0.008)	(0.009)
Gender Talent Stereotypes		0.034***
		(0.004)
Girl* Gender Talent Stereotypes		-0.074***
		(0.005)
Girl	-0.096***	0.086***
	(0.007)	(0.007)
Constant	0.046***	-0.038***
	(0.005)	(0.009)
Observations	498,526	492,652
R-squared	0.007***	0.010***

Notes: The Table shows the results of the regression on the whole sample of a variable measuring self-confidence in face of difficulty on a dummy for female, general performance and gender interacted with performance in the first specification, adding gender talent stereotypes as well as their interaction with general performance in the second specification. The index of self-confidence in face of difficulty is based on students' agreement with the assertions "When I'm in a difficult situation, I can usually find my way out of it" as well as "My belief in myself gets me through hard times" and is standardized at the country level. General performance is the unweighted mean of performance in math, reading and science and is also standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. Gender Talent Stereotypes are measured by the gender gap in perceived lack of talent taken in a reference group that comprises all students of the country that are in the same decile of general performance. All estimates and standard errors are based on plausible values for math, reading and science ability and account for measurement error in these abilities on top of standard sampling error. Standard errors in parentheses. **** p<0.01, ** p<0.05, * p<0.1*

Table S14: Impact of performance on boys' and girls' choice of IT fields on the whole sample, with and without the mediation of gender talent stereotypes

	Dependent variable: Choice of IT fields	
	(1)	(2)
General performance	0.026***	0.015***
	(0.0012)	(0.0016)
Girl*General performance	-0.021***	-0.012***
	(0.0012)	(0.0013)
Gender Talent Stereotypes		0.0144***
		(0.0006)
Girl*Gender Talent Stereotypes		-0.0137***
		(0.0007)
Girl	-0.044***	-0.011***
	(0.001)	(0.002)
Constant	0.053***	0.018***
	(0.001)	(0.0016)
Observations	437,057	428,403
R-squared	0.028***	0.036***

Notes: The Table shows the results of the regression on the whole sample of a variable measuring the choice of IT fields on a dummy for female, general performance and gender interacted with performance in the first specification, adding gender talent stereotypes as well as their interaction with general performance in the second specification. Choice of IT fields is measured by students' answers about what occupation they expect to be working in when they are 30 years old. (see Appendix A for more details). General performance is the unweighted mean of performance in math, reading and science and is also standardized to have a weighted mean equal to 0 and a weighted standard deviation equal to 1 in each country. Gender Talent Stereotypes are measured by the gender gap in perceived lack of talent taken in a reference group that comprises all students of the country that are in the same decile of general performance. All estimates and standard errors are based on plausible values for math, reading and science ability and account for measurement error in these abilities on top of standard sampling error. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S15: Self-confidence in reading and other measures possibly related to perceived talent: Gender gaps and Cross-country correlations with GTS

i. Levels of gender gaps

Fear of failure in math (G-B)	0.0219**
Self-efficacy reading a map (B-G)	0.203***
Perceived quickness to understand things (B-G)	0.153***
Confidence understanding difficult texts, controlling for reading performance (B-G)	0.0872***
Confidence understanding difficult texts, controlling for reading performance and confidence to be a good reader (B-G)	0.166***
Ambition (B-G)	-0.307***
Confidence being a good reader, controlling for reading performance (B-G)	0.149***
Confidence ability to succeed at school (B-G)	-0.0433***

ii. Cross country correlations with GTS

gender gap in …	***GTS***
Fear failure in math (G-B)	0.527***
Self-efficacy reading a map (B-G)	0.636***
Perceived quickness to understand things (B-G)	0.320**
Confidence understanding difficult texts, controlling for reading performance (B-G)	0.429***
Confidence understanding difficult texts, controlling for reading performance and confidence to be a good reader (B-G)	0.507***
Ambition (B-G)	-0.119
Confidence being a good reader, controlling for reading performance (B-G)	-0.142
Confidence ability to succeed at school (B-G)	0.012

Notes: Table i. shows the gender gaps in various measures that might be related to perceived talent. We control for performance in math, in reading and in science for all gender gaps, except those dealing with reading abilities, for which we control for reading performance only, or both reading performance and the belief to be a good reader. All variables are standardized at the country level. Table ii. shows the cross-country correlations of the measure of Gender Talent Stereotypes (GTS) with country level gender gaps in the various measures considered in Table i. All variables are described in Appendix A. **** p<0.01, ** p<0.05, * p<0.1*