

DOI: [10.20472/TE.2018.6.1.007](https://doi.org/10.20472/TE.2018.6.1.007)

THE GENDER GAP IN MATH. EVIDENCES OF A STUDY IN THE PRIMARY SCHOOL IN THE SWISS CANTON OF TICINO

GIOVANNA ZANOLLA

Abstract:

This work analyzes the gender gap in math results of an entire cohort of students attending the fifth grade of the primary school in the Swiss canton of Ticino. Three evaluation contexts have been considered: the mark itself assigned by the teacher at the end of the school year, the score obtained in a standardised math test and the evaluation provided by teachers in the same school year. Multivariate analysis allows to conclude that all things being equal, gender plays a significant effect on the mark in mathematics. Males are more likely than females to achieve a high mark and they are perceived as more active and participative, while females are seen as more conform to school rules, but less involved and more insecure.

Keywords:

gender gap, primary school, evaluation, gender stereotypes, achievement, mathematics

JEL Classification: I24

Authors:

GIOVANNA ZANOLLA, Department of Education and Learning, University of Applied Sciences and Arts of Southern Switzerland, Switzerland, Email: giovanna.zanolla@supsi.ch

Citation:

GIOVANNA ZANOLLA (2018). The gender gap in math. Evidences of a study in the primary school in the Swiss canton of Ticino. *International Journal of Teaching and Education*, Vol. VI(1), pp. 103-125., [10.20472/TE.2018.6.1.007](https://doi.org/10.20472/TE.2018.6.1.007)

1. Introduction

1.1 Boys' underachievement at school and girls' under-representation in STEM fields

In 'Some Thoughts Concerning Education' of 1693, John Locke deplored the poor language skills of males from the upper classes compared to females and laid responsibility on school education for that. This gender gap in school performance continues even nowadays and the international literature shows that boys typically do worse at school than girls and are more frequently exposed to failure and dropping out (Mickelson, 1989; NESSE, 2009; Bradley & Renzulli, 2011; Mc Neal, 2011; OECD, 2015; Borgna & Struffolino, 2016). Greater attention has been paid to this disparity in recent years due to the fact that the economic changes in many Western countries over the last ten years have led to a drop in the number of manufacturing jobs and generally speaking of medium- and low-skilled jobs, thus raising tangible concerns over the possibility that a certain percentage of males might be 'at risk' as a result of being poorly qualified and hence not apt to be readily integrated into the labour market (Erikson et al., 2005; Balfanz et al., 2007; Oreopoulos, 2007; Rumberger, 2011).

Why do girls do better? Although gender differences in IQ appear to be negligible (Strand et al., 2006), girls are often reported to possess better social skills, to be more motivated, to find it easier to comply with school regulations and to be prepared to work harder, gaining greater satisfaction from success at school (Di Prete & Buchmann, 2013). Girls are said to be more conscientious (Kling et al., 2015) and self-disciplined (Duckworth and Seligman, 2006) and care more about their teachers' opinions (Bray et al., 1997). By contrast, boys seem more likely to fall prey to anti-school male role models according to which they are unpopular if they work too hard at school. Allegedly males are more aggressive and behave more frequently in anti-social, deviant ways, suffering more often from attention deficit disorders and learning disabilities (Van Houtte, 2004; Di Prete & Buchmann, 2013).

However, although on average girls are more likely to achieve better results, choose to continue their education and seek qualified jobs (Schoon & Eccles, 2014), they are still under-represented in the fields of science, technology, engineering and mathematics, i.e. those areas where more jobs are available and remuneration is higher on average (OECD, 2012).¹ Math-gender stereotypes, math anxiety, math self-concepts and attributions and expectations for success and failure in math, which are influenced by

¹ The PISA 2015 survey, which refers to fifteen year olds, shows better performance by males than females in standardised maths tests, in Ticino as in many other OECD countries (Salvisberg, Zampieri and Egloff, 2016; OECD, 2016). The same conclusion was also reached by previous PISA surveys (OECD, 2010, 2014; PISA Consortium. ch 2011, 2014).

parents and teachers' math gender stereotypes, have been shown to contribute to the pursuit of math-related career paths (Eccles et al., 1990; Jacobs et al. 2005).

The higher risk of males under-achieving at school and the low numbers of females in the most dynamic areas of the labour market that are key to a country's competitiveness suggest that males and females have different school experiences which are responsible for the way the two sexes enter the labour market and build their career (OECD, 2015) and, in more general terms, for the gender-based occupational segregation in certain professions (Zanolla, 2015).

1.2 Objectives

This work intends to investigate if there are any gender disparities in school performance and compare the achievement of boys and girls in an entire cohort of fifth grade students in Ticino, the Italian-speaking canton of Switzerland, by considering three different types of assessment: a standardised maths test taken by fifth graders at the end of the school year, the marks scored in maths by fifth graders during the same school year as the test and the student progress report prepared by the incumbent teacher at the end of the first term of that same year.

Our decision to focus on maths is linked first of all to the fact that many prestigious jobs require good mathematical skills and that, as previously noted, males are over-represented in these jobs also because girls tend to interiorise the stereotype according to which they have less flair for mathematics (Steele, 1997). Secondly, maths is crucial for students' individual fate at school. In Ticino, from the eighth grade on (third year of middle school) classes are divided according to the level at which mathematics and the German language are taught. The level of difficulty of the curriculum followed by the students determines which course of study they can choose after compulsory education.

1.3 Three types of assessment

Let us consider the three types of assessment included in the study. Based on a combination of factors both linked to learning (results of class tests, exams, etc.) and not linked to learning (engagement, homework, classroom participation, conduct and tendency to improve or worsen) (Cizek et al., 1996; Guskey, 2002 and Randall & Engelhard, 2010), marks are considered to be good predictors of an individual's success at school. However, the above-mentioned literature suggests that marks involve a teacher-student relationship, which is likely to put boys at a disadvantage, unlike standardised tests which are not biased by such relational factor.

On the other hand, standardised tests might prove less female-friendly due to the 'stereotype threat' according to which an individual, notwithstanding personal cognitive skills, fares badly in a task if the social group to which they belong is the subject of a negative stereotype regarding the skills required by the task in question (Steele &

Aronson, 1995). To find oneself in a stereotype threat situation - such as a standardised test in a subject like maths, where girls are held to be inferior to males - could cause stress and anxiety in female students, which could affect the results of the test, thus confirming the negative stereotype (Ashcraft & Ridley, 2005; Zeidner & Matthews, 2011).

Lastly, student progress reports prepared by the teachers at the end of the first-term are a sort of snapshot of the student's performance with regard to their engagement at school in general, their attitude during lessons, care taken with homework and relationships with teachers and classmates. These make it possible to bring to light both the hidden curriculum the school requires of students, thus exposing the factors which are included in student assessment, and eventual gender-based stereotypes and gender-biased expectancies.

The decision to look at the combination of these three very diverse types of assessment is the novel and distinctive feature of this study whose aim is to shed light on how assessment is made and why a given set of students can achieve greater success at school than another. A further aim of this work is to provide teachers with food for thought about assessment and what is expected of students of different genders, two elements that strongly affect the educational and occupational fate of young people. It is also meant to act as a starting point for devising strategies that can reduce failure at school for all minority groups and promote equity. In fact, it is worth noting that in many countries disparities in school achievement based on social status are greater than those linked to gender, and in some countries ethnic origin is the feature which strongly conditions school performance, second only to social status (Archer et al., 2010).

2. Method

As mentioned above, three different types of assessment were considered for the same cohort of students representing Ticino's entire fifth grade student population in school year 2014/15: a standardised maths test, the maths marks assigned by the teacher at the end of the school year and the student progress report prepared by the teacher at the end of the first term.

2.1 The standardised maths test

In 2012 the Ticino Department of Education, Culture and Sport appointed a team of researchers, local experts in maths and teachers from primary and lower secondary schools to create a standardised test² to be taken by all fifth grade students (2929) in the

² Analyses were conducted by Miriam Salvisberg and Giovanna Zanolla: after assessing the answers to the items included in the pre-test, it was determined that data are best described by a one-dimensional model (competence in maths) with sub-domains. Therefore, for the purposes of the test, items were chosen which covered the entire scale of difficulties, had greater discrimination capacity and were more consistent with the model. For further details see Crescentini (2017).

Canton. The aim was twofold: to provide political decision-makers with information for monitoring the education system and school teachers, head teachers and inspectors with detailed information regarding trends in their classrooms and schools.

In drawing up the test, six competence areas were considered, taken from the compulsory education study plan (Divisione della scuola, 2015):

Dimensions and measures – Executing and use (GM_EA);

Dimensions and measures – Knowing, recognising and describing (GM_SRD);

Dimensions and measures – Mathematizing and modelling (GM_MT);

Numbers and calculating – Executing and use (NM_EA);

Numbers and calculating – Knowing, recognising and describing (NC_SRD);

Numbers and calculating – Mathematizing and modelling (NC_MT).

Although these areas do not cover the entire fifth grade curriculum, they do represent a significant part of it. For each area different items were developed which had to meet two requirements: they had to refer to just one dimension and have good discrimination capacity (i.e. they should trigger right answers from the good performing students and wrong answers from poor performers). The items were first tried out in a pre-test and each item was tested on at least 300 pupils. After collecting the questionnaires, the answers were entered into a database. The individual exercises, scales and groups of exercises were assessed using the Rasch (1961) model so as to identify the exercises to be included in the final test. Analyses identified 90 exercises, 15 for each of the six areas. The 'general mathematics' dimension considered here was normalised so that values of 0 to 100 can be used to represent the total score obtained by adding up the scores in the six areas.

2.2 Marks in mathematics

For each of the 2929 fifth grade pupils who sat the standardised test, the maths marks given to the pupil by the teacher in the same school year as the test (2013/14) were available. Such information is contained in the database of the GAGI application (Gestione Allievi Gestione Istituti - Pupil Management, Institute Management) run by the Ticino Department of Education, Culture and Sport. GAGI contains relevant social and personal details for all primary and secondary school students in Ticino, as well as the training they take part in for each school year, the marks achieved in each subject, number of absences, end-of-year results, etc.

2.3 First term overall assessment

At the end of the first term each pupil receives a progress report which shows not only the marks scored in the different subjects, but also an overall assessment that considers parameters such as student's attitude in the classroom, ability to work independently, relationships with classmates and teachers, obedience to school regulations and care of their own school material and that of others. Like marks, assessments are entered in the above-mentioned GAGI application database.

3. Results

3.1 Variables affecting maths marks

In all the maths fields considered in the test, males achieved average marks significantly higher than females (figure 1).

Figure 1: Average normalised marks in the maths fields included in the standardised test, broken down by gender.

Maths fields	Females			Males		
	Test average	N	Std. deviation	Test average	N	Std. deviation
GM_EA	45.91	1373	13.44	48.24	1448	14.25
GM_MT	45.64	1343	14.75	47.53	1430	15.77
GM_SRD	58.05	1379	13.45	62.10	1455	14.35
NC_EA	45.22	1367	12.53	47.10	1446	14.07
NC_MT	42.40	1369	13.79	44.65	1436	14.73
NC_SRD	50.61	1365	13.66	52.40	1448	14.29
General mathematics	61.54	1336	10.44	63.72	1422	11.20

Notwithstanding better results in the test, it can be seen that the average marks assigned by teachers to boys are not significantly different from those achieved by girls ($F = 0.733$, $Sign. = 0.392$, $Eta\ squared = 0.000$).

A logistic regression model was then used (figures 2 and 3) in which the dependent variable is a maths mark lower than or equal to 5.0³ (value of variable = 0) or higher than 5.0 (value of variable = 1). The model predictors include the 'general mathematics' score, whose deviation from the average score achieved by the pupils tested was assumed to

³ In Ticino, as in the rest of Switzerland, the marking scale ranges from a minimum of 1.0 to a maximum of 6.0. The pass threshold is 4.0 and 5.0 equals 'good'.

provide an indication of mathematical skills, as well as gender, social status and nationality⁴ .

Figure 2: Variables included in the logistic regression models and their frequency.

Variable	Method
Dependant	
Achieve a mark in maths of ≤ 5.0 or > 5.0 in fifth grade	0 = maths mark ≤ 5.0 (1619 cases) 1 = maths mark > 5.0 (960 cases)
Predictors	
Normalised deviation from average marks in general maths (Deltamatgen)	0 to 1, average = 0.50, std. deviation = 0.14
Gender	0 = Male (1239) 1 = Female (1189)
Social status	0 = Lower or middle class (highly skilled and unskilled blue collars and lesser qualified white collars) (1145) 1 = high (highly skilled white collars) (1283)
Nationality	1 = Swiss (1850) 2 = Italian (281) 3 = other nationalities (297)
Conduct marks	1 = up to 4.5 (190) 2 = 5.0 (835) 3 = 5.5 (1044) 4 = 6.0 (846)
Teacher gender	0 = Male (975 pupils) 1 = Female (1875 pupils)

Figure 3: Logistic regression model

⁴ A logistic regression model was preferred to a multilevel one due to the fact that when breaking down the variance of the 'maths mark' variable into its components within the classes and across classes, the inter-class variance was found to be just 5.3% of the total variance; for this reason a multilevel analysis was not considered necessary.

Variables	B	S.E.	Wald	gl	Sign.	Exp(B)
Gender (a)	.165	.109	2.299	1	.129	1.179
Social status (b)	.255	.112	5.171	1	.023	1.290
Nationality (c)			6.198	2	.045	
Italian	.099	.171	.333	1	.564	1.104
Other	-.449	.193	5.395	1	.020	.638
Deltamatgen	13.537	.589	527.519	1	.000	756665.390
Constant	-7.884	.339	539.624	1	.000	.000

Cox and Snell R-squared = 0.357; Nagelkerke R-squared = 0.488

Reference categories: (a) Male; (b) Low or middle; (c) Swiss

The model shown in figure 3 above, which explains 49% of the variance, reveals that above average mathematical skills and upper class status have a positive, significant effect on the likelihood of achieving maths marks above 5.0 in fifth grade. Gender, on the other hand, does not appear to have any significant effect. When fifth grade conduct marks are included in the model, the effect of gender becomes apparent and being a male increases the probability of achieving higher maths marks: in short, ability at school, conduct marks and social and national status being equal, boys are more likely than girls to achieve maths marks above 5.0 (figure 4).

Figure 4: Logistic regression model with conduct marks added

Variables	B	S.E.	Wald	gl	Sign.	Exp(B)
Gender (a)	-.271	.121	5.029	1	.025	.762
Social status (b)	.234	.115	4.135	1	.042	1.263
Nationality (c)			6.893	2	.032	
Italian	.197	.176	1.260	1	.262	1.218
Other	-.433	.197	4.837	1	.028	.649
Deltamatgen	13.495	.607	494.446	1	.000	725712.645
Conduct mark in fifth grade			82.243	3	.000	
(d)						
5.0	.582	.315	3.428	1	.064	1.790
5.5	1.382	.312	19.572	1	.000	3.983
6.0	1.876	.319	34.467	1	.000	6.525
Constant	-8.911	.457	379.689	1	.000	.000

Cox and Snell R-squared = 0.380; Nagelkerke R-squared = 0.520

Reference categories: (a) Male; (b) Low or middle; (c) Swiss; (d) up to 4.5

Figure 5 illustrates five ‘ideal-type’ cases obtained by including the coefficient B values shown in figure 4 in the regression equation. It can be seen that a male from a lower or middle class background, placed in the upper quartile on the basis of test results, but with a conduct mark of 4.5 (case 1) has 87% probability of achieving a maths mark higher than 5.0: in practice, bright pupils receive high marks even when their behaviour is not of the best. For the same type of pupil, but with a conduct mark of 6.0 the probability increases to 98% (case 2). A pupil having the same characteristics as case 2, but only average test results, is 43% more likely to achieve marks above 5.0 (case 3), which goes up to 49% if he comes from an upper class background (case 4) and drops to 36% if the pupil is female (case 5). While maths skills are the most important variable for determining marks, a good conduct mark, privileged social status and Swiss or Italian nationality are shown to increase the probability of achieving marks above 5.0 in maths.

Figure 5: Some ideal-type cases taken from the logistic regression model in figure 4

	Case 1	Case 2	Case 3	Case 4	Case 5
Gender	Male	Male	Male	Male	Female
Social status	Lower or middle class	Lower or middle class	Lower or middle class	Upper class	Lower or middle class
Nationality	Swiss	Swiss	Swiss	Swiss	Swiss
Deltamategen	Above average skills (0.8)	Above average skills (0.8)	Average skills (0.5)	Average skills (0.5)	Average skills (0.5)
Conduct marks	4.5	6.0	6.0	6.0	6.0
Probability of achieving a mark over 5.0 in maths	0.87	0.98	0.43	0.49	0.36

The assumption behind this work was that males were penalised in teachers’ assessments due to their behaviour. Actually, penalisation linked to behaviour exists, but the analysis also shows that, all other things being equal, it is girls who are placed at a disadvantage. The same conclusion was reached when the teacher’s gender was included in the model, although it should be noted that the latter variable operates on a different level than the pupil’s characteristics, as it is pertinent to the entire class and might therefore yield a distorted estimate (figure 6). Certain factors, which appear to escape quantitative analysis and go beyond conduct marks, could be characteristic of the female or male gender and affect the teacher’s assessment. It might be that teachers

expect less from girls in terms of mathematical skills, or that certain types of behaviour or attitude by males in the classroom while not raising their conduct mark let them gain the teacher's approval against female classmates with the same skills.

Figure 6: Logistic regression model with conduct marks and teacher gender added

Variables	B	S.E.	Wald	df	Sign.	Exp(B)
Gender (a)	-.293	.120	5.972	1	.015	.746
Social status (b)	.245	.114	4.645	1	.031	1.278
Nationality (c)			5.321	2	.070	
Italian	.114	.174	.431	1	.512	1.121
Other	-.409	.194	4.426	1	.035	.664
Deltamategen	13.382	.601	495.744	1	.000	648459.899
Conduct marks in fifth grade (d)			86.826	3	.000	
5.0	.510	.301	2.869	1	.090	1.664
5.5	1.345	.299	20.309	1	.000	3.840
6.0	1.837	.306	36.046	1	.000	6.277
Teacher's gender (e)	.105	.117	.806	1	.369	.901
Constant	-8.772	.445	388.025	1	.000	.000

Cox and Snell R-squared = 0.378; Nagelkerke R-squared = 0.516

Reference categories: (a) Male; (b) Lower or middle; (c) Swiss; (d) up to 4.5; (e) Male

3.2 Student progress report according to pupil gender

Up to now the conduct mark assigned by the teacher, which has been seen to affect maths marks, has been used as an indicator of a pupil's behaviour in the classroom.

Student progress reports (narrative reports in fact) prepared by the teachers at the end of the first term were analysed to understand which parameters teachers use to formulate their pupils' profiles and determine which character and behavioural traits are appreciated by the teachers and which are not.

Not surprisingly, test results were rather consistent with the marks assigned by the teacher: pupils with the best marks scored significantly higher in the test (figure 7). However, compared to the results of the standardised test, 60 pupils appear to have been heavily penalised in the teacher's assessment since their maths marks did not exceed 4.5, while in the test they rank in the group of best performers. By contrast, 62 pupils seem to have been clearly over-rated: in spite of ranking among the low-performing pupils in the test, they achieved maths marks of over 5.0 (figure 8). For the sake of convenience, hereinafter the former group will be referred to as the 'under-rated' and the latter as the 'over-rated'. It is worth noting however that also the test score includes a component not linked to the pupil's ability (Boaler, 2003; Buck et al., 2010)

and in no way do we wish to claim that the standardised test is more valid and reliable than the teacher's assessment. Our claim is that, compared to the 'over-rated', the 'under-rated' receive an assessment which more frequently includes characteristics that teachers view negatively and less frequently characteristics that teachers view positively.

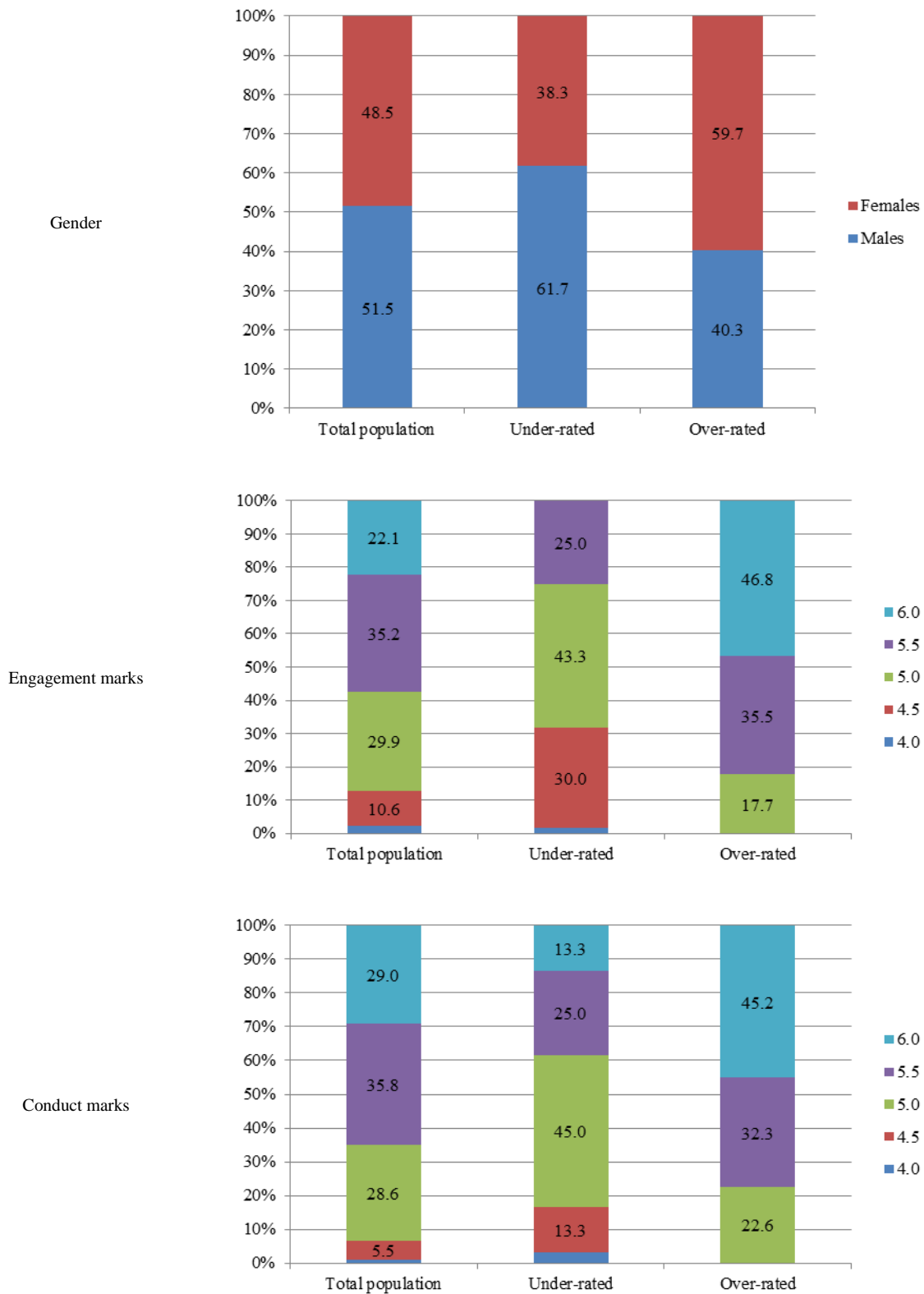
Figure 7: Average marks for the general maths test vis à vis fifth grade maths marks (F = 426.7, Sign. = 0.000, Eta squared=0.482).

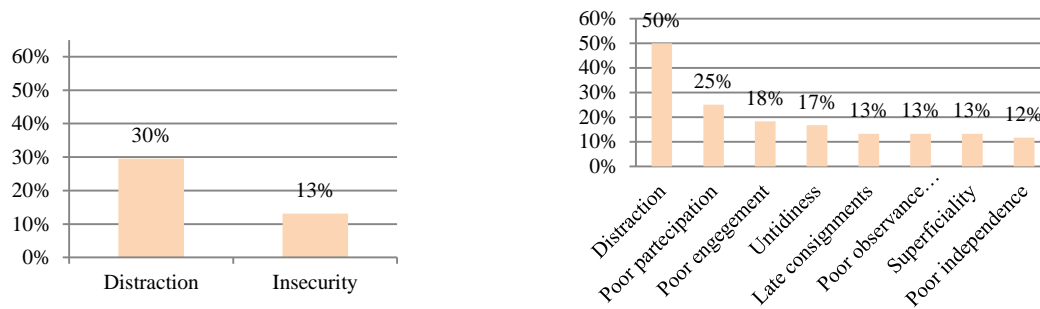
Maths marks	Average in general maths	N	Std. deviation
3.0	49.20	16	5.81
3.5	50.42	2	8.21
4.0	50.01	338	7.54
4.5	56.88	551	7.55
5.0	61.75	845	7.41
5.5	69.33	753	8.31
6.0	76.42	252	8.91
Total	62.67	2757	10.89

Figure 8: Standardised test scores and maths marks

		General maths average score in standardised test			Total
		<58.01	58.01 to 66.99	>66.99	
Maths marks	3.0 - 4.5	612	234	60	906
	5.0	240	419	186	845
	>5.0	62	269	674	1005
Total		914	922	920	2756

Figure 9: Comparison of ‘under-rated’ and ‘over-rated’ pupils and the entire cohort, by gender, engagement and conduct marks – percentage values.



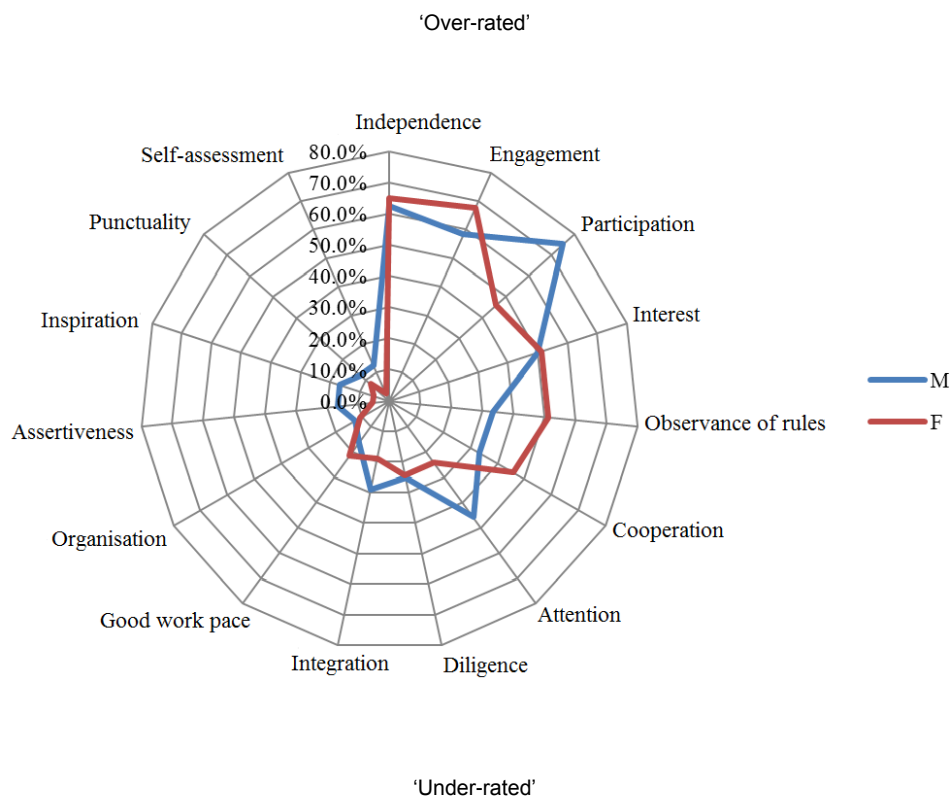


For the total number of 'over-rated' (on the left of the figure) and 'under-rated' pupils (on the right of the figure), Figure 10 shows the negative and positive aspects mentioned in teachers' reports that exceed a percentage frequency of 10%. 'Over-rated' pupils are often described as able to work independently and show engagement (64% of the cases); they are reported to participate in the classroom (57%), show interest (51%), obey rules (44%) and cooperate with their class mates (41%). Some of these characteristics also appear in the profiles of 'under-rated' pupils, but with a markedly lower incidence. Compared to 'over-rated' pupils, 'under-rated' pupils are more likely to be considered distracted (50% against 30% of over-rated) and are reported to show poor participation (25%), make little effort (18%), be untidy (17%), be late in submitting assignments (13%), disregard school regulations (13%) and be superficial (13%). After distraction, the main flaw in 'over-rated' pupils is insecurity (13%).

Positive and negative aspects often have gender connotations. While participation, inspirational contributions, attention, good integration in the class and assertiveness are characteristics found more often in the group of 'over-rated' male pupils, 'over-rated' females are distinguished by obeying rules (51% against 33% for males), engagement (68% against 58% for males) and a cooperative attitude towards class mates (46% against 23% for males) (figure 11). In the 'under-rated' group the only positive characteristic in which males exceed females is quietness (11% against 4% for females), whereas interest, independence, obeying rules and cooperative attitude remain female prerogatives (e.g. interest refers to 57% of under-rated females against 19% of males). With regard to the negative aspects (figure 12), it can be seen that females penalised by their teacher's assessment are distinguished by distraction (44%), insecurity (17% against 5% for males) and poor independence (13% against 11% for males), and whereas 'under-rated' males appear to be more distracted than females (54%), they exceed females in all other negative traits. Untidiness, for example, is quite rare in under-rated girls (4% of cases) but affects almost one under-rated male in four.

Our results seem quite coherent with what found by authors who put into evidence that in the primary school girls receive from teachers more positive feedback about nonintellectual performance than boys, such as neatness or behaviour, which ends up devaluing the positive feedback to their intellectual performance; boys on the contrary receive more negative feedback on their nonintellectual performance, which minimizes academic criticism (Dweck at al. 1978; Gunderson et al., 2012). Moreover according to some authors teachers have the tendency to perceive their best male students as more logical, competitive, independent and passionate about math than their best female students. While the latter’s success is often attributed to their effort and their failures to lack of ability, teachers tend to ascribe boys’ success to their ability and their failures to lack of effort (Fennema et al., 1990, Tiedemann, 2000). Quite analogously we have found that teachers perceive good female students as engaged and good male students as participative, while under-rated female students are defined insecure and poorly independent and under-rated male students untidy and less conform to school rules.

Figure 11: Most frequently mentioned positive aspects in assessments of ‘over-rated’ and ‘under-rated’ pupils, sorted by gender.



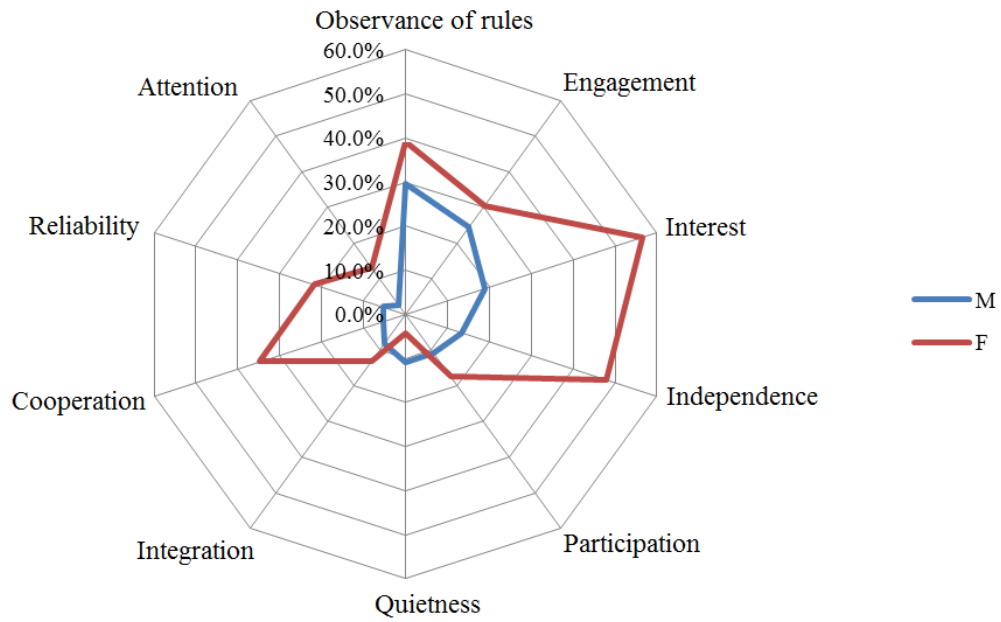
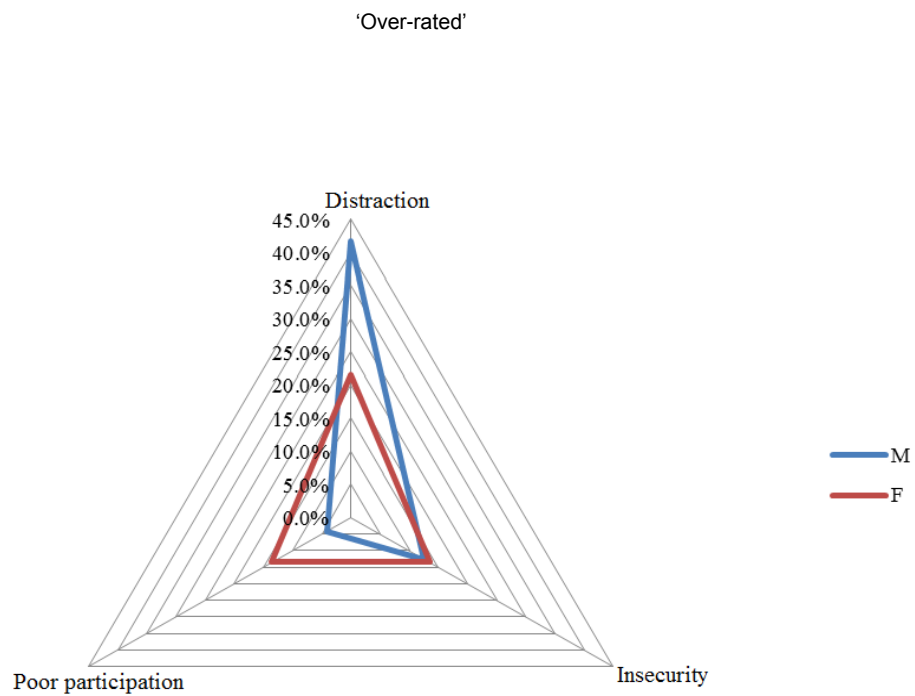
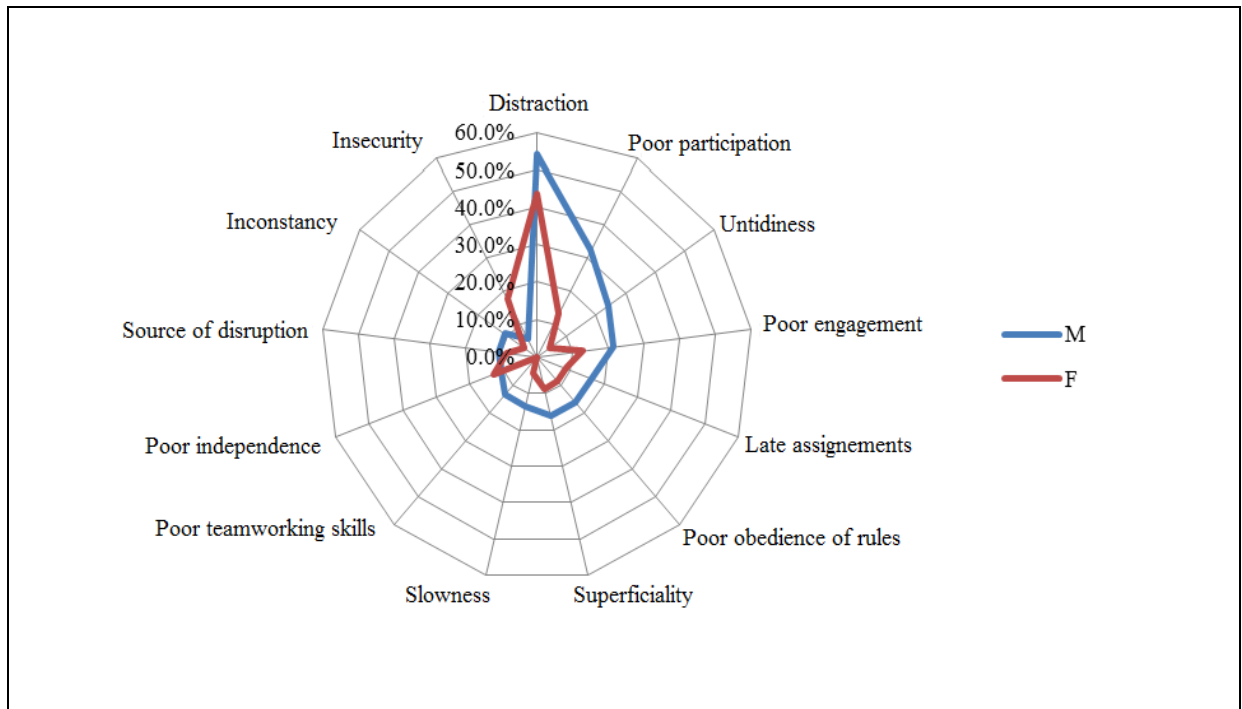


Figure 12: Most frequently mentioned negative aspects in assessments of 'over-rated' and 'under-rated' pupils, sorted by gender.



'Under-rated'



3 Conclusion

In this work, the results of a standardised maths test administered to an entire cohort of fifth grade pupils, the maths marks assigned by teachers to the same cohort and an analysis (albeit partial) of the first term assessments made by teachers of the same pupils have been put together to identify which variables affect maths marks and to determine whether gender differences exist.

The score achieved in the standardised test seems to be consistent with the marks assigned by the teacher, but even though males achieved a significantly higher average score in the test than females, no significant gender differences were observed in the average teacher-assigned marks. This is coherent with studies which put into evidence that boys on average earn lower grades than expected from their performance on ability or achievement test (Duckworth and Seligman, 2006) and the gap in academic success seems to be due to non-cognitive variables which are crucial for earning high grades. If our initial assumption was that males were penalised in teachers' assessments due to their behaviour, and in fact a certain penalisation linked to behaviour exists, a more in-depth analysis allows us to conclude that, all other things being equal, it is girls who are placed at a disadvantage in teachers' assessment. More specifically, a multivariate analysis revealed that although mathematical ability (expressed by the standardised test score) is the most important predictor variable of the teacher's marks, the latter are also positively influenced by the pupil's conduct marks, social status, native nationality and male gender. In other terms given equal maths ability, social status, nationality and conduct marks, males are more likely to achieve maths marks of over 5.0 (very good marks).

Since a large body of research converges on the fact that math ability gender stereotype is in favour of males (Schmader et al. 2004, 2008; Kiefer & Sekaquaptewa, 2006), we have analysed 120 progress reports concerning students whose standardised test results were very inconsistent with the teacher-assigned marks with the goal of putting into evidence the main parameters used by teachers to define their pupils' behaviour profile and to verify if they differ according to gender.

Looking at the 62 students who achieved maths marks of over 5.0 and ended up in the tertile of those less able in the test, it can be seen that males more frequently than females are defined as participating, capable of inspiring, attentive, well integrated in the class and assertive, whereas females appear to be more inclined to obey rules, are seen as more cooperative and more willing to engage, but less participative and more insecure. These findings seem quite coherent with the typical stereotypes according to which boys are more capable in math, whereas girls succeed because they put much effort (Espinoza et al. 2014).

For the purpose of shedding light on the hidden curriculum schools require of pupils, the criteria used for assessment and the mechanisms of transmission of teachers' gender stereotypes, it would be useful, on the one hand, to extend the analysis of assessments contained in the student progress report to the entire cohort of pupils, on the other hand, to carry out a qualitative study by interviewing teachers about what they mean by assessment and students about their experience at school.

Moreover, a qualitative research that explicitly studies how teachers bias plays out in the classroom would serve as a basis for devising strategies that remove prejudice regarding the poor maths skills of females, which end up holding up the glass ceiling (Valian, 1999), and that transmit to pupils of both sexes valuable cross competences that they can use to unleash their potential in their next education cycles and on the labour market when they are adults. If existing research has proven that parents' and teachers' expectancies and attributions of children' math success can influence children' math attitude, performance and, at a further step, career paths (Eccles et al. 1990), it is important to consider that adults' stereotypes are particularly effective and long lasting if presented when children are in the peak age of their gender rigidity, when they tend to believe that only boys or only girls are associated with certain traits, and that only boys can do male-stereotyped activities and only girls can do female-stereotyped activities. (Martin & Ruble, 2004). Pre-school and primary school constitute hence an important field of intervention.

References

- ARCHER, L., HOLLINGWORTH, S., AND MENDICK, H. (2010). *Urban youth and schooling: The aspirations and identities of educationally 'at risk' young people*. Open University Press, UK.
- ASHCRAFT, M.H. AND K.S. RIDLEY (2005). Math anxiety and its cognitive consequences. In J.I.D. Campbell (ed.), *Handbook of Mathematical Cognition*, Psychology Press, New York, NY, pp. 315-327.
- BALFANZ, R., L. HERZOG, L AND D.J. MAC IVER (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist*, Vol. 42/4, pp. 223-235. <https://doi.org/10.1080/00461520701621079>
- BOALER, J. (2003). When learning no longer matters: Standardized testing and the creation of inequality. *Phi Delta Kappan*, 84(7), 502. <https://doi.org/10.1177/003172170308400706>
- BORGNA, C., E STRUFFOLINO, E. (2016). Pushed or pulled? Girls and boys facing early school leaving risk in Italy. *Social Science Research*, XXX, pp. 1-16.
- BRADLEY, C. L., & RENZULLI, L. A. (2011). The complexity of non-completion: Being pushed or pulled to drop out of high school. *Social Forces*, 90(2), 521-545. <https://doi.org/10.1093/sf/sor003>
- BRAY, R., GARDNER, C., PARSONS, N., DOWNES, P., & HANNAN, G. (1997). *Can boys do better?* Leicester: Secondary Heads Association.
- BUCK, S., RITTER, G. W., JENSEN, N. C., & ROSE, C. P. (2010). Teachers say the most interesting things—An alternative view of testing. *Phi Delta Kappan*, 91(6), 50-54. <https://doi.org/10.1177/003172171009100613>
- CIZEK, G. J., FITZGERALD, S. M., AND RACHOR, R. E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179. https://doi.org/10.1207/s15326977ea0302_3
- CONSORZIO PISA.CH (2011). *PISA 2009: Risultati regionali e cantonali*. Berna e Neuchatel: UFFT/CDPE e Consorzio PISA.ch.
- CONSORZIO PISA.CH (2014). *PISA 2012 : Études thématiques*. Berne et Neuchâtel : SEFRI/CDIP et Consortium PISA.ch.
- CRESCENTINI, A. (2017). *Prove standardizzate ticinesi. Matematica nella classe V Scuola Elementare*. Locarno: Centro Innovazione e Ricerca sui Sistemi Educativi.
- DI PRETE, T. A., AND BUCHMANN, C. (2013). *The rise of women: The growing gender gap in education and what it means for American schools*. Russell Sage Foundation.
- DIVISIONE DELLA SCUOLA (2015). *Piano di studio della scuola dell'obbligo ticinese*. Repubblica e Cantone del Ticino, Dipartimento dell'educazione, della cultura e dello sport, Bellinzona. Online: <http://www.pianodistudio.ch/>.

- DUCKWORTH, A. L., AND SELIGMAN, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of educational psychology*, 98(1), 198. <https://doi.org/10.1037/0022-0663.98.1.198>
- DWECK, C. S., DAVIDSON, W., NELSON, S., AND ENNA, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Developmental psychology*, 14(3), 268. <https://doi.org/10.1037/0012-1649.14.3.268>
- ECCLES, J.S.; JACOBS, J.E.; HAROLD, R.D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of social issues*, 46 (2),183-201. <https://doi.org/10.1111/j.1540-4560.1990.tb01929.x>
- ESPINOZA, P., DA LUZ FONTES, A. B. A., & ARMS-CHAVEZ, C. J. (2014). Attributional gender bias: teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, 17(1), 105-126. <https://doi.org/10.1007/s11218-013-9226-6>
- FENNEMA, E., PETERSON, P. L., CARPENTER, T. P., AND LUBINSKI, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, 21(1), 55-69. <https://doi.org/10.1007/BF00311015>
- GUNDERSON, E. A., RAMIREZ, G., LEVINE, S. C., AND BEILOCK, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, 66(3-4), 153-166. <https://doi.org/10.1007/s11199-011-9996-2>
- GUSKEY, T. R. (2002). Computerized gradebooks and the myth of objectivity. *Phi Delta Kappan*, 83(10), 775-780. <https://doi.org/10.1177/003172170208301013>
- JACOBS, J. E., DAVIS-KEAN, P., BLEEKER, M., ECCLES, J. S., AND MALANCHUK, O. (2005). I can, but I don't want to. The impact of parents, interests, and activities on gender differences in math. In A. Gallagher & J. Kaufman (Eds.), *Gender difference in mathematics*, 246-263.
- KIEFER, A. K., AND SEKAQUAPTEWA, D. (2006). Implicit stereotypes, gender identification, and math-related outcomes: A prospective study of female college students. *Psychological Science*, 18(1), 13-18. <https://doi.org/10.1111/j.1467-9280.2007.01841.x>
- KLING, K. C., NOFTLE, E. E., AND ROBINS, R. W. (2012). Why do standardized tests underpredict women's academic performance? The role of conscientiousness. *Social Psychological and Personality Science*, 4(5), pp. 600-606. <https://doi.org/10.1177/1948550612469038>
- MARTIN, C. L., AND RUBLE, D. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current directions in psychological science*, 13(2), 67-70. <https://doi.org/10.1111/j.0963-7214.2004.00276.x>
- MCNEAL, R.B. (2011). Labor market effects on dropping out of high school: Variation by gender, race, and employment status. *Youth & Society*, 305–332. <https://doi.org/10.1177/0044118X10363776>
- MICKELSON, R. (1989). Why does Jane read and write so well?: The anomaly of women's achievement. *Sociology of Education* 62 (1), 47-63. <https://doi.org/10.2307/2112823>
- NESSE (2009). *Gender and Education (and Employment): Gendered imperatives and their implications for women and men: Lessons from research for policy makers*. European Commission, Directorate-General for Education and Culture.

- OECD (2010). PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I), online: <http://dx.doi.org/10.1787/9789264091450-en>.
- OECD (2014). PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014), PISA, OECD Publishing, online: <http://dx.doi.org/10.1787/9789264201118-en>.
- OECD (2015). The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence, PISA, OECD Publishing. Online: <http://dx.doi.org/10.1787/9789264229945-en>.
- OECD (2016). PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA, OECD Publishing, Paris, online: <http://dx.doi.org/10.1787/9789264266490-en>.
- OREOPOULOS, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, Vol. 91/11-12, pp. 2213-2229. <https://doi.org/10.1016/j.jpubeco.2007.02.002>
- RANDALL, J., & ENGELHARD, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372-1380. <https://doi.org/10.1016/j.tate.2010.03.008>
- RASCH, G. (1961). On general laws and meaning of measurement in psychology. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, (4), 321-333. Berkeley, CA: University of California Press.
- RUMBERGER, R.W. (2011), *Why Students Drop Out of High School and What Can Be Done About It*, Harvard University Press, Cambridge, MA. <https://doi.org/10.4159/harvard.9780674063167>
- SALVISBERG, M., ZAMPIERI, S., AND EGLOFF (2016). PISA 2015: primi risultati per il Canton Ticino. Rapporto presentato alla conferenza stampa del Dipartimento dell'educazione, della cultura e dello sport il 6.12.2016 a Bellinzona. Online: <http://www3.ti.ch/DECS/cartellastampa/pdf-cartella-stampa-84991392517.pdf>
- SCHMADER, T., JOHNS, M., AND BARQUISSAU, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex roles*, 50(11-12), 835-850. <https://doi.org/10.1023/B:SERS.0000029101.74557.a0>
- SCHMADER, T., JOHNS, M., AND FORBES, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological review*, 115(2), 336. <https://doi.org/10.1037/0033-295X.115.2.336>
- SCHOON, I., AND ECCLES, J. S. (edited by). (2014). *Gender differences in aspirations and attainment: A life course perspective*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139128933>
- STEELE, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, 52(6), 613. <https://doi.org/10.1037/0003-066X.52.6.613>
- STEELE, C. M., AND ARONSON, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of personality and social psychology*, 69(5), 797. <https://doi.org/10.1037/0022-3514.69.5.797>
- STRAND, S., DEARY, I. J., AND SMITH, P. (2006). Sex differences in cognitive abilities test scores: A UK national picture. *British Journal of Educational Psychology*, 76(3), 463-480. <https://doi.org/10.1348/000709905X50906>

- TIEDEMANN, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics*, 41(2), 191-207. <https://doi.org/10.1023/A:1003953801526>
- VALIAN, V. (1999). *Why so slow?: The advancement of women*. MIT press.
- VAN HOUTTE, M. (2004). Why boys achieve less at school than girls: The difference between boys' and girls' academic culture. *Educational Studies*, Vol. 30/2, pp. 159-173. <https://doi.org/10.1080/0305569032000159804>
- ZANOLLA, G. (2015). Percorsi scolastici e certificazioni nella formazione professionale iniziale. In: Cirse (2015) (edited by) *Scuola a tutto campo. Indicatori del sistema educativo ticinese*, SUPSI-DFA, Locarno.
- ZEIDNER, M., AND G. MATTHEWS (2011). *Anxiety 101*, Springer, New York, NY.