



# Do boys and girls learn the same way? A preliminary study in Primary Education analyzing gender differences

**Estíbaliz Aragón<sup>1</sup>, Noemí Serrano<sup>2</sup>, Jose Ignacio Navarro<sup>1</sup>**

---

<sup>1</sup> Department of Psychology, University of Cádiz, Puerto Real  
<sup>2</sup> Department of Didactic, University of Cádiz, Puerto Real

---

**Spain**

*Correspondence:* Estíbaliz Aragón Mendizábal. Avenida República Saharaui, s/n. CP. 11.510, Puerto Real-Cádiz (Spain). E-mail: [estivaliz.aragon@uca.es](mailto:estivaliz.aragon@uca.es)

---

© Universidad de Almería and Ilustre Colegio Oficial de la Psicología de Andalucía Oriental (Spain)

## Abstract

**Introduction.** Differences in learning mathematics between boys and girls are being controversial according to different studies. However, those divergences are not always supported by empirical studies. One of the specific sources of difficulties on this topic is the fluency calculation learning, where performance differences between boys and girls need be clarified by new research, in order to gain more differentiated insights in possibly emerging gender differ, because the social and gender consequences. The current study analysed the gender differences in fluency calculation.

**Method.** Participants were students of second cycle of primary education ( $n = 138$ ), aged from 101-133 months old (8 year and 4 months to 11 year). Participants were 50 girls and 88 boys.

**Results.** Statistical analyses did not showed differences between boys and girls in the 3rd grade of primary education. However, in the 4th grade the differences did reveal significant differences in the fluency calculation task. These results were supported by effect size and mutivariant analysis calculation.

**Discussion and Conclusion.** Data suggests a gender gap around aged 9 between boys and girls, that would benefit boys in fluency calculation scores. Results, limitations and future potential research of the topic are discussed.

**Keywords:** Fluency Calculation, Gender, Primary Education, Mathematical Skills

## Resumen

**Introducción.** Las diferencias en el aprendizaje de las matemáticas entre niños y niñas son controvertidas según diferentes estudios. Sin embargo, esas divergencias no siempre son respaldadas por estudios empíricos. Uno de los aspectos específicos de dificultad es el aprendizaje en la fluidez de cálculo: ser capaz de calcular con precisión y rapidez. El presente estudio analizó las diferencias de género en la fluidez de cálculo.

**Método.** Los participantes fueron estudiantes de segundo ciclo de educación primaria ( $n = 138$ ), con edades comprendidas entre 101-133 meses (8 años y 4 meses a 11 años). Los participantes fueron 50 niñas y 88 niños.

**Resultados.** Los análisis estadísticos no mostraron diferencias entre niños y niñas en el 3er grado de educación primaria. Sin embargo, en el 4º grado, las diferencias se mostraron significativas en la tarea de fluidez de cálculo. Estos resultados fueron respaldados por el análisis multivariante y el cálculo del tamaño del efecto.

**Discusión and Conclusión.** Los datos sugieren una brecha de género en torno a los 9 años entre niños y niñas, que beneficiaría a los niños en las puntuaciones de fluidez de cálculo. Se discuten los resultados, las limitaciones y las futuras líneas de investigación.

**Palabras Clave:** Fluidez de cálculo, Género, Educación Primaria, Habilidades Matemáticas.

## Introduction

Adequate knowledge of mathematics is related to progress and future of society, and institutions are therefore concerned with implementing actions to cope with difficulties that may arise, among other reasons, because of the high cost involved the low mathematical competence in the population (Hudson, Price, & Gross, 2009; Schrøter-Joensen & Skyt-Nielsen, 2010).

In order to compare student's performance between different countries, one of the most international prestigious educational report is PISA report (OECD, 2014). The last PISA version provides information on students' knowledge and skills in dealing with problems and activities of specific content, such as those related to mathematics and problem solving. For some countries, the results are not truthfully satisfactory, and describes a condition that can be worse over time. For example, the average score achieved in mathematical competence in for the Spanish students do not reach the average values of OECD.

The importance of mathematical competence lies not only in the economic implications (productivity, employability, incomes...) it has on society (Schrøter-Joensen & Skyt-Nielsen, 2010). These competences are fundamental in every daily life, to carry out tasks as simple as, for example, make a purchase. However, there is a lot of diversity in the students according to what the level of mathematical competence is concerned. There are studies that indicate that these differences among students appear in Early Childhood Education and that basic competencies predict the performance in school some years later (Cerdeira, Pérez, Navarro, Aguilar, Casas, & Aragón, 2015; Morgan, Farkas, Aunio, Heiskari, Van Luit, & Vuorio, 2015; Jordan, Kaplan, Ramineni and Locuniak & Wu, 2009). Currently, it is estimated that between 3 and 8% of primary school children have some Math Learning Difficulty (MLD) (González-Castro, Rodríguez, Cueli, Cabeza, & Álvarez, 2014). There are several factors that should have some empirically identified impact on children mathematics performance. Those variables were identified as specific and general cognitive predictors of early mathematics. Over the past decades relevant research could demonstrate students' gender to contribute to increase the differences in mathematical learning. First, those studies that highlights on the existence of gender differences (Bedard & Cho, 2010; Dickerson, McIntosh & Valente, 2015, Doris, O'Neill & Sweetman, 2013, Fryer & Levitt, 2010; Freeman, 2004, Geiser & Lehmann, 2008, Husain & Millimet, 2009, Jordan, Kaplan, Nabors-Olah & Locuniak, 2006 Levine et

al., 2005 Penner & Paret 2008 Robinson & Lubienski, 2011). These studies locate the gender gap between 9 and 10 years of age (Doris et al., 2013, Fryer & Levitt, 2010, Freeman, 2004, Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, Lubienski, 2011) with children showing the best skills in mathematics. According to this research, the gap would continue to grow and, over time, would further differentiate boys from girls as development progressed (Husain & Millimet, 2009; Penner & Paret, 2008; Robinson & Lubienski 2011). These differences would also be significant in later stages, i.e. during Secondary Education and in adulthood (Dee, 2007, Geary and Lehmann, 2008, Penner & Paret, 2008, Robinson & Lubienski, 2011) when differences would stabilize and the gap would remain constant and stable (Bedard & Cho, 2010).

Second, other studies support the lack of gender differences in mathematics throughout the life cycle (Aragón, Delgado, Aguilar, Araújo & Navarro, 2013; Aragón & Navarro, 2016, Hyde, Lindberg, Linn, Ellis & Williams, 2008; Lachance & Mazzocco, 2006; Lindberg, Hyde, Petersen & Linn, 2010; Navarro et al., 2010; Worell & Goodheart, 2006). These studies would be in the line of work that Hyde (2005) called as hypothesis of gender similarity. Her meta-analysis and a subsequent one (Lindberg et al., 2010) showed that gender was not a differentiating factor between males and females in terms of mathematical abilities, and this condition is extensive at any age or vital stage. Symbolic representation of magnitude is a critical predictor of academic performance and is the basis of higher order skills, such as calculation (Honoré & Noël, 2016). Consequently, algorithm proper use leads to success in mathematics (Lyons, & Ansari, 2015). Also, the calculation is thought to provide a foundation for solving problems (Fung, Swanson, & Orosco, 2014) that are present in daily life. Some results differences should be consequence of methodological issues. In a well known study by Tiedemann & Faber (1994) found decreasing gender differences in elementary math performance variables, as well as varying gender differences in math subskills (e.g., Carr & Davies, 2001).

Regarding the out-findings on the literature about the existence or not of gender differences in mathematical competence, the present work aims to verify the existence of a gender gap in the second cycle of Primary Education (children age average is around 9 years), in fluency calculation. Calculation was selected as the main aim because it is considered one of the most important skills in carrying out an arithmetic operation, in a flexible, precise, efficient and appropriate way (Kilpatrick, Swafford and Findell 2001). For this reason, fluency

calculation is presented as critical in problems arithmetic solving (Fuchs et al., 2006, 2010). Problem solving is essential in everyday life, and often difficult for children of school age as shown by PISA international performance assessment reports (OECD, 2014).

Consequently, if gender is seen as a differentiating factor between boys and girls in fluency calculation, it should be appropriate to address this issue in order to propose actions to decrease these differences, and provide the necessary educational support to girls to prevent possible difficulties.

### *Aims and hypothesis*

The main target of this study was analysing the impact of age (grade level) and gender on students' calculation skills. Related with this, the hypothesis was that differences between girls and boys in mathematic performance should recue with grade level in Primary Education.

## **Method**

### *Participants*

The participants were a total of 138 students whose ages ranged from 101 to 133 months ( $M = 113.36$ ;  $sd = 7.06$ ) and were distributed in 6 classrooms of Primary Education (3<sup>rd</sup> and 4<sup>th</sup> grades). The school is located in Cádiz, (Spain) in a middle class socio-economic level. Participants 64 students were in third grade ( $M = 107$ ;  $sd = 3.35$ ) and 74 were in fourth grade ( $M = 118.86$ ;  $sd = 7.06$ ). Third year students, 20 were girls ( $M = 106.05$ ,  $sd = 3.96$ ) and 44 boys ( $M = 107.43$ ,  $sd = 2.98$ ). The fourth year, 30 participants were girls ( $M = 117.77$ ,  $sd = 3.64$ ) and 44 were boys ( $M = 119.61$ ,  $sd = 4.45$ ) (table 1). Those students who showed special educational needs were removed from the total sample. Student participation was voluntary and there was an informed consent of parents and teachers.

Table 1. Age distribution according to grade and gender

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>sd</i>
3 <sup>rd</sup> Grade	64	101	112	107	3.35
Girls	20	101	112	106.05	3.96
Boys	44	101	112	107.43	2.98
4 <sup>th</sup> Grade	74	113	133	118.86	4.22
Girls	30	113	127	117.77	3.64
Boys	44	113	133	119.61	4.45
Total sample	138	101	133	113.36	7.06
Total girls	50	101	127	113.08	6.89
Total boys	88	101	133	113.52	7.19

### *Instruments*

*Fluency Calculation.* Test. For the present study, a Woodcock-Johnson III-Test of Achievement (WJ-III-ACH) subtest was used to evaluate intellectual abilities and academic achievement. A Spanish validated version named Spanish Version of the Woodcock-Johnson Mathematics Achievement Tests (Battery III APROV) (Diamantopoulou, Pina, Valero-García, González-Salinas & Fuentes, 2012) was completed. This test consists of 160 simple arithmetic operations distributed in 16 rows with 10 operations each. All operations are presented on a sheet with two printed sides. On one side we find 80 operations, and on the back side other 80 operations. The test assesses fluency in addition, subtraction and multiplication operations of a single digit. The time of administration was 3 minutes. The evaluated student must complete as many operations as possible. . As a criterion of reliability (internal consistency), the test showed an alpha Cronbach value of  $\alpha = .90$ .

### *Procedure*

Parents gave informed written consent for their child's participation. An assessment session was planned by class, as the activity was applied collectively. The total of the session lasted ten minutes taking into account the duration of the test, the explanation of the instruction and the possible issues that arose before administration. The assessment session was carry out in the ordinary classroom in which the students received regular classes during school hours, taking into account the rest periods of the students. The instructions were that students have to complete the tasks as quickly as possible and without mistakes; in case they had problems solving any math facts they were encouraged to continue with the next one. Attention was also paid to mathematical signs to avoid errors, according to the test instructions.

### *Data Analysis*

The aim of this study was to analyze gender and grade differences in fluency calculation at Primary Education. The SPSS.22 version was used for data analysis. Through this software, descriptive statistics of the sample were calculated and inferential analysis' (Student's *t*-Test, and MANOVA test was calculated for comparisons grade and gender was carried out to study differences between boys and girls. Also, the most used in Psychology research Cohen's effect size was calculated as complementary information.

## **Results**

The main objective of the study was to analyze the existing gender and grade differences in fluency calculation at the beginning of the second cycle of Primary Education (3<sup>rd</sup> and 4<sup>th</sup> grades), in order to analyse if these differences were maintained or even increased at the end of the cycle.

First, the students' performance in the provided fluency calculation test was descriptively analysed. For this purpose, we studied the distribution of correct and incorrect responses and efficacy in the fluency calculation test, being this the number of correct responses obtained considering the correction due to the incorrect response (table 2).



Table 2. Means (M), standard deviations (sd) for the students according to efficacy, correct and incorrect responses.

	Correct Responses <i>M (sd)</i>	Incorrect Responses <i>M (sd)</i>	Efficacy (Correct-Incorrect) <i>M (sd)</i>
3 <sup>rd</sup> Grade	50.98 (14.83)	2.28 (3.89)	48.70 (16.54)
4 <sup>th</sup> Grade	52.59 (12.63)	.99 (2.14)	51.60 (13.37)

Then, when the scores are presented using gender as a differentiating criterion, there were important differences between boys and girls. With regard to success and measure of effectiveness, both boys and girls improve similarly (about two or three points a year). However, the data show how the performance of boys is higher than girls, even the performance of boys in third was better than that of girls in fourth grade.

On the other hand, the incorrect responses followed a similar pattern that did not differ according to gender. Both boys and girls made a similar proportion of errors that were significantly reduced in the fourth year, decreasing them by half (table 3).

Table 3. Mean and standard deviations for participants based on gender and grade.

	Boys		Girls	
	3 <sup>rd</sup> grade <i>M (sd)</i>	4 <sup>th</sup> grade <i>M (sd)</i>	3 <sup>rd</sup> grade <i>M (sd)</i>	4 <sup>th</sup> grade <i>M (sd)</i>
Correct	53.32 (15.44)	55.52 (12.87)	45.85 (12.23)	48.30 (11.13)
Incorrect	2.32 (4.16)	1.07 (2.54)	2.20 (3.31)	.87 (.40)
Efficacy (C-I)*	51 (17.39)	54.45 (13.88)	43.65 (13.55)	47.43 (11.58)

To determine if the differences were significant on the fluency calculation test results, a hypothesis test was performed for independent samples, the Student's *t*-Test, and the effect size was calculated. The results are shown in table 4.

Table 4. Results of hypothesis contrast and effect size (Cohen's  $d$  and  $r$ ) for gender differences in fluency calculation.

	<i>3<sup>rd</sup> grade</i>			<i>4<sup>th</sup> grade</i>		
	$p$	$r$	$d$	$p$	$r$	$d$
Correct	.061	.25	.53	.015*	.28	.60
Incorrect	.911	.01	.03	.695	.05	.10
Efficacy	.100	.22	.47	.026*	.26	.54

\* $p < .05$ 

Results of hypothesis contrast show the differences between boys and girls in the descriptive analysis. This contrast was not statistically significant in the third graders ( $p < .061$ ) in correct responses and efficacy ( $p < .100$ ). However, the differences in incorrect responses were significant in fourth graders ( $p < .015$ ) and in efficacy ( $p < .026$ ). With regard to incorrect responses, there were no significant differences between boys and girls. The girls were similar to the boys in third grade, and even more accurate than boys in fourth grade although the number of operations performed was significantly lower.

Likewise, the size of the effect was calculated in order to analyse the results obtained and the impact of the gender on the performance of fluency calculation. Cohen's  $d$  (1988), and the Pearson's correlation coefficient were used as measures of the standardized effect size. For its correct interpretation, Cohen (1988) introduced a series of ranges explaining the effect size. According to the value  $d$ , the effect can be classified in: small effect (from  $r = .10$ ;  $d = .20$ ), medium (from  $r = .24$ ;  $d = .50$ ), and large (from  $r = .37$ ;  $d = .80$ ). The 3rd grade results showed a moderate effect size for incorrect responses in the fluency calculation test ( $r = .25$ ;  $d = .53$ ) and small in efficacy ( $r = .22$ ;  $d = .47$ ). However, in 4th grade results were medium in correct responses ( $r = .28$ ;  $d = .60$ ) and efficacy ( $r = .26$ ;  $d = .54$ ). No effect size was observed on incorrect responses in any of the grade.

To obtain more information on whether the variable gender and grade was able to differentiate the performance of boys and girls, a MANOVA analysis was carried out. This statistical calculation allows verifying if there were significant differences between the levels or groups of a categorical-factor, by comparing of equality of means. For this calculations, we used more than one factor (table 5).

Table 5. Interaction gender and grade MANOVA analysis results

Effect		<i>F</i>	<i>Sig.</i>	Partial Eta Squared
Gender	Pillai's Trace	5.578	.005*	.077
Grade	Pillai's Trace	2.698	.071	.039
Gender-Grade	Pillai's Trace	.003	.999	.000

(\*)  $p < .001$

Lévene's test homoscedasticity was computed obtaining equal variance ( $p > 0.05$ ) for correct answers ( $F_{(3, 134)} = 1.18$ ,  $p = .317$ ); incorrect answers ( $F_{(3, 134)} = 1.77$ ,  $p = .154$ ), and efficacy ( $F_{(3, 134)} = 1.27$ ,  $p = .286$ ). In this study, the Pillai's trace was selected (table 4) as a statistically robust criterion to assess differences between dimensions of the dependent variables. For this MANOVA, the multivariate contrasts obtained with the Pillai's trace found differences, based on the gender individually taken, but not grade. In addition, no significant differences were founded according to the interaction of gender versus grade ( $F_{(gender)} = 5.578$ ,  $p < .05$ ;  $F_{(grade)} = 2.698$ ,  $p > .05$ ).

Tabla 6. Gender and grade MANOVA interaction effects.

Source	Dependent Variable	<i>F</i>	<i>Sig.</i>	Partial Eta Squared
Gender	Correct	9.426	.003*	.066
	Incorrect	.082	.775	.001
	Efficacy	7.491	.007*	.053
Grade	Correct	.946	.332	.007
	Incorrect	5.379	.022*	.039
	Efficacy	1.900	.170	.014
Gender - Grade	Correct	.003	.959	.000
	Incorrect	.006	.940	.000
	Efficacy	.004	.950	.000

(\*)  $p < .05$

Intersubjects MANOVA effects indicated that based on the method, there were some statistical significant differences. The gender showed a significant effect in correct answers ( $F = 9.42$ ;  $p = .003$ ), and efficacy ( $F = 7.49$ ;  $p = .007$ ). The grade showed a significant effect

only in incorrect answers ( $F = 5.37$ ;  $p = .022$ ). No significant differences were found in gender-grade interaction. Table 6 shows the effect size calculating means of the Statistic partial eta Squared. This establishes the effect differences in the following way:  $\eta_p^2 = 0.02$  small;  $\eta_p^2 = 0.13$  medium and  $\eta_p^2 = 0.26$  large (Bakeman, 2005). Consequently, the effect of gender was small in correct responses ( $\eta_p^2 = .066$ ) and efficacy ( $\eta_p^2 = .053$ ). Considering grade, the difference in errors ( $\eta_p^2 = .039$ ) had also a small effect size value.

### Discussion or Conclusions

Skills in mathematics, and specifically in fluency calculation, are a convenient resource in people's lives, because they are involved in many daily activities (Fuchs et al., 2010). This feature is one of the main reasons for improving calculation skills. Calculation is not limited to the memorization of numerical facts, but require the acquisition and management of rules and concepts that result in an adequate and fluency calculation (Baroody, Bajwa, & Eiland, 2009). Therefore, the interest of researchers is analysing what factors may influence a higher or lower performance in calculation.

Overall, the descriptive analyses presented in this study showed an improvement in fluency calculation related to the academic grade, and age. Specifically, it was observed how from one grade to the next, the correct responses' mean and efficacy in calculation significantly increased, while at the same time the numbers of incorrect responses decreased. Therefore, fluency calculation become more precise.

Other concern about calculation is the controversial about gender differences existing throughout third and fourth grade of Primary Education in fluency calculation. Both school grade are considered essential to this line of research that maintains differences between boys and girls, as the gap would be around age 9 (Doris et al, 2013; Fryer & Levitt, 2010; Freeman, 2004 Levine et al., 2005; Robinson & Lubienski, 2011). Results in this study support gender differences in 3<sup>rd</sup> grade ( $M = 107$  months, i.e. 8.9 years), but do not become statistically significant ( $p > .05$ ). However, in the 4<sup>th</sup> grade ( $M = 118.86$  months, i.e. 9.9 years) significant differences were found ( $p < .05$ ). These results were maintained by the multivariate analysis. Considering gender a significant effect in correct answers and efficacy in fluency calculation task was found, but effect size was small. However, no significant differences were found in efficacy and correct answers considering grade, and gender-grade interaction. That would

be explained as at that age girls begin to show a characteristic and distinctive profile in mathematic.

On the other hand, the size of the effect also showed quantitative information on the magnitude of these existing differences between boys and girls. First, the effect size (Cohen's  $d$ ) for the correct responses in the test was moderate in both grades, increasing in the latter. Second, with regard to efficacy, the size of the effect (Cohen's  $d$ ) changed from small, in the third year, to moderate effect in the fourth year. These results support the existence of a gender gap that it seems to increase as the age advances. Thirdly, with respect to incorrect response, it appears that there were no differences between boys and girls. Possibly, the differences were focus on fluency rather than accuracy when performing the calculation. Other potential explanative perspective comes up from prominent research lines, concerning the role of social-cognitive or cognitive-motivational factors (self-beliefs) to explain differences (e.g., Frenzel, Pekrun & Goetz, 2007; Stipek & Gralinski, 1991; Ross, Bruce & Scott, 2012; Weidinger, Steinmayr & Spinath, 2018).

Acquisition of such fluency calculation in basic math facts is critical for more complex tasks solving, such as multi-digit operations (Kilpatrick et al., 2001). Consequently, if girls show less fluency when carrying out simple computation, the more complex calculations are, the more the difficulty will be. Consequently an increasing gender difference should be expected.

Although the analyses carried out in this study emphasized the differences in fluency calculation in 4<sup>th</sup> grade of Primary Education among boys and girls, we would like to point out some of the limitations of the study. Regarding to methodological weaknesses, we can mention those related to randomness and sample size. In fact, this study is presented as preliminary, with the purpose of expanding the sample of participants. In addition, a comparison with other grades in primary education should be convenient, specially between 1<sup>st</sup> and 6<sup>th</sup> grades, using a longitudinal design. Thus, it would confirm whether differences persist or even increase as school grades and the age of students grows. This should be expected according to previous research (Bedard & Cho, 2010; Dickerson, McIntosh & Valente, 2015; Doris, O'Neill & Sweetman, 2013), and results obtained in this study.

## References

- Aragón, E. L., Delgado, I., Aguilar, M., Araújo, A., & Navarro, J. I. (2013). Estudio de la influencia de la inteligencia y el género en la evaluación matemática temprana. *European Journal of Education and Psychology*, 6(1), 5–18.
- Aragón, E., & Navarro, J.I. (2016). Exploring gender differences in general and specific-domain predictors of early math skills. *Suma Psicológica*, 23 (2), 71-79. [doi:10.1016/j.sumpsi.2016.04.001](https://doi.org/10.1016/j.sumpsi.2016.04.001)
- Aunio, P., Heiskari, P., Van Luit, J.E., & Vuorio, J. (2015). The development of early numeracy skills in kindergarten in low, average and high performance groups. *Journal of Early Childhood Research*, 13(1), 3-16. doi:10.1177/1476718X14538722
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, 37(3), 379-384
- Baroody, A. J., Bajwa, N. P., & Eiland, M. (2009). Why can't Johnny remember the basic facts? *Developmental Disabilities Research Review*, 15, 69-79. doi:[10.1002/ddrr.45](https://doi.org/10.1002/ddrr.45)
- Bedard, K., & Cho, I. (2010). Early gender test score gaps across OECD countries. *Economics of Education Review*, 29, 348–363. doi:10.1016/j.econedurev.2009.10.015
- Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary educational psychology*, 26(3), 330-347. doi:10.1006/ceps.2000.1059
- Cerda, G., Pérez, C., Navarro, J.I., Aguilar, M., Casas, J.A., & Aragón, E. (2015). Explanatory model of emotional-cognitive variables in school mathematics performance: a longitudinal study in primary school. *Frontiers in Psychology*, 6 (1363). doi:10.3389/fpsyg.2015.01363
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence.
- Diamantopoulou, S., Pina, V., Valero-García, A. V., González-Salinas, C., & Fuentes, L. J. (2012). Validation of the Spanish version of the Woodcock-Johnson mathematics achievement tests for children aged six to thirteen. *Journal of Psychoeducational Assessment*, 30, 466-477.
- Dickerson, A., McIntosh, S. & Valente, C. (2015). Do the maths: An analysis of the gender gap in mathematics in Africa. *Economics of Education Review*, 46, 1–22. [doi:10.1016/j.econedurev.2015.02.005](https://doi.org/10.1016/j.econedurev.2015.02.005)

- Doris, A., O'Neill, D. & Sweetman, O. (2013). Gender, single-sex schooling and maths achievement. *Economics of Education Review*, 35, 104–119. [doi:10.1016/j.econedurev.2013.04.001](https://doi.org/10.1016/j.econedurev.2013.04.001)
- Freeman, C. E. (2004). *Trends in educational equity of girls and women: 2004 (NCES 2005-016)*. Washington, DC: U. S. Government Printing Office: U.S. Department of Education National Center for Education Statistics.
- Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics—A “hopeless” issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, 22(4), 497. doi:10.1007/BF03173468
- Fryer, R., & Levitt, S. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2), 210–240. [doi:10.1257/app.2.2.210](https://doi.org/10.1257/app.2.2.210)
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., ... Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology*, 46(6), 1731–1746.
- Fuchs, L. S., Fuchs, D., Compton, D.L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*. 98, 29–43. doi:[10.1037/0022-0663.98.1.29](https://doi.org/10.1037/0022-0663.98.1.29)
- Fung, W.W., Swanson, H. L., & Orosco, M.J. (2014). Influence of reading and calculation on children at risk and not at risk for word problem solving: Is math motivation a mediator?, *Learning and Individual Differences*, 36, 84-91. doi.org/10.1016/j.lindif.2014.10.011.
- Geiser, W., & Lehmann, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence*, 36(6), 556–563. [doi:10.1016/j.intell.2007.12.003](https://doi.org/10.1016/j.intell.2007.12.003)
- González-Castro, P., Rodríguez, C., Cueli, M., Cabeza, I., & Álvarez, I. (2014). Competencias matemáticas y control ejecutivo en estudiantes con Trastorno por Déficit de Atención con Hiperactividad y Dificultades de Aprendizaje de las Matemáticas, *Revista de Psicodidáctica*, 19, 125-143. doi:10.1387/RevPsicodidact.7510
- Honoré, N., & Noël, M.P. (2016). Improving Preschoolers' Arithmetic through Number Magnitude Training: The Impact of Non-Symbolic and Symbolic Training. *PLoS ONE*, 11(11). doi.org/10.1371/journal.pone.0166685
- Hudson, C., Price, D., & Gross, J. (2009). *The long-term costs of numeracy difficulties*. London, UK: Every Child a Chance Trust.

- Husain, M., & Millimet, D. (2009). The mythical ‘boy crisis’? *Economics of Education Review*, 28, 38–48. doi:10.1016/j.econedurev.2007.11.002
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6), 581–592. doi:10.1037/0003-066X.60.6.581
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321, 494–495. doi:10.1126/science.1160364
- Jordan, N. C., Kaplan, D., Nabors-Olah, L., & Locuniak, M. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77(1), 153–175. doi:10.1111/j.1467-8624.2006.00862.x
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45, 850–867. doi:10.1037/a0014939
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington DC: National Academies Press.
- Lachance, J., & Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences*, 16(3), 195–216. doi:10.1016/j.lindif.2005.12.001
- Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science*, 16, 841–845. doi:10.1111/j.1467-9280.2005.01623.x
- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136, 1123–1135. doi:10.1037/a0021276
- Lyons, I. M., & Ansari, D. (2015). Foundations of Children's Numerical and Mathematical Skills. *Advances in Child Development and Behavior*, 48, 93–116. doi:10.1016/bs.acdb.2014.11.003.
- Morgan, P., L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, 42, 306–321. doi:10.1177/0022219408331037
- Navarro, J. I., Aguilar, M., García, M., Menacho, I., Marchena, E., & Alcalde, C. (2010). Diferencias en habilidades matemáticas tempranas en niños y niñas de 4 a 8 años. *Revista Española de Pedagogía*, 68, 85–98.



- OECD (2014). Programme for international student assessment PISA 2012. Results in Focus. Retrieved from <http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf>
- Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research*, 37, 239–253. doi:10.1016/j.ssresearch.2007.06.012
- Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, 48(2), 268–302. doi:10.3102/0002831210372249
- Ross, J. A., Scott, G., & Bruce, C. D. (2012). The gender confidence gap in fractions knowledge: Gender differences in student belief–achievement relationships. *School Science and Mathematics*, 112(5), 278–288. doi:10.1111/j.1949-8594.2012.00144.x
- Schröter-Joensen, J., & Skyt-Nielsen, H. (2010). Is there a causal effect of high school math on labor market outcomes. *The Journal of Human Resources*, 44(1), 171–198. doi:10.1353/jhr.2009.0004
- Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, 83(3), 361.
- Tiedemann, J., & Faber, G. (1994). Ist Mathe nichts für Mädchen? Mädchen unterschätzen sich im Mathematikunterricht. *Die Grundschulzeitschrift*, 8(94), 74.
- Weidinger, A. F., Steinmayr, R., & Spinath, B. (2018). Changes in the relation between competence beliefs and achievement in math across elementary school years. *Child development*, 89(2), e138–e156. doi:10.1111/cdev.12806
- Worell, J., & Goodheart, C. (2006). *Handbook of Girls' and Women's Psychological Health*. New York, NY: Oxford University Press.