International Journal of Educational Development 104 (2024) 102974

Contents lists available at ScienceDirect

International Journal of Educational Development

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

The effect of philosophy on critical reading: Evidence from initial teacher education in Colombia

Alejandro Farieta [1,*], Marcos Delprato [2]

Centre for International Education and Development, University of Sussex, Hove, Brighton, United Kingdom

ARTICLE INFO

Keywords:
Philosophy
Critical reading
Educational policy
Initial teacher education
Standardized tests
Propensity score matching

ABSTRACT

Teacher quality, its effect on students' outcomes, and the association of these with economic growth, is the core of recent discussions in Latin America given the region's weak results in international learning assessments. This paper investigates whether there is an effect of philosophy on the outcomes of critical reading for students in B. Ed. programs in Colombia. Relying on exact matching combined with propensity score matching with regression adjustment, we use national data from Colombia to show that students in B.Ed. in philosophy outperformed students in other B.Ed. in critical reading test (0.401–0.124 SD), and, importantly, with higher effects observed for students with lower prior academic achievement (0.44 SD). This suggests that philosophy can help to narrow educational outcomes of students whose socioeconomic conditions are disadvantageous, contributing to social justice in education.

1. Introduction

Teacher quality has become one of the most important concerns in the educational policy domain because of their link to economic growth (Hanushek and Woessmann, 2015a). According to the human capital theory (Becker, 2009; Hanushek, 2011), improving the quality of teachers will enhance learning outcomes (with literacy being one of the most important), and in turn it will boost social and economic conditions of their students in the long run. The cognitive skills of a population are a crucial factor for a country's economic growth, as shown in a series of longitudinal studies about the association between PISA and TIMMS international tests and national GDP growth (Hanushek and Woessmann, 2012a, 2015a, 2015b; Hanushek, Piopiunik, and Wiederhold, 2019). A clear example of the effects of educational outcomes on economic growth can be found across the East Asian countries, which have been in the top bracket of performance in PISA results over the last few years and they have also significantly increased their economies compared to two generations ago (Hanushek and Woessmann, 2016).

A vital component behind this economic expansion process is higher teacher quality. Indeed, better teachers in the classroom can boost students' learning and so help them to get better-paid jobs, thereby boosting economic growth (Chetty, Friedman, and Rockoff, 2014a, 2014b; Hanushek, Piopiunik, and Wiederhold, 2019). Teacher experience, specific education, and cognitive skills are determining factors of school performance (Bardach and Klassen, 2020; Burroughs et al., 2019; Coenen et al., 2018; OECD, 2018a).

Due to the importance of teacher quality, initial teacher education has become central to educational policy in recent years. According to Cochran-Smith, since the end of the 1990 s, educational policies about teacher education have become "the public policy problem", trying to identify which elements can be manipulated from public policy about initial teacher education to improve the quality of education at a large scale (Cochran-Smith, 2005, 2021; Furlong et al., 2008; Tatto and Menter, 2019). The underlying argument is that improving undergraduate teacher education will have an impact on future teacher quality which, as a result, will have an impact on the economy and development at a national level (Barber and Mourshed, 2007; Rockoff, 2004).

Latin America is a region where low teacher quality and cognitive skills have a bearing on the observed stagnation of GDP growth. In this region, despite the increase in enrolment in the last 50 years, the economic effect has been significantly lower than other regions in the world, especially compared to southeast Asia, with similar growth in enrolment but considerably higher economic outcomes (Bruns and Luque, 2015). Consequently, several educational policy

* Correspondence to: University of Sussex, Essex House 128, Falmer, Brighton BN1 9QR, UK.
E-mail address: R.Farieta-Barrera@sussex.ac.uk (A. Farieta).
[1] ORCID: 0000–0001-6895–3113
[2] ORCID: 0000–0001-9333–8331

https://doi.org/10.1016/j.ijedudev.2023.102974
Received 22 May 2023; Received in revised form 4 December 2023; Accepted 9 December 2023
Available online 19 December 2023
0738-0593/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

recommendations for Latin America advise to improve teacher education, recruitment, and remuneration to boost economic growth (Hanushek and Woessmann, 2012b, OREALC/UNESCO, 2013, 2016). This is the case of Colombia, where students in initial teacher education programs —equivalent to a Bachelor in Education, B.Ed.—, have lower academic outcomes compared with students in other undergraduate majors (Barrera-Osorio et al., 2012; Bonilla-Mejía, 2018; García J. et al., 2014). Recently, this has motivated important national reforms of educational policy to improve the educational outcomes of future teachers (Colombia, Ministry of Education [MEN], 2016, Feb. 3; 2017, Sep. 15; Arias G. et al., 2018).

In this paper, we investigate a novel unexplored mechanism that can improve the outcomes of B.Ed. students; that is, what is the impact on teaching quality if future teachers study more philosophy. This is our main research question. Our motivation is to add to the body of empirical evidence about the effectiveness of philosophy undergraduate majors, which is almost null. The only exception is the study of Farieta (2022) showing that, in Colombia, more philosophy credits and subjects in B.Ed. are associated with better outcomes in reading, citizenship education, and writing. Our study further contributes to this for the Colombian case, relying on a robust methodology approach, that is: matching with regression adjustment.

Philosophy has been a valuable knowledge area for western culture and history since its rise in ancient Greece. Its influence and importance are undeniable for developing other knowledge areas, such as mathematics, physics, medicine, politics, economics, psychology, and sociology, among many others. Nevertheless, in the last fifty years, the inclusion of philosophy modules or credits in initial teacher majors has declined, especially in countries like the UK, where most of them are optative (Barrow, 2019). Additionally, many philosophy faculties, departments, and undergraduate programs have been closed or threatened to be closed, or they are experiencing spending cuts. According to www.dailynous.com, a website focused on the diffusion of philosophy, leading universities around the world have experienced some of these menaces in the last five years in countries such as the United States, the United Kingdom, Australia, Serbia, Hungary, and Spain (Weinberg, 2021, Oct. 6). In Colombia, specifically, ten philosophy programs were closed in the last ten years (Farieta, 2018; Farieta et al., 2015); the number of students in these programs has been significantly reduced (Farieta et al., 2024), and the hours spent in high school teaching philosophy have also been reduced (Herrera, 2022; Prada-Dussán, 2023).

This article empirically shows that philosophy has a national economic potential because of its impact on literacy, which is seen as one of the most valuable skills in economic development and democratic stability (OECD, 2019). Literacy is deeply linked to democratic values and the health of political institutions because it is necessary to dialogue and to solve differences between humans in a non-violent way (Morais, 2018). The importance of philosophy for democracy, along with the arts and humanities, has been defended vehemently by authors like Nussbaum (2012) and others (Belfiore, 2015; Machura, 2018; O'Brien, 2015; Olmos-Peñuela et al., 2015). Plainly, the leading thread of the paper is to show that philosophy, in a global south country context, is not only important for maintaining democratic stability, but also it could have a significant economic potential through boosting educational quality. We attempt to answer the following three research questions (RQ):

- RQ1: Do students in B.Ed. in philosophy obtain higher scores in critical reading test than students attending a different B.Ed.?
- RQ2: Do students who have attended B.Ed. in philosophy obtained better scores than students in B.Ed. more related to literacy, such as literature, Spanish language, linguistics, etc.?
- RQ3: Is the effect of philosophy homogeneous or, rather, are there gaps in critical reading outcomes depending on students' socioeconomic and contextual characteristics or prior academic achievement?

Beyond the generic arguments on teaching quality leverage impacts on economic growth and democratic values mentioned above, the Colombian context is particularly suitable for answering these questions. It has a long tradition in B.Ed. programs that instruct future teachers for elementary and high schools in specific subjects like mathematics, natural sciences, social sciences, child education, language, philosophy, special education, etc. (Colombia, MEN, 2016, Feb. 3; 2017, Sep. 15). Moreover, Colombia has mandatory standardised tests at the end of high school (Colombia, Congress of the Republic [CRC], 2009, Jul. 13) and at the end of undergraduate programs (Colombia, Presidency of the Republic [PRC], 2009, Oct. 14). So, it is possible to estimate the added value of the undergraduate programs controlling from prior academic achievement —a leading drive of educational performance— and net of an array of observable characteristics, making the empirical contribution of the paper more robust. Additionally, it is possible to assess of whether the impact of philosophy is heterogenous and varies by students' socioeconomic conditions, thereby highlighting possible mechanisms behind how philosophy can translate into boosted educational outcomes among the most disadvantaged students.

The structure of the paper is as follows. In Section 2 we present a review of the effects and consequences of a better teacher education for students, with a greater focus on a review of the effects of philosophy and related matters on student outcomes. In Section 3, we show the characteristics of the Colombian B.Ed. programs and how they are appropriate to answer the paper's research questions. In Section 4, we introduce the methodological design, i.e., the data analysis and the matching strategy. Section 6 includes the results and Section 7 offers conclusions and policy recommendations from the analysis.

2. Conceptual framework

2.1. Initial teacher education and student outcomes

Several studies show how the performance of students depends on the training and quality of teachers, which, beyond socioeconomic, familiar, and contextual conditions, are the main determinants of student achievement (Hanushek and Woessmann, 2015a; Burroughs et al., 2019; Coenen et al., 2018). Teachers with better training —a significant malleable entry point for an educational policy design— and with more experience have equally positive effects on students' education trajectories —such as more chances to enter higher education— and their social outcomes —e.g., a lower risk of teenage pregnancy, better-paid jobs, etc.— and, in the long term, on living in areas where people have higher incomes and better retirement plans (Chetty et al., 2014b).

According to a meta-analysis done by Hanushek and Rivkin (2010), the added value of a teacher one SD above the mean of the teacher effectiveness distribution could improve student outcomes in reading tests by 0.10–0.26 SD. The long-term economic effect of student outcomes' increase is that one standard deviation in scores translates into 10–20% higher annual earnings throughout the lifetime of students (Hanushek, 2011). Likewise, in a study across 31 countries, Hanushek et al. (2019) found that teacher cognitive skills have effect on student performance of 0.145–117 SD in mathematics and 0.092–148 SD in reading scores, controlling by parents' cognitive skills, student and school characteristics, and country variations.

Professional development and experience of in-service teachers is often considered one of the most central strategies to improve students' outcomes (Didion et al., 2020), but there is also evidence of the effects of the quality of pre-service teacher training programs on student outcomes (Goldhaber et al., 2013; Koedel et al., 2015; Mihaly et al., 2013). Additionally, academic achievement of future teachers predicts teacher performance and, therefore, it has a positive effect on future students' cognitive outcomes (Corcoran and O'Flaherty, 2018).

2.2. Philosophy and effectiveness

Research about the effectiveness of philosophy undergraduate programs is scarce. A notable exception is the Colombian study of Farieta (2022), which through multilevel regression analysis, found that there is an association between more credits in philosophical modules and higher outcomes in critical reading, writing, and citizenship competencies (but lower outcomes in quantitative reasoning).

Nevertheless, critical thinking and argument processing are more widely studied, usually with positive effects in different educational levels (Abrami et al., 2015). Logic, argumentation, and critical thinking have been rooted in philosophy since the ancient Greek philosophers (Castagnoli and Fait, 2022), and nowadays they are mandatory in any philosophy undergraduate program. Logic and argumentation are considered central in teaching critical thinking (Hausman et al., 2021; Salmon, 2012). A study carried out at a university in the United States by Quintana and Schunn (2019) concluded that logic modules in first-year students of undergraduate STEM-related majors improve their academic achievement, especially for students with lower academic background. According to Quintana and Schunn (2019), the mediator for obtaining critical thinking skills is the argumentation linked to the logic modules. A recent study shows that argument processing can be improved with online programs, compensating student's lack of previous formal instruction in scientific literacy skills (Münchow et al., 2023). In a different context, a study by Sultan et al. (2017) with pre-service language teachers in Indonesia, showed that a critical literacy training approach had a significant effect on reading skills.

Even though philosophy has not been widely researched in terms of its impact on efficiency in higher education, there is evidence on its positive effects for different educational levels. For instance, in programs and projects as "Philosophy for Children" [P4C] (Bynum and Lipman, 1976; Lipman and Sharp, 1978), and "Philosophy with Children" [PwC] (Cassidy and Christie, 2013; Kennedy and Kennedy, 2011; Vansieleghem and Kennedy, 2011). PwC has shown to be very effective when applied in schools in terms of improving critical and creative thinking skills. The meta-analysis of García-Moriyón et al. (2005) reported positive effects on students' cognitive skills, with an average difference between treated and untreated groups of about half a standard deviation ($d = 0.58$). Likewise, important effects on basic cognitive skills, personality traits and academic achievement of students attending a PwC programme in Madrid (Spain) were found in a long longitudinal study ten years after the intervention (Colom et al., 2014). A study in the UK also found that the results obtained by PwC are better for students from disadvantaged backgrounds (Ventista, 2019).

Some of the evaluations of PwC programmes have been questioned for not being rigorous enough or for having some conflict of interest (García-Moriyón et al., 2005; Colom et at, 2014). However, more recent studies show better and more reliable evidence not only in critical thinking and cognitive skills (Işiklar and Öztürk, 2022), but also in noncognitive, social, and ethical skills (Karadağ and Demirtaş, 2018; Ventista, 2019) like self-regulation, engaged participation in the school, and collaborative dialogue with their peers in children with social, emotional, and behavioural needs (Cassidy et al., 2018).

3. The Colombian Context

3.1. Factors associated with student academic achievement

One of the main drivers of students' performance in Colombia is their socioeconomic status, whose explanatory power is very high in Latin American countries, accounting for almost 30% of the variation of the scores (Avendaño et al., 2016). Latin American education systems are highly segregated, and of the total variation of students' achievement, around 40% is attributable to school wealth composition and a further 10% to additional individual and school factors (Delprato et al., 2015). This shows the high levels of inequality in the region's educational systems, especially in Colombia (OECD, 2016, 2018b; García V. et al., 2013; García J. 2015; Rodríguez et al., 2014; Timarán-Pereira et al., 2016). Socioeconomic status is related to parental education, another important factor explaining student performance in the Colombian context (García-González and Skrita, 2019; Rangel and Lleras, 2010). Students with more educated parents have more support when it comes to assignments and other academic tasks and they have higher income, better house infrastructure and more books and other academic resources at home (Camacho et al., 2016; Hernández G and Padilla G, 2019). Generally, students of B.Ed. came from families with lower educational levels than other undergraduate programs (García J. et al., 2014).

An additional essential contextual characteristic is where a student lives (region and city, rural/urban location) and the neighbourhood income level where they come from. Colombia is a highly unequal and segregated country, and students' educational outcomes are a result of this embedded and cumulative geographical inequalities (Arias-Velandia et al., 2021; García J. et al., 2015). Teachers with lower standardised scores —a proxy for teacher quality— are mostly located in the poorest regions of the country, which makes inequalities due to education more difficult to tackle (Bonilla-Mejía et al., 2018). Students in richer regions have better educational outcomes, since these regions have universities with better educational infrastructure, higher-paid staff, and a better reputation (Cvecic et al., 2019; Gibbons y Vignoles, 2012; Helbig et al., 2017). Education segregation even happens within big cities, where students' achievement is associated with socioeconomic inequalities, and students with higher economic resources are the ones who can attend universities with improved resources and better staff (Rojas, 2019).

Gender is also associated with student performance, with gender inequality persisting due to culturally stereotyped behaviour (Cárcamo et al., 2020; Morris, 2012; OECD, 2018b). According to the literature, gender performance in high school is positive for women in literacy, but negative in mathematics (Abadía and Bernal, 2016; Cárcamo, and Mola, 2012; Correa, 2016; Woessmann, 2010). In higher education, the gap increases in mathematics especially in STEM programs (Gómez et al., 2020). Teacher education in Colombia has prominently more female students (63%), more than other undergraduate courses, whose female student percentage is 54% (Arias G. et al., 2018). However, and contrary to the rest of the B.Ed., in philosophy courses the number of female students and their outcomes in reading are lower than those of male students (Farieta, 2022). Further, age is linked to educational performance (Castro et al., 2018; Rodríguez et al., 2014) because of the opportunity costs of studying for worse-off students' groups, and this means that students in B.Ed. are usually 1–2 years older in than students in other undergraduate programs (Rodríguez et al., 2014).

At the undergraduate level, one of the leading factors explaining educational outcomes is student's prior achievement (Schneider and Preckel, 2017). In Colombia it can be obtained from the student score on the *Saber 11* test, a mandatory test to obtain a high school diploma and the main benchmark to measure the added value of undergraduate majors (Rodríguez-Revilla and Vallejo-Molina, 2022; Sarmiento et al., 2019). Students in B.Ed. have lower scores on the *Saber 11* test than students in other undergraduate programs (García J. et al., 2014; Rodríguez et al., 2014).

In addition, modality of education (on/campus/distance) is found to be an important determinant of undergraduate student performance, with distance students having lower performance than on-campus students (Aguilera-Prado, 2017; Arias-Velandia et al., 2018; Rodríguez, Gómez, and Ariza, 2014). Students in distant programs mostly come from rural areas, smaller and poorer towns of lower socioeconomic status and, compared with students in similar conditions in on-campus programs, they attain lower scores (Arias-Velandia et al., 2021; Pineda and Celis, 2018; Timarán-Pereira et al., 2016).

Other quality indicators associated with student outcomes is the high-quality accreditation. Colombian higher education system has two

types of accreditations: institutional and program accreditation (Colombia, MEN/Comisión Nacional de Acreditación [CNA], 2013). Usually, the institutional accreditation is associated with better student outcomes (Bayona et al., 2018; Cayón et al., 2020), nevertheless, the evidence about the program accreditation is not conclusive, and some studies evidence positive outcomes (Camacho et al., 2016) but others don't (Sarmiento et al., 2015). In philosophy, the association hasn't been found (Farieta, 2020). In 2015, the Colombian government compelled al the B.Ed. programs to obtain high-quality accreditation (Colombia, CRC, 2015, Jun. 9) to improve the student achievement. Other institutional conditions related to student scores in undergraduate programs are teachers' characteristics (educational level, teaching experience, evaluation, etc.) (Ordóñez et al., 2019; Sáenz-Castro et al., 2021).

3.2. Philosophy programs in Colombia

There are two different kinds of philosophy majors in Colombia: the traditional B.A. in philosophy, but also a different type of program called "*Licenciatura*", equivalent to a Bachelor of Education (B.Ed.) in the anglophone educational systems, which are aimed at teaching and training schoolteachers. Some universities offer both programs. Usually, the first type is shorter (4 years) and the second is longer (5 years) (Farieta et al., 2015) because it includes school practices or modules related to learning, teaching, and pedagogy. In this study, we will focus only on the B.Ed. in philosophy to compare them with different B.Ed. programs. Some of the B.Ed. in philosophy include other subjects like language, religious studies, history, political science, or other humanities (Farieta et al., 2015; Farieta, 2018). There are slight differences in curricular terms, and there are big discussions about the relevance of their strong European tradition in the Latin American context and other curricular issues (Bernal-Ríos, 2022). Additionally, these programs are strongly masculinised, with only an average 24.8% of professors and lecturers of philosophy programs being female (Acevedo-Zapata and Rivera-Sanín, 2023).

4. Methodological design

4.1. Data and Sample

The educational system in Colombia is a fitting case to assess the hypothesis that students who attend a philosophy program have better outcomes. First, Colombia has a mandatory test called *Saber Pro* for all students in the final year of their undergraduate studies (Colombia, PRC, 2009, Oct. 14), and one of the modules in the test is "Critical Reading." Second, since the 1960 s Colombia has had a very broad range of initial teacher education programs in all disciplines and knowledge areas which are mandatory in basic and middle education: mathematics, natural sciences, social sciences, philosophy, arts, early childhood, special education, etc. (Colombia, MEN, 2016, Feb. 3; 2017, Sep. 15). There are currently 558 B.Ed. in the country, 29 of which are in philosophy or philosophy with another discipline —such as theology, religious sciences, political theory, humanities— (Farieta et al., 2015; Farieta, 2018). Some of these are inactive or in process of being closed, due to new regulations for Colombian B.Ed. (Arias G. et al., 2018) and other issues related to financing through tuition fees in private universities combined with the lack of getting enough students to open cohorts (Farieta et al., 2024; Herrera, 2022; Prada-Dussán, 2023; Valderrama--Leongómez et al., 2019). This allows a comparison of the outcomes of the students in a B.Ed. in philosophy with those of students in other B. Ed. programs. The main data used came from the ICFES, a governmental organisation in charge of the design and assessment of all the national educational tests. All the data is public and is available at www.icfes.gov.co.

Table 1 describes total population that presented the *Saber Pro* test during the 2012–2021 period by year and program type. In total, 225,984 B.Ed. students presented the *Saber Pro* test over this period. Out of this total population, 5240 were students in B.Ed. in philosophy [ph.] or with the noun "philosophy" in the program name: ph. and theology; ph. and religious studies; ph. and ethics; ph. and humanities, etc. (Farieta et al., 2015; Farieta, 2018). Table 1 shows that the number of B.Ed. in Philosophy students has reduced throughout the years, reaching a maximum of 728 in 2014, and a minimum of 302 in 2021, with a reduction in its proportion compared to the rest of B.Ed. students (from 3.16% in 2012 to 1.41% in 2021). Similarly, there has been a decreasing uptake for students in B.Ed. in literature (including linguistics, literature, Spanish, and other related areas presented the test), dropping from 15.4% in 2012 to 10.53% in 2021.

The *Saber Pro* test changed the scoring system in 2016 (ICFES, 2017). The previous version (2012–2015) was designed to have a standard mean of 10 points and a standard deviation of 1, with a range between 5.2 and 15.8. In the more recent version (2016–2021), scores' interval is 0 to 300 (ICFES, 2018). Since both scores are normally distributed, we standardise the results for comparison so that to avoid any bias due to changes in the test. Methodologically, we use exact matching for the year of the presentation of the exam (see subsection 4.2).

Most of the control variables we employ in our analytical approach (i.e., matching) are drawn from earlier studies (Arias-Velandia et al., 2021; Avendaño et al., 2016; García J. et al., 2014, 2015; Timarán--Pereira et al., 2016). The control variables can be classified in two groups: (1) student and context, and (2) institution and program. The most crucial control variable we use is students' prior academic achievement, which comes from the *Saber 11* test in the reading module. As mentioned above, all students in the last year of high school (11th grade, according to the Colombian educational system) must present the *Saber 11* test (Colombia, CRC, 2009, Jul. 13). The test is run twice a year. We merged the databases of *Saber Pro* and *Saber 11* to include students' prior academic achievement in the analysis.

Table 1
Saber Pro Critical Reading std. scores by year and type of B.Ed.

Year	B.Ed. in Philosophy			Other B.Ed.					B.Ed. in Literature				
	N	Mean	SD	N	Mean	SD	Diff	t test	N	Mean	SD	Diff	t test
2012	487	0.035	1.077	14,947	-0.326	1.066	-0.361	* **	2377	-0.218	1.068	-0.183	* **
2013	702	0.053	1.104	21,619	-0.276	1.052	-0.329	* **	3233	-0.010	1.060	0.043	
2014	728	0.070	1.120	19,974	-0.306	1.030	-0.376	* **	3307	-0.175	1.068	-0.105	* **
2015	563	0.269	1.095	18,467	-0.233	1.009	-0.502	* **	2648	0.002	1.053	0.271	* **
2016	503	0.372	0.946	25,141	-0.278	1.020	-0.651	* **	4257	-0.108	1.038	0.264	* **
2017	566	0.454	1.057	27,224	-0.245	1.013	-0.699	* **	4202	-0.054	1.053	0.400	* **
2018	519	0.427	0.980	25,043	-0.271	1.017	-0.697	* **	3120	0.103	1.046	0.530	* **
2019	460	0.535	0.940	25,600	-0.166	0.995	-0.701	* **	3163	0.246	0.986	0.780	* **
2020	410	0.465	1.013	21,602	-0.215	1.004	-0.681	* **	2735	0.083	1.037	0.549	* **
2021	300	0.577	0.905	21,012	-0.192	0.977	-0.769	* **	2243	0.169	1.001	0.746	* **
Total	5238			220,629					31,285				

Note: * $p < .05$, * * $p < .01$, * ** $p < .001$

Table 2
Saber 11 std. critical reading students score by period and type of program.

Period	B.Ed. in Philosophy			Other B.Ed.					B.Ed. in Literature				
	N	Mean	SD	N	Mean	SD	Diff	t test	N	Mean	SD	Diff	t test
2006-1	35	0.758	0.935	742	0.165	0.931	-0.593	* **	69	0.459	0.984	-0.299	
2006-2	202	0.471	0.795	8846	0.301	0.820	-0.171	* *	1437	0.390	0.813	-0.081	
2007-1	21	0.091	0.899	510	0.352	0.855	0.261		48	0.349	0.856	0.258	
2007-2	222	0.307	0.933	10,497	0.119	0.852	-0.188	* *	1637	0.204	0.846	-0.103	
2008-1	55	0.738	0.882	1108	0.305	0.988	-0.433	* *	110	0.620	1.070	-0.118	
2008-2	224	0.312	0.840	11,271	-0.007	0.798	-0.319	* **	1793	0.094	0.764	-0.217	* **
2009-1	30	0.133	0.918	637	0.216	0.975	0.084		58	0.236	0.908	0.103	
2009-2	222	0.084	0.820	12,249	-0.024	0.760	-0.108	*	1889	0.003	0.779	-0.081	
2010-1	14	0.727	0.770	343	0.417	1.014	-0.310		32	0.353	1.218	-0.374	
2010-2	285	1.128	1.258	15,706	0.630	1.166	-0.498	* **	2168	0.811	1.181	-0.317	* **
2011-1	7	0.413	1.119	367	0.407	0.903	-0.006		26	0.397	0.789	-0.016	
2011-2	252	0.573	1.022	15,443	0.219	1.276	-0.354	* **	2318	0.441	1.253	-0.132	
2012-1	8	0.593	0.954	386	0.250	0.929	-0.344		28	0.481	1.086	-0.112	
2012-2	249	0.671	0.878	15,049	0.181	0.931	-0.490	* **	2075	0.415	0.940	-0.256	* **
2013-1	15	0.906	1.014	353	0.214	0.920	-0.692	* *	35	0.439	1.119	-0.467	
2013-2	218	0.705	0.886	13,073	0.198	0.948	-0.506	* **	1675	0.456	0.938	-0.249	* **
2014-1	2	0.422	1.248	229	0.255	1.031	-0.167		15	0.389	0.962	-0.033	
2014-2	163	0.802	0.829	10,713	0.188	0.865	-0.614	* **	1372	0.479	0.845	-0.323	* **
2015-1	3	1.141	0.879	205	0.330	1.088	-0.811		21	0.791	1.146	-0.350	
2015-2	97	0.791	0.915	7776	0.186	0.846	-0.605	* **	998	0.504	0.863	-0.286	* *
2016-1	1	2.382	.	73	0.723	0.843	-1.660		3	0.781	0.299	-1.601	
2016-2	60	0.961	0.761	4634	0.475	0.844	-0.486	* **	503	0.818	0.834	-0.143	
2017-1	0	-	-	32	0.701	0.922	-		2	1.255	2.010	-	
2017-2	9	0.705	1.113	868	0.488	0.883	-0.217		89	0.814	0.824	0.109	
2018-1	1	1.304	.	3	1.958	0.591	0.653		1	1.892	.	0.588	
2018-2	1	0.422	.	70	0.202	1.000	-0.220		4	-0.583	0.562	-1.005	
Total	2396			131,183					18,406				

Note: * $p < .05$, * * $p < .01$, * ** $p < .001$

Table 2 shows the working sample after merging, by period of presentation of the *Saber 11* test and the type of program. This test had significant changes in 2014 to align all the national tests, and the previous scores from 2012 to 2014 were recalculated with the new structure and scoring system (ICFES, 2013). To avoid bias due to changes in the test, we apply exact matching by period of *Saber 11* presentation, as we explain in the next subsection. Also, to improve comparability, we only keep in the sample students whose tests time gap were more than three years and less than ten. Using this data, we can determine if students with similar *Saber 11* test scores (and other similar factors associated with student performance) have better outcomes in the *Saber Pro* reading module when they attend a B.Ed. in philosophy, in comparison with those who went to a different B.Ed.

When students register to present the *Saber Pro* test, they must complete a socioeconomic survey and the ICFES includes this information with students' scores. From the ICFES survey, we select the following control variables: gender, age, Neighbourhood Socio-Economic Level [NSEL] —also known as socioeconomic strata, (*estrato socioeconómico* in Spanish)—. We also include socioeconomic indicators by national departments (Colombian administrative regions) and municipalities of students to account for socioeconomic disparity within the country at a more aggregated level. Specifically, we use the Index of Multidimensional Poverty [IMP] from the data from the 2018 national census, calculated by the National Department of Statistics [*Departamento Nacional de Estadística* – DANE] (2018). The IMP considers the conditions of the population regarding health, housing, access to basic services (water, electricity, telephone, internet, and domiciliary gas), educational conditions, and conditions of children and youth. The IMP is calculated for the 33 administrative departments of Colombia. Additionally, we use the Index of Unsatisfied Basic Needs [IUBN] calculated by the DANE from the national census of 2018 for the 1103 Colombian municipalities (DANE, 2018). The IUBN measures the living conditions in aspects like inadequate or overcrowded housing, children and youth not attending school, economic dependency, and access to basic public services. We do not use the "rural/urban living place" dichotomous variable because the population of students living in rural areas is quite low, hindering balancing the matching groups. Instead, we split the sample into students living in big cities, capitals of departments (or metropolitan areas) and students living in small towns or rural areas, which with social inequality measured at different levels is enough to capture socio-economic factors behind B.Ed. choices and supply of courses.

Besides, we include variables of the institutions and programs, such as the accreditation status of the institution and the program, since they have been reported by the literature having better added value for student outcomes (Bayona et al., 2018; García J. et al., 2014). We also use the program modality (on-campus or distance) as students in distance courses have shown lower scores in *Saber Pro* (Aguilera-Prado, 2017; Farieta, 2020). Descriptive statistics are presented in Table 3.

Summary statistics for control variables by the treatment and control groups are shown the Table 3. Statistics show that students in B.Ed. in philosophy come from less disadvantaged background in terms of living area, education of parents, NSEL, IMP, IUBN, and prior academic achievement (*Saber 11* scores) in comparison with students of other B.Ed., as well as students in language B.Ed. Also, they are more likely to be male, (63%) in comparison with students in other B.Ed. (32%) or language B.Ed. (24%). Additionally, more of them come from programs and institutions with accreditation and on-campus programs. These differences support the necessity to implement a matching technique to narrow the characteristics' differences across the two groups, offering a better control of differences on observables. Since the *Saber 11* and the *Saber Pro* tests have had several changes across the years, we use exact matching for the year of presentation of both tests. This, too, offers a better control by weakening the impact of changing cohorts across time.

Table 3
Descriptive statistics of control variables and comparison tests.

Variable	B.Ed. In Philosophy		Other B.Ed.				B.Ed. In Literature			
	N	μ or %	N	μ or %	Diff		N	μ or %	Diff	
Control variables										
Individual or context										
Age	2200	24.109	124,193	23.773	-0.336	* **	17,502	23.662	-0.447	* **
		(2.911)		(2.69)				(2.465)		
Female (†)	821	37.3%	84,677	68.18%	-30.88%	* **	13,304	76.01%	-38.71%	* **
Rural or small city (†)	555	25.22%	30,239	24.35%	0.87%		4374	24.99%	0.23%	
NSEL (†)	2120		121,042			* **	17,062			* **
Rural or no NSEL	12	0.57%	917	0.76%	-0.19%		87	0.51%	0.06%	
NSEL 1	471	22.22%	35,517	29.34%	-7.12%		5511	32.3%	-10.08%	
NSEL 2	840	39.62%	49,725	41.08%	-1.46%		7037	41.24%	-1.62%	
NSEL 3	632	29.81%	29,229	24.15%	5.66%		3775	22.13%	7.68%	
NSEL 4	93	4.39%	4335	3.58%	0.81%		529	3.1%	1.29%	
NSEL 5	39	1.84%	962	0.79%	1.05%		90	0.53%	1.31%	
NSEL 6	33	1.56%	357	0.29%	1.27%		33	0.19%	1.37%	
Education of Parents (years)	2201	16.378	124,198	17.199	0.821	* **	17,504	16.848	0.47	*
		(8.82)		(8.6)				(8.509)		
IMP	2201	0.195	124,198	0.209	0.014	* **	17,504	0.213	0.018	* **
		(0.031)		(0.038)				(0.037)		
IUBN	2201	7.732	124,198	11.741	4.009	* **	17,504	12.409	4.677	* **
		(9.261)		(14.314)				(13.835)		
Institution or program										
Institutional Accreditation (†)	1418	64.43%	50,043	40.29%	24.14%	* **	9123	52.12%	12.31%	* **
Program Accreditation (†)	1006	45.71%	49,687	40.01%	5.70%	* **	5913	33.78%	11.93%	* **
On campus (†)	489	22.22%	25,744	20.73%	1.49%		4195	23.97%	-1.75%	

Note: Categoric variables with (†) using the chi square test. Factor variables using t-test. μ: mean. Standard deviations (SD) in parentheses. * ** $p < 0.001$, * * $p < 0.01$, * $p < 0.05$

4.2. Analysis strategy

We use *Propensity Score Kernel Matching with Regression Adjustment*, combined with *exact matching* for the categorical variables to improve the quality of the matching (Jann, 2017a; Rubin and Thomas, 2000). The technique allows us to estimate the *Average Treatment of the Treated* [ATT] after pairing students in the B.Ed. in philosophy (treatment group) with students with same or extremely similar characteristics in different B.Ed. (control group). The regression adjustment adds double robustness property. The matching double robust estimator (Nguyen et al., 2017) uses the same specification for estimating the propensity score and the outcome model. This estimator has the theoretical advantage that it yields unbiased estimates if either or both the propensity and outcomes are correctly specified —the double robust property which offers more protection against misspecification (Jann, 2017a; Rubin and Thomas, 2000). To avoid bias due to the different years of the test, we use exact matching for the year of the *Saber Pro* test, but also for the *Saber 11* test, because, as mentioned above, these tests have had significant changes through the time. The exact matching by year also avoids any bias caused by the time changes or contextual issues such as the pandemics occurred during 2020 and 2021 which caused the tests to be presented virtually in 2020. We also use exact matching for all the dichotomous variables: gender, living area (capital city or metropolitan area/rural or small city), institutional accreditation, program accreditation, and program modality (on-campus/distance).

The aim of the matching process is to obtain a counterfactual of what the average score would have been if students had attended a program different to the B.Ed. in philosophy. We define the average treatment effect of attending the program as:

$$E(Y_i|D=1) - E(Y_0|D=0) =$$

$$E(Y_{1i}|D=1) - E(Y_{0i}|D=1) + E(Y_{0i}|D=1) - E(Y_{0i}|D=0) \quad (1)$$

Where Y_i denotes student scores, and D is a dummy for program selection, which equals 1 when is a B.Ed. in philosophy and 0 otherwise. The outcome variable Y_{1i} is the outcome of a philosophy student and Y_{0i} the outcome of a student in a different program. $E(Y_i \mid D=1) - E(Y_0 \mid D=0)$ is the average treatment effect (ATE), and $E(Y_0 \mid D=1)$ is the unobserved counterfactual that has to be constructed using matching estimators. The estimation of the ATE requires two assumptions: the conditional independence assumption (CIA) and common support (Rosenbaum and Rubin, 1983).

CIA assumption implies that potential outcome variables (Y_0, Y_1) are independent of the treatment (here, the decision of being enrolled in a B. Ed. in philosophy) when conditioning on a set of observable covariates X; in other words: CIA: $(Y_i, Y_0) \perp\!\!\!\perp D \mid X$. The second assumption entails common support (or overlap), which means that, for each student in a B. Ed. in philosophy, there is a positive probability of a match within a group of non-philosophy students with a similar set of covariates X; that is: $0 < Pr(D=1 \mid X) < 1$.

If both assumption holds, meaning that $F_i(Y_i \mid X, D=1) = F_0(Y_0 \mid X, D=0)$, the counterfactual $E(Y_{0i} \mid D=1)$ can be built independent of the decision of the program attended after conditioning for the propensity score. The propensity score ($Pr(D=1 \mid X = p(X))$ is the conditional probability of choosing a B.Ed. in philosophy, given the pre-treatment variables (X) (Rosenbaum and Rubin, 1983):

$$p(X) = \Pr\{T=1|X\} = E\{T|X\} \quad (2)$$

The next step is to match students in B.Ed. in philosophy with peers in a different B.Ed. based on the propensity score of receiving the treatment calculated from the covariates. If CIA holds, the ATT is defined as

$$ATT = E(Y_{1i} - Y_{0i} \quad | \quad D=1) = E(Y_{1i} - Y_{0i}) \quad (3)$$

To determine the matching weight, we use propensity score kernel matching, where the matching weight is defined as:

$$W_{ij} = \frac{K(P_j - P_i)/a}{\sum_{k\varepsilon D_k=0} G(P_k - P_i)/a)} \tag{4}$$

Where K is the Epanechnikov kernel density function: $K(x) = \frac{3}{4}(1 - x_2)$, for $|x| < 1$, a is the bandwidth parameter, P_j is the propensity score of the case j in the control group, P_i denotes the propensity score of case i in the treated group, and $P_j - P_i$ represents the distance between propensity scores. The bandwidth is calculated by weighted cross-validation with respect to Y (Frölich, 2004, 2005; Galdo et al., 2008). The weights are normalized so, for each student that was enrolled in a B.Ed. in philosophy, $\sum_{j\varepsilon D=0} W_{ij} = 1$. The counterfactual estimate is the weighted average of the observed outcome of the students in B.Ed. different to philosophy:

$$(Y_j \quad | \quad D = 0) = \sum_{i\varepsilon D=1} \sum_{j\varepsilon D=0} W_i(Y_j \quad | \quad D = 0) \tag{5}$$

We estimate the ATT with the weight from Eq. (3) as:

$$\widehat{ATT} = \frac{1}{N^{D=1}} \sum_{j\varepsilon D=1} \left[(Y_i \quad) \quad D = 1) - \sum_{j\varepsilon D=0} W_i(Y_j \quad) \quad D = 0) \right] \tag{6}$$

The ATT equals the weighted average of the differences between observed and potential outcomes, which represents the difference in scores of a student in a B.Ed. in philosophy compared with the counterfactual that attended a different B.Ed.

Entropy balancing is applied to achieve the covariate balance of the groups, adjusting differences in standard deviation, variances, and skewness, and therefore reducing model dependence for the estimation of ATT (Galdo et al., 2008; Hainmueller, 2012). Additionally, as previously mentioned, regression adjustment adds the double robustness property to the analysis strategy, reducing confounding if there is an imbalance after PSM and avoiding selection bias (Huber et al., 2013; Jann, 2017a; King and Nielsen, 2019; Rubin and Thomas, 2000).

The equation for the regression adjustment uses all the covariates X for the matching, and can be defined as:

$$Y_i = \begin{pmatrix} \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \\ \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} + \varepsilon_i \end{pmatrix} \tag{7}$$

Where β_0 is the regression intercept, x_1 is the prior academic achievement (i.e., standardised score of the *Saber 11* test), x_2 denotes gender (male/female), x_3 age, x_4 the education of the parents, x_5 the neighbourhood socioeconomic level [NSEL], x_6 the living area (rural area or small town/capital of region or metropolitan area), x_7 the IUBN, x_8 the IMP, x_9 if the institution has or not high-quality accreditation, x_{10} if the program has or not high-quality accreditation, and finally, x_{11} the modality of the program (on-campus/distance). For the estimation we use Stata© 18 and the Stata command *kmatch* (Jann, 2017b).

To answer the three RQs, we carried out estimations with different treatment and comparison groups. For RQ1, we compare B.Ed. in philosophy students with students in different B.Ed. For RQ2, the control group is constrained to students in B.Ed. in literature. RQ3 is the most complex question to answer as it deals with sub-groups heterogeneity driving critical reading scores. We estimate the ATT by subpopulations, according to key control variables and assessing differences by Wald tests. If the variable is dichotomous (gender, program modality, living place, institutional and program accreditation), we split it by categories. If the variable is continuous, we split it into two levels (a, b): standardised score in the *Saber 11* test ($a \leq 0.5$ SD $< b$), NSEL ($a < 2 \leq b$) IMP ($a < 0.2 \leq b$), IUBN ($a < 5.2 \leq b$), years of education of the parents ($a \leq 16 < b$). Additionally, since the number of philosophical credits within the B.Ed. in philosophy varies from 40 to 126 (Farieta, 2018), and it is associated with higher student scores (Farieta, 2022), we divide the treatment group into two different subgroups (40–80 and 81–126 credits) to estimate if there is a significant difference in ATT according to the intensity of the treatment. Usually, programs with lower number of credits are more multidisciplinary, including philosophy and literature, religious studies, or other disciplines (Farieta, 2018; Farieta et al., 2015).

4.3. Tests of group balancing and common support

We verify the common support condition to test the validity of the estimation in four different steps. First, we conduct balancing tests of the covariates between the treated and untreated groups before and after the matching. We evaluate the balance of each covariate across the treatment groups in each matched sample, calculating the standardized mean difference (SMD):

$$SMD = \frac{|\overline{W_{i1}} - \overline{W_{i0}}|}{\sqrt{\frac{S_{i1}^2 + S_{i0}^2}{2}}} \tag{7}$$

Where $\overline{W_{i0}}$ and $\overline{W_{i1}}$ denote the means, and S_{i0}^2 and S_{i1}^2 the variances of the control and treated groups, respectively (Nguyen et al., 2017). As shown by Table 4 and Fig. 1, before the matching there are significant differences between covariates, which are then corrected after the matching, with the Standard Mean Difference (SMD) and the skewness difference leading to zero, and the variance ratio to 1, which means that covariates biases are corrected (Nguyen et al., 2017). Therefore, the balancing property by covariates is satisfied because the treated and untreated groups are similar enough to corroborate the common support condition, reaching a balance loss near to zero (2.02e-16 for the first model). Details of the balance loss for each model are shown in Table 5, (last row), where we also present the ATT for the balanced groups. Also, for comparison, we present the naïve average treatment (NATE) estimate, without balance corrections.

In the Fig. 1 we show graphically how the matched sample corrects the SMD, variance ratio, and skewness compared to the raw sample, for the models estimated in the next section. For reference, the Fig. 1a displays the results in the Table 4. In a second step, we also visually check that the density of the propensity score in the treatment and the control group are as similar as possible (Fig. 2).

In the third and four step, we check that there are no significant differences between the original treated group and the matched to verify there is not selection bias after the matching in terms of observables. We check that the SMD between treated and matched group across the covariates is below the threshold of < 0.2 SMD (Rubin and Thomas, 2000) (Fig. 3) and, hence, there is not significant changes in terms of observables after the matching.

Lastly, we compare the density of the treated group. Densities of the entire treated group and the matched group are similar, ensuring there is no significant changes or selection bias after matching (Fig. 4). In summary, we argue that empirical checks carried out entail robustness for the matching approach we implement.

5. Results

5.1. RQ 1 – Students in B.Ed. in philosophy compared to all other B.Ed

Students who attend a B.Ed. in philosophy obtain, in average, 0.401 SD (SE 0.02; $p < 0.001$) higher scores than students who attend others B.Ed. (see: Table 5, column 2; and Fig. 5). Plainly, philosophy students obtain improved critical reading scores and the effect of attending B.Ed. in philosophy is rather high. For the treatment group, 2012 students were matched, and 107 students were discarded since they did not have similar counterparts in the control group. In the control group, 71,444 out of 121,006 students were used for the estimation. The bandwidth was estimated at 0.019 with a very low balance loss (2.46e-16) (Table 5, column 2), which means that after the matching the selection imbalance was very close to zero.

After matching, the covariate balancing between the raw data and the matched data improve significantly, reducing covariate bias, since

Table 4
Comparison of covariates before and after matching.

Means		Raw			Matched (ATT)		
		Treated	Untreated	StdDif	Treated	Untreated	StdDif
	Saber 11 std. score in reading	0.6103	0.2370	0.3813	0.6072	0.6072	0.0000
	Female	0.3756	0.6818	-0.6443	0.3941	0.3941	0.0000
	Age	24.1132	23.7563	0.1272	24.0448	24.0448	0.0000
	Parents education	16.8065	17.5088	-0.0832	16.9115	16.9115	0.0000
	NSEL	2.2695	2.0403	0.2360	2.2366	2.2366	0.0000
	Big city or metropolitan area	0.7461	0.7570	-0.0252	0.7495	0.7495	0.0000
	IUBN	7.8122	11.7088	-0.3226	7.7073	7.7073	0.0000
	IMP	0.1955	0.2092	-0.3945	0.1954	0.1954	0.0000
	Institutional Accreditation	0.6413	0.4004	0.4969	0.6322	0.6322	0.0000
	Program Accreditation	0.4535	0.3987	0.1111	0.4697	0.4697	0.0000
	On-campus program	0.7740	0.7933	-0.0471	0.8042	0.8042	0.0000
Variances		Raw			Matched (ATT)		
		Treated	Untreated	Ratio	Treated	Untreated	Ratio
	Saber 11 std. score in reading	0.9536	0.9636	0.9897	0.9322	0.9317	1.0005
	Female	0.2346	0.2169	1.0816	0.2389	0.2388	1.0005
	Age	8.6520	7.0946	1.2195	7.5516	7.5480	1.0005
	Parents education	72.7019	69.9016	1.0401	72.4626	72.4276	1.0005
	NSEL	1.0610	0.8256	1.2851	0.9714	0.9709	1.0005
	Big city or metropolitan area	0.1895	0.1840	1.0303	0.1878	0.1878	1.0005
	IUBN	88.3944	203.3391	0.4347	85.4018	85.3606	1.0005
	IMP	0.0009	0.0015	0.6396	0.0009	0.0009	1.0005
	Institutional Accreditation	0.2301	0.2401	0.9585	0.2326	0.2325	1.0005
	Program Accreditation	0.2480	0.2397	1.0343	0.2492	0.2491	1.0005
	On-campus program	0.1750	0.1639	1.0676	0.1576	0.1575	1.0005
Skewness		Raw			Matched (ATT)		
		Treated	Untreated	Diff	Treated	Untreated	Diff
	Saber 11 std. score in reading	0.1298	0.1098	0.0200	0.1240	0.1240	0.0000
	Female	0.5135	-0.7808	1.2943	0.4333	0.4333	0.0000
	Age	4.8312	3.6013	1.2299	4.4206	4.4206	0.0000
	Parents education	0.1452	0.0016	0.1436	0.1289	0.1289	0.0000
	NSEL	0.9140	0.6371	0.2769	0.7788	0.7788	0.0000
	Big city or metropolitan area	-1.1309	-1.1985	0.0675	-1.1516	-1.1516	0.0000
	IUBN	4.6458	2.9870	1.6588	4.7179	4.7179	0.0000
	IMP	0.3163	0.2341	0.0822	0.3000	0.3000	0.0000
	Institutional Accreditation	-0.5894	0.4065	-0.9959	-0.5483	-0.5483	0.0000
	Program Accreditation	0.1867	0.4139	-0.2272	0.1215	0.1215	0.0000
	On-campus program	-1.3099	-1.4490	0.1391	-1.5330	-1.5330	0.0000

Note: Total balance loss= 2.02e-16

there are no significant differences between the untreated and the treated after matching (Fig. 1a). Also, there is enough common support, since the untreated group has a similar density that the treated group after the matching (Fig. 2a). Also, there is no selection bias after the matching as the SMD between covariates is below the threshold (Fig. 3a), and densities are practically the same for the raw treatment group and the matched group (Fig. 4a).

Estimates strongly suggest a positive and considerably high impact on critical reading scores for students attending B.Ed. in philosophy in comparison to the control group (i.e., all other B.Ed. students) because of the quality of matching. That is, the treatment and the control groups after the matching are extremely similar, not only in terms of the propensity score density but also according to the covariates, which reduces substantially residual confounding. Likewise, the treatment group after matching is like the original one, since there are no differences in propensity score density before and after matching, and the differences in terms of the covariates are low, which ensures that there are no significant changes in the treatment group after the matching.

5.2. RQ 2 – Philosophy compared with Literature and language B.Ed

The estimated ATT for philosophy students compared to language students is 0.124 SD and statistically significant (SE 0.026; $p < 0.001$) (Table 5, column 3; Fig. 5). From the treatment group, 1836 out of 2119 students were matched, whereas for the control group 11,132 students out of 19,176 in language B.Ed. were matched (Table 5, column 3). Even though estimates are comparatively lower than the one we obtained for all B.Ed., a significant difference still holds which suggests a positive impact of philosophy for literacy and, especially, for critical reading (compared with focusing only on learning language from a linguistic approach).

As for RQ1, group balancing was achieved as well since the SMD and skewness difference for all the covariates between both groups after matching are zero and the variance ratios are one (Fig. 1b). Similarly, the density distribution for treatment and control groups after matching are indistinguishable (Fig. 2b). There are also no significant differences between the original treated group and the matched group after matching, since the standard mean differences by covariates are no higher than 0.1 (Fig. 3b), and there is also no difference in terms of propensity score distribution (Fig. 4b), and all of which ensures that there is no selection bias driven by observables after matching.

5.3. RQ 3 – Effect of philosophy by subpopulations

As far as the heterogeneity analysis is concerned, we find significant differences between the philosophy students according to the number of philosophical credits (Table 5, columns 4–9), prior academic achievement and gender. The group balancing (treated/untreated) fits well after matching for all the comparison groups in terms of covariates (Fig. 1c–d), as well as in terms of the propensity score kernel density (Fig. 2b c–d). Also, the selection bias is minimum for the treated groups after matching, since these are extremely akin to the original treatment group in covariates SDM (Fig. 3c–d), but also in propensity score density (Fig. 4c–d). Conversely, we do not find statistically significant differences in estimates for the ATT in terms of education of the parents, living region, NSEL, IMP, IUBN, program modality, or institutional or program

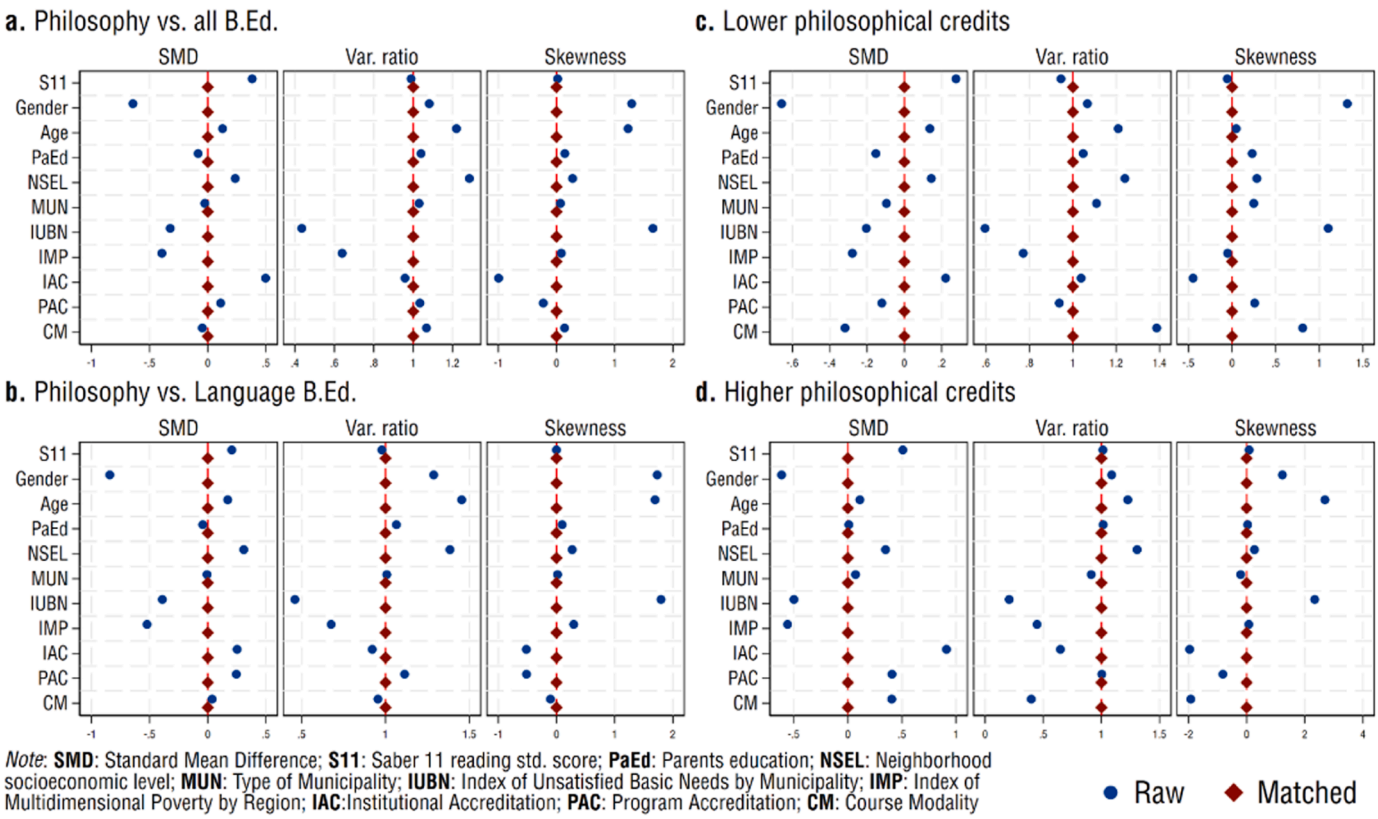

Fig. 1. Covariate balancing before and after matching.

Table 5
Effect of attending a B.Ed. in philosophy.

Critical reading std. scores	Comparison group		Philosophical credits		Saber 11 score		Gender		Program Accreditation	
	All B.Ed.	Language B.Ed.	Low (40-80)	High (80-120)	Low (≤ 5 SD)	High (> 5 SD)	Male	Female	No	Yes
Wald test			$F = 7.548$ * **		$F = 6.782$ * *		$F = 3.862$ *		$F = 3.717$ *	
ATT	0.401 * **	0.124 * **	0.344 * **	0.444 * **	0.440 * **	0.334 * **	0.417 * **	0.337 * **	0.436 * **	0.357 * **
	(0.020)	(0.026)	(0.027)	(0.030)	(0.031)	(0.026)	(0.026)	(0.032)	(0.029)	(0.029)
NATE	0.680 * **	0.398 * **	0.471 * **	0.937 * **	0.646 * **	0.444 * **	0.590 * **	0.724 * **	0.720 * **	0.564 * **
	(0.021)	(0.023)	(0.028)	(0.031)	(0.030)	(0.026)	(0.027)	(0.035)	(0.030)	(0.029)
N	123,125	19,176	123,125	123,125	75,779	47,346	39,824	83,301	73,923	49,202
Matched										
Yes	2012	1836	1131	899	920	1075	1248	779	1101	901
No	107	283	65	24	55	69	75	17	57	60
Total	2119	2119	1196	923	975	1144	1323	796	1158	961
Controls										
Total	121,006	17,057	121,929	122,202	74,804	46,202	38,501	82,505	72,765	48,241
Used	71,444	11,132	61,348	45,861	33,174	26,095	27,428	39,310	42,801	22,726
Unused	49,562	5925	60,581	76,341	41,630	20,107	11,073	43,195	29,964	25,515
bwidth	0.019	0.211	0.024	0.008	0.032	0.023	0.050	0.004	0.063	0.003
loss	2.02e-16	1.78e-15	2.56e-15	5.48e-15	1.23e-16	4.23e-13	4.95e-16	4.71e-17	1.43e-16	1.83e-16

Note: * ** $p < 0.001$, * * $p < 0.01$, * $p < 0.05$.

accreditation (for details, see the Appendix).

5.3.1. *Effect by philosophical credits*

Students in B.Ed. in philosophy with more philosophical credits surpass students in lower credit programs by 0.1 SD (Tables 5, columns 4 and 5; Fig. 5). This means the effect of philosophy relies not only in attending the major but also in the intensity of the treatment. The effect of philosophy on critical reading is not linear, contrary to what previous studies show (i.e., Farieta, 2022) as the difference of the more exposed group is not twofold when compared to the less exposed group (0.444 SD vs. 0.344 SD), perhaps indicating a ceiling effect on terms of credits for critical reading scores operating within the higher credits' interval sub-group (81−126). Nevertheless, even a lower exposition to philosophy has a large impact on student achievement, which can be worthwhile for a student interested in a multidisciplinary higher education and with a strong philosophical core.

5.3.2. *Effect by prior academic achievement*

We find contrasting effects according to students' prior academic achievement. Weaker students (i.e., with lower outcomes in the *Saber 11* test) outperform stronger students in 0.11 SD (Table 5, columns 6 and 7). *Saber 11* test scores are the main predictor of higher education

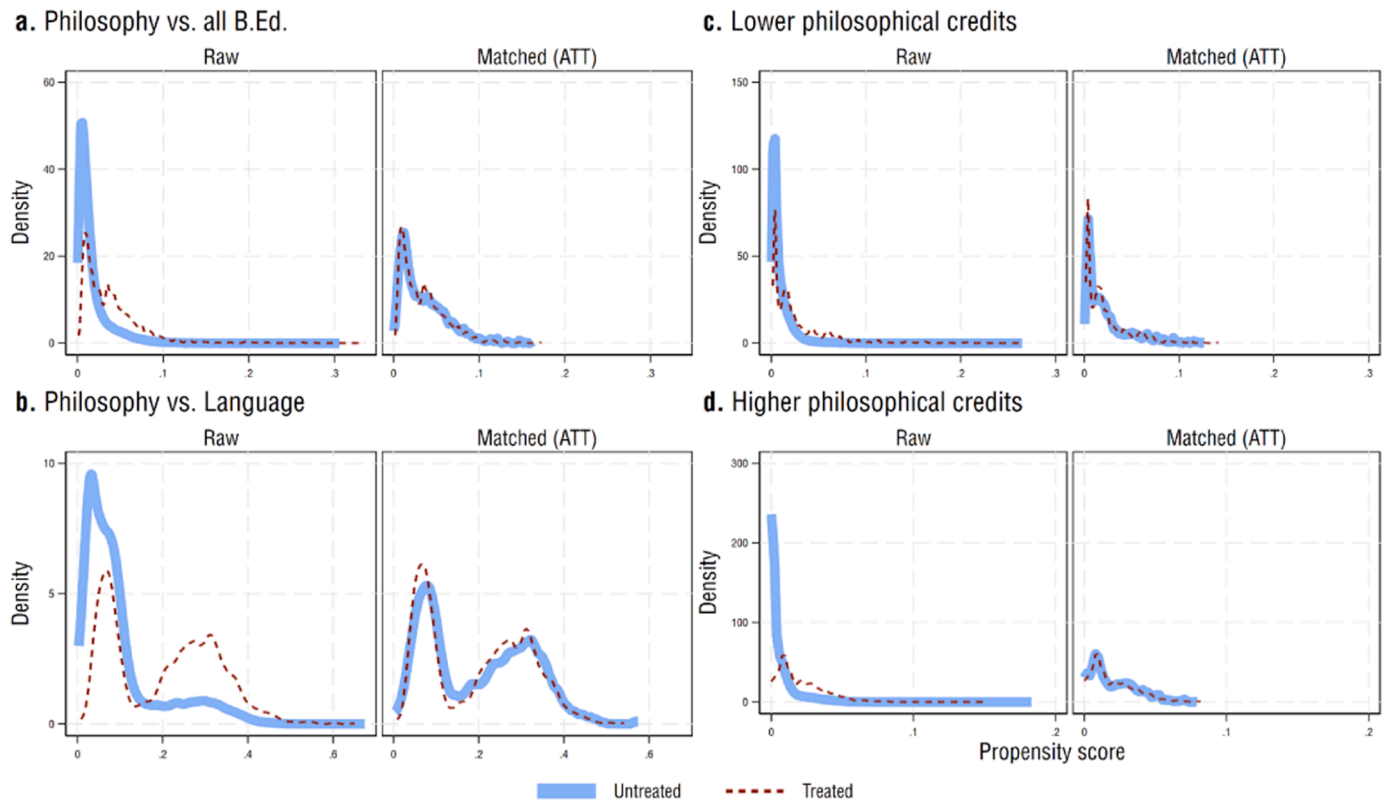

Fig. 2. Kernel density plot before and after matching.

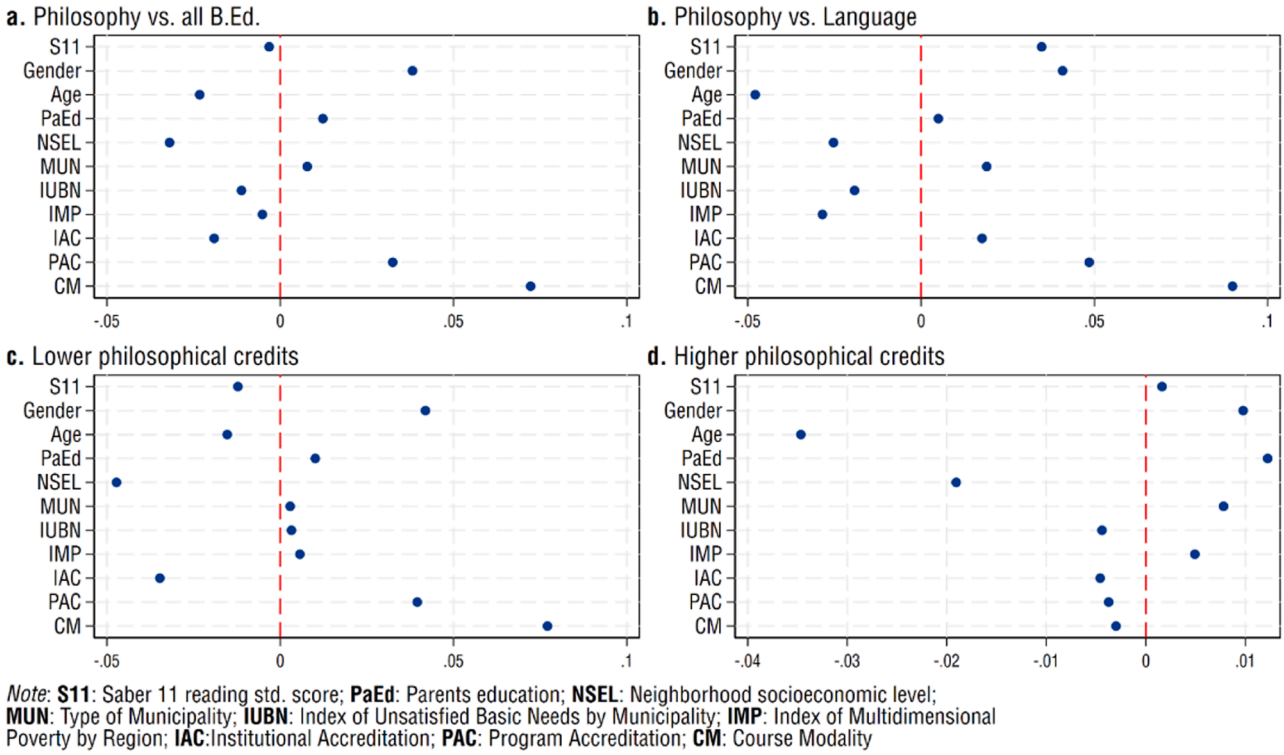

Fig. 3. Standardized Mean Difference (SMD) between the entire treated group and the matched by covariates.

outcomes (Sarmiento et al., 2019). This means that for students with learning deficits at the end of high school, philosophy has the potential to allow them to catch up with students that had higher scores at that point and, therefore, to remedy some of the gaps accumulated by a deficient basic education or other contextual disadvantages (e.g., lower parental education or adverse socioeconomic conditions). Philosophy therefore has a vital role on reducing existing learning-driven ability gaps.

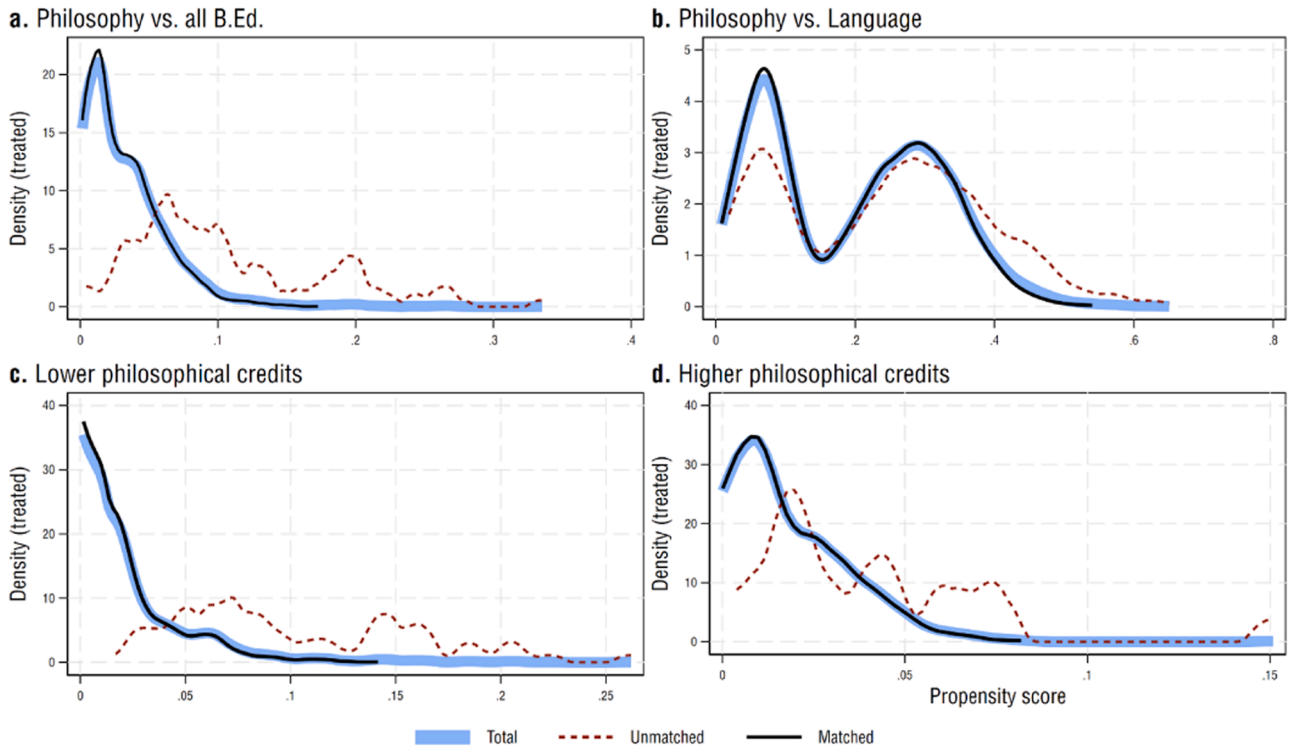

Fig. 4. Kernel density of the treated group.

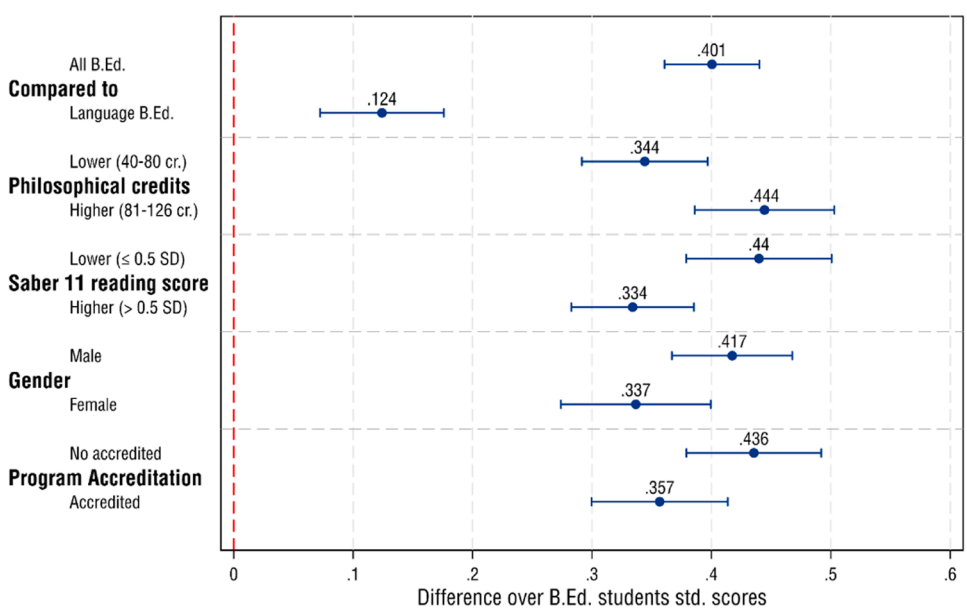

Fig. 5. ATT Estimates.

5.3.3. *Effect by gender*

We also find differences on the estimated ATT by gender. Our estimates indicate that male students achieve 0.08 SD higher than female students (Table 5, columns 8 and 9). This result is concerning since it shows a gender gap inside philosophy programs. In Colombia, women tend to overperform men in reading and literacy at the high school level (Abadía and Bernal, 2016; Cárcamo et al., 2020; Correa, 2016), but the results show that after majoring in philosophy, the situation changes drastically against women. It is important to consider the low percentage of women in B.Ed. in philosophy (32.03%) compared to the rest of B. Ed. (68.41%) (Table 3). In Colombia, 75% of the professors and staff in philosophy departments are male (Acevedo-Zapata and Rivera-Sanín, 2023), and this low female representation is likely to cause women's underperformance, despite other reasons related to a highly masculinized environment (Bernal-Ríos, 2022).

5.3.4. *Effect by program accreditation*

Students in non-accredited programs outperformed students in accredited ones by 0.079 SD (Table 5, columns 10 and 11). This result is at odds with the accreditation criteria, according to which the added value of the programs is a criterion to obtain the high-quality accreditation status (Colombia, MEN/CNA, 2013). A possible explanation for

this is the fact that students with low prior academic achievement are more likely to attend programs without accreditation ($n = 620$), in comparison with students with high prior academic achievement attending the same type of courses ($n = 383$). This is aligned with the earlier finding about the relatively higher impact on learning of studying philosophy for the lower prior academic achievement group.

6. Discussion

We found conclusive evidence about the advantages of studying philosophy in terms of critical reading for initial teacher education. Our results are in line with previous studies that applied multilevel fixed-effects linear regression (Farieta, 2022), and with studies about the effect of programmes like Philosophy with Children (García-Moriyón, Rebollo, and Colom, 2005), or the effect of logic and argumentation modules in undergraduate programs (Quintana and Schunn, 2019), or the research related to the effects of critical thinking in student achievement. The most remarkable point is that the effect size found is considerably large (0.401 SD) since it's rare to find in the literature a higher effect of interventions of critical thinking in reading or literacy, as the meta-analysis of Abrami et al. (2015) shows. The large sample we employed, and the fact that the analysis was done with the entire population of students who graduated from a B.Ed. in the country during the last 10 years, give additional validity to our estimates, allowing us to answer affirmatively RQ1: students in philosophy B.Ed. are in fact having better scores than students in a different B.Ed.

Regarding RQ2 —if students in B.Ed. in philosophy are obtaining higher scores than students in other B.Ed. programs focused specifically on language—, the answer is equally positive. It is worthy to consider that some of the programs with lower levels of philosophical credits combine philosophy and language, most of them with literature (Farieta, 2018; Farieta et al., 2015). This shows that in curricular terms, it would be a good idea to introduce more credits or subjects of philosophy in the B.Ed. in language, especially if these are related to philosophy or language, semiotics, argumentation, or other similar topics. It is likely that some of these programs already have these modules in their study plans, but a significant increment of these would have an impact on improving student outcomes in critical reading.

There are some important findings from RQ3 —i.e., if there are differences in outcomes according to student and program conditions. The main one is that students with lower scores on the *Saber 11* test have a higher improvement in their outcomes at the end of the undergraduate program. The findings are consistent with the literature, supporting the idea that those with lesser prior academic achievement can benefit from philosophy (Quintana and Schunn, 2019; Ventista, 2019). It is also very rare to find an intervention that faces lower prior academic achievement with the impact that philosophy have in our study (Schneider and Preckel, 2017). This a key finding of the paper from a policy perspective, namely, Philosophy has a remarkable potential to reduce academic gaps driven by socioeconomic conditions previously to being enrolled in an undergraduate program, and in consequence, to contribute to social justice.

There are some warnings from our analysis about how philosophy impacts on sub-populations, however. The gender gap is a critical concern, despite being a common issue inside the philosophy courses around the world (Antony, 2012; Beebee, 2021; Dougherty et al., 2015; Hutchison and Jenkins, 2013). Issues like stereotyped threats or implicit bias reported in the literature (Saul, 2013), seem to be affecting women's scores. It is worth noting that the gender gap is like the one found in STEM programs in Colombia for mathematics (Gómez S. et al., 2020), with similar issues like the lack of female role models and a highly masculinized environment (Acevedo-Zapata and Rivera-Sanín, 2023; Bernal-Ríos, 2022). This means that universities should take the staff gender gap seriously in these programs. All in all, this suggests that the gender gap in outcomes deserves more detailed research to gauge more precisely how wide the gap is, what its causes are, and therefore what are possible ways to address it.

The second important warning has to do with the high-quality accreditation. According to the literature, institutional accreditation is usually associated with higher student scores (Bayona et al., 2018; Cayón et al., 2020). Yet, the evidence is not conclusive about program accreditation, and some studies show positive effects on student outcomes (Camacho et al., 2016), but others show no association (Sarmiento et al., 2015), especially in philosophy (Farieta, 2020). As we said before, the possible reason of this negative effect has to do with the nature of self-selection into the program, that is: the fact that more students with low prior academic achievement attend non-accredited programs. Other explanation is that, since the accreditation is a peer review very demanding process in terms of resources and time (Arias G. et al., 2018; Rodríguez-Ávila, 2021), the programs could be investing more time in the bureaucracy involved and management tasks instead of teaching and research (Hanushek and Woessmann, 2015a). The main concern is that if this is true, the accreditation is favouring the ability of the programs to select students with higher student performances, but at the expense of taking into account their added value. This is especially concerning since the Colombian government imposed mandatory accreditation to the B.Ed. programs (Colombia, CRC, 2015, Jun. 9) to improve student achievement, but as it is shown, this has not been the case. However, since program accreditation was not the focus of this study, the conclusions should be considered with caveats as deeper research about the association between program accreditation and student outcomes is needed.

This research was limited to a specific context in a low-middle income country. Further research is needed to determine if in a different context or country the philosophy has the same effect in students' outcomes. Other of the limitations of the study are different factors reported by the literature that can improve the outcomes of philosophy students, like the class size or the teacher experience, educational level, or other characteristics of the programs (Ordóñez R et al., 2019; Sáenz-Castro et al., 2021).

7. Policy Recommendations

Recent changes in policy regulations for B.Ed. courses in Colombia (Colombia, CRC, 2015, Jun. 9) compelled these to obtain high-quality accreditation, but the results show a negative effect for courses with this award, not necessarily improving student outcomes. A revision of the accreditation criteria is needed to revise this issue, along with more research about its association with student outcomes. Additionally, these policy regulations asked the programs to increase the school practices (Colombia, MEN, 2016, Feb. 3; 2017, Sep. 15) which made them reduce the disciplinary credits (Arias G. et al., 2018; Valderrama-Leongómez et al., 2019; Farieta et al., 2024). This is concerning since the reduction can negatively impact student outcomes, along with other negative effects such as the reduction of institutional autonomy and a higher risk of closing the programs in regions more in need (Rodríguez-Ávila et al., 2021). The policy regulations were cancelled (Colombia, CRC, 2019, May 25), but our findings are proof that these policies could be more harmful than beneficial, and a wider discussion with more evidence about teacher education is needed.

On the other side, many philosophy programs have either been closed or threatened to be closed in the last ten years around the world. Sometimes these threats occur on a national level, like in Brazil, where a former President announced the defunding of philosophy programs in all public universities (Bolsonaro, 2019; Apr. 26), asserting that these do not have impact in the national economy. There were reactions against this announcement from academics in Brazil, Latin America, and the entire world (Weinberg, 2019, Apr. 30), and finally the programs weren't closed. Our results show that these reactions were not only conceptual correct but empirically sounded, and the widespread opinion that philosophy do not have any potential impact on economic growth is false.

The risks of eliminating the arts and humanities, and especially philosophy, are imminent for democracies. Some authors claim that due to a lack of critical reading skills, some political situations affecting democracy worldwide have occurred due to the radicalization of social or cultural differences and the rise of intolerant or politically polarized environments, which happened, for example, during Brexit or the last elections in the US (Barton, 2019; Oelkers, 2017). We agree with the philosophical tradition that reading, critical thinking, and philosophy, have value per se and are fundamental for human flourishing, and are additionally important values for a democratic, diverse, and pluralistic society (Nussbaum, 2012). But besides that, it is important to note how many of the main critics of these are also misled and biased.

This study has proved how philosophy can be a vital mechanism to improve student performance in critical reading. If the premises coming from the economics of education are accepted and better teachers have positive effects on the economic future of the students (Chetty et al., 2014a, 2014b; Hanushek, 2011; Hanushek and Rivkin, 2010; Hanushek et al., 2019), there is no doubt that philosophy has an important role to improve national economies in the long term. This is particularly relevant in the Colombian context, where teacher quality and academic achievement have strong effects on future student outcomes (Bonilla--Mejía et al., 2018). The number of students in B.Ed. in philosophy has been reduced year by year (Farieta et al., 2024; cf. Table 1) which should be a big concern because, for the sake of educational quality, they should be increasing and the enrolling on these programs should be promoted. In terms of educational public policy, this is a central argument to support the programs at risk of being closed, but also to promote the opening of more philosophy programs in regions where they do not exist, not only in Colombia, but also in other countries with emerging and poor economies. The recommendation is also to increase the content and credits of philosophy, critical thinking, and argumentation in other B.Ed. and other programs whose students have lower prior academic achievement.

CRediT authorship contribution statement

Farieta Alejandro: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Delprato Marcos:** Formal analysis, Investigation, Methodology, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. However, the authors are thankful to the Commission on Education of the Colombian Society of Philosophy, whose members read a draft of this paper and made valuable comments to improve the final version, the discussion, and the policy recommendations. The authors are also grateful for the comments provided by two anonymous referees. Also, thanks to the philosophical community in Colombia for their valuable comments and support, but also for keeping on doing, teaching, and spreading philosophy, just for the sake of wisdom and a better life for Colombians, humans, living things, and the planet, even in the face of unfavorable circumstances.

Appendix. Estimates by subgroups without significant differences

Estimates by subgroups without significant differences

Critical reading std. scores	Parents Education		Living area		NSEL		IUBN		IMP		Program Modality		Institutional Accreditation	
	Lower	Higher	Big city	Small/rural	Lower	Higher	Lower	Higher	Lower	Higher	Distance	On-Campus	No	Yes
Wald test	$F = 0.67$		$F = 0.931$		$F = 1.444$		$F = 0.01$		$F = 0.246$		$F = 0.411$		$F = 0.103$	
ATT	0.410 * **	0.376 * **	0.369 * **	0.415 * **	0.424 * **	0.373 * **	0.397 * **	0.393 * **	0.394 * **	0.371 * **	0.369 * **	0.406 * **	0.418 * **	0.404 * **
	(0.028)	(0.030)	(0.042)	(0.023)	(0.026)	(0.034)	(0.029)	(0.030)	(0.026)	(0.036)	(0.053)	(0.021)	(0.035)	(0.024)
NATE	0.692 * **	0.691 * **	0.602 * **	0.712 * **	0.699 * **	0.568 * **	0.583 * **	0.647 * **	0.598 * **	0.614 * **	0.586 * **	0.724 * **	0.569 * **	0.531 * **
	(0.030)	(0.030)	(0.045)	(0.023)	(0.027)	(0.034)	(0.028)	(0.031)	(0.026)	(0.036)	(0.045)	(0.022)	(0.036)	(0.025)
N	59,728	63,397	29,942	93,183	87,451	35,674	43,638	79,487	57,634	65,491	25,485	97,640	73,314	49,811
Matched														
Yes	1052	924	514	1507	1264	744	1000	996	1324	640	416	1622	737	1270
No	73	70	24	74	59	52	54	69	76	79	63	18	23	89
Total	1125	994	538	1581	1323	796	1054	1065	1400	719	479	1640	760	1359
Controls														
Total	58,603	62,403	29,404	91,602	86,128	34,878	42,584	78,422	56,234	64,772	25,006	96,000	72,554	48,452
Used	26,847	30,149	11,170	59,813	44,522	17,413	24,378	29,311	32,546	15,893	9230	60,208	37,521	33,805
Unused	31,756	32,254	18,234	31,789	41,606	17,465	18,206	49,111	23,688	48,879	15,776	35,792	35,033	14,647
Bandwidth	0.026	0.019	0.038	0.017	0.021	0.071	0.052	0.024	0.041	0.012	0.057	0.015	0.042	0.012
Loss	1.94e-16	6.26e-16	6.21e-16	8.24e-17	1.19e-16	1.33e-13	1.18e-14	5.97e-15	3.56e-14	4.17e-15	1.72e-15	1.12e-13	1.22e-14	3.59e-15

Note: * ** $p < 0.001$, * * $p < 0.01$, * $p < 0.05$.

References

Abadía, L.K., Bernal, G., 2016. Brechas de género en el rendimiento escolar a lo largo de la distribución de puntajes: evidencia pruebas Saber 11. Vniversitas Económica 16 (9) https://cea.javeriana.edu.co/vniversitas-economica#2016.

Abrami, P.C., Bernard, R.M., Borokhovski, E., Waddington, D.I., Wade, C.A., Persson, T., 2015. Strategies for teaching students to think critically: a meta-analysis. Rev. Educ. Res. 85 (2), 275–314. https://doi.org/10.3102/0034654314551063.

Acevedo-Zapata, D., Rivera-Sanín, M., 2023. Gender Equality in Colombia's Philosophy Programs: Faculty Participation. In: Hagengruber, R. (Ed.), Women Philosophers on Economics, Technology, Environment, and Gender History: Shaping the Future, Rethinking the Past. De Gruyter, pp. 295–322. https://doi.org/10.1515/9783111051802-027.

Aguilera-Prado, M., 2017. Diferencias en las competencias específicas entre los educadores de formación presencial y no presencial. In: Aguilera-Prado, M., Farieta-Barrera, A. (Eds.), Evaluación de la educación superior: una mirada desde Latinoamérica. Editorial Uniagustiniana, pp. 163–215. https://doi.org/10.28970/9789585639539.03.

Antony, L., 2012. Different voices or perfect storm: why are there so few women in philosophy? J. Soc. Philos. 43 (3), 227–255. https://doi.org/10.1111/j.1467-9833.2012.01567.x.

Arias G, D.H., Díaz F., O.C., Garzón B, I., León P, A.C., Rodríguez A, S.P., & Valbuena U, E.O. (2018). *Entre las exigencias de calidad y las condiciones de desigualdad: Formación inicial de profesores en Colombia*. Universidad Pedagógica Nacional. http://hdl.handle.net/20.500.12209/7826.

Arias-Velandia, N., Rincón-Báez, W.U., Cruz-Pulido, J.M., 2018. Desempeño de mujeres y hombres en educación superior presencial, virtual y a distancia en Colombia. Panorama 12 (22). https://doi.org/10.15765/pnrm.v12i22.1142.

Arias-Velandia, N., Rincón-Báez, W.U., Cruz-Pulido, J.M., 2021. Diferencia de logro geolocalizado en educación presencial y a distancia en Colombia. Rev. Electrónica De. Invest. Educ. 23, 1–22. https://doi.org/10.24320/redie.2021.23.e18.3711.

Avendaño, R., Barrera-Osorio, F., Nieto-Parra, S., Vever, F., 2016. Understanding Student Performance Beyond Traditional Factors: Evidence from PISA. OECD Development Centre Working. OECD Publishing. https://doi.org/10.1787/18151949.

Barber, M., & Mourshed, M. (2007). *How the World's Best-Performing School Systems Come Out on Top*. McKinsey. https://www.mckinsey.com/industries/education/our-insights/how-the-worlds-best-performing-school-systems-come-out-on-top.

Bardach, L., Klassen, R.M., 2020. Smart teachers, successful students? A systematic review of the literature on teachers' cognitive abilities and teacher effectiveness. Educ. Res. Rev. 30, 100312 https://doi.org/10.1016/j.edurev.2020.100312.

Barrera-Osorio, F.; Maldonado, D., & Rodríguez, C. (2012). *Calidad de la educación básica y media en Colombia: diagnóstico y propuestas*. Documentos CEDE 010321. Universidad de los Andes. Facultad de Economía. https://ideas.repec.org/p/col/000089/010321.html.

Barrow, R., 2019. The Decline of Philosophy in Educational Study and Why It Matters. In. In: Colgan, A., Maxwell, B. (Eds.), The Importance of Philosophy in Teacher Education: Mapping the Decline and its Consequences, 2019. Routledge. https://doi.org/10.4324/9780429426827.

Barton, C.C., 2019. Critical literacy in the post-truth media landscape. Policy Futures Educ. 17 (8), 1024–1036. https://doi.org/10.1177/1478210319831569.

Bayona-Rodríguez, H., Bedoya, J., Pulido, X., & Sánchez, F. (2018). Efectos (Algunos No Esperados) De La Acreditación De Alta Calidad En Las Universidades Colombianas [Effects (Some Non-Expected) of High Quality Accreditation on Colombian Universities]. Documento CEDE No. 2018–35 (SSRN Scholarly Paper 3215465). https://doi.org/10.2139/ssrn.3215465.

Becker, G.S., 2009. Human capital: A theoretical and empirical analysis, with special reference to education. University of Chicago press.

Beebee, H., 2021. Women in Philosophy: what's changed? Philosophers' Mag. (*93*), 50–56. https://doi.org/10.5840/tpm20219340.

Belfiore, E., 2015. Impact', 'value' and 'bad economics': making sense of the problem of value in the arts and humanities. Arts Humanit. High. Educ. 14 (1), 95–110. https://doi.org/10.1177/1474022214531503.

Bernal-Ríos, L.P., 2022. Cuatro injusticias epistémicas en los currículos universitarios de filosofía en Colombia: Anglo-eurocentrismo, racismo, sexismo y humanismo (Article). Cuad. De. Filos. Latinoam. 43 (126), 126. https://doi.org/10.15332/25005375.7604.

Bolsonaro, J. (2019, Apr. 26) O Ministro da Educação @abrahamWeinT estuda descentralizar investimento em faculdades de filosofia e sociologia (humanas). Alunos já matriculados não serão afetados. O objetivo é focar em áreas que gerem retorno imediato ao contribuinte, como: veterinária, engenharia e medicina [Tweet]. https://twitter.com/jairbolsonaro/status/1121713534402990081.

Bonilla-Mejía, L., Londoño-Ortega, E., Cardona-Sosa, L., Trujillo-Escalante, L.D., 2018. ¿Quiénes son los docentes en Colombia? Características generales y brechas regionales. Banco De. la Repúb. De. Colomb. Doc. De. Trab. sobre Econ. Reg. Y. Urbana 276. https://doi.org/10.32468/dtseru.276.

Bruns, B., & Luque, J. (2015). *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. World Bank. https://doi.org/10.1596/978–1-4648–0151-8.

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., Schmidt, W., 2019. A Review of the Literature on Teacher Effectiveness and Student Outcomes. In: Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., Schmidt, W. (Eds.), *Teaching for Excellence and Equity: Analyzing Teacher Characteristics, Behaviors and Student Outcomes with TIMSS*. Springer International Publishing, pp. 7–17. https://doi.org/10.1007/978-3-030-16151-4_2.

Bynum, T.W., Lipman, M., 1976. Philosophy for children. Blackwell.

Camacho, A., Messina, J., & Uribe Barrera, J. (2016). *The Expansion of Higher Education in Colombia: Bad Students or Bad Programs?* (SSRN Scholarly Paper No. 2921965). https://doi.org/10.2139/ssrn.2921965.

Cárcamo, C., Moreno, A., Del Barrio, C., 2020. Diferencias de género en matemáticas y lengua: Rendimiento académico, autoconcepto y expectativas. Suma Psicológica 27 (1), 27–34. https://doi.org/10.14349/sumapsi.2020.v27.n1.4.

Cárcamo V, C., Mola Á, J.A., 2012. Diferencias por sexo en el desempeño académico en Colombia: Un análisis regional. Econ. Región 6 (1), 133–169 https://revistas.utb.edu.co/economiayregion/article/view/137.

Cassidy, C., Christie, D., 2013. Philosophy with children: talking, thinking, and learning together. Early Child Dev. Care 183 (8), 1072–1083. https://doi.org/10.1080/03004430.2013.773509.

Cassidy, C., Marwick, H., Deeney, L., McLean, G., 2018. Philosophy with children, self-regulation and engaged participation for children with emotional-behavioural and social communication needs. Emot. Behav. Difficulties 23 (1), 81–96. https://doi.org/10.1080/13632752.2017.1388654.

Castagnoli, L., Fait, P., 2022. Cambridge Companion to Ancient Logic. Cambridge University Press.

Castro, M., Ruiz, J., Guzmán, F., 2018. Cruce de las pruebas nacionales Saber 11 y Saber Pro en Antioquia, Colombia: una aproximación desde la regresión geográficamente ponderada (GWR). Rev. Colomb. De. Educ. (74), 63–79. https://doi.org/10.17227/rce.num74-6898.

Cayón, E., Correa, J.S., Sarmiento, J.A., 2020. Does quality make a difference for higher education graduates in Colombia? Measuring differences in accredited and non-accredited institutions with similar financial characteristics. WSEAS Trans. Bus. Econ. 17, 859–868. https://doi.org/10.37394/23207.2020.17.84.

Chetty, R., Friedman, J.N., Rockoff, J.E., 2014a. Measuring the impacts of teachers i: evaluating bias in teacher value-added estimates. Am. Econ. Rev. 104 (9), 2593–2632. https://doi.org/10.1257/aer.104.9.2593.

Chetty, R., Friedman, J.N., Rockoff, J.E., 2014b. Measuring the impacts of teachers ii: teacher value-added and student outcomes in adulthood. Am. Econ. Rev. 104 (9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633.

Cochran-Smith, M., 2005. The new teacher education: for better or for worse? Educ. Res. 34 (7), 3–17. https://doi.org/10.3102/0013189X034007003.

Cochran-Smith, M., 2021. Rethinking teacher education: the trouble with accountability. Oxf. Rev. Educ. 47 (1), 8–24. https://doi.org/10.1080/03054985.2020.1842181.

Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., Van Klaveren, C., 2018. Teacher characteristics and their effects on student test scores: a systematic review. J. Econ. Surv. 32 (3), 848–877. https://doi.org/10.1111/joes.12210.

Colom, R., Moriyón, F.G., Magro, C., Morilla, E., 2014. The Long-term Impact of Philosophy for Children: A Longitudinal Study (Preliminary Results). Anal. Teach. Philos. Prax. 35 (1). https://journal.viterbo.edu/index.php/atpp/article/view/1129.

Colombia, CRC. (2009, Jul. 13). Law 1324. Diario Oficial, 145(47,409). https://funcionpublica.gov.co/eva/gestornormativo/norma.php?i=36838.

Colombia, CRC. (2015, Jun. 9). Law 1753. Diario oficial, 151(49,538). http://www.secretariasenado.gov.co/senado/basedoc/ley_1753_2015.html.

Colombia, CRC. (2019, May 25). Law 1955. Diario Oficial, 155(50964). http://www.secretariasenado.gov.co/senado/basedoc/ley_1955_2019.html.

Colombia, MEN. (2016, Feb. 3). Resolution 2041. Diario Oficial, 152(49,776). https://www.mineducacion.gov.co/1759/articles-356982_recurso_1.pdf.

Colombia, MEN. (2017, Sep. 15). Resolution 18583. Diario oficial, 153(50,357). https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=71384&dt=S.

Colombia, PRC. (2009, Oct. 14). Decree 3963. Diario Oficial, 145(46,502). https://funcionpublica.gov.co/eva/gestornormativo/norma.php?i=37606.

Colombia, MEN/CNA. (2013). *Lineamientos para la acreditación de programas de pregrado*. MEN/CNA. https://www.mineducacion.gov.co/1621/articles-342684_recurso_1.pdf.

Corcoran, R.P., O'Flaherty, J., 2018. Factors that predict pre-service teachers' teaching performance. J. Educ. Teach. 44 (2), 175–193. https://doi.org/10.1080/02607476.2018.1433463.

Correa F, J., 2016. Desempeño académico y diferencias de género en Colombia: un análisis con base en las pruebas TIMSS 2007. Rev. Soc. Y. Econ. (30), 15–42. https://doi.org/10.25100/sye.v0i30.3899.

Cvecic, I., Sokolic, D., Mrak, M., 2019. Higer education and economic prosperity at regional level. Rev. Port. De. Estud. Reg. 50, 9–25 http://www.apdr.pt/siteRPER/numeros/RPER50/50.1.pdf.

DANE. (2018). Censo Nacional de Población y Vivienda 2018. Microdatos. https://microdatos.dane.gov.co/index.php/catalog/643.

Delprato, M., Köseleci, N., Antequera, G., 2015. Educación para Todos en América Latina: evolución del impacto de la desigualdad escolar en los resultados educativos. Rev. Latinoam. De. Educ. Comp.: RELEC 6 (8), 45–75 https://www.saece.com.ar/relec/revistas/8/mon2.pdf.

Didion, L., Toste, J.R., Filderman, M.J., 2020. Teacher professional development and student reading achievement: a meta-analytic review of the effects. J. Res. Educ. Eff. 13 (1), 29–66. https://doi.org/10.1080/19345747.2019.1670884.

Dougherty, T., Baron, S., Miller, K., 2015. Why do female students leave philosophy? The story from Sydney. Hypatia 30 (2), 467–474. https://doi.org/10.1111/hypa.12150.

Farieta, A., 2018. Análisis comparativo de los planes de estudio de los programas de pregrado en filosofía en Colombia, segunda parte: contenidos. Folios (48), 167–191. https://doi.org/10.17227/folios.48-8142.

Farieta, A., 2020. Initial education of philosophy teachers in colombia: association between new public policy requirements and national standardized tests. Aust. J. Teach. Educ. 45 (6) https://doi.org/10.14221/ajte.2020v45n6.3.

Farieta, A., 2022. La incidencia de la filosofía en la lectura crítica y otras competencias genéricas en los programas de formación de docentes en Colombia. W. Herrera.

Actualidad y Defensa de la Filosofía. Universidad del Rosario, Universitaria Agustiniana, pp. 241–290. https://doi.org/10.12804/urosario9789587848182.
Farieta, A., Murillo L, L.A., 2024. Reformas a las licenciaturas en Colombia: efectos en programas de filosofía. Perfiles Educ. (in press).
Farieta, A., Gómez, M., Almeida, S., 2015. Análisis comparativo de los planes de estudio de filosofía en Colombia, parte 1: generalidades y flexibilidad. Folios (42), 87–104. https://doi.org/10.17227/01234870.42folios87.104.
Frölich, M., 2004. Finite-sample properties of propensity-score matching and weighting estimators. Rev. Econ. Stat. 86 (1), 77–90 https://www.jstor.org/stable/3211661.
Frölich, M., 2005. Matching estimators and optimal bandwidth choice. Stat. Comput. 15, 197–215. https://doi.org/10.1007/s11222-005-1309-6.
Furlong, J., Cochran-Smith, M., Brennan, M., 2008. Politics and policy in teacher education: international perspectives. Teach Teach 14 (4), 265–269. https://doi.org/10.1080/13540600802037660.
Galdo, J.C., Smith, J., Black, D., 2008. Bandwidth selection and the estimation of treatment effects with unbalanced data. Ann. D. 'Économie Et. De. Stat. 91/92, 89–216 https://www.iza.org/publications/dp/3095/bandwidth-selection-and-the-estimation-of-treatment-effects-with-unbalanced-data.
García J, S., Maldonado C, D., Perry R, G., Rodríguez O, C., & Saavedra C, J.E. (2014). Tras la excelencia docente: cómo mejorar la educación de todos los colombianos. Fundación Compartir, Punto Aparte. https://www.fundacioncompartir.org/pdf/Tras%20la%20excelencia%20docente%20-%20estudio%20final.pdf.
García J, S., Rodríguez Orgales, C., Sánchez Torres, F.J., & Bedoya Ospina, J.G.. (2015). *La lotería de la cuna: La movilidad social a través de la educación en los municipios de Colombia.* Documentos CEDE 013816. Universidad de los Andes. Facultad de Economía. https://ideas.repec.org/p/col/000089/013816.html.
García V, M., Espinosa R, J.R., Jiménez A, F., & Parra H, J.D. (2013). *Separados y desiguales. Educación y clases sociales en Colombia.* Dejusticia. https://www.dejusticia.org/wp-content/uploads/2017/04/fi_name_recurso_591.pdf.
García-González, J.D., Skrita, A., 2019. Predicting academic performance based on students' family environment: evidence for Colombia using classification trees. Psychol., Soc. Educ. 11 (3), 299–311. https://doi.org/10.25115/psye.v11i3.2056.
García-Moriyón, G., Rebollo, I., Colom, R., 2005. Evaluating philosophy for children. A meta-analysis. Think.: J. Philos. Child. 17 (4), 14–22. https://doi.org/10.5840/thinking20051743.
Gibbons, S., Vignoles, A., 2012. Geography, choice, and participation in higher education in England. Reg. Sci. Urban Econ. 42 (1–2), 98–113. https://doi.org/10.1016/j.regsciurbeco.2011.07.004.
Goldhaber, D., Liddle, S., Theobald, R., 2013. The gateway to the profession: assessing teacher preparation programs based on student achievement. Econ. Educ. Rev. 34, 29–44. https://doi.org/10.1016/j.econedurev.2013.01.011.
Gomez S, S.C., Abadía A, L.K., Bernal N, G.L., 2020. Women in STEM: does college boost their performance? High. Educ. 79, 849–866. https://doi.org/10.1007/s10734-019-00441-0.
Hainmueller, J., 2012. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. Political Anal. 20 (1), 25–46. https://doi.org/10.1093/pan/mpr025.
Hanushek, E.A., 2011. The economic value of higher teacher quality. Econ. Educ. Rev. 30 (3), 466–479. https://doi.org/10.1016/j.econedurev.2010.12.006.
Hanushek, E.A., Rivkin, S.G., 2010. Generalizations about using value-added measures of teacher quality. Am. Econ. Rev. 100 (2), 267–271. https://doi.org/10.1257/aer.100.2.267.
Hanushek, E.A., Woessmann, L., 2012a. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. J. Econ. Growth 17 (3), 267–321. https://doi.org/10.1007/s10887-012-9081-x.
Hanushek, E.A., Woessmann, L., 2012b. Schooling, educational achievement and the latin american growth puzzle. J. Dev. Econ. 99 (2), 497–512. https://doi.org/10.1016/j.jdeveco.2012.06.004.
Hanushek, E.A., Woessmann, L., 2015a. The knowledge capital of nations: education and the economics of growth. MIT Press.
Hanushek, E.A., Woessmann, L., 2015b. Universal Basic Skills: What Countries Stand to Gain. OECD. https://doi.org/10.1787/9789264234833-en.
Hanushek, E.A., Woessmann, L., 2016. Knowledge capital, growth, and the east asian miracle. Science 351 (6172), 344–345. https://doi.org/10.1126/science.aad7796.
Hanushek, E.A., Piopiunik, M., Wiederhold, S., 2019. The value of smarter teachers: international evidence on teacher cognitive skills and student performance. J. Hum. Resour. 54 (4), 857–899. https://doi.org/10.3368/jhr.54.4.0317.8619R1.
Hausman, A., Boardman, F., Kahane, H., 2021. Logic and Philosophy. A Modern Introduction, 13th ed. Hackett.
Helbig, M., Jähnen, S., Marczuk, A., 2017. Does place of residence matter? How the geographical proximity of universities affects enrollment in higher education in Germany. Z. fur Soziol. 46 (1), 55–70. https://doi.org/10.1515/zfsoz-2017-1004.
Hernández G, O.E., Padilla G, L.E., 2019. Expectativas de los estudiantes hacia la educación superior: influencia de variables familiares, personales y escolares. Socio (México) 34 (98), 221–251 http://www.sociologicamexico.azc.uam.mx/index.php/Sociologica/article/view/1546/1448.
Herrera, W., 2022. Actualidad y Defensa de la Filosofía. Bogotá. Universidad del Rosario, Universitaria Agustiniana https://doi.org/10.12804/urosario9789587848182.
Huber, M., Lechner, M., Wunsch, C., 2013. The performance of estimators based on the propensity score. J. Econ. 175 (1), 1–21. https://doi.org/10.1016/j.jeconom.2012.11.006.
Hutchison, K., Jenkins, F., 2013. Women in Philosophy: What Needs to Change?. Oxford University Press.
ICFES. (2013). *Sistema Nacional de Evaluación Estandarizada de la Educación: Alineación de* Saber 11. ICFES. https://www.icfes.gov.co/documents/39286/1645749/Alineacion+examen+Saber+11.pdf.
ICFES. (2017). Saber Pro desde 2016. Boletín Saber en Breve, (17). http://www.icfes.gov.co/edicion-17-boletin-saber-en-breve.
ICFES. (2018). ¿Cómo se generan los puntajes en las pruebas saber del ICFES? Saber al detalle (1): 2–5. http://www.icfes.gov.co/divulgaciones-establecimientos/boletin-saber-al-detalle.
Işiklar, S., Öztürk, Y.A., 2022. The effect of philosophy for children (P4C) curriculum on critical thinking through philosophical inquiry and problem-solving skills. Article 1 Int. J. Contemp. Educ. Res. 9 (1). https://doi.org/10.33200/ijcer.942575.
Jann, B. (2017a). Why propensity scores should be used for matching. German Stata Users' Group Meetings 2017 01. Stata Users Group. https://www.stata.com/meeting/germany17/slides/Germany17_Jann.pdf.
Jann, B. (2017b). KMATCH: Stata module for multivariate-distance and propensity-score matching, including entropy balancing, inverse probability weighting, (coarsened) exact matching, and regression adjustment. *Statistical Software Components S458346*, Boston College Department of Economics, revised 19 Sep 2020. https://ideas.repec.org/c/boc/bocode/s458346.html.
Karadağ, F., Demirtaş, V.Y., 2018. The effectiveness of the philosophy with children curriculum on critical thinking skills of pre-school children. Educ. Sci. 43 (195) https://doi.org/10.15390/EB.2018.7268.
Kennedy, N., Kennedy, D., 2011. Community of philosophical inquiry as a discursive structure, and its role in school curriculum design. J. Philos. Educ. 45 (2), 265–283. https://doi.org/10.1111/j.1467-9752.2011.00793.x.
King, G., Nielsen, R., 2019. Why propensity scores should not be used for matching. Political Anal. 27 (4), 435–454. https://doi.org/10.1017/pan.2019.11.
Koedel, C., Parsons, E., Podgursky, M., Ehlert, M., 2015. Teacher preparation programs and teacher quality: are there real differences across programs? Educ. Financ. Policy 10 (4), 508–534. https://doi.org/10.1162/EDFP_a_00172.
Lipman, M., Sharp, A.M., 1978. Growing up with Philosophy. Temple University Press.
Machura, P., 2018. Flourishing vs. market: towards the aristotelian concept of education. Filozofia 73 (2), 145–157 http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.cejsh-51a222aa-7e18-477b-8a93-dc797007754c.
Mihaly, K., McCaffrey, D., Sass, T.R., Lockwood, J.R., 2013. Where you come from or where you go? Distinguishing between school quality and the effectiveness of teacher preparation program graduates. Educ. Financ. Policy 8 (4), 459–493. https://doi.org/10.1162/EDFP_a_00110.
Morais, J., 2018. Literacy and democracy. Lang., Cogn. Neurosci. 33 (3), 351–372. https://doi.org/10.1080/23273798.2017.1305116.
Morris, E., 2012. Learning the Hard Way. Masculinity, Place, and the Gender Gap in Education. Rutgers University Press.
Münchow, H., Tiffin-Richards, S.P., Fleischmann, L., et al., 2023. Promoting students' argument comprehension and evaluation skills: Implementation of two training interventions in higher education. Z. Erzieh. 26, 703–725. https://doi.org/10.1007/s11618-023-01147-x.
Nguyen, T.-L., Collins, G.S., Spence, J., Daurès, J.-P., Devereaux, P.J., Landais, P., Le Manach, Y., 2017. Double adjustment in propensity score matching analysis: choosing a threshold for considering residual imbalance. BMC Med. Res. Methodol. 17 (1), 78 https://doi.org/10.1186/s12874-017-0338-0.
Nussbaum, M.C., 2012. Non for profit. Why Democracy Needs the Humanities, (2nd. ed.).,. Princeton University Press. With an afterword by the author.
O'Brien, D., 2015. Cultural value, measurement and policy making. Arts Humanit. High. Educ. 14 (1), 79–94. https://doi.org/10.1177/1474022214533892.
OECD (2018a). *Effective Teacher Policies: Insights from PISA* . OECD Publishing. https://doi.org/10.1787/9789264301603-en.
OECD. (2018b). *PISA 2015 Results in Focus.* OECD Publishing. https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf.
OECD, 2016. PISA 2015 Country Note for Colombia. OECD Publishing. https://www.oecd.org/pisa/PISA-2015-Colombia.pdf.
OECD, 2019. PISA 2018 Assessment and Analytical Framework. OECD Publishing. https://doi.org/10.1787/b25efab8-en.
Oelkers, J., 2017. The European Crisis and Education for Democracy. Eur. Leg. 22 (7–8), 832–843. https://doi.org/10.1080/10848770.2017.1368781.
Olmos-Peñuela, J., Benneworth, P., Castro-Martínez, E., 2015. Are sciences essential and humanities elective? Disentangling competing claims for humanities' research public value. Arts Humanit. High. Educ. 14 (1), 61–78. https://doi.org/10.1177/1474022214534081.
Ordóñez R, P., Guarnizo V, M., Fetecua S, I., Gómez A, J., & Rodriguez, E. (2019). El Factor Docentes en las Resultados de Valor Agregado de las Instituciones de Educación Superior Colombianas. *Proceedings of the 17th LACCEI International Multi-Conference for Engineering, Education, and Technology: "Industry, Innovation, and Infrastructure for Sustainable Cities and Communities."* https://doi.org/10.18687/LACCEI2019.1.1.173.
OREALC/UNESCO. (2013). *Antecedentes y criterios para la Elaboración de Políticas Docentes en América Latina y el Caribe.* OREALC/UNESCO. http://unesdoc.unesco.org/images/0022/002232/223249S.pdf.
OREALC/UNESCO. (2016). Perspectivas sobre políticas docentes en América Latina y el Caribe. Aprendizajes de la Estrategia Regional de Docentes de la OREALC/UNESCO 2011–2016. OREALC/UNESCO. http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/Perspectivas-Politicas-Docentes-LAC.pdf.
Pineda, P., Celis, J., 2018. The maelstrom of online programs in colombian teacher education. Educ. Policy Anal. Arch. 26 (114) https://doi.org/10.14507/epaa.26.3873.
Prada-Dussán, M., 2023. Reducción de la filosofía: dos décadas de políticas públicas en la construcción de la enseñanza de la filosofía en Colombia. Estud. De. Filos. 69 https://doi.org/10.17533/udea.ef.353222.

Quintana, R., Schunn, C., 2019. Who benefits from a foundational logic course? Effects on Undergraduate Course Performance. J. Res. Educ. Eff. 12 (2), 191–214. https://doi.org/10.1080/19345747.2018.1543372.

Rangel, C., Lleras, C., 2010. Educational inequality in Colombia: family background, school quality and student achievement in Cartagena. Int. Stud. Sociol. Educ. 20 (4), 291–317. https://doi.org/10.1080/09620214.2010.530855.

Rockoff, J.E., 2004. The impact of individual teachers on student achievement: evidence from panel data. Am. Econ. Rev. 94 (2), 247–252. https://doi.org/10.1257/0002828041302244.

Rodríguez, A., Gómez L, V., Ariza D, M., 2014. Calidad de la educación superior a distancia y virtual: un análisis de desempeño académico en Colombia. Invest. Desarro. 22 (1), 58–99. https://doi.org/10.14482/indes.22.1.6079.

Rodríguez, A., G., Ariza D, M., Ramos R, J.L., 2014. Calidad institucional y rendimiento académico. El caso De. las Univ. Del. Caribe Colomb. *Perf. Educ.* 34 (143), 10–29. https://doi.org/10.1016/S0185-2698(14)70607-5.

Rodríguez-Ávila, S.P., Díaz-Flórez, O.C., Arias-Gómez, D.H., 2021. Los efectos de las políticas de calidad en las licenciaturas en Colombia: Balance y alternativas. Rev. Colomb. De. Educ. (*81*), 35–59. https://doi.org/10.17227/rce.num81-10688.

Rodríguez-Revilla, R., Vallejo-Molina, R.-D., 2022. Valor agregado y las competencias genéricas de los estudiantes de educación superior en Colombia. Rev. Iberoam. De. Educ. Super. 13 (36), 44–62. https://doi.org/10.22201/iisue.20072872e.2022.36.1183.

Rojas O, D.E., 2019. Relaciones de Clase en el Sistema Universitario y su Efecto sobre el Rendimiento Académico: El Caso de Bogotá. Article 1 Multidiscip. J. Educ. Res. 9 (1). https://doi.org/10.17583/remie.2019.3999.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70 (1), 41–55. https://doi.org/10.1093/biomet/70.1.41.

Rubin, D., Thomas, N., 2000. Combining propensity score matching with additional adjustments for prognostic covariates. J. Am. Stat. Assoc. 95 (450), 573–585. https://doi.org/10.1080/01621459.2000.10474233.

Sáenz-Castro, P., Vlachopoulos, D., Fàbregues, S., 2021. Exploring the relationship between Saber Pro test outcomes and student teacher characteristics in Colombia: Recommendations for improving bachelor's degree education. Educ. Sci. 11 (9) https://doi.org/10.3390/educsci11090507.

Salmon, M., 2012. Introduction to Logic and Critical Thinking, 6th ed. Cengage Learning.

Sarmiento E, J.A., Silva A, A.C., van Gameren, E., 2015. Quality differences of higher education and its determinants in a less-developed country. J. High. Educ. Policy Manag. 37 (2), 204–221. https://doi.org/10.1080/1360080X.2015.1019118.

Sarmiento E, J.A., Silva A, A.C., van Gameren, E., 2019. Evolution of the inequality of educational opportunities from secondary education to university. Int. J. Educ. Dev. 66, 193–202. https://doi.org/10.1016/j.ijedudev.2018.09.006.

Saul, J., 2013. Implicit Bias, Stereotype Threat, and Women in Philosophy. In: Hutchison, K., Jenkins, F. (Eds.), Women in Philosophy. What needs to change? Oxford University Press, pp. 39–60.

Schneider, M., Preckel, F., 2017. Variables associated with achievement in higher education: a systematic review of meta-analyses. Psychol. Bull. 143 (6), 565–600. https://doi.org/10.1037/bul0000098.

Sultan, S., Rofiuddin, A., Nurhadi, M., Priyatni, E.T., 2017. The effect of the critical literacy approach on pre-service language teachers' critical reading skills. Eurasia J. Educ. Res. 17 (71), 159–174. https://doi.org/10.14689/ejer.2017.71.9.

Tatto, M., Menter, I., 2019. The significance of teacher education. In: Tatto, M., Menter, I. (Eds.), Knowledge, policy, and practice in teacher education: A cross-national study. Bloomsbury Academic, pp. 9–17.

Timarán-Pereira, S.R., Hernández-Arteaga, I., Caicedo-Zambrano, S.J., Hidalgo-Troya, A., Alvarado-Pérez, J.C., 2016. Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional. Universidad Cooperativa de Colombia. https://doi.org/10.16925/9789587600490.

Valderrama-Leongómez, M., Murillo, L. A., Farieta-Barrera, A., Vela, M. (2019). Licenciaturas en filosofía: disputas y articulaciones entre lo disciplinar y lo pedagógico. M. L. Cárdenas-Beltrán, C. R. Cáceda y L. A. Murillo. Formación de docentes en universidades latinoamericanas. Editorial Uniagustiniana, pp. 71–112. doi:10.28970/9789585498273.02.

Vansieleghem, N., Kennedy, D., 2011. What is philosophy for children, what is philosophy with children-after Matthew Lipman. J. Philos. Educ. 45 (2), 171–182. https://doi.org/10.1111/j.1467-9752.2011.00801.x.

Ventista, O.M. (2019). *An evaluation of the "Philosophy for Children" programme: the impact on cognitive and non-cognitive skills*. Doctoral Dissertation, Durham University. http://etheses.dur.ac.uk/13121/1/Ventista_Thesis.pdf?DDD29⟨+=.

Weinberg, J. (2019, Apr. 30). Open Letter Regarding President Bolsonaro's Recent Pronouncements on Defunding Philosophy and Sociology (updated). *Daily Nous*. http://dailynous.com/2019/04/30/open-letter-regarding-president-bolsonaros-recent-pronouncements-defunding-philosophy-sociology/.

Weinberg, J. (2021, Oct. 6). Philosophy Departments Under Threat: Information, Pro-Active Strategies, Defense. *Daily Nous*. ⟨https://dailynous.com/2021/10/06/philosophy-departments-threat-information-pro-active-strategies-defense/.

Woessmann, L., 2010. Families, schools, and primary-school learning: evidence for argentina and colombia in an international perspective. Appl. Econ. 42 (21), 2645–2665. https://doi.org/10.1080/00036840801964617.