International Journal of Engineering Education Vol. 34, No. 3, pp. 1073–1084, 2018
Printed in Great Britain
0949-149X/91 $3.00+0.00

Factors Influencing the Academic Performance in Standardized Tests of Computer Science/Engineering Students in Colombia*

LUCY GARCÍA
Department of Systems and Computing Engineering. E-mail: lucyr@uninorte.edu.co

VICTOR CANTILLO
Department of Civil and Environmental Engineering, Universidad del Norte, Km 5 Vía a Puerto Colombia, Barranquilla, Colombia.
E-mail: victor.cantillo@uninorte.edu.co

There is worldwide concern about the presence of factors that affect the performance of students in the computer science/engineering (CS) education process. Gender, socioeconomic and institutional factors may create gaps that affect not only the quality of education but also the entire society. This paper examines the influence of these factors on the performance of two tests used in Colombia for assessing the quality of education: The SABER-11, applied to senior high-school students, and SABER-PRO, applied to senior CS students (N = 2,964). We used effect size to compare mean difference by gender, and correlation and multiple regression analysis to examine the influence of socioeconomic and institutional factors, and the results of the SABER-11 on SABER-PRO. The results suggest the existence of moderate gender differences, favouring men, at the high-school level; however, when compared to the undergraduate level, we found evidence of significant gaps in all evaluated topics and of men performing better than women do. Results also suggest the existence of gaps associated with income, parental education, and the nature and quality of the university and the academic program where the student studies.

Keywords: gender gap; computer science/engineering; outcomes-based assessment, socioeconomic factors

1. Introduction

During the last decade, university enrolment rates in developing countries have been increasing; however, significant disparities in quality persist. Gender, institutional and socioeconomic gaps remain, and reducing them is considered as a needed pedagogical and social objective. Considering that education is an essential right for all social groups, the educational system must ensure equal opportunities. However, in many developing countries, such as Colombia, there are still vast disparities in access to high-quality universities that are ruled mostly by socioeconomic variables [1, 2]. For computer science/engineering (CS) related programs, this is a particularly complex issue, as students coming from poor schools are at a disadvantage as they did not have access to information and communications technologies (ICT) at the same level as their medium and high-income peers. Engineering education for all is a priority considering that engineers are to be key players in solving current and future technological challenges. Modern societies require the most talented and motivated engineers regardless of their socioeconomic background [3].

Modern pedagogy recognises that diversity in the field of CS is necessary as it favours more creative, efficient and practical solutions to real-life problems. Surviving the world that is becoming more complex every day, necessitates the joining of diverse styles of thinking and points of view. The shortage of women and other social groups in CS has severe consequences for the loss of productivity and innovation necessary for the advancement of science and technology. The existence of exclusion for any circumstance raises serious questions about equity, human development, economic welfare and democratic values in society [4]. In sum, society loses out on potential innovations when it does not have a diverse workforce—regarding gender, culture, race, and socioeconomic background—fully participating in the creation of technology [5].

In this context, it is relevant to propose strategies to satisfy the social need for an innovative, creative, skilled computing workforce, but at the same increase diversity. An initial step to achieve it is to know how gender, socioeconomic and institutional variables influence the behaviour and the performance of CS students.

Computer science/engineering education in Colombia (best known as Systems and Computer Engineering or Informatics Engineering), has about 50 years of tradition. A typical CS program in Colombia is five years long. The curricular structure includes basic training in mathematics, basic

* Accepted 10 February 2018.

sciences, humanities, computer basics and informatics. Although most Colombians are mestizo, in CS schools indigenous students, Afro-Colombians and women are minority groups: a huge gap in income exists, correlated with ethnicity, also exists.

In Colombia, like in most countries, the scarcity of female students in computer science and computer and informatics engineering education is perceived as a problem. Fig. 1 shows that from 1985 to 2012, the participation of women in CS programs in Colombia has been decreasing dramatically from 46% to 16%. Therefore, it is a major challenge to look for ways to increase the participation of women in these programs. Unfortunately, lack of progress persists in the enrolment of women in CS programs, a situation that requires investigation of the structural obstacles that would inhibit the CS major from enjoying all the benefits of diversity.

The purpose of this paper is to evaluate the factors influencing the performance of CS students in standardised tests in Colombia. The emphasis is put on gender, socioeconomic and institutional variables. Specifically, the article analyses the influence of these aspects on the outcome of two tests: one to final-level high-school students (SABER-11) and the other to senior computer engineering students (SABER-PRO).

This investigation aims to address the following set of questions:

Q1. What patterns among items in both tests evidence gender differences among CS students in Colombia? How do genders differ on SABER-11 and SABER-PRO tests?

Q2. Which other socioeconomic and institutional gaps exist affecting the performance of CS students in SABER-11 and SABER-PRO tests?

Q3. How do the results of the SABER-11 test correlate to the results of the SABER-PRO test?

Although this paper focuses specifically on that data which is pertinent to the study located in Colombia, the findings have broad relevance, particularly in the design of strategies to achieve more inclusive CS programs. The results are also useful to the design of instruments for measuring the performance of CS students.

2. Literature review

Considering the global interest to achieve a more extensive and diverse labour force in the broad field of CS, it is relevant to know the influence that gender, socioeconomic and institutional factors have on the performance and persistence of CS students. The previous studies related to encouraging diversity have addressed the following three major issues: (a) gender gaps (b) socioeconomic gaps (c) influence of secondary education in CS students as outlined below.

2.1 Gender gaps

Considering that perceptions of programming may differ among men and women, teaching and learning should consider these differences. In a study conducted in Spain, Rubio et al. [6] analysed gender differences in an introductory programming course. Their results suggested that perceptions and learning outcomes differ: male students who feel programming is easier are more likely to program in the future and demonstrate higher learning outcomes than their female counterparts. However, when they used the principles of the physical computing approach, applying computational concepts to the real world, male and female students had an equal performance in perceptions and learning outcomes.

Based on interviews with female and male students who were members of five major ethnic/racial

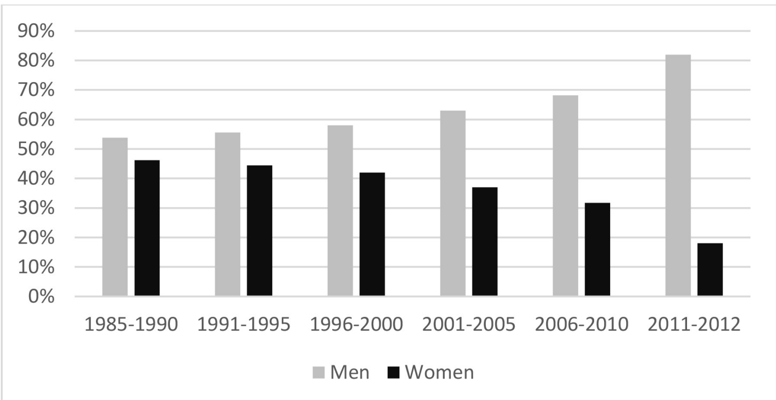

Fig. 1. Evolution of women in CS programs in Colombia. Source: Own elaboration using data from the Colombian Ministry of Education.

groups, Cribbs et al. [7] evaluated the experiences and perceptions of undergraduate computer science and computer engineering students. Results show that significant differences exist between female and male students concerning their perceptions of classes, teachers, and advisors.

The review performed by Singh et al. [4] about women's enrollment and persistence in computer-related majors indicates that, in general, they did not find gender differences in quantitative measures of students' abilities and performances. In contrast, the literature overwhelmingly revealed a critical gap in women's confidence and self-assessment of ability; that is, women show less confidence in their computing abilities. The review also showed strong inconsistencies and contradictions among international studies. For instance, the evaluation of gender disparities in science and engineering executed by Guo et al. [8] at the undergraduate level in China showed that females exhibit better performance than males in general courses.

On the other hand, several studies described the existence of significant differences in the results of standardised tests, in which men used to perform better than women. However, the level of such differences was less significant in those cases involving larger samples [9–13]. Other kinds of studies and instruments, when comparing by gender, show more similitude than differences in performance [14, 15]. It is interesting to point out that several investigations among undergraduate students concluded the existence of a positive GPA gap for women [8, 12]. In their research, Ceci et al. [9] found a higher dispersion and asymmetries among male scores. Results indicate that men are much more frequently in the top 1% than women concerning the results of spatial reasoning and mathematics tests.

Hohlfeld et al. [16] examined gender differences related to digital literacy using two tests applied to eighth-grade students from Floridian public schools. The results regarding demonstrated and perceived ICT skills showed significant differences in all areas favouring females. However, when adding predictors to the model and using more advanced statistical methods, results indicated that gender was not highly significant. These results raise doubts about the importance of gender differences associated with ICT skills.

As noted before, results from international studies about gender gaps are, in several aspects, contradictory. For instance, the study conducted by Fan and Li [17] about gender differences for university computer science majors in Taiwan, did not find significant gender differences for most of the college admission test scores, prior computer experience and the prediction models of university performance. Nevertheless, females obtained significantly higher scores in the language component. Women also outperformed men in academic achievements at both, the high-school and university levels.

It is probable that local cultural conditions strongly influence the differences reported in the literature on the possible existence of gender gaps in CS students. In the specific case of Latin America, as far as it is known, no studies were found to analyse the effect of such gaps in CS students. As a result, more evidence is needed to assess the existence of gender gaps and to propose policies and actions to increase the successful presence of women in CS programs.

2.2 Socioeconomic gaps

Diverse socioeconomic factors influence high-school students' intentions and motivations towards and against pursuing academic studies in CS; for instance, the family and the scholastic environment influence a students' choice of major. Also, there is extensive evidence about the influence that parental education and socio-economic background have on the students' achievement [18–21].

Social origin and income have also had a substantial impact on students' likelihood of earning a bachelor in STEM [22]. The influence of family and social background on high achievement is significant as an appropriate environment instils intellectual values [23–25]. Factors related to the characteristics of the institution in which the student is enrolled also influence his/her performance. It is expected that colleges which are more exacting in the high-school GPA and SAT scores required for aspirants to be enrolled will exhibit greater success and retention among their students [26].

A discussion arises about the relationships between engineering undergraduate student demographic characteristics and the social capital these students utilise when deciding their major. Unlike other social capital work in education, the research conducted in the United States by [27] concluded that gender and race/ethnicity are not significant or adequate for characterising the social capital of engineering undergraduates.

In Chile, García et al. [28] assessed whether seventh-grade students' use of information and communication technology (ICT) was related to performance on working memory tasks. They did not find an interaction effect of gender and ICT use, or an interaction effect of socioeconomic level and ICT use. The results suggest that the most important factors in developing digital skills are having access to a computer and experience using it, linguistic capital, and socioeconomic condition.

Cantillo and García [29] analysed the existence of gender, socioeconomic and institutional gaps in standardised tests applied to Civil Engineering students in Colombia. Their results suggest that there are no significant gender differences except in physics (positive for males) and verbal (positive for females) at the end of high school. However, senior civil engineering male students perform better than women in all topics, except in verbal. The study also shows that institutional and socioeconomic variables affect the results.

In sum, the existing literature shows evidence that the presence of socioeconomic and institutional gaps influence the academic performance of CS students. This situation may be even more dramatic in developing countries, such as Colombia, characterised by sharp social and economic inequalities, which limit access to quality education to broad sectors of the population. In Colombia, academic performance is strongly influenced by institutional quality [30]. In this research, it is of particular interest to evaluate if the performance of standardised tests used to assess the quality of CS students could be affected by such factors.

2.3 Influence of secondary education

The effect of the formation at secondary level on undergraduate students is another topic of interest. Zhang et al. [31] concluded that the High-School Grade Point Average (HSGPA) and Scholastic Aptitude Test (SAT)—Mathematics scores have a positive correlation with graduation rates for all the cases analysed. Socioeconomic variables such as ethnicity and gender also presented significant correlations, but the signs were not always consistent across the entire sample. The investigation conducted by French et al. [32], using data collected from engineering undergraduate students in the USA, concluded that two pre-university variables (i.e., SAT Math scores and high-school rankings) were good predictors of university GPA. Also, they found that the GPA obtained by women was higher than that obtained by men. According to this research, academic training during high-school affects the level of enrollment and persistence in undergraduate programs. In the same way, Laefer [33] concluded that achievement in physics courses is a differential factor as an explanatory element of female engineering enrollment levels, and achievement in physics courses influences long-term success.

Mathematics identity impacts the choice of engineering majors for male and female students. Moreover, the study conducted by Cribbs et al. [7] demonstrates that the interaction between recognition in mathematics and gender is significant and indicates that being recognised in mathematics has a stronger positive impact on the choice of an engineering major for females than for males. Aptitude and personality traits are predictors of retention and success. According to Hall et al. [34], the overall high school GPA and SAT math scores are significant aptitude predictors for retention in engineering students.

To summarize, most previous research supports the hypothesis that secondary education, particularly in mathematics and physics, has a high incidence on enrollment, persistence, and success in CS programs. However, little research has been done on the correlation between the performance of different basic areas in secondary schools, including sciences, language and humanities, and the performance of CS students through standardised tests of broad coverage.

3. Data collection and analysis

The Colombian Institute for Educational Evaluation (ICFES) is a government agency that has as a mission to evaluate the Colombian education system at all levels and conduct research on factors affecting educational quality in order to provide information to improve it. The ICFES designs, develops and implements tools for evaluating the quality of education, aimed at students of primary, secondary and higher education levels, by the guidelines defined by the Ministry of Education.

We studied the outcomes of two instruments applied by the ICFES. The first is the SABER-11, a test applied to final-year high-school students (in 11th grade, the final of secondary education). The second is the SABER-PRO, a national standardised exam that the students take during the last two semesters of their undergraduate program. Both tests collect personal, institutional and socioeconomic information (e.g., gender, age, race, personal status, income, parental education, household characteristics, working status, funding of studies). The SABER-Pro has information about the institution and the program (i.e., whether or not the program is accredited and the nature of the institution: public or private). Additionally, SABER-11 includes a question about which major the student wants to choose.

The goal of the SABER-11 test is to assess basic skills in students, just before finishing high school [35]. It has two sections: a common core and a flexible component. The first includes fundamental knowledge areas and basic standards of competence. These core questions are mandatory. The topics to be assessed are language (verbal ability in Spanish), mathematics, history, geography, biology, chemistry, physics, philosophy and English. Meanwhile, the flexible component assesses the areas of interest of students, who should choose one from among the following five options according to their convenience: mathematics, language, biology, history, or social sciences.

The scale of the SABER-11 test ranges from 1 to 100. If the score is under 30, it ranks as low level, from 30 to 70 the score is average, and beyond 70 it is high. The test has a multiple-choice question format, usually scheduled on Sunday. The SABER-PRO test aims to evaluate some basic competencies of final-year undergraduate students. The design of the components included in the test assesses skills and competencies considered essential for future professionals. The first set of modules assesses generic skills (e.g., skills required in any major such as reading comprehension). The other modules evaluate specific skills, related to the core elements of specific major disciplines (e.g., modelling processes and phenomena; solving problems; modelling systems, and sizing and evaluating alternatives).

All undergraduate students must take the SABER-PRO when they have completed more than 75% of their program credits as it is a mandatory test for obtaining a bachelor degree. SABER-PRO also has a multiple-choice question format with a unique answer. Like the SABER-11, the SABER-PRO is also scheduled on a particular Sunday every semester/every year.

The SABER-PRO test for CS students [36] assesses five specific components which are in concordance with several ABET [37] Student Outcomes (SO) (see Table 1). Additionally, the test includes two generic skills: English and reading or verbal comprehension (in Spanish). For analysis purposes, the current study did not consider the English component. It is important to clarify that CS programs in Colombia are better known as Systems Engineering in most cases, but also as Informatics Engineering, Software Engineering or Telematics Engineering. The test is the same for all these denominations.

The referential content covers basic sciences (mathematics and physics), engineering basic sciences (operations research, statistics, and numerical methods) and applied engineering (discrete mathematics, informatics, software engineering, communications and information management, computer architecture and computer networks). Besides all that, the test includes complementary formation topics (i.e., economics and management).

Each component weighs on the overall score proportionally to the number of questions. The two common components—English skills and reading comprehension (Spanish)—are not part of the overall score. The overall score (OS) scale is normalized with mean 100 and standard deviation 10. Meanwhile, each specific component was normalized with mean 10 and standard deviation 1.

In this study, two databases were used that cover the whole country. The first is the data collected by the ICFES concerning the SABER-PRO test applied in 2011 to CS undergraduate students. The ICFES also collected the second data about the SABER-11 test that had been applied to the same individuals when they were studying in their final year of high-school. A total of 7,847 CS students (30.9% women) took the SABER-PRO test, corresponding to the entire population of senior CS students in Colombia. After depuration and matching, the data used in this study has 2,964 registers (31.4% women), each one with information

Table 1. Specific components of the SABER-PRO test

SABER-PRO component	Number of questions	Equivalency to ABET student outcomes
1. Modeling processes and phenomena	30	(a) "an ability to apply knowledge of mathematics, science, and engineering." (k) "an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice."
2. Solving problems by applying natural science and mathematics, using logical and symbolic language	30	(a) (e) "an ability to identify, formulate, and solve engineering problems."
3. Using theory, practical and appropriate tools for solving programming problems	20	(e) (k)
4. Modeling systems, components or computer processes meeting the desired specifications.	20	(c) "an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability."
5. Sizing and evaluation of alternative informatics solutions	20	(b) "an ability design and conduct experiments, as well as to analyze and interpret data." (e) (h) "the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context."

Table 2. Variables used in SABER-11 and SABER-PRO tests

SABER-11	SABER-PRO
Gender (G): females = 1 and males = 0	Gender (G): females = 1 and males = 0
Language, verbal ability in Spanish (LANG)	Overall Score (OS): A weighted outcome of the specific components
Mathematics (MATH)	Modeling processes and phenomena (MPP)
History (HIST) Geography (GEO)	Solving problems by applying the natural sciences and mathematics using logical and symbolic language (SPAS)
Philosophy (PHIL)	Using theory, practical and appropriate tools for solving programming problems (TSPP)
Biology (BIO)	Modeling systems, components or computer processes meeting the desired specifications. (DSCP)
Chemistry (CHE)	Sizing and evaluation of alternative informatics solutions (SEAIS)
Physics (PHYS)	Reading or verbal comprehension in Spanish (RCS)

about both tests for each student. The sample used for analysis purpose is 37.8% of the population. It includes data from all the CS undergraduate programs around the country. The ICFES has verified the statistical validity of the instruments. The size and coverage guarantee that the sample is representative of the population.

It is interesting to point out that the proportion of women studying CS in Colombia is higher than in The United States, where they represented only 20.5% in 2006 [38] and 18% in 2010 [5]. However, the tendency is that women's representation in CS is stagnant. Table 2 presents the set of variables considered from both tests for analysis purposes.

4. Results

This chapter presents the most relevant results of the statistical analyses to answer the investigation questions defined before. First, gender gaps were analysed for each specific component in both tests. Then, by using correlation analysis and linear regression models, we study the socioeconomic and institutional factors influencing the SABER-PRO test, which led to seeing how the results of the SABER-11 test correlate to performance on the SABER-PRO test.

4.1 Gender disparities

Table 3 summarises descriptive data (means and standard deviation) scores by gender for the SABER-11 test for future CS students. The table presents the weighted effect size d [39] defined as Equation (1)

$$d = (\bar{x}_m - \bar{x}_f)/s, \tag{1}$$

where $\bar{x}_m$ is the mean male score, $\bar{x}_f$ is the mean female score and s is the pooled standard deviation. In essence, the value of d indicates the distance between the male and female means in terms of standard deviation units. Note that positive values for d indicate gender differences favouring males.

For interpretation of the effect size, the scale is as follows: if $|d| \geq 1.0$ the effect is very large, if $1.00 < |d| \leq 0.66$ it is large, if $0.66 < |d| \leq 0.35$ it is moderate and when $0.35 < |d| \leq 0.10$ it is small. Close to zero values (i.e., $|d| < 0.10$) indicate an equal performance for both genders [14].

A detailed analysis of the results presented in Table 3 indicates positive differences for men for all components of the SABER-11 test. However, the effect sizes are small for all cases except for PHYS and CHEM, where they are moderate. Furthermore, in the component Philosophy (PHIL) males and females performed almost equally.

Table 4 summarises the equivalent analyses of the SABER-PRO test. Results suggest similar patterns of gender differences to the SABER-11 test with males obtaining higher scores for all components over females. In this test, the measure of the effect size (d = 0.553) on the overall score indicates that males scored more than half a standard deviation higher than females; this value indicates a medium effect size. The effect size was also moderate for components SPAS and DSCP. Meanwhile, the size effect was small for all the other components. These results are consistent with those obtained for Civil Engineering Students [29]; however, in that case, women performed better in reading comprehension (RCS) although the size effect was less than 0.1.

Fig. 2 shows the cumulative frequency distribution for the overall score (OS) obtained on SABER-PRO for men and women. The curve for men is shifted to the right of that for women, indicating that male CS students scored higher than their female peers across most components evaluated in the test. An exhaustive analysis of the data shows

Table 3. Mean scores and standard deviations on SABER-11 test

Measures	Gender	N	Mean	Minimum	Maximum	Standard deviation	Pooled mean	Pooled standard deviation	d*
LANG	Males	2032	55.82	28.00	86.69	8.40	55.45	8.25	0.14
	Females	932	54.66	29.00	103.00	7.92			
MATH	Males	2032	46.76	17.00	102.00	7.94	46.17	7.53	0.25
	Females	932	44.88	27.00	73.00	6.54			
HIST	Males	2032	47.38	25.00	67.00	5.98	47.06	5.85	0.17
	Females	932	46.37	25.00	61.49	5.58			
GEO	Males	2032	50.50	25.00	76.54	7.71	49.88	7.66	0.26
	Females	932	48.52	21.00	80.04	7.56			
PHIL	Males	2032	47.27	12.00	80.80	6.51	47.24	6.31	0.01
	Females	932	47.18	23.00	84.10	5.87			
BIO	Males	2032	51.18	29.00	80.00	6.44	50.51	6.36	0.34
	Females	932	49.03	30.00	75.00	6.18			
CHEM	Males	2032	49.46	26.87	92.00	6.84	48.64	6.59	0.39
	Females	932	46.86	29.89	71.00	6.00			
PHYS	Males	2032	50.18	22.00	85.00	7.43	49.18	7.18	0.44
	Females	932	47.02	26.70	72.90	6.59			

* Positive values for *d* indicate that males scored higher.

Table 4. Mean scores and standard deviations on SABER-PRO test

Measures	Gender	N	Mean	Minimum	Maximum	Standard deviation	Pooled mean	Pooled standard deviation	d*
OS	Males	2032	102.77	71.07	146.17	8.88	101.308	8.391	0.553
	Females	932	98.13	80.44	122.58	7.20			
MPP	Males	2032	10.25	6.1	14.0	0.86	10.171	0.818	0.310
	Females	932	10.00	7.9	12.1	0.72			
SPAS	Males	2032	10.06	5.3	14.7	1.01	9.921	0.986	0.462
	Females	932	9.61	5.3	12.9	0.93			
TSPP	Males	2032	10.29	6.2	14.3	1.11	10.184	1.086	0.307
	Females	932	9.96	6.2	12.8	1.02			
DSCP	Males	2032	10.15	6.8	13.6	0.90	10.050	0.872	0.362
	Females	932	9.83	6.8	12.7	0.82			
SEAIS	Males	2032	10.07	6.3	13.0	0.84	10.005	0.837	0.250
	Females	932	9.86	6.3	12.1	0.83			
RCS	Males	2032	10.19	7.1	13.2	0.74	10.155	0.740	0.165
	Females	932	10.07	7.1	12.4	0.74			

* Positive values for *d* indicate that males scored higher.

that from the top 5% of scores, only 7% corresponds to females. This fact indicates significant asymmetries at the right extreme, with females underrepresented at the top. Ceci et al. [9] obtained similar conclusions which were explained mainly through socio-cultural factors rather than biological causes.

The joint analysis of the two tests applied to CS students in Colombia suggests that by the end of high school there are slight differences in boys when compared to girls on the SABER-11 test, and these differences widen throughout their higher education according to the results of SABER-PRO. The above results should be contrasted by the CS student's grade point average (GPA) obtained by men and women. In Colombia, the grading scale goes from 0 to 5, with 3.0 being the minimum level for approval.

Fig. 3 shows that the averages for women are slightly higher than those for men. It is interesting that although women on average obtain higher grades than men, the opposite occurs on the standardised SABER-PRO test. This raises the question of whether or not the test is biased towards males.

4.2 Impact of socioeconomic and institutional factors on the SABER-PRO

To examine socioeconomic and institutional factors influencing the overall score on the SABER-PRO test taken by CS students, we conducted a multiple linear regression analysis (MLR). The model is as presented in Equation (2):

$$y = \theta_0 + \Sigma_{k=1}^{K} \theta_k x_k, \quad (2)$$

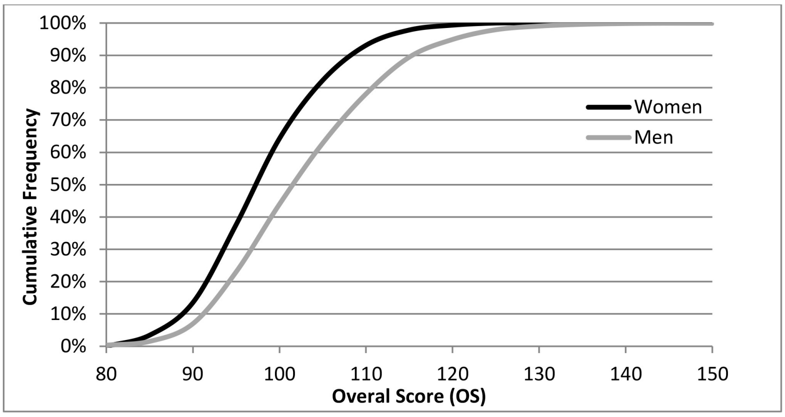

Fig. 2. Cumulative frequency by gender for overall score on SABER-PRO test.

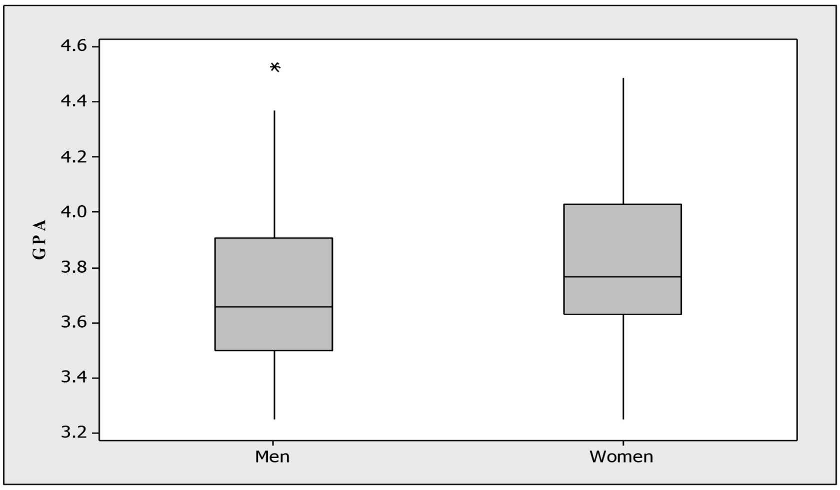

Fig. 3. CS student's grade point average by gender in Colombia.

The dependent variable *y* is the overall score. The vector of predictor variables (**x**) contains personal characteristics of the student, socioeconomic characteristics of the student's household, and characteristics of the institution and the program where the student studies. All independent variables were categorical and then they are dummy coded (0, 1).

Table 5 presents the calibrated model. It shows both the estimators with the original values of the OS variable and the estimators normalising such variables (mean = 0, standard deviation = 1). In accordance with the previous analysis, gender is significant for performance on the overall score. The results suggest that, ceteris paribus, female senior students taking a major in CS score significantly worse than males by about 0.36 standard deviations in their OS.

Other individual characteristics, such as age, also had a significant impact on the test. Interestingly, younger students (i.e., age < 22) got better scores than ones aged within the range used as a reference (22–26) by about 0.23 standard deviations. At the same time, older students (i.e., age > 26) had worse outcomes when compared to the reference age range by about 0.2 standard deviations; however, the difference was not significant at the 5% level. Also, the individual's marital status had a significant impact on the overall score, resulting in married students scoring on average 0.17 standard deviations less than those who are single. Another significant result is the fact that working students got better overall scores than ones who were not. On the other hand, as expected, scholarship students scored better than non-scholarship do.

Several socioeconomic characteristics of the student's household were also relevant to the outcome. The estimated model demonstrates that the higher the household income is, the higher the score obtained. Similar tendency occurs if the student's parents are university graduates. These results suggest the existence of socioeconomic disparities influencing the quality of education in CS programs.

The analysis backs the thesis that social capital has significant influences on the performance of students. That is because the support and resources derived from the family and the immediate social

Table 5. Socioeconomic factors influencing performance on overall score in SABER-PRO test

Variable	Original Estimates	Standardised Estimates	t-test	p-value
Intercept	99.295	–0.215	–5.52	0
Gender (Female = 1, Male = 0)	–3.347	–0.358	–10.5	0
Older than 26 (1)	–1.937	–0.207	–1.06	0.29
Younger than 22 (1)	2.174	0.232	6.34	0
Married (1)	–1.598	–0.171	–2.2	0.028
Working (1)	0.864	0.092	2.92	0.004
Scholarship	2.606	0.278	5.5	0
Student's father has a professional degree (1)	0.774	0.083	2.29	0.022
Student's mother has a professional degree (1)	1.011	0.108	2.98	0.003
Student household with median income (1)	1.060	0.113	3.24	0.001
Student household with high income (1)	2.368	0.253	3.55	0
Private university (1)	–2.808	–0.300	–9.14	0
Accredited program by CNA (1)	7.914	0.846	23.1	0
Evening program (1)	–0.323	–0.035	–0.21	0.836
Distance learning program (1)	–2.799	–0.299	–3.46	0.001

R-squared = 0.30. Model F-Value = 90.1. Dependent Variable (OS) Mean = 101.3.

environment may play a relevant role in the use and exploitation of ICTs [40]. Moreover, developing ICT competence involves a social process that needs a supportive learning environment (e.g., family, neighbourhood) [41]. Institutional factors (i.e., characteristics of the university and the program) also have a substantial impact on the results obtained in the test.

The negative sign of the parameter for private university indicates that students at public universities got better OS in the SABER-PRO test than those studying in private ones by about 0.3 standard deviations on average. A detailed analysis of the data shows that the dispersion of data is higher among private universities which is indicative of substantial differences between them. In Colombia, the quality of several low-cost private universities has been questioned. In fact, a survey conducted amongst students at 12 low-fee private universities [42] shows that the majority of undergraduate students perceived the quality of their education as poor and wanted to attend another university. They also expressed uncertainty about finding a job and anticipated having difficulties as they felt they lacked essential skills. The research concluded that due to relatively unhindered privatisation and marketisation of the higher education sector, the low-quality private provision in Colombia demonstrates this structural inequality.

On the other hand, the condition of being a student enrolled in a CS high-quality program (accredited by the National Accreditation Council, CNA) had a positive and significant impact on the overall score outcome by about 0.85 standard deviations. According to the model, the condition of being or not being in an accredited program was the variable with the highest impact on the SABER-PRO test; this result backs the process led by the CNA.

The condition of studying in an evening program did not have a significant influence on the test outcomes. Meanwhile, it is clear that students registered in distance learning programs had lower results than those studying in traditional classroom course programs by about 0.3 standard deviations in the OS.

4.3 Correlation between SABER-11 and SABER-PRO

Table 6 presents the multiple linear regression model for explaining the overall score on the SABER-PRO test as a function of performance

Table 6. Regression analysis of SABER-11 outcomes on OS of the SABER-PRO test

Component	Original Estimate	Standardised Estimates	t-test	p-value
Intercept	38.958	–	–	–
LANG	0.154	0.136	7.27	0
MATH	0.202	0.164	10.08	0
HIST	0.088	0.055	3.13	0.002
GEO	0.147	0.121	6.75	0
PHIL	0.028	0.019	1.34	0.181
BIO	0.154	0.106	5.68	0
CHEM	0.309	0.221	11.8	0
PHYS	0.180	0.141	8.66	0

R-squared = 0.50. Model F-Value = 371.2. Dependent Variable (OS) Mean = 101.3.

obtained by the student on the different components of the SABER-11 test. The model shows all parameters to be positive and significant at 5%; the only exception was PHIL. The majors that turned out to be the best predictors of overall score in the SABER-PRO test were chemistry, mathematics, and physics, in said order. According to the model, a student obtaining a higher score in the SABER-11 test will obtain a higher score in the SABER-PRO test. It is interesting to note that there are high positive correlations not only in technical areas such as physics and mathematics but also in humanities such as geography and history.

4.4 Discussion

Several previous investigations suggest that the design of standardised achievement tests had been found to be male-biased, mainly in quantitative sections [43–47], generating a gender gap favouring men. At the same time, literature also shows that women tend to perform better than men in verbal aptitude tests [48, 49]. The two tests applied to CS students confirm these tendencies, but more experimentation is required to obtain conclusive results. The analysis of the SABER-11 test applied to future students in CS allowed for the conclusion that in most of the topics assessed small to medium gender positive differences exist for males. The largest magnitude of the difference is in physics, a subject that according to Laefer [33] is a critical differential factor as one explanatory element of female engineering enrollment levels.

It is also interesting to point out that, according to previous research [12], women have higher grade point averages (GPA) than their male peers. The same trend is followed among CS students in Colombia where, on average, women have a higher GPA than men. The contradiction between the results on the possible existence of gender gaps between the SABER-PRO test and GPA requires further analysis. A hypothesis may be that the design of the test is biased towards men, but further experimentation is required to verify it.

The analysis of socioeconomic factors affecting the SABER-PRO test leads to exciting implications. In particular, the influence of the home background is interesting because Colombia is a highly segregated society. As a consequence of the extremely high concentration of wealth and regional differences, frequently only high-income students have access to the best schools, exacerbating gaps. [50, 51]. In response, during the last few years, the government has implemented policies to facilitate the access of low-income students to the best universities. These programs are essentially subsidies to demand: as is the case of the national-level subsidized loan programs ACCES (Access with Quality to Higher Education), and "SER PILO PAGA" (which roughly translates as "It pays to be smart"). They have been effective regarding increasing the potential number of low- and medium- income students from the margin who can enrol in a university, and in turn decreasing the number of students who drop out, and increasing their academic outcomes [52].

It is necessary to design appropriated public policies regarding education offered to CS students. During the last years, the priority policy was to increase coverage, favouring access to universities. However, recently, the discussion about the quality of education has become the main concern. The results of both tests show evidence of significant differences and gaps in education processes and closing them requires appropriate strategies.

5. Conclusions

This investigation analyses personal, socioeconomic and institutional factors which influence the academic performance of students majoring in computer science/engineering in Colombia by taking the results of the SABER-PRO, designed to assess academic achievement at the university level, and contrasting them with the SABER-11, applied to final-year, high-school students. The sample used for analysis purpose (N = 2,964) is representative of the population of senior computer science/engineering students in Colombia, covering about 40% of a cohort.

Even considering the long tradition and experience that the country has about their design and implementation, the tests have the well-known limitations of multiple-choice techniques. Multiple-choice is not always the best and most unique instrument for assessing problem-solving and higher-order reasoning skills. Other complementary instruments are necessary for a more rigorous and equitable assessment.

Given that the SABER-PRO test is used to assess academic performance of higher education in computer science/engineering, it is suggested that the ICFES, which is responsible for the test, revise the design by looking for a more inclusive evaluation. A comparison with the results of other majors in engineering could provide more elements for design appropriate policies.

The results of the analysis suggest the existence of moderate gender differences, favouring men, in most topics evaluated by SABER-11. However, for all evaluated topics in the SABER-PRO test, evidence of significant gaps was found of men performing better than women do; in other words, according to the results of the test, during the CS major, gender gaps increase. In contrast, women's

GPAs are slightly higher than men's. This fact supports the hypothesis that the instrument used, a standardised multiple-choice test, tends to favour males.

The evaluation of socioeconomic and institutional factors affecting performance on the SABER-PRO test suggests the presence of other gaps related to income, civil status, age, parents' education and type and quality of the university where the student studies. A huge challenge for regulatory authorities and institution managers is to design strategies for closing such gaps.

A strong correlation between the results of both tests exists. The analysis demonstrated that students who achieve a higher score on the SABER-11 test would obtain a higher score on the SABER-PRO test. In response, computer science/engineering academic programs must design strategies oriented towards first-year students who obtained the lowest scores in the SABER-11 test to try to avoid increasing the gap throughout the major.

In sum, the research adds evidence about the presence of gaps. Although the data is specific to the Colombia context, most results and conclusions can be extended to countries with similar education systems. This is the case of developing and middle-income countries, characterized by the existence of profound social and economic gaps. In most cases, they have educational systems in the process of consolidation, where there are also wide gaps in the quality offered by universities. Policies aimed at improving the quality, such as the design and implementation of adequate accreditation systems for academic programs, are going in the right direction. Finally, the findings of this study may be also useful for establishing pedagogical strategies oriented to build more friendly and inclusive environments in CS schools.

References

1. A. Gaviria, G. N. Páez and J. R. Toro, Cobertura y calidad en la educación superior en Colombia, in *La educación superior:Retos y perspectivas*, 1st ed., Universidad de Los Andes, Colombia, 2013, pp. 79–120.
2. OECD, IBRD, and The World Bank, *Reviews of National Policies for Education: Tertiary Education in Colombia 2012*. Paris: Organisation for Economic Co-operation and Development, 2013.
3. A. Diaz Lantada, J. Manuel Munoz-Guijosa, E. Chacon Tanarro, J. Echavarri Otero and J. L. Munoz Sanz, Engineering Education for All: Strategies and Challenges, *Int. J. Eng. Educ.*, **32**(5), 2016, pp. 2155–2171.
4. K. Singh, K. R. Allen, R. Scheckler and L. Darlington, Women in computer-related majors: A critical synthesis of research and theory from 1994 to 2005, *Rev. Educ. Res.*, **77**(4), 2007, pp. 500–533.
5. W. M. DuBow, Diversity in Computing: Why It Matters and How Organizations Can Achieve It, *Computer*, **46**(3), 2013, pp. 24–29.
6. M. Angel Rubio, R. Romero-Zaliz, C. Manoso and A. P. de Madrid, Closing the gender gap in an introductory programming course, *Comput. Educ.*, **82**, 2015, pp. 409–420.
7. J. D. Cribbs, C. Cass, Z. Hazari, P. M. Sadler and G. Sonnert, Mathematics Identity and Student Persistence in Engineering, *Int. J. Eng. Educ.*, **32**(1), 2016, pp. 163–171.
8. C. Guo, M. C. Tsang and X. Ding, Gender disparities in science and engineering in Chinese universities, *Econ. Educ. Rev.*, **29**(2), 2010, pp. 225–235.
9. S. J. Ceci, W. M. Williams and S. M. Barnett, Women's Underrepresentation in Science: Sociocultural and Biological Considerations, *Psychol. Bull.*, **135**(2), 2009, pp. 218–261.
10. L. Dickson, Race and Gender Differences in College Major Choice, *Ann. Am. Acad. Pol. Soc. Sci.*, **627**, 2010, pp. 108–124.
11. N. M. Else-Quest, J. S. Hyde and M. C. Linn, Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis, *Psychol. Bull.*, **136**(1), 2010, pp. 103–127.
12. G. Sonnert and M. F. Fox, Women, Men, and Academic Performance in Science and Engineering: The Gender Difference in Undergraduate Grade Point Averages, *J. High. Educ.*, **83**(1), 2012, p. 73.
13. Z. Zhu, Gender differences in mathematical problem solving patterns: A review of literature, *Int. Educ. J.*, **8**(2), 2007, pp. 187–203.
14. J. S. Hyde, The gender similarities hypothesis, *Am. Psychol.*, **60**(6), 2005, pp. 581–592.
15. E. S. Spelke, Sex differences in intrinsic aptitude for mathematics and science? A critical review, *Am. Psychol.*, **60**(9), 2005, pp. 950–958.
16. T. N. Hohlfeld, A. D. Ritzhaupt, and A. E. Barron, Are gender differences in perceived and demonstrated technology literacy significant? It depends on the model, *Educ. Technol. Res. Dev.*, **61**(4), 2013, pp. 639–663.
17. T. S. Fan and Y. C. Li, Gender issues and computers: college computer science education in Taiwan, *Comput. Educ.*, **44**(3), 2005, pp. 285–300.
18. J.-T. Liu, S.-Y. Chou and J.-L. Liu, Asymmetries in progression in higher education in Taiwan: Parental education and income effects, *Econ. Educ. Rev.*, **25**(6), 2006, pp. 647–658.
19. J. Erola, S. Jalonen and H. Lehti, Parental education, class and income over early life course and children's achievement, *Res. Soc. Stratif. Mobil.*, **44**, 2016, pp. 33–43.
20. J. Huang, B. Guo, Y. Kim and M. Sherraden, Parental income, assets, borrowing constraints and children's post-secondary education, *Child. Youth Serv. Rev.*, **32**(4), 2010, pp. 585–594.
21. R. S. Wells and C. M. Lynch, Delayed college entry and the socioeconomic gap: Examining the roles of student plans, family income, parental education, and parental occupation, *J. High. Educ.*, **83**(5), 2012, pp. 671–694.
22. S. Hurtado, M. K. Eagan and B. Hughes, *Priming the pump or the sieve: institutional contexts and URM STEM degree attainments*, presented at the Association for Institutional Research, 2012.
23. R. S. Albert and M. A. Runco, The achievement of eminence: A model based on a longitudinal study of exceptionally gifted boys and their families, *Concept. Gift.*, 1986, pp. 332–357.
24. C. P. Benbow and O. Arjmand, Predictors of High Academic Achievement in Mathematics and Science by Mathematically Talented Students: A Longitudinal Study, *J. Educ. Psychol.*, **82**(3), 1990, pp. 430–441.
25. B. S. Bloom, 1913-, and L. A. Sosniak, *Developing talent in young people*. Ballantine Books, 1985.
26. M. S. DeBerard, G. I. Spielmans and D. L. Julka, Predictors of Academic Achievement and Retention among College Freshmen: A Longitudinal Study, *Coll. Stud. J.*, **38**(1), 2004, p. 66.
27. J. P. Martin, S. Brown, M. K. Miller and S. K. Stefl, Characterizing Engineering Student Social Capital in Relation to Demographics, *Int. J. Eng. Educ.*, **31**(4), 2015, pp. 914–926.
28. L. Garcia, M. Nussbaum and D. D. Preiss, Is the use of information and communication technology related to performance in working memory tasks? Evidence from seventh-grade students, *Comput. Educ.*, **57**(3), 2011, pp. 2068–2076.
29. V. Cantillo and L. Garcia, Gender and Other Factors

Influencing the Outcome of a Test to Assess Quality of Education in Civil Engineering in Colombia, *J. Prof. Issues Eng. Educ. Pract.*, 2014, **140**(2).
30. A. Rodríguez, D. Ariza and R. Ramos, Institutional quality and academic performance for universities Colombian Caribbean, *Perfiles Educ.*, **36**(143), 2014, pp. 10–29.
31. G. L. Zhang, T. J. Anderson, M. W. Ohland and B. R. Thorndyke, Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study, *J. Eng. Educ.*, **93**(4), 2004, pp. 313–320.
32. B. F. French, J. C. Immekus and W. C. Oakes, An examination of indicators of engineering students' success and persistence, *J. Eng. Educ.*, **94**(4), 2005, pp. 419–425.
33. D. F. Laefer, Gender Disparity in Engineering as a Function of Physics Enrollment and Its Implications for Civil Engineering, *J. Prof. Issues Eng. Educ. Pract.*, **135**(3), 2009, pp. 95–101.
34. C. W. Hall et al., Aptitude and Personality Traits in Retention of Engineering Students, *J. Eng. Educ.*, **104**(2), 2015, pp. 167–188.
35. ICFES, *Orientaciones para el examen de Estado de la educación media ICFES SABER 11*, 2011.
36. ICFES, *Exámenes de calidad de la educación superior en ingeniería de sistemas: Guía de Orientación*, 2009.
37. ABET, Engineering Accreditation Commission, *Criteria for Accrediting Engineering Programs, 2016–2017 | ABET*.
38. C. Hill, Christianne Corbett, and Andresse St. Rose, Why So Few? Women in Science, Technology, Engineering and Mathematics, *AAUW: Empowering Women Since 1881*.
39. J. Cohen, A power primer, *Psychol. Bull.*, **112**(1), 1992, pp. 155–159.
40. D. Y. Wohn, N. B. Ellison, M. L. Khan, R. Fewins-Bliss and R. Gray, The role of social media in shaping first-generation high school students' college aspirations: A social capital lens, *Comput. Educ.*, **63**, 2013, pp. 424–436.
41. A. Clement and L. R. Shade, The Access Rainbow: Conceptualizing Universal Access to the Information/Communications Infrastructure, *Httpservicesigi-Glob.-1-878289-69-8ch002*, pp. 32–51, 2000.
42. P. N. Páez and C. Teelken, Perceived quality of private education and fears of stratification: Investigating the propositions of human capital theory by exploring the case of Colombia, *Policy Futur. Educ.*, **14**(6), 2016, pp. 597–616.
43. R. Azen, S. Bronner and N. Gafni, Examination of Gender Bias in University Admissions, *Appl. Meas. Educ.*, **15**(1), 2002, pp. 75–94.
44. M. J. Clark and J. Grandy, Sex Differences in the Academic Performance of Scholastic Aptitude Test Takers, *ETS Res. Rep. Ser.*, **1984**(2), 1984, p. 1–27.
45. F. Gándara and M. Silva, Understanding the Gender Gap in Science and Engineering: Evidence from the Chilean College Admissions Tests, *Int. J. Sci. Math. Educ.*, **14**(6), 2016, pp. 1079–1092.
46. C. Good, J. Aronson and M. Inzlicht, Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat, *J. Appl. Dev. Psychol.*, **24**(6), 2003, pp. 645–662.
47. P. O. Saygin, Gender differences in preferences for taking risk in college applications, *Econ. Educ. Rev.*, **52**, 2016, pp. 120–133.
48. E. Khodadady and B. S. Dastgahian, Gender and field of study and performance on an English language proficiency test, *Theory Pract. Lang. Stud.*, **2**(11), 2012, pp. 2322–2329.
49. G. D. Kuh, J. L. Kinzie, J. A. Buckley, B. K. Bridges and J. C. Hayek, *What Matters to Student Success: A Review of the Literature*, National Postsecondary Education Cooperative, 2006.
50. A. Cuenca, Inequality of opportunities in Colombia: Impact of social origin on academic achievement and income level among university graduates, *Estud. Pedagog.*, **42**(2), 2016, pp. 69–93.
51. Q. Loaiza and V. Hincapié, A study of regional gaps in education quality in Colombia: 2000–2012, *Ens. Sobre Polit. Econ.*, **34**(79), 2016, pp. 3–20.
52. T. Melguizo, F. Sanchez and T. Velasco, Credit for Low-Income Students and Access to and Academic Performance in Higher Education in Colombia: A Regression Discontinuity Approach, *World Dev.*, **80**, 2016, pp. 61–77.

Lucy García obtained her BSc in Systems Engineering from Universidad del Norte, Colombia. She received a PhD in Engineering Sciences (Computer Sciences) at Pontificia Universidad Católica de Chile. Her research interest areas are related to technology and education. Dr García is currently an associate professor of the Department of Systems and Computing Engineering at Universidad del Norte.

Víctor Cantillo received his BS Engineering in Civil Engineering in 1987, from the Universidad del Norte, Colombia, MSc in Traffic and Transport Engineering in 1990, from the Universidad del Cauca, Colombia and PhD in Engineering Sciences (Transport) in 2004, from the Pontificia Universidad Católica de Chile. He is a Professor at the Universidad del Norte, Colombia since 1990. His research interests include transport planning, transport economics, econometric, transport modelling and logistics. Education in engineering is other of his relevant topics of interest.