

At following, the six aspects of scientific reasoning that are measured through the modified LCTSR (Jensen et al., 2015; Piraksa et al., 2014): (i) Conservation of physical magnitudes (items 1 to 4): It seeks to evaluate what is the relationship of magnitudes such as mass and volume when their shape is manipulated. (ii) Proportional thought (items 5 to 8): It evaluates the relationship between two data series, which can be mathematical or scientific. (iii) Identification and control of variables (items 9 to 14): It seeks to have the ability to identify and isolate variables, and to reason in an experiment to conclude which was the cause of the problem. (iv) Probabilistic thought (items 15 to 18): The notion of probability, related to the understanding of chance and causality, is related to notions of proportion as well as combinatorial schemes and would be useful both for the solution of mathematical problems and for the understanding of non-deterministic scientific phenomena. (v) Correlational thought (items 19 and 20): The notion of correlation is linked to both proportion and probability and would be necessary for the analysis of data and scientific experimentation in complex tasks or before probabilistic phenomena. (vi) Combinatorial thought (items 21 to 24): Combinatorial operations, given a series of variables or propositions, make it possible to exhaust all possible combinations among them to achieve a certain effect. Operations of this type would be combinations, variations and permutations.

It is important to note that dominate the last three aspects of scientific reasoning requires a higher skill level of the student. Besides, the development of these last three aspects is essential for the student to reach a level of formal thought. Finally, according to the overall score obtained by a student in the LCTSR, which has a minimum score of 0 and a maximum score of 12, the student can be classified in one of the three categories of thought: (i) Concrete (0-4 points); (ii) Transition (5-8 points); and (iii) Formal (9-12 points).

Experiment, results and discussion

Experimental design

As previously explained, this work aims to measure and detect significant changes in the scientific reasoning skills of university engineering students. In particular, it seeks to determine if the curriculum map of courses belonging to the core of mathematics and physics, which is typically seen in the first two years of the curriculum of engineering programmes, contributes significantly to the academic education and learning of students. Specifically, it is expected that once the engineering students complete the first four semesters, they will have the mathematical logic skills required to show adequate academic performance in the basic core courses of the engineering programmes.

A case study was developed in the College of Engineering of the Universidad del Norte, Colombia to achieve the objective of this research. It is composed of two important chronological stages. Stage 1: In the first semester of 2015, it was applied the modified LCTSR to all students who will complete their first semester in the engineering programmes (more than 800 students). Stage 2: In the first semester of 2017, the same test was applied to a large group of students. As a result, it was obtained that 126 students presented the same LCTSR in both 2015 and 2017. Finally, for each of the 126 students, the scores obtained in the following seven categories were recorded: (i) conservation of physical magnitudes (PM); (ii) proportional thinking (PR); (iii) identification and control of variables (IV); (iv) probabilistic thinking (PT); (v) correlational thinking (CT); (vi) combinatorial thought (CM); and (vii) overall (OV). The last category represents the total score obtained by the student in the LCTSR, which is the result of adding the scores obtained in the six aspects of scientific reasoning.

Results of the case study

The results and discussion of this case study are divided into the following three subsections: Characteristics of the samples; Lawson test; and Lawson Test vs Academic performance.

Characteristics of the samples

The purpose of this subsection is to provide detailed information about the individuals that make up the two data samples that were used in this case study. Note that, by definition, these samples are not independent and therefore are considered paired, since the 126 experimental objects are the same in both samples.

It was observed that 99% of the population in 2015 and more than 40% in 2017 is in a range of age less than 19 years. Thus, as stated by Piaget and various authors of constructivist theories, it can expect that their levels of reasoning have not yet wholly reached formal thought (Aguilar Villagrán et al., 2002; Linares, 2009; Rodríguez et al., 2010). In addition, Figure 1a indicates that in the samples there is a majority of male students. Although it is not the object of work, previous studies carried out by (Piraksa et al., 2014), have concluded that there is no significant correlation between gender and the ability to reason scientifically. Finally, Figure 1b shows that, in the samples, there are students belonging all the engineering programmes offered by Universidad del Norte.

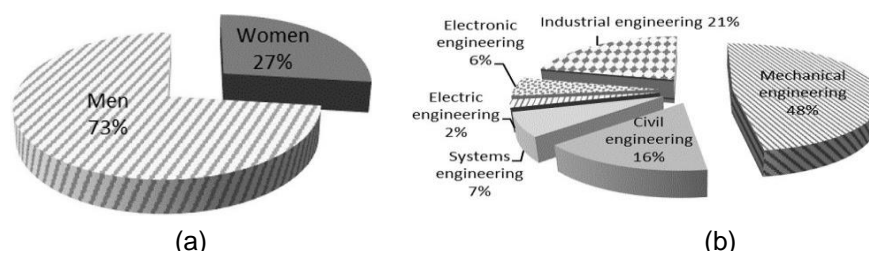


Fig 1: Distribution of students by (a) gender and (b) academic programme.

Lawson test

This subsection presents the results of the LCTSR using two types of statistical analysis.

Descriptive analysis

Table 1 presents several statistics that allow describing numerically the characteristics of the studied samples. Note that, the statistics presented in the third column require a previous arithmetic operation between the data of the 2015 and 2017 samples. Specifically, for each student is calculated the difference between the overall score obtained in 2017 minus the overall score obtained in 2015. It is observed that the average score obtained by the students in 2017 was higher than in 2015, specifically there is an average increase of 1.42 points in the LCTSR. Similarly, the analysis of the 25, 50, and 75 percentiles for 2017 indicates that the three percentiles show an increase of 2 points compared to the results obtained in 2015. For example, the 75 percentile for the year 2017 states that 75% of students scored at or below 9 points, while for 2015 the 75 percentile says that 75% of students scored a maximum of 7 points. Note also that the 2015 and 2017 samples have an almost identical variability since their respective standard deviations are very similar.

Table 1: Descriptive measures on the samples.

	2015	2017	(2017 - 2015)
Mean	5.18	6.6	1.42
Percentile 25	3	5	-
Percentile 50	5	7	-
Percentile 75	7	9	-
Desv. Standard	2.64	2.63	2.07

Additionally, Figure 2a shows the percentage of students who improved, worsened or maintained the same score on the LCTSR and Figure 2b gives more details about the students who obtained an improvement in their score. For example, Figure 2a indicates that 65% of students improved their score. In turn, Figure 2b shows that 22% of students improved their score by 1 point. It is interesting to note that 55% (22 + 16 + 15 + 2) of the students showed an improvement in their score of maximum 4 points. That is, few students

achieved an exceptional jump in their logical reasoning skills since only 10% of students achieved an improvement of 5 points or more. This result is especially valuable since only an improvement equal to or greater than 5 points would allow a student who in 2015 was classified in Concrete Thought could be classified in 2017 in Formal Thought. It must be remembered that this last category represents the ideal state of thought. It is worrisome that 19% of students will not improve their scores and that 16% of them will present a setback.

However, according to Piaget and Vygotsky, learning is a slow and qualitative process that is presented in stages. They also explain that learning requires an intellectual, cultural and historical development that depends on the experiences lived by each student. In addition to this, Vygotsky suggests that the development of scientific concepts (“non-spontaneous”) can be achieved, focusing attention on the processing related to the context of the concept. Consequently, it seeks to promote spontaneous thinking and therefore the understanding of the concept of science, in a period that may be slow, but in the long term, students will develop high levels of thinking Ramos Serpa & López Falcón, 2015; Severo, 2012; Vygotsky, 1978).

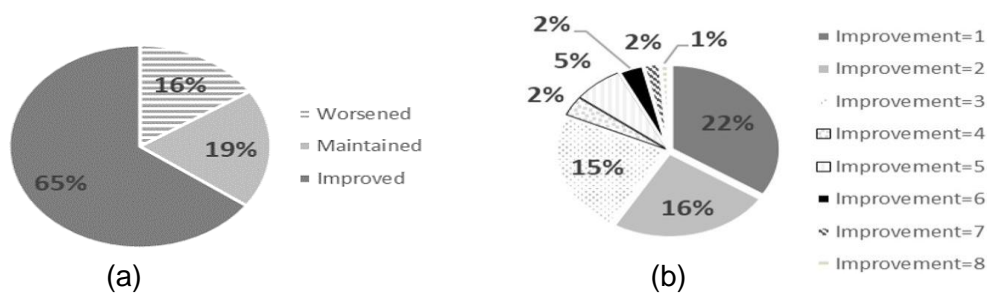


Fig. 2: Student performance: 01-2015 vs 01-2017.

Finally, for the years 2015 and 2017, Figure 3 presents the percentage of students that were classified in each category of thought, according to the scores obtained in the LCTSR. The results are intuitive since it shows how the percentage of students in concrete thought decreases in 2017, but also how the percentage of students in formal thought increases in 2017.

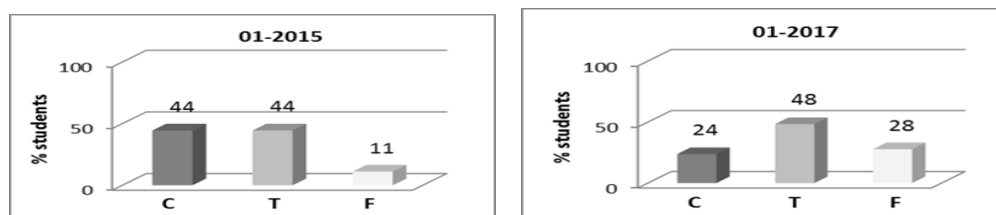


Fig. 3: Percentage of students by category of thought in 01-2015 and 01-2017. (C: concrete; T: transition; F: formal).

Inferential analysis

The purpose of the inferential analysis is to determine if there is a significant improvement when comparing the average scores obtained from the LCTSR in 2015 and 2017. That is, we seek to prove that the average scores for the year 2017 are higher than the average scores of the year 2015. In order to achieve this objective, it is proposed to apply tests of hypotheses of the difference of means that determine the validity or falsity of the statements made. In particular, seven tests of unilateral hypotheses (right) are proposed; one for the overall scores of the LCTSR and the other six for the scores of the six aspects of scientific reasoning that are also evaluated in the test. Note that this study considers two random samples dependent (i.e., paired). Furthermore, since the sample size is large enough (i.e., $n = 126$), the Central Limit Theorem can assume that the distribution of the means follows an approximately normal distribution. Therefore, the paired t-test statistic can be applied.

For each of the seven proposed hypothesis tests, the null hypothesis H0 expresses that the difference of the average scores between both populations is equal to zero. On the other hand, the alternative hypothesis H1 expresses that the average score for the year 2017 is strictly higher than the average score of the year 2015. Table 2 presents the results for the seven tests of right unilateral hypotheses.

Given that several hypothesis tests are being carried out simultaneously on the same data set, false rejections of the null hypotheses should be avoided. Therefore, it was used the Bonferroni method to adjust the level of significance (α) about the number of statistical tests performed simultaneously. The method says that, if it seeks to guarantee a level of significance for the set of M pairwise comparisons, it is enough to take a corrected significance level that can be expressed as $\alpha^* = \alpha/M$. In this way, the possible error that can be made by making many comparisons in pairs is compensated. For this study, a level of significance of $\alpha = 0.05$ was considered, obtaining a corrected significance level $\alpha^* = 0.05/7 = 0.0071$. Note that, if for any of the seven tests the calculated P-value is less than α^* , this means that H0 is rejected in favour of H1. Otherwise, H0 is not rejected.

Table 2: Results of the comparison test of means: 01-2015 vs 01-2017.

	Aspects of Scientific Reasoning						Overall
	PM	PR	IV	PT	CT	CM	OV
Statistic T	3.81	6.96	3.96	2.38	2.38	1.69	7.70
Value-P	1.09E-04	1.02E-07	6.24E-05	0.0095	0.0095	0.0467	1.04E-07

Table 2 shows that in four of the seven tests there is statistical evidence that there is a significant improvement in the average scores of 2017 concerning the average scores of the year 2015. However, for the scientific reasoning aspects Probabilistic (PT), Correlational (CT) and Combinatorial (CM), the hypothesis test says that there is no significant difference between the average scores of the years 2015 and 2017.

Lawson Test vs Academic performance

The purpose of this subsection is to detect if the results of the LCTSR have a positive association with the academic status of the students, the latter measured by the academic average. That is, through simple exploratory analysis, we evaluate whether there are indications that students with a high overall score on the LCTSR also have a high academic average, and vice versa. For the year 2017 and the 126 students sampled, Table 3 shows in percentage, how students are distributed concerning the LCTSR score and the Thought Category. In turn, the table groups the students according to their academic status: (i) Trial period (2.95 - 3.24); (ii) Normal (3.25 - 3.94); and (iii) Distinguished (3.95 - 5.00). Consider that in the sample there are three students in a trial period, 92 are in a normal academic state, and 31 are in the distinguished state. Note also that, in Colombian universities, the academic grade is measured in a range of 0.0 to 5.0, such that 2.95 is the minimum passing score.

Table 3 presents fairly intuitive results; that is, all students who are in the trial period are in the lowest thought category. In turn, most students who have a normal academic status are classified in the Transition thought category, which represents a state of intermediate thinking. Finally, students who are in a distinguished academic state are classified in the categories of Transition thought and mostly in Formal, 42% and 52% respectively. This analysis supports the idea that there is a positive correlation between the LCTSR score and the academic average.

Table 3: Percentage of students according to the LCTSR test score and the Thought Category: 01- 2017.

Academic average	Concrete					Transition					Formal				
	1	2	3	4	Sub-total	5	6	7	8	Sub-total	9	10	11	12	Sub-total

(2,95 - 3,24)	0.0	33.3	33.3	33.3	100	0.0	0.0	0.0	0.0	0	0.0	0.0	0.0	0.0	0
(3,25 - 3,94)	0.0	8.7	9.8	8.7	27	10.9	16.3	13.0	12.0	52	12.0	5.4	2.2	1.1	21
(3,95 - 5,00)	0.0	3.2	0.0	3.2	6	6.5	12.9	16.1	6.5	42	12.9	25.8	9.7	3.2	52
Total	0.0	7.9	7.9	7.9	24	9.5	15.1	13.5	10.3	48	11.9	10.3	4.0	1.6	28

Conclusions

This study focused on answering two key questions:

(i) What is the contribution of basic science education to the development and enhancement of formal thought in Engineering students? The descriptive analysis showed that in 2015 there were 44% of students in concrete thought, 44% in transitional thought, and 11% in formal thought. For the year 2017, there were 24% of students in concrete thought, 48% in transitional thought, and 28% in formal thought. These descriptive results indicated that even when there was an improvement in 2017, this improvement does not seem to be large enough for most students to develop formal thinking. On the other hand, the results obtained by the inferential analysis show a partially positive result. That is, with respect to the overall scores of the LCTSR and the first three aspects of scientific reasoning (i.e., Physical Magnitudes, Proportional Thinking, Identification and Variable control), the students presented significant improvements. However, in the aspects of scientific reasoning that require a higher level of skill in the student (i.e., Probabilistic Thought, Correlational Thought, and Combinatorial Thought), there was no significant improvement. These results indicate that, on average, the training received by the students from Engineering programmes in the core courses of mathematics and physics, actually develops their logical thinking/reasoning skills. However, the mathematical logic skills attained in said core seem not to be sufficient for these students to show an ideal academic performance in the basic core courses of engineering programmes. It should be taken into consideration that in 2017 at least 40% of students are still under 19 years old. It is important to emphasise since Piaget's postulates explain that individuals must wait until 20 years to consolidate their formal thinking.

(ii) Is there a correlation between the academic performance of the students and the level of thought measured to the students through the LCTSR? The results of the LCTSR indicated that, students with a high overall score on the LCTSR also have a high academic average, and vice versa. This result supported the idea that there is a positive correlation between the LCTSR score and the academic performance of students.

Finally, the overall results of this study show an opportunity for improvement in engineering programs. That is, it makes sense to make a more significant effort so that at an early stage of the engineering programme (first four semesters), students will develop the six aspects of scientific reasoning with a higher level.

References

- Acevedo, J., & Oliva Martínez, J. (1995). Validación y aplicación de un test de razonamiento lógico. *Validación y Aplicación de Un Test de Razonamiento Lógico*, 48(3), 339–351.
- Aguilar Villagrán, M., Navarro Guzmán, J., López, J. M., & Alcalde, C. (2002). Pensamiento formal y resolución de problemas matemáticos. *Psicothema*, 14(2), 382–386.
- Daniels, H. (Ed.). (2017). *An introduction to Vygotsky*. Psychology Press, 2005. (3rd ed.). Routledge.
- Delgado, A. (2001). *Formación Valoral a nivel universitario*. Universidad Iberoamericana.
- Fulmer, G. W., Chu, H.-E., Treagust, D. F., & Neumann, K. (2015). Is it harder to know or to reason? Analyzing two-tier science assessment items using the Rasch measurement model. *Asia-Pacific Science Education*, 1(1), 1–16. <https://doi.org/10.1186/s41029-015-0005-x>
- Guanipa Marquez, J., Nava Díaz, J., & Dávila Cazzato, S. (2007). La disciplina escolar: aportes de las teorías psicológicas. *Revista de Artes y Humanidades UNICA*, 8(18), 126–148.
- Huang, P.-S., Peng, S.-L., Chen, H.-C., Tseng, L.-C., & Hsu, L.-C. (2017). The relative influences of domain knowledge and domain-general divergent thinking on scientific creativity and mathematical creativity. *Thinking Skills and Creativity*, 25, 1–9. <https://doi.org/10.1016/j.tsc.2017.06.001>

- Inhelder, B., & Piaget, J. (2013). *The Early Growth of Logic in the Child: Classification and Seriation*. Routledge. <https://www.routledge.com/The-Early-Growth-of-Logic-in-the-Child-Classification-and-Seriation/nhelder-Piaget-Jean/p/book/9780415868853>
- Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE Life Sciences Education*, 14(1), 1–12. <https://doi.org/10.1187/cbe.14-08-0129>
- Lawson, A. E. (1976). Formal operations and field independence in a heterogeneous sample. *Perceptual and Motor Skills*, 42(3 I), 981–982. <https://doi.org/10.2466/pms.1976.42.3.981>
- Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15(1), 11–24. <https://doi.org/10.1002/tea.3660150103>
- Lawson, A. E. (2000). Classroom Test of Scientific Reasoning. Revised Edition *Journal of Research in Science Teaching*, 15(1), 11–24. <http://www.public.asu.edu/~anton1/AssessArticles/Assessments/MathematicsAssessments/ScientificReasoningTest.pdf>
- Lawson, A. E. (2010). Basic inferences of scientific reasoning, argumentation, and discovery. *Science Education*, 94(2), 336–364. <https://doi.org/10.1002/sce.20357>
- Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, 44(5), 706–724. <https://doi.org/10.1002/tea.20172>
- Linares, A. R. (Ed.). (2009). *Desarrollo cognitivo: Las teorías de Piaget y de Vygotsky*. Universitat Autònoma de Barcelona.
- López Martínez, O., & Ramón Martín, B. (2010). Creative intelligence and thinking styles. *Anales de Psicología*, 26(2), 254–258.
- Opitz, A., Heene, M., & Fischer, F. (2017). Measuring scientific reasoning—a review of test instruments. *Educational Research and Evaluation*, 23(3–4), 78–101. <https://doi.org/10.1080/13803611.2017.1338586>
- Picquart, M., Guzmán, O., & Sosa, R. (2010). Razonamiento científico e ideas previas en alumnos de ciencias básicas de la UAM- Iztapalapa. *Latin-American Journal of Physics Education*, 4(1), 1056–1064.
- Piraksa, C., Srisawasdi, N., & Koul, R. (2014). Effect of Gender on Student's Scientific Reasoning Ability: A Case Study in Thailand. *Procedia - Social and Behavioral Sciences*, 116, 486–491. <https://doi.org/10.1016/j.sbspro.2014.01.245>
- Ramos Serpa, G., & López Falcón, A. (2015). La formación de conceptos: Una comparación entre los enfoques cognitivista y histórico-cultural. *Educacao e Pesquisa*, 41(3), 615–628. <https://doi.org/10.1590/S1517-9702201507135042>
- Rodríguez, M. D., Mena, D. A., & Rubio, C. M. (2010). Razonamiento Científico y Conocimientos Conceptuales de Mecánica : Un Diagnóstico de Alumnos de Primer Ingreso a Licenciaturas en Ingeniería Scientific Reasoning and Conceptual Knowledge in Mechanics : A Diagnosis of Freshmen to Undergraduate Engineering. *Formación Universitaria*, 3(5), 37–46. <https://doi.org/10.4067/S0718-50062010000500006>
- Severo, A. (2012). *Teorías del Aprendizaje: Jean Piaget y Lev Vigotsky*.
- Stecconi, C. (2010). Inteligencias múltiples y el cuestionario de autoevaluación (CAIM). *Calidad de Vida y Salud*, 3(2), 147-164,.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes* (E. Rice (Ed.)).

Acknowledgements

The authors are grateful to the Center for Teaching Excellence – CEDU at the Universidad del Norte for their support in the implementation LCTSR

Copyright statement

Copyright © 2021 Virginia Paredesa, Nestor Durango, Jonathan González Ospino, César Augusto Henao, Germán Jiménez, Mario Alberto Gómez Villadiego, and Julián Yepes-Martínez: The authors assign to the Research in Engineering Education Network (REEN) and the Australasian Association for Engineering Education (AAEE) and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to REEN and AAEE to publish this document in full on the World Wide Web (prime sites and mirrors), on Memory Sticks, and in printed form within the REEN AAEE 2021 proceedings. Any other usage is prohibited without the express permission of the authors.