Vol 13, no 3
FALL 2021

# The Calculus Concept Inventory Applied to the Case of Large Groups of Differential Calculus in the Context of the Program "Ser Pilo Paga" in Colombia.

Gabriel Villalobos-Camargo

Universidad de Bogotá Jorge Tadeo Lozano, Área de Ciencias Básicas y Modelado

Carrera 4 Número 22 - 61, Mód. 6, Of. 501. 110311. Bogotá, Colombia

gabriel.villalobosc@utadeo.edu.co

*Abstract: The Calculus Concept Inventory (CCI), Epstein (2013) aims to test the understanding of calculus ideas, rather than ability to perform calculations. In this paper the CCI is used to measure the effect of the undergraduate calculus cohort over the understanding of calculus in a heterogeneous population including recipients of the program Ser Pilo Paga (Pilos). There is a global positive gain of 0.10 (3), a weak correlation between gain and the percentage of Pilos, a negative correlation between initial score and the gain, and no correlation between class size and gain. The values hopefully would provide a baseline for comparing future interventions on the teaching of calculus.*

The prime goal of teaching calculus for non math-major students is to achieve an understanding of the mathematical concepts and their relations. In this paper, I aim to gauge the outcome of the education process in different groups of students that took a Differential Calculus class at the *Universidad de Bogotá Jorge Tadeo Lozano* during the second semester of 2016. I have characterized the results of students within different subgroups: by gender, recipients of "ser pilo paga" (a state-sponsored program aimed at low-income students), session size, and initial score. The results contribute to the understanding of the differences between *Pilos* (recipients of ser pilo paga") and the general population of this university.

Firstly, a definition of the understanding of first-year calculus is needed. (Sofronas, et al., 2011) interviewed 24 experts and book authors and concluded with a set of goals and sub-goals, organized in a framework (a hierarchically organized list). For each goal of the framework, the authors reported the percentage of experts who think it is a key element of understanding. The basic general goals in the survey are: "(a) mastery of the fundamental concepts and-or skills of the first-year calculus, (b) construction of connections and relationships between and among concepts and skills, (c) the ability to use the ideas of the first-year calculus, and (d) a deep sense of the context and purpose of the calculus" (Sofronas, et al., 2011). Let me point out to the fact that there was consensus only on one goal: "Mastery of Fundamental Concepts and-or Skills", thus there is no agreement among experts on what constitutes understanding of mathematics.

Furthermore, what has been considered "fundamental" has changed over time. Historically, math education has swung back and forth between two extremes. On one end, there is have traditional education, also called procedural teaching, which puts a strong focus on the teaching of basic skills. Their proponents argue that the acquisition of those skills lowers the cognitive load needed to understand higher concepts, as well as provides basic facts that are important in the construction of higher-order concepts and their interpretation (NCTM, 2021). On the other end, there are proponents of *high order understanding* (Österman & Bråting, 2019). They claim that high-level skills can and should be taught, thus the focus of the education effort should lie in, for instance, problem-solving and creative thinking. There is a large divide between the two communities, for instance, Berry and Nyman (2003), report results of interviews highlight a better understanding of the algebraic symbolic view of calculus over the graphical representation (presumably a traditionalist approach); Habre and Abboud (2006) called the traditional approach to calculus teaching, as "teaching techniques for solving drill problems", (p. 57); while comparing to a reformed calculus class, which emphasizes visualization (Habre & Abboud, 2006).

A particular goal of the calculus reform movement was to give the students fluency in the use of multiple representations. Since computers can perform the procedural, pencil-and-paper algorithmic techniques (called pejoratively "symbol pushing"), mathematics should involve more visual work and real work scenarios (Zazkis, 2013). In contrast conceptual knowledge is described as including graph interpretation and creation skills, knowledge of various representations and how to translate between them, ability to derive procedures from basic principles, ability to tackle novel problems, physical interpretations of functions/graphs that describe motion, and the ability to translate word problems into calculus equations (Zazkis, 2013). Contemporary calculus textbooks like Stewart Calculus, used in this study include several exercises that aim to improve the understanding of representations.

Now, is there really a split between mastery of fundamental concepts or skills and a deep sense of context and purpose of calculus? It is common -in discussions among colleagues or the literature-to see it referred to as a dichotomy: a class could focus on either of them. I adhere to the view that there is no such dichotomy. On the contrary, there is a reinforcement loop between the acquisition of skills and the understanding of the purpose of a calculation. Thus, understanding of higher order concepts both requires "the mastery of fundamentals concepts and skills" and helps to achieve it. (Wu, 1999, p. 3)

Is in this context that Concept Inventories were created. They are specifically designed to test the comprehension of the conceptual base of a given subject. In order to solve the questions, students have to be able to apply the principles to simple but interesting situations; where usual calculations and algorithms are of little or no use (Epstein, 2007). The first Concept Inventory was the Force Concept Inventory (FCI), by Hestenes, et al., (1992). Here I use the Calculus Concept Inventory (CCI), (Epstein, 2006, 2007, 2013), in order to assess the students' conceptual understanding of basic Differential Calculus. The CCI has been used in different contexts to measure the understanding of calculus concepts (Rhea, 2008). One of the claims by Epstein and others about calculus understanding, in which the CCI has been used as an argument is this: the single most important factor increasing the understanding of calculus concepts is the use of the instructional style called Interactive Engagement (IE) (Epstein, 2013, Thomas & Lozano, 2013). Broadly speaking, an IE oriented teaching style looks for ways to have the students actively participating in the classroom activities; compared to the traditional lecture in which they are only recipients of the knowledge that is being given to them.

Within this framework, I wanted to test the level of understanding of calculus concepts by students of Differential Calculus at the Universidad de Bogotá Jorge Tadeo Lozano (UTADEO). The population consists of undergraduates, some of whom (about 1/3) come from a state-sponsored program aimed at low-income students. Another particularity of the present study is the use of a large session (about 100 students) to present the theory of mathematics; previously (until 2015) most calculus courses consisted in small groups (less than 30 students). First, I report how this particular cohort of students and this institution compared to other institutions who have taken the CCI. Secondly, I use the results of the CCI as a way to characterize the different subgroups of students.

Regarding the structure of this article, in the section *Materials and Methods* I characterize the student population and explain the methodology. In the section *Results* I discuss the overall results and findings. Finally, in the section *Conclusion* I have included the closing remarks and concussions.

# MATERIALS AND METHODS

## Population

This study was conducted with students of Differential Calculus at the *UTADEO*. This is a Colombian university, located in *Bogotá*. There are some particularities regarding the population:

- The Colombian government has a test for high school students, Saber 11 (ICFES, 2014). Since the year 2015, the Colombian government started a scholarship program "Ser Pilo Paga", aimed at those students who do well in this test (the specific threshold varied during the existence of the program), and cannot afford the cost higher education. Those students are referred to as *Pilos*. Some of the students in this study are *Pilos*. (ICETEX - MEN, 2015)
- While most Colombian universities use either the state test or internal knowledge testing in order to filter admission, the UTADEO does not. The point in case is to give the opportunity of higher education even to students whose previous background has not given them enough opportunities.
- The first session of the week is given by full-time professors. They present the concepts to a large class (about 100 students). In class, grading is done by using clickers. Both the large and the small sessions happen at the same time of the day, at different days of the week. There were four of those large classes, given by three different professors.
- In the second session, students participate in smaller groups (about 25 students), led by a part time teacher; there they solve questions, exercises, and problems in a hands-on approach. I obtained averaged gain for the small session groups, as well as the percentage of *Pilos* in the small sessions. There were 20 small sessions, led by 12 part time teachers.
- The book used throughout the class is "Single Variable Calculus: Early Transcendentals", by James Stewart (Stewart, 2015). As with most contemporary calculus books, this textbook has been influenced by the reform movement. Therefore, it includes, for instance, the graphical, analytical, and numerical treatment of functions.
- The first CCI test was administered during the first semester of 2016 (January through May).

Table 1 shows the characterization of the population of students in Differential Calculus. Causes of the difference in population between those in the CCI study and the universe of the Differential Calculus population (521 students) include students who did not attend class the day of either the first of the second application of the CCI test, students who withdrew the class, and students who did not fill the gender information in the questionnaire. I did include in the CCI study students who did not fill the fields related to the prior Differential Calculus knowledge, gender, or participation in the Ser Pilo Paga program.

| Quantity | Value |
|---|---|
| Number of Students who signed up for Differential Calculus | 521 |
| Percentage of signed up who did not cancel their enrollment | 90 % |
| Percentage of students who did not cancel their enrollment and later on approved | 64 % |
| Percentage of students who did not cancel their enrollment but later on attendance | 3 % |
| Percentage of students who did not cancel their enrollment and later on failed by grade | 33 % |

Table 1: Characterization of the Differential Calculus Students

Table 2 shows the characterization of the population in the CCI test. It is worth mentioning that *Pilos* make up 1/3 of it, and only 31 (10%) of the students claimed to have studied calculus before.

| Quantity | Value |
|---|---|
| Number of students | 305 |
| Males | 158 |
| Females | 147 |
| *Pilos* | 105 |

| | |
|---|---|
| Students who have studied calculus | 31 |
| Students who have studied pre-calculus | 49 |

Table 2. Characterization of the Students in this study

Fig. 1 is a histogram of the distribution of ages of the students within the study. It peaks at 18 years of age. Most of the population (254 students) is younger than 22.

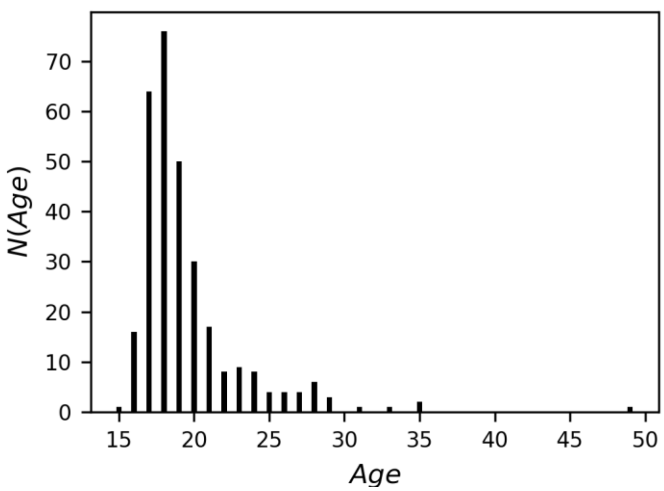


Figure 1. Distribution of ages of students in the study

**The Test**

In words of Epstein (Epstein, 2006): “The calculus concept inventory (CCI) is a test of conceptual understanding (and only that -there is essentially no computation) of the most basic principles of Differential Calculus”. The 2006 version consists of 22 multiple options - single choice questions. In this study, I translated the original English text to Spanish.

In the context of the CCI, the improvement of the understanding of calculus by a group of students is measured by the normalized gain. Epstein, (Epstein, 2013), defined it as:

$$\langle g \rangle = \frac{\mu_f - \mu_0}{s_{max} - \mu_0} \quad (1)$$

with $\mu_0$ the mean score of the class at the first test, $\mu_f$ the mean score of the class in the second test, and $s_{max}$ is the maximum possible value of the score (22 in this case). The gain measures: "the gain in the class' performance (...) as a fraction of the maximum possible gain" (Epstein, 2013).

In the present work, I have calculated the gain for each of the 20 small sessions, as well as for some subgroups of students characterized by different characteristics: gender, *Pilos*, and initial score.

## RESULTS

For each of the 305 students I have both the initial and final value of the score of the test. The gain defined by Equation 1 is interpreted as the improvement of the understanding of calculus by the whole population. To determine an uncertainty of the gain I interpreted the standard deviation of each of $\mu_0$ and $\mu_f$ as the uncertainty of each measurement. Therefore, the uncertainty of the gain is given by:

$$\delta_g = \frac{\mu_f - s_{max}}{(s_{max} - \mu_0)^2}\delta_{\mu_0} + \frac{\delta_{\mu_f}}{s_{max} - \mu_0} \qquad (2)$$

where $\delta_\mu$ is the standard deviation, $\delta_g$ uncertainty in the gain.

This gives a gain of $g = 0.10 \pm 0.03$, which is clearly statistically significant gain. Now, since the gain relies in average values this gain could be the result of a few individuals making strong improvement, but this is not the case here as can be seen in Figure 2. At the left, in black, the distribution of pre-class test scores; at the right, shaded, the post-class test scores. It is clear that after the class the distribution of test values is shifted towards higher score values. Thus, as measured by the CCI, I conclude that taking a one semester class of Differential Calculus resulted in a clear improvement in understanding.

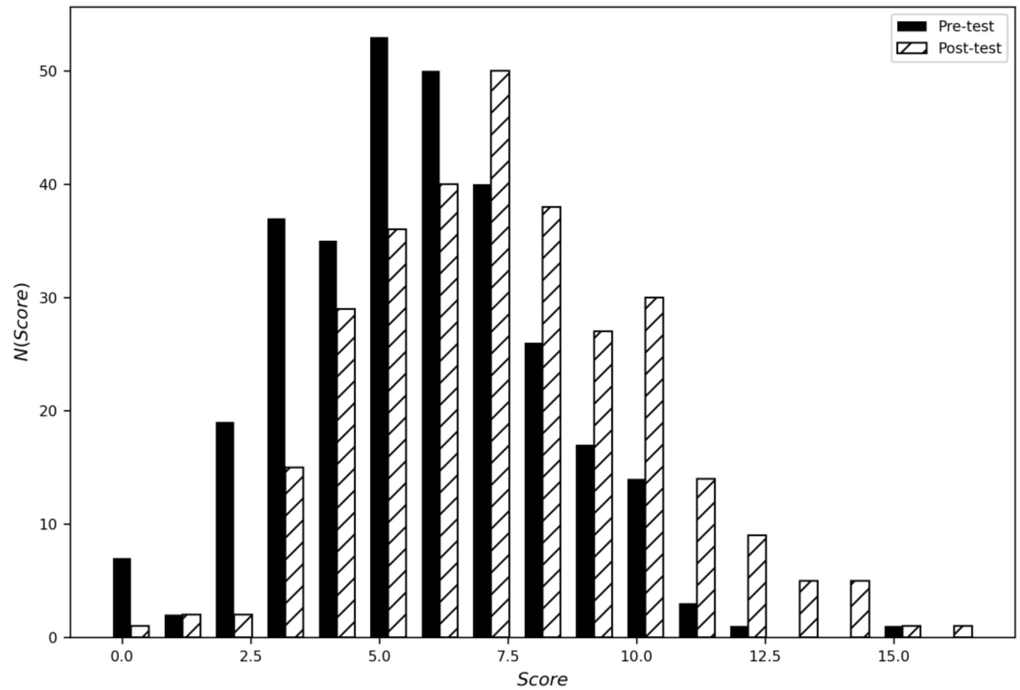


Figure 2. Distribution of raw scores, both for pre- and post-test

Since the gain of the whole population only gives a global characterization, I proceeded to calculate the gain that different sub-groups have had on the test. This was applied as an exploration of the data that could guide the improvement of the teaching of calculus and could help us find where I should focus my efforts. I did not explicitly start from a hypothesis of these groups to have had an advantage, on the contrary, I wanted the data to show the inequities. The result of this process is summarized on Table 3. Every value in this table is statistically significant, as can be seen comparing the respective values and the uncertainty. Of these groups, *Pilos* have a higher gain 0.12, compared to 0.09 of *no-Pilos*. Males also improved more than females, with 0.11 compared to 0.09. The best session, session 6.2, scored 0.27, a very impressive value compared with the rest of the students. The worst session, at 0.05, did half of the improvement than the general population.

| Gain | Sub-Group |
|---|---|
| $0.10 \pm 0.03$ | Whole Population |
| $0.11 \pm 0.03$ | Males |
| $0.09 \pm 0.04$ | Females |
| $0.12 \pm 0.04$ | *Pilos* |
| $0.09 \pm 0.03$ | *No-Pilos* |
| $0.09 \pm 0.01$ | Have studied Calculus |
| $0.08 \pm 0.01$ | Have studied pre-calculus |
| $0.27 \pm 0.17$ | Best session |
| $0.05 \pm 0.01$ | Worst session |

Students attended two sessions a week, and the second session was led by a part time teacher and held in smaller groups. Does the group size impact the understanding of calculus? From Figure 3, I concluded that there is no correlation between the second session group size and gain.

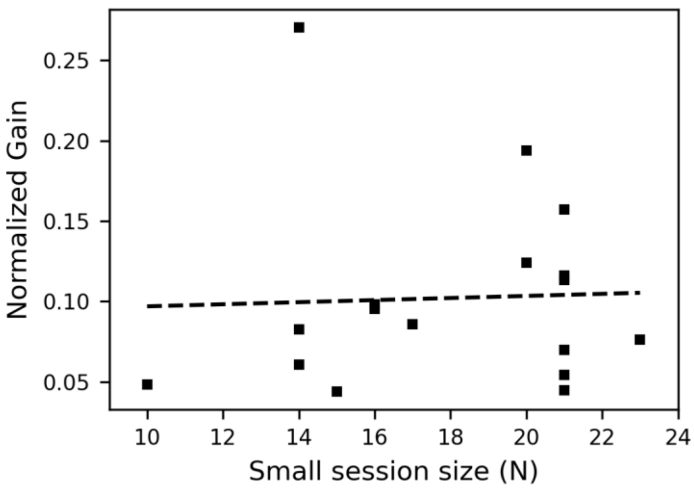


Figure 3. Normalized gain vs. complementary class size. The horizontal line corresponds to a linear regression giving a slope of 0.04, and p-value $5 \times 10^{-6}$. Therefore, one can say with significance that there is no correlation between second (small) session size and gain.

Furthermore, I wanted to find out whether the presence of *Pilos* in a group had an effect on the class performance. Indeed, as the percentage of *Pilos* increases, the gain tends to increase also. This is seen in Figure 4, in which I have plotted the normalized gain against the class percentage of *Pilos* in the second (small) session. My analysis yielded a positive correlation represented in a line with slope of $0.097 \pm 0.056$. Nonetheless, the p-value is 0.1, so the analysis is inconclusive.

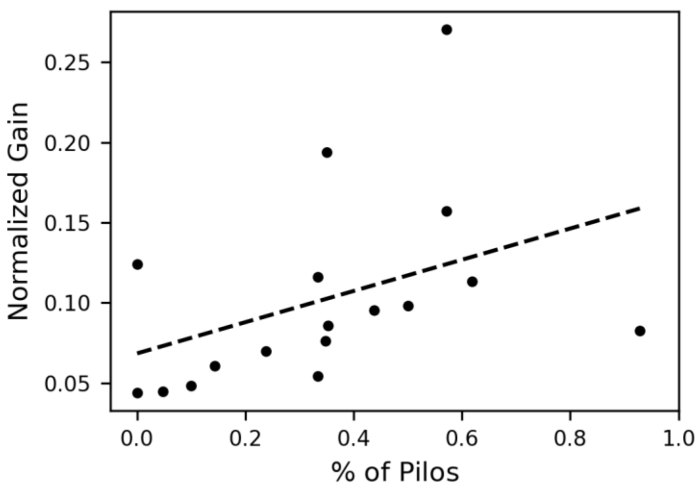


Figure 4. Normalized gain vs. percentage of *Pilos*. The dashed line corresponds to linear regression analysis with a slope of 0.097; the p-value is 0.10.

In Figure 5, I have plotted the normalized gain against the initial class score. Since I am averaging over the population of students who obtained a given score, I have included error bars. The last two points (12 and 14) correspond to the gain of single students; therefore, the concept of gain hardly can be applied to them and is safe to disregard them. A decreasing trend for low scores can be seen in the data, in which having a higher initial score leads to a smaller gain. One hypothesis is that having a higher initial score implies a lower chance of improvement, so gain decreases.

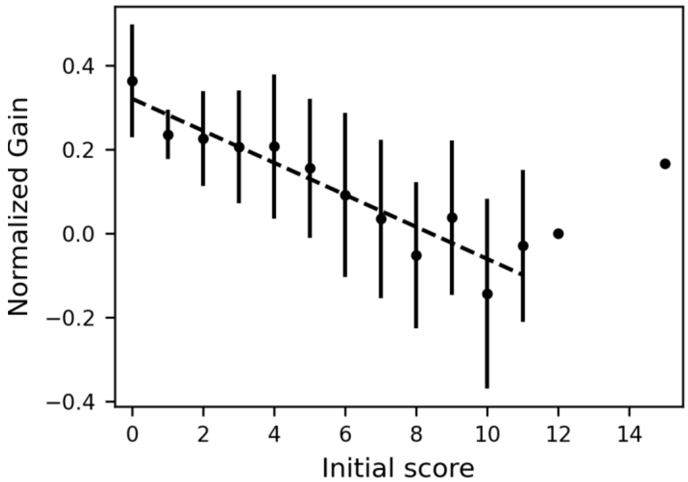


Figure 5. Normalized gain vs. initial raw score: the better the initial score, the lower the gain. The dashed line corresponds to linear regression analysis with a slope of -0.038; the p-value is $5 \times 10^{-6}$, therefore the correlation among the variables is significant.

Another question of interest was, is there a relationship between the time of the day the lessons were held and the gain? Additionally, is there a relationship between the years of experience of the professor in charge of the large sessions and the gain? For the data, these two questions are related, as the most senior professor also taught the early sessions. The data is shown in Figure 6. Average normalized gains of the small session groups that were held at 7:00 is significantly higher than those at 9:00, 11:00, or 18:00. Now, since students choose their schedule, time of the lesson could be a proxy to other characteristics of the population. The effect, however, could also be attributed to the experience of the professor in charge of the large class, as she has over 30 years of experience against 10 years of experience for the full-time professors in charge of the other large groups.

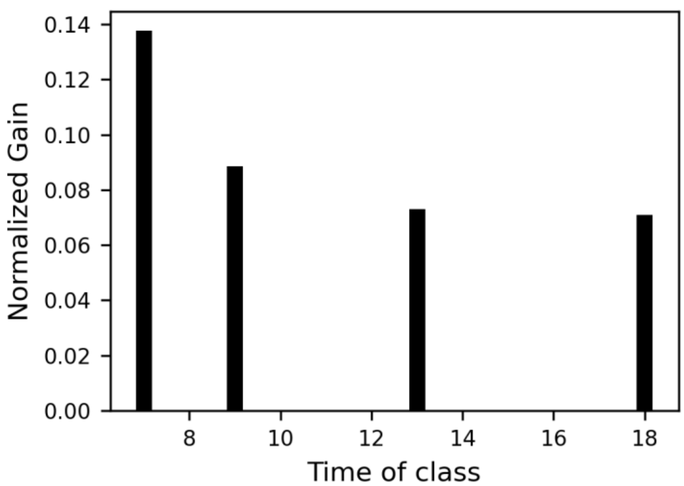


Figure 6. Average normalized gain of small sessions against time of the day the lecture was given.

**CONCLUSION**

I have applied CCI, a conceptual test of Differential Calculus, to 305 students of the subject, in Bogotá, who attend a class with both a large session (around 100 students in a lecture room) and a small session (between 14 and 25 students); each session is 2 hours, once a week. There is a positive gain of $0.10 \pm 0.03$ when taken over the whole population of the study, which indicates that there is a better understanding of basic concepts after taking the class. However, the gain is small, when compared to the reported data from University of Michigan (with an average of 0.35 (Epstein, 2013)).

Since students were split into 20 small sessions, and the percentage of *Pilos* (students belonging to a national excellence program) varied among small sessions, I have found evidence of a positive correlation between the number of *Pilos* and the normalized gain, however, the statistics is inconclusive. Having some *Pilos* in a classroom may have improved understanding of Differential Calculus concepts. I expect the results of the present report to be a useful baseline. Therefore, future changes to the curriculum or other teaching intervention can be compared against the values herein reported, and therefore, could provide useful information. Moreover, researchers in similar institutions can compare their results against ours.

There are open questions. I have obtained a negative linear correlation between the initial test score and the gain. This trend is different from what is usually stated in the literature: that there is no relationship between the initial score and the gain. It would be interesting to see if this is seen in different data sets, since the p-value of $5 \times 10^{-6}$ gives confidence in my analysis. A possible interpretation is that the students having a higher initial score diminishes the possibility of improvement. However, I have also been aware of the criticism of the CCI, according to the work of (Gleason, Thomas, et al., 2015) the CCI relies strongly on notation rather than on conceptual understanding relating multiple representations. I do not plan to address the shortcomings of the CCI in this work. Further studies could bring more light into this issue.

Finally, the effect of seniority of the teacher with the time of the day the class is taught cannot be disentangled, therefore it is impossible to conclude whether any of those variables has an impact on the outcome of the students. These are still an open questions that should be tackled in further research.

**Limitations of the CCI**

The main weakness of the study is its reliance on the CCI itself, which is not exempt of criticism. According to a factor analysis reported by (Gleason, Thomas, et al., 2015, Gleason, White, et al., 2015), the CCI can be used to explain one single factor, "overall knowledge of calculus content"; rather than distinguishing between three factors, functions, derivatives and "limits/ratios/the continuum". Furthermore, the internal reliability of the CCI (a Cronbach alpha of 0.7) is below the standard (0.8). On his work on the subject, Gleason does not call for a dismissal of the CCI, rather he wants it to be improved to overcome these difficulties. Such modification of the CCI is beyond the scope of this paper, but the author would gladly collaborate on such endeavor.

**Acknowledgments**

I would like to thank Professor Epstein for giving us access to the CCI. His article inspired me to try to understand calculus comprehension. Secondly, I would like to acknowledge the work of the full-time professors in charge of the large sessions, Adelina Ocaña and Javier Riascos. I would like to thank the teachers in charge of the small sessions who proctored and graded the test: Edgar Palacios, Manuel Ariza, Diego Cáceres, Orlando Aya, Carolina Cipagauta, Sandra González, Esperanza González, Irina Reyes, Bibiana Franco, Paola Clavijo, Veimar Rey, and Luis Lara. I would also like to thank Javier Riascos and Andrea

Franco for fruitful conversations about early versions of this work; their comments helped me to improve it.

## References

[1] Berry, J. S. & Nyman, M. A. (2003). Promoting students' graphical understanding of the calculus. *The Journal of Mathematical Behavior*, 22(4), pp. 479 – 495.

[2] Epstein, J. (2006, October 19-21). The calculus concept inventory. *National STEM Assessment Conference*, pp. 60–67. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.467.6024&rep=rep1&type=pdf#page=64

[3] Epstein, J. (2007, September 7-12). Development and validation of the calculus concept inventory [Paper Presentation]. *Proceedings of the ninth international conference on mathematics education in a global community*, pp. 165–170. http://math.unipa.it/~grim/21_project/21_charlotte_2007.htm

[4] Epstein, J. (2013). The calculus concept inventory - measurement of the effect of teaching methodology in mathematics. *Notices Amer. Math. Soc*., 60(8), pp. 1018–1026.

[5] Gleason, J., Thomas, M., Bagley, S., Rice, L., White, D., & Clements, N. (2015). Analyzing the calculus concept inventory: Content validity, internal structure validity, and reliability analysis. In Tim Fukawa-Connelly, Nicole Engelke Infante, K. K. M. Z., (Eds.), *Proceedings of the 37th International Conference of the North American Chapter of the Psychology of Mathematics Education*, pp. 1291–1297.

[6] Gleason, J., White, D., Thomas, M., Bagley, S., & Rice, L. (2015). The calculus concept inventory: A psychometric analysis and framework for a new instrument. In Tim Fukawa-Connelly, Nicole Engelke Infante, K. K. M. Z., (Eds.), *Proceedings of the 37th International Conference of the North American Chapter of the Psychology of Mathematics Education*, pp. 135–149.

[7] Habre, S. & Abboud, M. (2006). Students' conceptual understanding of a function and its derivative in an experimental calculus course. *The Journal of Mathematical Behavior*, 25(1), pp. 57 – 72.

[8] Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The physics teacher*, 30(3), pp. 141–158.

[9] Instituto Colombiano de Crédito Educativo y Estudios Técnicos en el Exterior & Ministerio de Educación [ICETEX & MEN] (2015). "Créditos condonables para la excelencia en educación superior" Ser pilo paga. Reglamento Operativo. https://www.eafit.edu.co/serpilopaga/Documents/SPP1.PDF

[10] Instituto Colombiano para la Evaluación de la Educación [ICFES] (2014). Sistema Nacional de Evaluación Estandarizada de la Educación. Alineación del examen SABER 11. Lineamientos generales 2014 – 2. ISBN:978-958-11-0630-1. http://www.icfes.gov.co/instituciones-educativas-y-secretarias/saber-11/documentos

[11] National Council of Teachers of Mathematics [NCTM]. (2021, July 9). *Procedural Fluency in Mathematics.* https://www.nctm.org/Standards-and-Positions/Position-Statements/Procedural-Fluency-in-Mathematics/

[12] Rhea, K. (2008) The calculus concept inventory at a large research university. [Unpublished report]. http://math.arizona.edu/~guada/F09/597T/docs/kr-CCIprelimrep.pdf.

[13] Sofronas, K. S., DeFranco, T. C., Vinsonhaler, C., Gorgievski, N., Schroeder, L., & Hamelin, C. (2011). What does it mean for a student to understand the first-year calculus? perspectives of 24 experts. *The Journal of Mathematical Behavior*, 30(2), pp. 131 – 148.

[14] Stewart, J. (2015). Single variable calculus: Early transcendentals (7th ed.). Cengage Learning.

[15] Österman, T & Bråting, K. (2019) Dewey and mathematical practice: revisiting the distinction between procedural and conceptual knowledge. *Journal of Curriculum Studies.* 51(4), pp. 457-470. https://doi.org/10.1080/00220272.2019.1594388

[16] Thomas, M. & Lozano, G. (2013). Analyzing calculus concept inventory gains in introductory calculus. In *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education*. http://pzacad.pitzer.edu/~dbachman/rume_xvi_linked_schedule/rume16_submission_95.pdf.

[17] Wu, H. (1999). Basic skills versus conceptual understanding. *American Educator*, 23(3), pp. 14–19.

[18] Zazkis, D. (2013). Calculus Students' Representation Use in Group-Work and Individual Settings. [Doctoral Dissertation, UC San Diego], https://escholarship.org/uc/item/77n1m243