EDVCATIO PHYSICORVM
QVO NON ASCENDAM?
ISSN 1870-9095

Thirty-six years of the Forced Concept Inventory and the Mechanics Baseline Test: is Aristotle still playing hide and seek in our classrooms?

C. Pérez Millán[1], S. **Otranto**[2]
[1]*Departmento de Física, Universidad Nacional del Sur, Av. Alem 1253, B8000CPB Bahía Blanca, Argentina.*
[2]*Departamento de Física, Instituto de Física del Sur (IFISUR), Universidad Nacional del Sur (UNS), CONICET, Av. Alem 1253, B8000CPB Bahía Blanca, Argentina.*

E-mail: cperezmillan@uns.edu.ar

(Received 23 May 2020, accepted 5 April 2021)

Abstract

In this work, we analyze the conceptual comprehension of Newtonian Physics topics by students of different engineering careers during the first Physics course at the Universidad Nacional del Sur, in Bahía Blanca, Argentina. The Mechanics Baseline Test introduced by Hestenes and Wells more than three decades ago, is complemented with semi-structured personal interviews. The categories identified during the interviews for the concept of force and the different study techniques employed are analyzed. Present results indicate low overall conceptual comprehension. Moreover, we find clear indications that a major fraction of the interviewed students still retain Aristotelian representations even after they have passed formal exams on the topic.

Keywords: Newtonian Physics, Conceptual learning, Mechanics Baseline Test.

Resumen

En este trabajo analizamos la comprensión conceptual de temas de Física newtoniana por parte de estudiantes de diferentes carreras de ingeniería durante el primer curso de Física de la Universidad Nacional del Sur, en Bahía Blanca, Argentina. La prueba *Mechanics Baseline Test* presentada por Hestenes y Wells hace más de tres décadas se complementa con entrevistas personales semiestructuradas. Se analizan las categorías identificadas durante las entrevistas para el concepto de fuerza y las diferentes técnicas de estudio empleadas. Los resultados actuales indican una comprensión conceptual general baja. Además, encontramos indicios claros de que una fracción importante de los estudiantes entrevistados aún conserva representaciones aristotélicas incluso después de haber aprobado exámenes formales sobre el tema.

Palabras clave: Física Newtoniana, Aprendizaje conceptual, Mechanics Baseline Test.

I. INTRODUCTION

Nearly 40 years ago, computers left laboratories and became household items. Probably unnoted at first, it was just the beginning of a drastic technological advance that completely changed our whole lifestyle, including the way in which we interact with information on a daily basis. In contrast to previous decades where public and personal libraries were the major source of information, nowadays information is widely available and at a thumb's pace from our mobiles. All these changes naturally led to seriously question the way in which we teach and learn and where the focus should be put on.

By the 1980s Physics and Natural Sciences education research started to focus in the possibility to measure the conceptual learning by students in contrast to the simple recitation of memorized statements [1]. Although the latter can nowadays hardly be considered a learning process, it is worth quoting Richard Feynman's words on his teaching experience in Brazil during the 1950s``...*I discovered a very strange phenomenon: I could ask a question, which the students would answer immediately. But the next time I would ask the question, as far as I could tell they couldn' t answer it at all!.....After a lot of investigation, I finally figured out that the students have memorized everything, but they didn't know what anything meant*.." [2]. Some of these conceptual learning studies even started at very early stages of the instruction and identified the previous representations of different natural sciences phenomena during the primary and secondary schools (ages 10-16) [3].

By 1992, Hestenes and Wells highlighted the dominant role that common-sense beliefs play in introductory Physics courses which focus on Newtonian mechanics. By then, it was found that those beliefs are not entirely in agreement with Newton's laws and are so firmly established that the conventional instruction does not change them. Moreover, these findings turned out to be course and instructor-independent. In other words, students who acquired a deep conceptual comprehension of the Newtonian mechanics most probably reached it thanks to their own learning devices and not to the pedagogical design of the course or a vital intervention by the instructor. These authors then introduced the Force Concept Inventory (FCI) and the Mechanics Baseline Test (MBT) as potential tools to assess student understanding of the Newtonian mechanics grounds and assist teachers to find out where their instruction can be improved [4, 5]. While the Force Concept Inventory was designed to be applied in students with no formal training in mechanics (pre-University stage), the Mechanics Baseline Test was designed mainly as a post-instruction test during or after formal courses. Even when most of the questions look simple, it has proven to be a difficult exam provided that answers strongly rely in exploring the students conceptual domain, avoiding the “plug-in” practice (i.e find a mathematical formula where the given data fit in order to get a number).

Thirty-six years from their introduction, these tests and other similar [6], have been widely used worldwide and are considered valid for large courses in which personal interaction to identify the previous representations of each and every student is difficult if not unfeasible in practice [7, 8, 9, 10].

In this work, we gain insight into the conceptual comprehension of Newtonian mechanics by Engineering students at the Universidad Nacional del Sur, in Bahía Blanca, Argentina. The MBT is complemented with subsequent personal interviews. Students from the Physics I course which is taught simultaneously by four lecture commissions volunteered for the test. In section II we introduce the context in which the MBT was performed. In section III we show and analyze the obtained results. The complementary views acquired during the personal interviews are presented and analyzed in section IV. Finally, in section V conclusions are drawn.

II. CONTEXT OF IMPLEMENTATION

The Universidad Nacional del Sur (UNS) [11] is the seventh National University created in Argentina in January 1956 out of the 55 created to date. In contrast to other major Argentinian National Universities like Buenos Aires, Córdoba or La Plata which are organized in terms of faculties, the UNS is organized in terms of Academic Departments. This quite unique context of implementation in the country is particularly relevant for the present study, provided that the Physics Department is in charge of all Physics courses taught in every degree program at UNS. Following the national policies in Argentina, National Universities base their functioning basis on unrestricted access and gratuity. These conform the common denominator to all their degree programs throughout the country and have led during the XX century to five Nobel laureates (Medicine (2), Peace (2) and Chemistry (1)).

The first physics course that engineering students take during the second semester of their first year is Physics I, which is calculus-based. The course is taught simultaneously by four lecture commissions, each one conformed by a Professor, a Chief of Teaching Assistants and a set of Teaching Assistants which can be either graduated or advanced students. In general terms, the professor provides the lectures and the Chief of Teaching Assistants coordinates the problem solving and laboratories which are performed separately according to a previously established agenda. In practice, it is common to find professors interacting with the students during the problem solving and laboratories spaces to gain insight first hand on students' development. All four professors have PhDs either in Physics or in Science and Technology of Materials. Three of them are Researchers for the National Research Council of Argentina (CONICET). Three of the chiefs of teaching assistants have PhDs and the remaining one is about to complete hers. The number of teaching assistants changes among commissions but a figure of one every 15 to 20 students closely resembles the Physics Department spirit during the last few years. Most of the graduate teaching assistants are graduate students transiting their PhDs programs with studentships from CONICET. Neither of the faculty members have formal studies in Didactics or Pedagogy. Classes are usually organized in terms of two weekly events which consist of a 2-hour lecture followed by 2-hours of problem solving.

The laboratory schedule changes among commissions but runs separately from the lectures and problem-solving spaces. During the semester the students need to pass 2 to 3 partial exams (depending on the commission) to gain the chance of taking the final examination that would lead to the approval of the course. Professors have the freedom to choose whether they provide a second chance for every partial exam or if they give a second opportunity at the end of the semester to those who overcome a global threshold score. Extra classes are usually added in the days that precede to the different partial exams and are mainly targeted to help students catch up with the problem-solving agenda. The grading of the exams is numerical from 0 to 100 allowing a direct percentage analysis.

The bibliography for this course mainly consists in the Spanish translation of authors worldwide recommended for the target level or books written in Spanish by Argentinean authors [12, 13, 14, 15, 16, 17].

III. FIRST STAGE: THE MBT

The MBT was implemented on 133 students that were taking the Physics I course with the different commissions right before their third exam (movement integrals and conservation theorems). All of them previously passed the first two partial exams (i.e., answered correctly more than 60% of the exams) corresponding to kinematics and dynamics or got over the required score of 50% to keep attending the course and take a second chance exam at the end of the semester.

The MBT results are shown in Fig. 1 and expressed in percentage terms of correct answers. The mean value was of 35% with a standard deviation of 14%. This value is 25% below the approval threshold. Only 5 out of 133 (4%) got over 60% and no one scored in the 90%-100% band. Three scored in the 0%-10% band.

According to the secondary school background, the universe of students was divided in Natural Sciences (30%), Technical (26%), Economics (24%), Social Sciences (15%) and 4% encompassing students from Arts and Communication. In other words, about 56% of this universe is expected to have tackled topics of Newtonian mechanics at some extent without elaborate manipulation of Calculus or Algebra. In Fig. 2 we show the MBT percentages in terms of the secondary school orientation. In terms of performance, students with Economics and Technical backgrounds obtained the higher scores (about 38%), closely followed by Social Sciences. Interestingly, students with background in Natural Sciences only outperformed those who have background in Arts and Communication. Overall, differences are subtle and we conclude that no clear correlation was identified between the secondary school orientation and the MBT results.

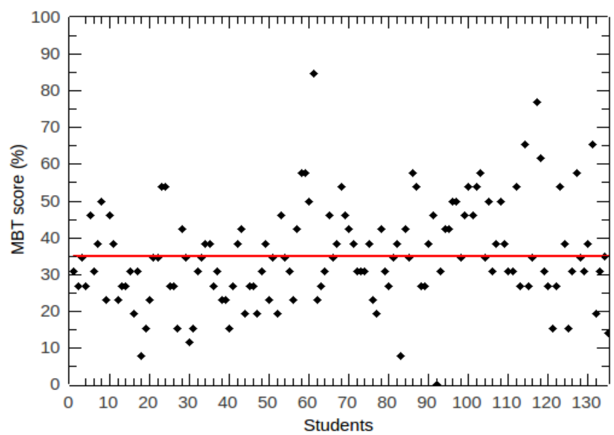

FIGURE 1. Distribution of MBT percentages. The general mean value is shown with a red-line.

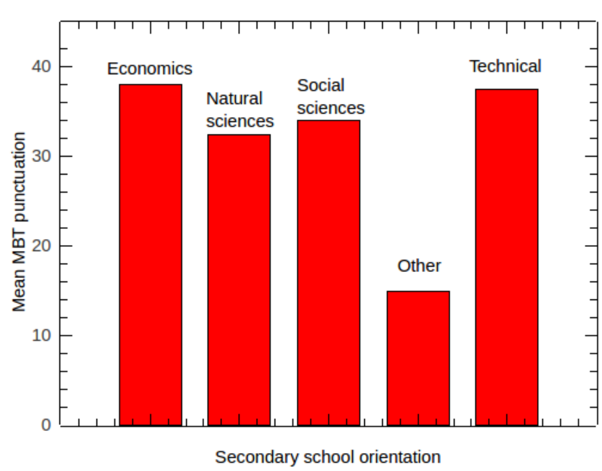

FIGURE 2. Histogram of MBT percentages according to the secondary school orientation.

Having provided the overall result, we now perform a more detailed analysis of the collected data. In table I we show the distribution of concepts evaluated along the MBT. Their partial weights are indicated in Fig. 3 (a). The overlaps indicate that some of the questions do not belong to a single category. In our case, a closer inspection of each test lets us conclude that more than 60% of the students answered correctly those MBT questions belonging to the category of Kinematics that looked very similar to those previously explored during the problem solving.

Regarding Newton's Second Law, some of the questions intended to explore the students' conceptions on the dependence on mass of the Second Law were answered correctly while other which mainly refer to the same physical situation collected very low successful answers. Hence, we find that the larger fraction of correct answers relate to physical situations or questions that have been formerly tackled in the problem agenda. This would not necessarily mean a deeper comprehension of the physical situations associated to those points. It could well be that students perform some sort of pattern recognition, having traced that route before.

TABLE I. Newtonian Concepts distribution on the MBT.

A. Kinematics
Linear Motion
Constant acceleration
Average acceleration
Average velocity
Integrated displacement
Curvilinear Motion
Tangential acceleration
Normal acceleration
$a = v^2 / r$
B. General Principles
First Law
Second Law
Dependence on mass
Third law
Superposition principle
Work-energy
Energy conservation
Impulse-momentum
Momentum conservation
C. Specific Forces
Gravitational free fall
Friction

Following Hestenes and Wells, we group the MBT questions in three categories: calculus, diagram and kinematics. Fig. 3 (b) shows the partial weights corresponding to these three categories. It can be seen that 10% of the exam requires the use of diagrams and calculus in order to solve a particular situation belonging to kinematics. Results for this categorization are shown in Fig. 4 (a)-(c). The average scores for the three categories read 29%, 27% and 33% with standard deviations of 17%, 18% and 17% respectively. If we now remove from the kinematics category those questions which refer to physical situations that have not been explicitly considered in the list of problems solved by the students, we note that the average score for this adapted kinematics category increases to 57% and 18% of the students get over the 60% score.

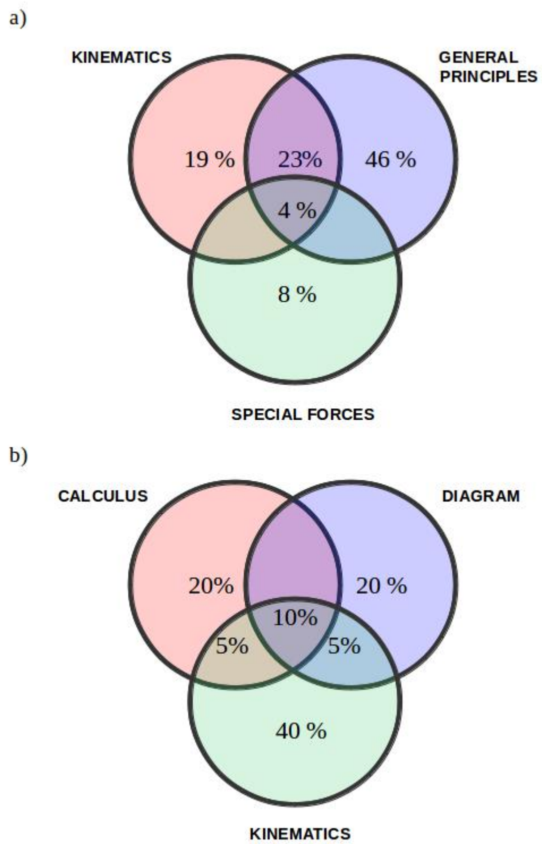

FIGURE 3. a) Partial weights for the Newtonian Concepts distributions and Categories in the MBT. b) Partial weights for the three categories within the MBT established by Hestenes and Wells [5].

This fact reinforces our previous statements regarding that students react mainly via pattern recognition when they face a new physical context.

Last, given the low performance exhibited in Fig. 1, we compare our present results against the random resolution of the MBT to gain insight up to which extent can the present analyses be considered reliable in terms of conceptual comprehension. For that purpose, we perform 10^5 MBT identical Monte Carlo simulations (i.e., 10^5 sets of 133 students answering 26 questions with 5 options each). No questions were left blank. The obtained results are shown in Fig. 5 (a) and clearly indicate a maximum at the expected score of 100% divided by the number of options. The random resolution over extremely large sets suggest a band-like structure containing the possible outcomes. The chance of approval in a random solution of the MBT is found negligible. Our present MBT results clearly overlap with the random simulation results. In Fig. 5 (b), we show the histogram for the percentage of individual exams exceeding our MBT mean value for each simulated test. From our simulation, we estimate that less than about 7% of the students could have reached the mean value of 35%

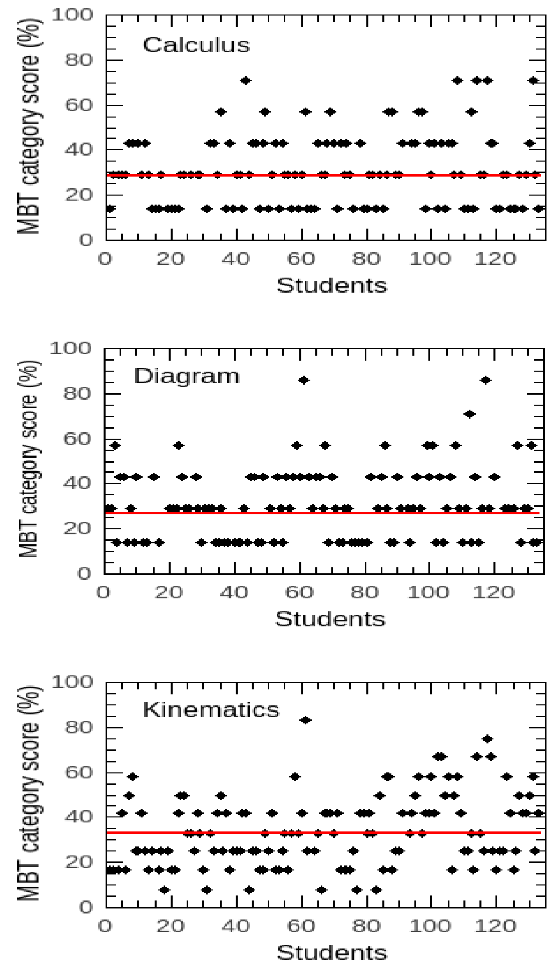

FIGURE 4. a) Results for the three categories within the MBT established by Hestenes and Wells [5]. The mean value is shown with a red line.

with no explicit knowledge on the topic.

We are then led to conclude that the MBT can be considered a valid instrument in the present low score situation.

IV. SECOND STAGE: INTERVIEWS

In a subsequent stage, 22 out of the 133 students were interviewed in order to define or characterize the concept of Force. In Fig. 6 we show their MBT performance and contrast it to their first partial exam scores. These results clearly show that the 22 interviews clearly reflect the same trends described in the previous section for the whole universe of students considered. The average score in the first partial exam was 72% with a standard deviation of 82%. Out of the 22 students, 13 were taking the course for the first time while 9 had failed to pass in previous semesters. 16 would end up passing the course while 6 would fail in their third exam.

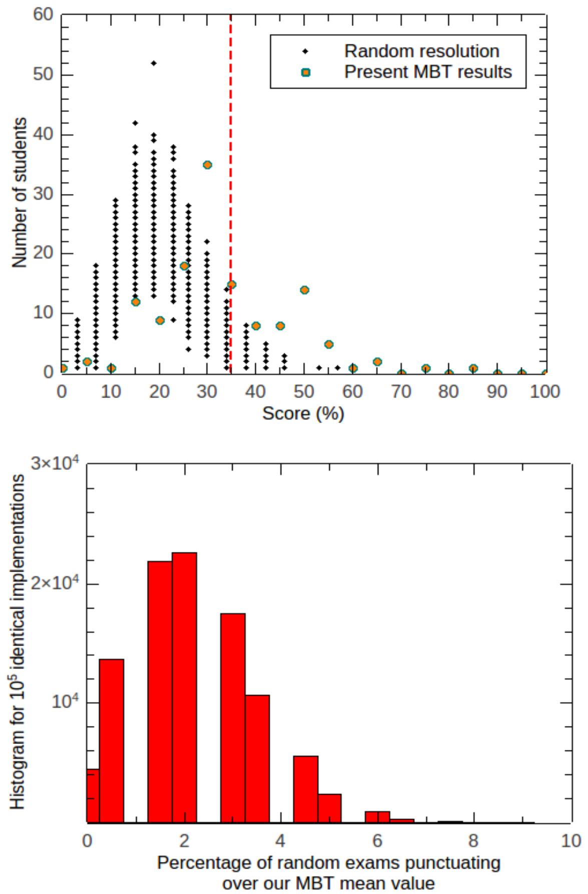

FIGURE 5. Comparison of our present results against 10^5 identical MBT simulations.

In average, they claimed a 92% of attendance to the lectures and an 82% attendance to the problem-solving classes.

The interviews were of qualitative character and semi-structured, so as to gain insight on the students' conceptions in a guided conversation scheme [18]. Each interview was carried out separately from the class and consisted of one meeting of approximately 1 hour which was recorded for subsequent analysis. The identities of the students were kept confidential by the present authors. In general terms, the interview structure can be described as follows:

- *Introductory stage*: Student personal data: age, sex, career, city and province of origin, if they have a job to support themselves or if they are supported by their families, secondary school orientation, performance at the MBT and first partial exam etc.

- *Physics*: focus is made on their general conceptions on Physics and its particular relevance for their career.
- *Classes*: Questions are driven to gain insight on their personal views regarding lectures, problem solving and laboratory classes and the role of the different actors (faculty members and students).
- *Study Techniques*: Students are asked on their conceptions about the learning-teaching process and if they developed/used particular study techniques for Physics I in contrast to those employed in other courses like Calculus, Algebra or Chemistry. If they have tools to self-control their level of meaningful learning of the topics covered in the course and other tools employed like sketches, graphic representations etc.
- *Resources*: In this point, students detail their preferred resources throughout the semester: books, class notes, older exams, web pages etc.
- *Concepts*: This represents the last stage of the interview. The students are asked on the meaning of the terms: reference frame, position, velocity, acceleration and force.

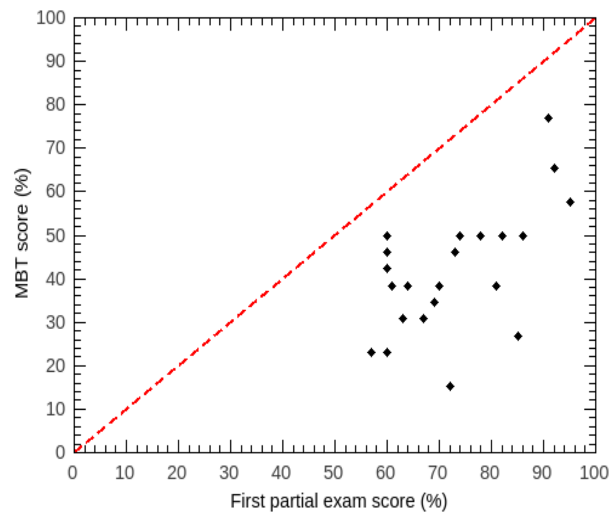

FIGURE 6. MBT results and their comparison to the corresponding first partial exam scores for the 22 students interviewed.

In what follows we will analyze in detail two of the topics covered during the interviews: the students' conception of the force term and their study techniques. Based on the definitions given by the students, 8 categories were identified for the concept of force which are shown in Fig. 7 and described in Table II. The histogram describing the frequencies of each category is shown in Fig. 8. Some interviews led to more than one category. The most frequent answer (8 repetitions) associated the force as the causal of movement. This is quite puzzling since one would expect that three months of formal training in Newtonian mechanics should have helped students migrate from any earlier Aristotelian view. Furthermore, 7 out of 22 explicitly stated they did not know what the concept of force meant.

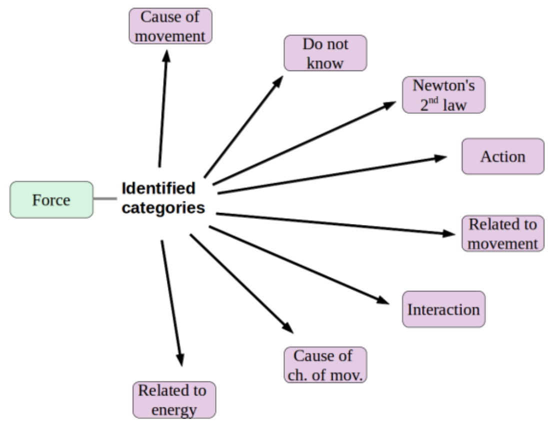

FIGURE 7. Categories identified during interviews regarding the concept of force.

Only 5 out of 25 associated it at least with the Second Law without further insight. Having worked in the explicit resolution of problems involving the concept of force, it is noteworthy that just a minor percentage of students experienced the need to find out what it represented. Maybe the clue to this puzzling scenario can be found in the words of one of the students that were interviewed: "*...I don't think that it is technically necessary to know what a force is in order to describe it*".

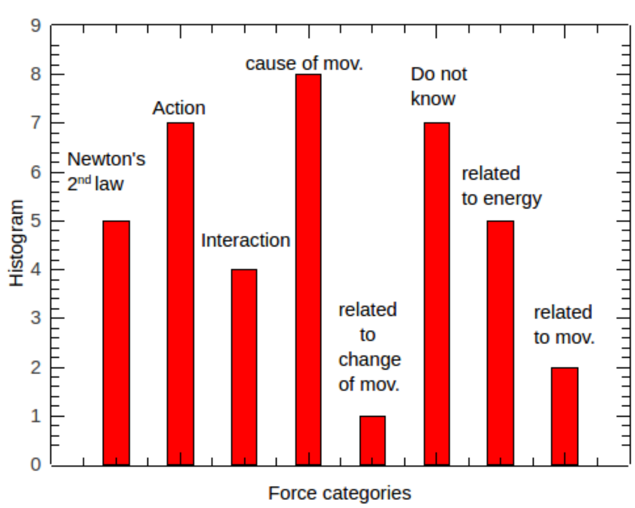

FIGURE 8. Histogram for the different force categories.

Regarding the study techniques used by the students, interviews revealed different categories that are shown in Fig. 9 and described in Table III. The histogram describing the frequencies of each category is shown in Fig. 10. According to the detected frequencies, the largest fraction of students bases their study of Newtonian mechanics on the resolution of the problem lists with no further theoretical insight than notes taken from the lectures they attended.

TABLE II. Description of the categories identified with respect to the students' concept of Force. Interview order numbers are referenced when quoting students' expressions

Category	***Force Description***	***Data***
Second Law	Defined according to the second law	"*... force is proportional to mass times the acceleration and I don't know how to explain it"* (I13)
Action	Identified as an action over the body under study	"*.. the force is an action performed on a body"* (I4)
Interaction	Identified as an interaction	"*...well, a force is... I am going to use the term force, but if I do a force then it induces a reaction, it is an interaction, I do not know how to describe it"* (I20)
Vector	Identified via its vector character	-----
Cause of movement	What makes the body move	"*`...is what makes a body move...I don't know well... it's what you apply on a body to make it move"* (I1)
Cause of change of movement	What induces a change in the body movement	"*...all what you apply to a body so that it modifies its movement or direction"* (I5)
Do not know	No answer or explicit acceptance of the ignorance of meaning	"*...I wouldn't know how to define it... "*(I13)
Related to energy	Related to energy concepts	"*...it is a way to measure amounts of energy*" (I19)
Related to movement	Relation between force and movement	"*...it is something that produces work, but most of all it is an opposition to movement...I wouldn't know how to explain it differently*" (I21)

The next category in terms of frequencies refers to students that read the theoretical notes of the lectures and the suggested books before moving to the problem-solving stage. Overall, the fraction of students that explicitly indicated that they revisit the theoretical concepts presented during lectures by reading the recommended bibliography represent only about 28% of the set of students interviewed.

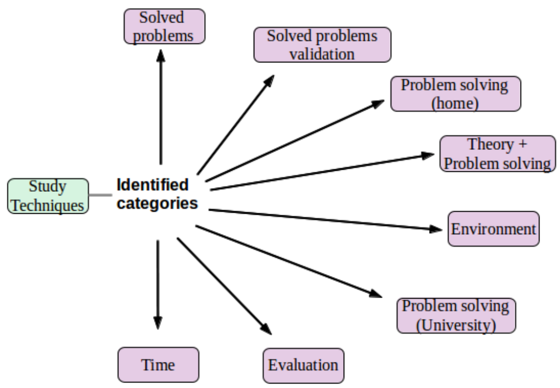

FIGURE 9. Description of the categories identified with respect to the students' studying techniques.

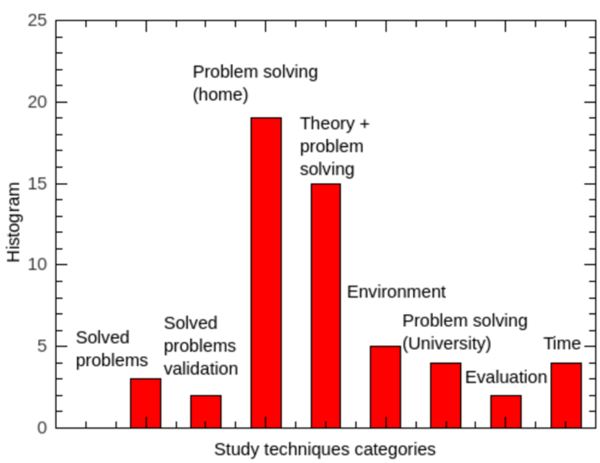

FIGURE 10. Histogram for the difference studying techniques categories.

V. CONCLUSIONS

In this work we have implemented the MBT at the Universidad Nacional del Sur, in Bahía Blanca, Argentina. The organization of this University in terms of academic departments allowed for a wide survey which encompassed all the engineering careers.

Present results suggest that the conceptual comprehension of engineering students is below the expectations. Since questions from the MBT that were already tackled during problem solving obtained more score than others, we are led to conclude that students evidence some mechanical approach to the problems and perform some pattern recognition while doing so.

Unfortunately, and a signal of alert for the faculty, the large contrast between the MBT and the scores of the partial exams seem to indicate that tests are probably conceived in terms of the list of problems solved by the students and not

TABLE III. Description of the categories identified with respect to the students' studying techniques.

Category	***Force Description***	***Data***
Solved problems	Look after the problems answer keys	*``...I study with books that explain the topics and contain example problems. I need to see a lot of examples."* (I4)
Solved problems-validation	Examples with numerical results are used as control	*".. I search the book for examples because I then have numerical results to compare to."* (I2)
Problem solving (home)	Studying Physics means problem solving.	*"...I take the problems lists and work on them"* (I19)
Theory + problem solving	Theory must be understood before moving to problem-solving	*"...I read a lot of theory and as the partial exam approaches, I then focus on the problems lists"* (I20)
Environment	The action of studying is related to the environment	*"`...I wasn't studying much at first…I am not used. I need to put some music and sit down with the book"* (I11)
Problem solving (University)	Intensive use of the problem-solving lapse	*"...I practically do everything during problem solving with the teaching assistants. That is what I find most helpful"* (I3)
Evaluation	Study is related to the evaluation scheme	*"...I try to keep up to date with the problem-solving agenda. The partial exam is based on those problems..."* (I1)
Time	Study is related to the time constrain	*"...at home I check how many exercises I need to solve from the list. Let´s say until number 10. As I have time until next Monday, I work on 2 problems on Thursday, 2 on Friday, 2 on Saturday and 2 on Sunday"* (I6)

in terms of their conceptual comprehension or their previous representations. This hypothesis is also supported by the fact that no correlation was found between the MBT results and the secondary school orientation of the students. One would naturally expect that those students who followed Natural Science and Technical orientations, which represents 56\% of the universe of students considered and which according to their curricula should have been faced to some extent to Newtonian mechanics, would be the ones exhibiting a deeper comprehension of the topics under study.

That scenario would have pushed up the MBT overall score to a value close to the 56% and that was not the case. A subsequent examination of the official secondary school curricula for the Natural Science orientation by the authors revealed a list of Physics contents (rigid bodies, moment of inertia, angular momentum, electro-weak interaction, just to cite a few) that is hardly achievable in terms of the mathematical tools students are supposed to handle at that stage of their instruction [19]. Hence, this context limits any specific expectations based on the secondary school orientation.

Personal interviews revealed a large fraction of students retaining an Aristotelian view of the concept of Force even after having solved a large number of problems related to Newton's second law and momentum conservation by themselves.

Regarding the study techniques, only about 28% of the interviewed students indicated that they went through the conceptual analyses presented in different books, and the larger fraction exhibited a pragmatic approach based solely in problem solving, since this is what they mostly face during partial exams.

Overall, present results suggest that the sole resolution of problems does not imply the comprehension of the involved concepts in the short term, but a mechanical training to plug-in numbers in pre-established formulas instead.

Problems lists and classroom organization should be re-designed to i) gain insight into the specific students universe and previous representations for the particular semester, ii) incorporate the discussion of physical situations that would force the student to confront any previous Aristotelian views and iii) implement already established (or design new) instruments specifically conceived to measure the conceptual comprehension at a post-instruction stage. By doing so, we will be helping students not only to acquire the basic concepts of Newtonian mechanics but also to shape their own critical thinking procedure, a not minor essential tool in future years.

ACKNOWLEDGEMENTS

This work has been supported by PGI 24/F069, Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur (Argentina). The authors would like to thank the faculty in charge of the four commissions that participated of this study and Dr. M. M. Pincelli for critical comments regarding this manuscript.

REFERENCES

[1] Halloun, I. and Hestenes, D., *The initial knowledge state of college physics students*, Am. J. Phys. **53**, 1043 (1985).
[2] Feynman, R. and Leighton, R., *Surely you're joking Mr. Feynman* (W. W. Norton & Company, New York 1985).
[3] Driver, R., Guesne, E. and Tiberghien, A., *Children's Ideas in Science* (Open University Press, Philadelphia 1985)
[4] Hestenes, D., Wells, M. and Swackhamer, G., *Force Concept Inventory*, The Physics Teacher **30**, 141 (1992).
[5] Hestenes, D. and Wells, M., *A mechanics baseline test*, The Physics Teacher **30**, 159 (1992).
[6] Thornton, R. K. and Sokoloff, D. R., *Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Laboratory Curricula*, Am. J. Phys. **66**, 338 (1997).
[7] Bao, L. and Redish, E., *Model Analysis: Representing and assessing the dynamics of student learning*, Phys. Rev. Special Topics: Phys. Ed. Res. **2**, 2010102 (2012).
[8] Hake, R., invited talk, "*National meeting on STEM Concept Inventories*", Washington DC, August 8 (2011).
[9] Scott, T. F. and Schumayer, D., *Central distractors in Force Concept Inventory data*, Phys. Rev. Phys. Educ. Res. **14**, 010106, (2018).
[10] DeVore, S., Stewart, J. and Stewart, G., *Examining the effects of testwiseness in conceptual physics evaluations*, Phys. Rev. Phys. Educ. Res. **12**, 020138 (2016).
[11] www.uns.edu.ar
[12] Alonso, M. and Finn, E., *Física vol. 1* (Addison-Wesley Iberoamericana, México, 1995).
[13] Serway, R. A. and Faughn, F. J., *Física* (Prentice Hall, México, 2001).
[14] Reese, R. L., *Física Universitaria vol. 1* (Thomson, México, 2002).
[15] Resnick, R. and Halliday, D., *Física vol. 1* (CECSA, México, 2004).
[16] Beer, F. P. and Johnston, E. R., Jr, *Mecánica Vectorial para Ingenieros* (McGraw-Hill, México, 2013).
[17] Roederer, J., *Mecánica Elemental* (Eudeba, Buenos Aires 2003).
[18] Taylor, S. J., Bogdan, R. and DeVault, M. L., *Introduction to qualitative research methods* (John Wiley & Sons, New York 1998).
[19]<http://servicios.abc.gov.ar/lainstitucion/organismos/consejogeneral/disenioscurriculares/secundaria/sexto/orientaciones/naturales/marco_naturales.pdf>, visited on February 4, 2019.