Correct answers with wrong justifications? Analysis of explanations in classical mechanics with FCI test

Carla Hernández1 and Silvia Tecpan1
1Physics Department, Universidad de Santiago, Chile
carla.hernandez.s@usach.cl

Abstract. This study presents the results of an analysis of the explanations constructed by future physics teachers about their chosen answers in a conceptual test in classical mechanics. Additionally, to selecting a multiple-choice answer in the Force Conceptual Inventory test [1] the teachers were required to give an explanation for their answers in order to determine the relationship between their level of conceptual knowledge and the quality of the explanations they gave using this knowledge. The results suggest that the 28 participants have difficulties in constructing adequate physical explanations. They suggest too that they have deficiencies in the development of their scientific thinking skills and in the use of their prior knowledge. We propose indicators for the construction of explanations in physics and we discuss the implications for the initial formation of physics teachers.

1. Introduction
Over the past two decades, various standardized tests have been developed with the intention of quantitatively measuring the level of disciplinary knowledge acquired by university and school students in a Physics course. Using a methodological design of an experimental type, several studies have compared the knowledge acquired by students after instruction, which is usually based on some methodology of the active learning, i.e., non-expository [2,3] establishing normalized gain percentages to determine the results [4].

In the case of classical mechanics content, one of the most used test has been the Force Conceptual Inventory (FCI) [1]. Although this test have served as a learning indicator in basic physics courses for engineering students [5,6], there have been few investigations into the in service teachers’ content learning and even fewer into that of pre-service physics teachers [7].

On the other hand, although the quantitative result is important to consider as evidence of the impact that the new teaching methodologies have on the learning of physics, this article focuses on the justifications that the students gave to their selected answers in the test. Below are shown the theoretical foundations that justify this study, followed by an explanation of the methodology of analysis performed, and finally the results and pedagogical implications are described.

2. Usefulness of FCI test
Internationally there is a consensus on the usefulness of the Force Concept Inventory as a tool to measure conceptual knowledge in classical mechanics [4,8]. And it is that the content associated with the concepts of force and movement is often difficult to learn for students from the rigorous point of view of science, since the interaction of each subject with the world around him generates a series of preconceived ideas on the movement of objects that are difficult to change with formal instruction.
Student’s ideas about force and movement have been extensively studied [9] and are known to condition the way they generate explanations for the phenomena of nature. The inventory includes the most common misconceptions that students present in different contexts, which have been determined through research processes related to: kinematics, impetus, active force, action/reaction pairs, concatenation of influences, other influences on motion like resistance and gravity. Several misconceptions present in the inventory are based on poorly differentiated concepts and function properly as distractors that discriminate the level of knowledge among students who have succeeded in overcoming them from those who do not.

However, there is debate on the implications of obtaining a high score in the FCI [10,11]. Studies have reported the discrepancies between the quantitative results obtained with the known Conceptual Physical Inventories, and explanations given by students to answer questions raised them [12,13].

In this way, a study by [13], where engineering students had to justify the answers chosen in the test, determined that a low percentage of students could develop the so-called "Newtonian thinking", having difficulties applying concepts consistently in their explanations. In the case of teachers, these difficulties are also present and affect their educational practices [14].

3. Importance of explanations in science
In the last two decades, it has been recognized in the sciences teaching the importance of considering the argumentation and explanations building as a central scientific activity that students must learn [15,16]. Consistently, there are numerous researches about the characteristics of student's scientific explanations [15,17,18].

In words of [16], “constructing explanations requires both a conceptual understanding of relevant theories and their application to a specific problem and the epistemic understanding of the criteria for a good explanation” (pp.32).

For its part, [19] highlight three important elements within an explanation:

a) Claim. Is a statement that responds directly to the question posed.

b) Evidence. Explaining scientific phenomena, students should be able to collect, select and use the data as evidence to support their claims.

c) Reasoning. It is the logic that allows to link the evidence with the scientific principles involved, to support their claims. Usually, students have more difficulty with the reasoning component of scientific explanations than with gathering evidence[19].

Although scientific explanations are important for science, they are often omitted in practice in the classroom [20, 21]. Moreover, when students must formulate an explanation or construct an argument, often they struggle to articulate and justify their claims [22].

However, several studies on the subject have suggested that this ability is not natural for most individuals, but is learned through practice [23]. To support students in building explanations and arguments have been used Toulmin model of argumentation (1958) [16,23,24].

Even so, student’s understanding of content may also influence their ability to effectively use the evidence in their explanations. Nevertheless, understanding the scientific content itself is not a sufficient condition for properly reasoning a scientific explanation [22].

4. Methodology
A qualitative study was carried out with exploratory scope and survey design. The instrument used was the Force Conceptual Inventory test [1] and applied to 28 pre-service physics teachers who had already approved the course of classical mechanics in a Chilean state university. The revised version of the FCI of 30 items was used [10] and translated into Spanish by [25], where for each item response it was requested to write an explanation regarding the situation of the question to justify and verify whether the choice of each option was accompanied by an appropriate reasoning. This is to ensure that the student performed an exercise of analysis of the situation presented when answering each question.

Explanations built by students were analyzed by content [26] which allowed, on one side, to
identify which are the concepts most used to explain the different situations, and other, how they relate these concepts to existing evidence to generate explanations consistent with theory and/or corresponding laws [27]. Given the characteristics of the study, should be mentioned that the students were not previously instructed on how to construct an explanation or in its characteristics according to the literature.

Based on the proposal of [28] the following analysis criteria were established to determine whether an explanation was considered correct or not:

1) Use the context of the question to extract the evidence and data that, linked to theory, support a claim
2) Establishes logical relationships between concepts and phenomena involved and the evidence, consistent with theories and/or laws of physics
3) Incorporates all the theoretical concepts involved accurately
4) Considers and establishes the limitations of validity of the theory involved and/or law of physics involved
5) Uses adequate and rigorous scientific language

However, it must be considered that the students do not have additional information or evidence than those that can be deduced from the situation presented. Furthermore, the choice item response in the FCI corresponds to what [19] consider the statement or conclusion directly answer the question.

Based on the proposed criteria that were validated by construct and content [29] with a group of university physics teachers, the explanations were classified into three categories:

- Correct: Those that meet all the proposed criteria.
- Incomplete: When you meet certain criteria, regardless of how many and which are. In this case, it is considered that the explanations are not necessarily incorrect, but could be improved upon the criteria already fulfilled.
- Wrong: In case you do not meet all the criteria suggested.

5. Results
Statistical analysis of student’s performance was performed in the complete test, on a scale of 1 to 100 (Table 1).

| Statistic          | Percentage (%) |
|--------------------|----------------|
| Average            | 73             |
| Median             | 73             |
| Standard deviation | 16             |
| Mode               | 67             |
| Range              | 70             |
| Minimum            | 30             |
| Maximum            | 100            |

Analysis of the frequency of correct responses for each item were also made. Graphic 1 shows the results obtained in the application of FCI.
The bars represent student performance for each of the 30 questions. Dark bars are those with 25 or more correct answers. Of total items was established as a cut-off criterion to analyze the explanations, those where at least 25 students selected the correct option in each item, representing an 83% efficiency in answering each. This decision is based on the hypothesis that a greater number of correct answers to the item could imply a greater number of correct explanations. Applying the defined categories to the students’ explanations, we identified differences between the number of correct answers obtained with the test [quantitative result] and the number of explanations considered correct (qualitative results). Table 2 shows the contrast of these results.

**Table 2.** Classification of the explanations for each item that exceeds the cut-off criterion for the analysis.

| Item | Q1 | Q2 | Q4 | Q6 | Q7 | Q11 | Q12 | Q13 | Q28 |
|------|----|----|----|----|----|-----|-----|-----|-----|
| Frequency of correct answers to the item \([n]\) | 28 | 25 | 26 | 28 | 27 | 25 | 25 | 25 | 25 |
| Correct explanation | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| Incomplete explanation | 1 | 3 | 2 | 7 | 6 | 5 | 3 | 2 | 4 |
| Wrong explanation | 27 | 22 | 23 | 20 | 19 | 20 | 22 | 22 | 21 |

**6. Analysis and discussion**

Details of the analysis performed on the explanations literals constructed by the students for two questions of the FCI are presented below, by way of example. On one side, the question 1 [Q1] where all students (S) selected the right choice in the test and yet none of the explanations satisfies all of the proposed criteria.

Q1. *Two metal balls are the same size, but one weighs twice as much as the other. The balls are dropped from top of a two-story building at the same instant of time. The time it takes the balls to reach the ground below will be: [...]*

In this case, the only explanation classified as incomplete was:
S3: "free-fall time is not dependent on the mass when located in the same space [gravity same]"

As explanation, according to the criteria, consideration of two physical concepts mainly "time" and "mass" is appreciated, establishing a relationship of independence between them that referred to the use of a theory and an adequate reasoning. In addition, a limitation of the theory is included to mention that this is valid if it occurs "in the same space".

However, it does not meet the criteria of using appropriate scientific language, since it refers to the concept of "gravity" and not "gravitational acceleration". Moreover, no reference to the context of the question, given by the fall of two metal spheres, in particular.

The other 27 explanations given for this question were classified as wrong, because they do not meet any of the criteria. Here are some cases:
S1: “The mass is negligible”
S10: “Because they have the same size”
S11: Bodies subject to the same acceleration, Galileo movement studies
S14: The mass does not depend on the time
S27: Sum of the forces on the Y axis equals zero

In general, conceptual difficulties related to the presence of misconceptions, consistent with Section II of this article is appreciated. In the case of S14, a difficulty related to the distinction between dependent and independent variables is observed, which is repeated in other students.

Regarding the criteria, explanations do not consider the evidence given by the context of the question nor the limitations of the theory involved. Nor are relations established between concepts such as “mass”, “free-fall time objects” and “gravitational acceleration”.

Table 3 shows the diversity of concepts used by students in their explanations where "mass" is distinguished as the most used concept. Table 4 shows the relations established between some of them.

### Table 3. Frequency $[f]$ of most used concepts in the explanations for Q1

| Concepts                     | $f$ |
|------------------------------|-----|
| Mass                         | 19  |
| Time                         | 5   |
| Acceleration                 | 6   |
| Gravity                      | 3   |
| Movement                     | 2   |
| Size                         | 1   |
| Friction                     | 1   |
| Gravitational field          | 1   |
| Weight                       | 1   |
| Gravitational acceleration   | 1   |
| Velocity                     | 1   |
Table 4. Frequency [f] of most used relationship between concepts in the explanations for Q1

| Relationship between concepts | f |
|------------------------------|---|
| Mass - Time                  | 5 |
| - Acceleration/gravitational acceleration | 4 |
| - Movement                   | 1 |
| - Velocity                   | 1 |
| Gravity – Friction           | 1 |

The diversity of concepts used suggests confusion about the content since for the same question (Q1), students build various explanations. Besides, lack of rigor is distinguished in the use of vocabulary, for example, only one student mentioned "gravitational acceleration" while the rest only refers to the concept "acceleration". Similarly, the use of the concept "gravity" to refer to the “force of gravity” or to “gravitational acceleration” is confusing.

On the other hand, there is the case of Q13 where 25 students marked the correct answer option, but only one (S17) could to construct an adequate explanation for it according to the established criteria.

Q13: A boy throws a steel ball straight up. Disregarding any effects of air resistance, the force (s) acting on the ball until it returns to the ground is [are]...

The correct answer to this question, and therefore, the claim of the explanation is … a constant downward force of gravity only. When requesting an explanation for this answer, the only one classified as correct was as following:

S19: “There is no other force acting on the ball since there is no friction and the ball is within the gravitational field of the Earth that is uniform”

In this case, the explanation satisfies the criteria set because it includes the necessary concepts, uses extracted evidence from the context of the question, also considers the limitations of physical theory indicating that the phenomenon is valid just near Earth and considering the uniformity of the gravitational field. From the point of view of theory, the explanation is adequate and indicates a right use of the language when using the corresponding terms without confusion.

On the other hand, the explanations classified as incomplete were the following:

S10: “Only the gravitational field interacts upon the object”

S17: “The child gives an impulse to the ball and then only affects the weight during the fall as it is on Earth”

In the case of S10, although it alludes to the gravitational field, the language used is not appropriate since it mentions that it acts "upon" the object, which is incorrect from a theoretical point of view. Moreover, it does not refer to the context since it mentions any object and does not establish limitations of validity of the theory that supports the explanation.

Among the explanations considered incorrect for Q13 are:

S2: “Gravity only depends on the falling objects”

S3: “Only gravity acts at that moment”

S14: “Action and reaction force”

S18: “The hand strength is not maintained”

These responses generally do not mention limits of validity to the theory, not use evidence or explanation to contextualize the case of question where the object was a steel ball. Particularly in the case of S3, the use of the term "gravity" is inappropriate and should refer to the weight. In this sense,
the language is not precise and to understand the meaning of the answer requires an interpretation that is in the opinion of the reader.

On the other hand, in the S2 response, a confusion in the determination of dependent and independent variables is again evident.

For S14, the use of Newton's third law, known as "action and reaction" would be inappropriate because the laws are descriptive but explanatory, therefore, not enough to sustain an explanation [27]. Finally, in the case of S18, also to not meeting the criteria is using the concept of "hand strength" is incorrect.

Table 5 shows the concepts most used by students to explain the answers and the relationship between concepts for Q13.

| Concepts                          | f |
|----------------------------------|---|
| Force                            | 6 |
| Weight                           | 6 |
| Gravity                          | 5 |
| Movement                         | 3 |
| Gravitational Field              | 2 |
| Gravitational Acceleration       | 2 |
| Gravitation                      | 1 |
| Action and reaction force        | 1 |
| Impulse                          | 1 |
| Hand strength                    | 1 |
| Inertia                          | 1 |

Table 6. Frequency [f] of most used relationship between concepts in the explanations for Q13

| Relationship between concepts     | f |
|----------------------------------|---|
| Weigh - Force                    | 1 |
| - Impulse                        | 1 |
| - Movement                       | 1 |
| Movement - Gravitational acceleration - inertia | 1 |
| Force – Gravity                  | 1 |
As in Q1, and all the questions whose answers were analyzed, the various relationships established by students between the concepts, offer diverse ways of relating the evidence provided by the question and previously learned theory. This is indicative that they have made distinct types of logical reasoning to construct their explanations suggesting the existence or persistence of misconceptions in future teachers.

In general, the results show an inconsistency between the quantitative results obtained with the test, which indicates a level of knowledge consistent with a Newtonian thinking in students [13], and the result of the qualitative analysis carried out on the explanations which suggests important conceptual confusion.

Moreover, regarding the proposed criteria for the analysis, the lack of consideration to the context of the questions and the absence of limitations for theories and/or laws used to explain, were the most frequent absences built explanations for all questions discussed.

7. Conclusions
In this paper, results of an application of the FCI test [1] with two variants have been presented. On one side, it was applied to pre-service physics teachers in the context of a teaching course, where the objective was to learn to teach classical mechanics, and other, was requested to explain each of the answers selected in the test.

Regarding the hypothesis, it is rejected because a high score in the test has not involved, in this case, a high number of correct explanations under the proposed criteria. The results suggest that the ability to explain is incipiently developed, which can happen because few spaces in the classroom are provided for pre-service teachers learn to build their own explanations. This is indicative of weaknesses in the level of reasoning and low understanding of the content by the participants, in coherence with other studies where these factors have been positively correlated [19].

In this line, it is proposed to use conceptual inventories as tools to promote the construction of scientific explanations, requesting to justify each response to the item and later to establish criteria for analysis of explanations such as those proposed in this study. The analysis of results in this type of interventions requires mixed methods to deepen qualitative results complementary to the quantitative ones.

Finally, the implemented strategy allows orienting training actions focused on the development of skills related to reasoning and the construction of scientific explanations, which are exercised infrequently in classes. Given the specific characteristics of the sample, it is not possible to generalize the results. It is suggested to replicate the methodology with a probabilistic sample, as well as to extend this type of studies to other areas of teacher education in science to strengthen the criteria for analysis of explanations.

Acknowledgment
The authors acknowledge financing from the Universidad de Santiago de Chile, through the projects DICYT 041531HS and Basal USA 1555.

References
[1] Hestenes D, Wells M and Swackhamer G 1992 Force concept inventory Phys Teach 30 141
[2] Meltzer DE and Thornton RK 2012 Resource Letter ALIP–1: Active-Learning Instruction in Physics Am J Phys 80 478.
[3] Redish EF 2003 Teaching Physics with the Physics Suite. (Somerset, NJ: John Wiley & Sons)
[4] Hake RR 1998 Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses Am J Phys 66 64
[5] Martin-Blas T, Seidel L, Serrano-Fernandez A 2010 Enhancing Force Concept Inventory diagnostics to identify dominant misconceptions in first-year engineering physics Eur J Eng Educ. 35 597–606.
[6] McKagan SB, Perkins KK, Wieman CE. Reforming a large lecture modern physics course for engineering majors using a PER-based design. *AIP Conf. Proc.* 2007 ed L McCullough, L Hsu and P Heron, 34–7

[7] Zavala G, Alarcón H and Benegas J 2010 A professional development course with an introduction of models and modeling science eds R Lesh, PL Galbraith, CR Haines and A Hurford *Modeling Students’ Mathematical Modeling Competencies* (NY: Springer) pp 491–500.

[8] Fagen AP, Crouch CH and Mazur E 2002 Peer instruction: Results from a range of classrooms *Phys Teach* **40** 206–9

[9] McDermott LC, Shaffer PS and Somers MD 1994 Research as a guide for teaching introductory mechanics: An illustration in the context of the Atwood’s machine *Am J Phys.* **62** 46–55

[10] Hestenes D and Halloun I 1995 Interpreting the force concept inventory: A response to March 1995 critique by Huffman and Heller *Phys Teach.* **33** 502.

[11] Huffman D and Heller P 1995 What does the force concept inventory actually measure? *Phys Teach.* **33** 138–43.

[12] Wood AK, Galloway RK, Donnelly R and Hardy J 2016 Characterizing interactive engagement activities in a flipped introductory physics class. *Phys Rev Phys Educ Res.* **12** 10140

[13] Corona A 2010 Opciones Newtonianas de estudiantes no-Newtonianos, análisis de alumnos Universitarios: FCI. *Latin-American J Phys Educ.:* **4** 422–8.

[14] Fernández M, Perez R, Peña S and Mercado S 2008 Concepciones de maestros de primaria sobre el planeta Tierra y gravedad. Implicaciones en la enseñanza de la ciencia. *Rev Electrónica Investig Educ.* **10** 1–25.

[15] Driver R, Newton P, and Osborne J 2000 Establishing the norms of scientific argumentation in classrooms. *Sci Educ* **84** 287–312

[16] Sandoval WA and Millwood KA 2005 The quality of students use of evidence in written scientific explanations. *Cogn Instr.* **23** 23–55.

[17] Kelly GJ and Takao A 2002 Epistemic levels in argument: An analysis of university oceanography students’ use of evidence in writing. *Sci Educ.* **86** 314–42

[18] Toth EE, Suthers DD and Lesgold AM 2002 Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry. *Sci Educ.* **86** 264–86.

[19] McNeill KL, Lizotte DJ, Krajcik J, and Marx RW 2006 Supporting students’ construction of scientific explanations by fading scaffolds in instructional materials. *J Learn Sci.;* **15** 153–91.

[20] Kuhn D 1993 Science as argument: Implications for teaching and learning scientific thinking. *Sci Educ.;* **77** 319–37.

[21] Newton P, Driver R and Osborne J 1999 The place of argumentation in the pedagogy of school science. *Int J Sci Educ.* **21** 553–76.

[22] Sadler TD 2004 Informal reasoning regarding socioscientific issues: A critical review of research. *J Res Sci Teach.* **41** 513–36

[23] Osborne J, Erduran S, and Simon S 2004 Enhancing the quality of argumentation in school science. *J Res Sci Teach.* **41** 994–1020

[24] Jimenez-Aleixandre MP, Rodriguez AB and Duschl RA 2000 Doing the lesson “or” doing science”: Argument in high school genetics. *Sci Educ.* **84** 757–92.

[25] Macia-Barber E, Hernández M V and Menéndez J 1995 Cuestionario sobre el concepto de fuerza.

[26] Merriam SB 2009 *Qualitative Research. A guide to design and implementation* (San Francisco, CA: Jossey Bass)

[27] Abd-El-Khalick F 2012 Examining the sources for our understandings about science: Enduring conflations and critical issues in research on nature of science in science education. *Int J Sci Educ.* **34** 353–74.

[28] McNeill KL and Krajcik J 2009 Synergy between teacher practices and curricular scaffolds to support students in using domain-specific and domain-general knowledge in writing arguments to explain phenomena. *J Learn Sci.* **18** 416–60
[29] Wiersma W 1986 *Research Methods in Education: an Introduction*. (Boston, MA, EE.UU: Allyn and Bacon)