

Modifying the Thermodynamic Concept Survey: Preliminary results

Pablo Barniol¹ and Genaro Zavala^{2,3}

¹*School of Humanities and Education, Tecnológico de Monterrey, Monterrey, Mexico*

²*School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico*

³*School of Engineering, Universidad Andres Bello, Santiago, Chile*

Abstract. The Thermodynamic Concept Survey is a multiple-choice test used in the Physics Education Research community. Analyzing this survey, we detected two issues that made it difficult to evaluate students' understanding on the survey's topics: (1) numerous items present only responses of the type: "increase", "decrease" or "remains unchanged" without including the reasoning that led to these answers, and (2) several questions have design problems. Considering these two issues, we decided to undertake a research project with the objective of modifying and refining this survey. In this article, we present preliminary results of this ongoing investigation regarding those two issues. In the first part, we illustrate the modifications made in some items, describing the modifications made in four of them. In the second part, we illustrate critical design problems in some items, describing in detail a problem in one of them. The results and discussion may be useful for researchers using the test as an assessment tool.

Keywords: Conceptual test, multiple choice test, misconceptions, design problem, thermodynamics.

PACS: 01.40.Fk, 01.40.

I. INTRODUCTION

Multiple-choice tests are highly used in PER because they allow for the evaluation of large populations [1]. Ding et al. [2] note that these instruments should use statistic tests to assert their reliability and discriminatory power, and recommend performing three item-focused statistical tests: difficulty index, discrimination index, and point biserial coefficient. Several studies have identified design problems in some items of these multiple-choice tests that did not satisfy the recommended values of these item-focused statistical tests [3]. Moreover, there have been modifications to some tests, for example, standardizing the number of options to five and modifying questions with some problems in their original design [4].

The Thermodynamic Concept Survey (TCS) [5] is a 35-item multiple-choice test used for conceptual evaluation that assesses the understanding of three topics: temperature and heat transfer, ideal gas law, and first law of thermodynamics. Analyzing the survey, we detected two issues that made it difficult to evaluate students' understanding on these topics: (1) numerous items have fewer than five possible responses and present only responses of the type: *increase*, *decrease* or *remains unchanged* without including the reasoning that led to these answers, and (2) several questions have design problems. Considering these two issues, we decided to undertake a research project with the aim of modifying the TCS.

In this article, we present preliminary results of this ongoing investigation. In the first part, we address the first objective, which is to illustrate the modifications made in some items, describing the modifications made in four of them. In the second part, we address the second objective, which is to illustrate critical design problems in some items, describing in detail one of them.

II. PREVIOUS RESEARCH

Two studies have used the TCS as an assessment tool in the past. In the first, it was used to measure learning gains in a course based on interactive lecture demonstrations [6]; in the second, it was used to measure correlations with another assessment tool [7]. In addition, the test is available at PhysPort (physport.org) to be used in a generalized way.

There is a study [8] in which researchers took 9 of the 35 items of the TCS to design a new test with 15 two-tier multiple-choice questions, the first tier of content-based and the second of reasoning-based questions. It is important to identify the main differences between their study and this work to justify the need for our study. The first difference is that they focused on designing another test instead of modifying the TCS. The second is that they only converted nine of the 35 items of TCS, and they used the two-tier format, which is not widely used in PER. (Note that this conversion shows evidence of the need of a study like ours that focuses on modifying the responses of some items of the TCS). Finally, the third difference is that they did not identify the design problems that some items of the test had.

III. ITEMS' MODIFICATIONS

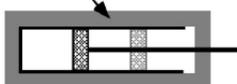
In this section, we address the first objective: to illustrate the changes made to some of the items, describing the modifications done in four of them (items 16-19). We chose these four items because they show different types of modifications. Figure 1 shows these items that evaluate the application of the first law of thermodynamics in an adiabatic compression (i.e. no exchange of heat). To construct these items the designers of the TCS used an interview question of Loverude et al. [9].

Items 16-19 of the TCS: A cylindrical pump contains one mole of an ideal gas. The piston fits tightly so that no gas escapes, and friction is negligible between the piston and the cylinder walls. The piston is quickly pressed inward so the volume of gas reduces instantly.



- 16. How does the temperature of the gas change?**
 A) Increase B) Decrease C) Remains unchanged
- 17. How does the total work done by the system (gas) change?**
 A) Increase B) Decrease C) Remains unchanged
- 18. How does the heat transferred into the system (gas) change?**
 A) Increase B) Decrease C) Remains unchanged
- 19. How does the internal energy of the gas change?**
 A) Increase B) Decrease C) Remains unchanged

Modified version of these items (new order of questions and new figure): A cylindrical pump *thermally isolated* contains one mole of an ideal gas. The piston fits tightly so that no gas escapes, and friction is negligible between the piston and the cylinder walls. The piston is quickly pressed inward so the volume of gas reduces instantly. Insulation



- 17. How is the total work done by the gas?**
 A) Positive, because the pressure of the gas increases.
 B) Negative, because the force of the gas on the piston is opposite to the movement of the piston.
 C) Zero, because the gas does not do work, the one who does work is the piston.
 D) Positive, because the force of the piston on the gas goes in the direction of the movement of the piston.
 E) Positive, because the volume of the gas decreases.
 F) Negative, because the pressure of the gas increases.
- 18. How is the heat transferred into the gas?**
 A) Positive, because the piston transfers it.
 B) Zero, because the temperature of the gas does not change.
 C) Positive, because the temperature of the gas increases.
 D) Zero, because the walls are insulating.
 E) Positive, because work is done on the gas.
- 16. Which of the following statements presents the reasoning that justifies what happens to the temperature of the gas?**
 A) The volume of the gas decreases so that the pressure increases, and then the temperature increases.
 B) The volume of the gas decreases then the temperature decreases.
 C) There is no heat transfer but as work is done on the gas as its volume decreases the temperature increases.
 D) The volume of the gas decreases and the pressure increases in such a way that the temperature does not change.
 E) Being in an insulated container there is no heat transfer so the temperature cannot change.
- 19. How does the internal energy of the gas change?**
 A) Increases, because the heat transferred is greater than the work done by the gas.
 B) Remains unchanged, because there is no heat transfer.
 C) Increases, because the pressure of the gas increases.
 D) Remains unchanged, because the temperature of the gas does not change.
 E) Increases, because although there is no transfer of heat, the piston does work on the gas.

FIG.1. Original and modified items 16-19 of the TCS.

A. Modifications to cover design issues

Analyzing the original items 16-19, we first observe a major design problem. Item 17 asks how does the total work done by the gas change, item 18 ask how does the heat transferred into the gas change, and in both questions the responses are “increase”, “decrease” and “remains unchanged”. As we can see, work and heat are not being treated as energy transferred into or out of the system (work being the transfer of mechanical energy and heat being the energy transferred thermally), but as energy of the system. The way the questions are written could lead students to acquire or strengthen some conceptual errors they might have. For example, students could think that work or heat is something that the system possesses and that something is transferable. To cover this design problem, we decided to modify the questions and responses of these items as it is usually asked and answered in physics: We modified the questions to “how is the total work done by the gas?”, “how is the heat transferred into the gas?”, and in both items we included the responses “positive”, “negative” or “zero”.

Moreover, analyzing these items we detected two features of their design that are not critical but that to a certain extent hinder the evaluation of students' understanding of an adiabatic process. The first feature is the order of the questions. To answer these questions, students should follow the next reasoning path established by Loverude et al. [9]: (1) *because the compression is carried out quickly, the process is approximately adiabatic and, therefore, $Q \approx 0$* , (2) *because the force that the piston exerts on the gas and the displacement of its point of application are parallel, the work done on the gas is positive*, (3) *therefore, according to the first law, ΔU is positive and the temperature of the ideal gas will increase*. As noted by this reasoning, students should first reflect on heat and work before answering the question about temperature change. When observing the current design of the items, we noticed that the order does not help students to follow that reasoning path since the first question regards the change of temperature. We modified the order of the questions asking first for heat and work and then for change in temperature since we had an interest in evaluating sequential reasoning of students.

The second feature of design of these items that is not critical, but could also hinder the evaluation, is the way in which the situation is described. As noted in the reasoning established by Loverude et al. [9], students must assume that the process is approximately adiabatic, because it is quickly compressed. However, students can raise questions about the insulating capabilities of the pump or the speed with which the handle was pushed inward. To cover this issue, we modified the statement and figure of the items specifying that the pump is thermally isolated. It is important to mention that Loverude et al. [9] included this specification in the statement and figure of one version of their interview question.

B. Modification of responses to include reasoning

Analyzing the original items, we also observe that these have only three possible responses, not five, as is the common number of options used in PER, and present only responses of the type: *increase*, *decrease* or *remains unchanged*, without including the reasoning that led to these answers. This latter issue makes it difficult to analyze students' understanding in detail, since students can choose the correct answer based on correct or incorrect reasoning, and choose an incorrect answer based on different reasoning. Thus, we decided to design five responses with answers and reasoning. We administered items 16-19 with the modifications described above and asking for reasoning to 96 Mexican university students finishing a thermodynamics module in an introductory level course. We categorized students' reasoning and constructed the possible responses considering these reasoning and the misconceptions identified previously [9, 10].

C. Preliminary results of the modified items

Figure 1 shows the modified items 16-19 with the modifications described above and five or more responses with answers and reasoning. We administered these items to other 90 Mexican university students finishing the same thermodynamics module. Table I shows the percentage of students selecting each of the options for these items. Next, we analyze students' performance on each of the items.

In modified item 17, we observe that 52% of students choose the correct answer with the right reasoning: "*(work done by the gas during adiabatic compression is) negative, because the force of the gas on the piston is opposite to the movement of the piston*" (option B). Moreover, we note that in this item the most frequent error (D, 19%) is to establish that "*(work is) positive, because the force of the piston on the gas goes in the direction of the movement of the piston.*"

In modified item 18, we detect that the percentage of students selecting the correct answer with correct reasoning was low (D, 23%): "*(the heat transferred into the gas is) zero, because the walls are insulating*". We also found two frequent incorrect responses: one (B, 27%), with correct answer and incorrect reasoning "*zero, because the temperature of the gas does not change*" (identified by Kautz et al. [10]); and the other (E, 33%), with incorrect answer and reasoning: "*positive, because work is done on the gas*" (in which students have a difficulty discriminating heat and work, which was identified by Loverude et al. [9]).

In modified item 16, 31% chooses the correct response (option C): "*There is no heat transfer but as work is done on the gas as its volume decreases the temperature increases*". The two most frequent errors, one with the correct answer "*increases*" (option A, 22%) and the other with the incorrect answer "*remains unchanged*" (option D, 28%) are based on an incorrect reasoning grounded on the ideal gas law. The incorrect use of this law in this situation was also identified by Loverude et al [9].

TABLE I. Percentage of students selecting each option in the modified items 16-19. (Correct answer is in boldface.)

Modified items (in new order)	%					
	A	B	C	D	E	F
17	11	52	9	19	3	6
18	7	27	10	23	33	--
16	22	5	31	28	14	--
19	7	18	33	11	31	--

In modified item 19, 31% of students selects the correct answer and reasoning (option E) based on the first law of thermodynamics; however, 33% chooses option C, which is the correct answer with an incorrect reasoning: "*(the internal energy of the gas) increases, because the pressure of the gas increases.*"

Finally, note that in this preliminary version of modified item 17 had six options because, in our administration in which we asked for reasoning, we found five frequent incorrect answers. In the final version, this item will only have five options, eliminating option E, the least frequent.

IV. CRITICAL DESIGN PROBLEMS IN ITEMS

In this section, we address the second objective: to illustrate critical design problems in some items of the TCS, describing in detail one of them (item 5). Figure 2 shows the item and item 14 of the Thermal Concept Evaluation (TCE) [11] on which item 5 is based.

<p>Item 14 of the TCE on which item 5 of the TCS is based. Jan announces that she does not like sitting on the metal chairs in the room because "they are colder than the plastic ones."</p> <p>A) Jim agrees and says: "They are colder because metal is naturally colder than plastic."</p> <p>B) Kip says: "They are not colder, they are at the same temperature."</p> <p>C) Lou says: "They are not colder, the metal ones just feel colder because they are heavier."</p> <p>D) Mai says: "They are colder because metal has less heat to lose than plastic."</p> <p>Who do you think is right?</p>
<p>Item 5 of the TCS. Jan announces that she does not like sitting on the metal chairs in the room because "when touching it, they are colder than the plastic ones." Which statement do you strongly agree with?</p> <p>A) Jim agrees and says, "The metal chairs feel colder because metal is naturally colder than plastic."</p> <p>B) Kip says, "The metal chairs are not colder because they are at the same temperature."</p> <p>C) Lou says, "The metal chairs are not colder, the metal ones just feel colder because they are heavier."</p> <p>D) Mai says, "The metal chairs are colder because metal absorbs the heat from body faster."</p>

FIG. 2. Original item 5 of the TCS and item 14 of the TCE on which item 5 of the TCS is based.

As mentioned in the Introduction, several studies [3, 4] have identified design problems in items that did not satisfy

the recommended values established by Ding et al. [2] on the item-focused tests: difficulty index, discrimination index, and point biserial coefficient. To reveal the problematic aspects of these critical items, the researchers of these studies analyzed the wording of the item and the proportion of students selecting each option.

In the TCS original article [5], the authors presented histograms with the values of these three indexes for all the items. When analyzing the indexes, we detected that item 5 had two indexes with values much lower than the recommended ones: a discriminatory index of 0.08 that is well below the recommended value of 0.3, and a point biserial coefficient of 0.11, that is also below the recommended value of 0.2. Those values show that item 5 has two characteristics: (1) absence of adequate discrimination between students with high scores and those with low scores on the entire test, and (2) a low correlation between these items' scores and entire test scores. Note that authors of the TCS did not elaborate on the problematic aspect of this critical item.

Analyzing together the wording of items 5 of the TCS and 14 of the TCE (on which the item 5 is based), as well as the proportion of students selecting each option in item 5, it is possible to reveal the critical design problem. Item 14 of the TCE (Fig. 2) asks students to choose the right answer for Jan, who points out that metal chairs are colder than plastic chairs in a room. The correct answer is option B: “*They are not colder, they are at the same temperature.*”

TCS item 5 is similar, the only difference is that it asks students to choose the right answer for Jan who points out that when she *touches the chairs*, metal chairs are colder than plastic chairs in a room (Fig. 2). In this case, according to the authors, the correct answer is option D: “*The metal chairs are colder because metal absorbs the heat from body faster.*” The point that draws much attention is that the designers of the TCS left the original correct option B, as an option that according to them is now incorrect, without explaining the reasons behind their modifications to item 14. With the above description, item 5 has now two possible answers that can be considered in certain way correct: options B and D. The existence of these two possible answers seems to explain the percentages for these options reported by the designers of the TCS for a population of 349 Australian undergraduate students

finishing a thermodynamics module: 16% chooses option B and 81% chooses option D. Moreover, the existence of these two possible correct answers seems to be a critical design problem that explains the fact that the discriminatory index and the point biserial coefficient are lower than the recommended values.

Finally, it is important to mention two points about item 5. The first point is that, in a general overview of the TCS, we can notice that item 6 (not shown) evaluates exactly the same concept as item 5 without the critical design problem described above. Due to this, our general recommendation is not to include item 5 in the modified version of the test. The second is that, strictly speaking, option D, which is, according to the designers, the correct one, has the drawback that the metal chairs *are* not colder, but *feel* colder.

V. CONCLUSION

In this article, we first illustrate the modifications made in some items of the TCS, describing the modifications made in four of them (items 16-19) to include reasoning in their responses and to solve design issues of the original items. In our ongoing investigation, we are modifying in this way items 8-15, 20, 22 and 23 of the TCS. Then, in this article, we illustrate critical design problems in items of the test, describing in detail one with item 5. It is interesting to briefly note that three others items of the TCS (items 26-28) also exhibit critical indexes according to the article in which the TCS was presented. This might be due to a design problem in the general statement (shared by the three items) which may cause students to have problems about the characteristics of the cyclic process described in this general statement.

Finally, it is important to mention that the purpose of the discussions on the design issues of items 5 and items 16-19 of the TCS, and the modifications presented in this article, is mainly to improve the assessment tools we have in the PER community. These discussions may also be useful for researchers using the test as an assessment tool. In a future article, we will present a final modified version of the TCS that will consider the issues shown in the present study and other issues we have encountered.

-
- [1] E. Redish, Am. J. Phys. **67**, 562 (1999).
 [2] L. Ding, R. Chabay, B. Sherwood & R. Beichner, Phys. Rev. ST Phys. Educ. Res. **2**, 010105 (2006).
 [3] P. Barniol, E. Campos & G. Zavala, Physics Education Research Conference 2016, 44 (2016).
 [4] P. Barniol & G. Zavala, Phys. Rev. Phys. Educ. Res. **12**, 010107 (2016).
 [5] P. Wattanakasiwich, P. Taleab, M. Sharma & I. D. Johnston, Int. J. Innovation in Sci. Math. Educ. **21**, 29 (2013).
 [6] H. Georgiou & M. Sharma, Eur. J. Phys. **36**, 1 (2015).
 [7] J. Hatami, M. Farrokhnia & M. Hassanzadeh, Commun. Comput. Inf. Sci. **635**, 169 (2016).
 [8] C. Kamcharean & P. Wattanakasiwich, Int. J. Innovation in Sci. Math. Educ. **24**, 14 (2016).
 [9] M. E. Loverude, C. H. Kautz, & P. R. L. Heron, Am. J. Phys. **70**, 137 (2002).
 [10] C. H. Kautz, P. R. L. Heron, M. E. Loverude & L. C. McDermott, Am. J. of Phys. **73**, 1055 (2005).
 [11] S. Yeo & M. Zadnik, Physics Teacher **39**, 496 (2001).