

Modifying the test of understanding graphs in kinematics

Genaro Zavala,^{1,2,*} Santa Tejada,¹ Pablo Barniol,³ and Robert J. Beichner⁴

¹*Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Monterrey 64849, Mexico*

²*Facultad de Ingeniería, Universidad Andres Bello, Santiago 7500971, Chile*

³*Escuela de Humanidades y Educación, Tecnológico de Monterrey, Monterrey 64849, Mexico*

⁴*Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA*

(Received 14 March 2017; published 31 August 2017)

In this article, we present several modifications to the Test of Understanding Graphs in Kinematics. The most significant changes are (i) the addition and removal of items to achieve parallelism in the objectives (dimensions) of the test, thus allowing comparisons of students' performance that were not possible with the original version, and (ii) changes to the distractors of some of the original items that represent the most frequent alternative conceptions. The final modified version (after an iterative process involving four administrations of test variations over two years) was administered to 471 students of an introductory university physics course at a large private university in Mexico. When analyzing the final modified version of the test it was found that the added items satisfied the statistical tests of difficulty, discriminatory power, and reliability; also, that the great majority of the modified distractors were effective in terms of their frequency selection and discriminatory power; and, that the final modified version of the test satisfied the reliability and discriminatory power criteria as well as the original test. Here, we also show the use of the new version of the test, presenting a new analysis of students' understanding not possible to do before with the original version of the test, specifically regarding the objectives and items that in the new version meet parallelisms. Finally, in the PhysPort project (physport.org), we present the final modified version of the test. It can be used by teachers and researchers to assess students' understanding of graphs in kinematics, as well as their learning about them.

DOI: [10.1103/PhysRevPhysEducRes.13.020111](https://doi.org/10.1103/PhysRevPhysEducRes.13.020111)

I. INTRODUCTION

A complete comprehension of kinematics concepts requires students to have an adequate understanding of graphs of position, velocity, and acceleration versus time in one dimension. Several researchers have investigated students' difficulties with understanding kinematics graphs [1–5]. In 1994, Beichner [2] presented the Test of Understanding of Graphs in Kinematics (TUG-K), the most widely used test to date designed to evaluate university students' understanding in this subject (see, for example, Refs. [6–9]). In that test all graphs relate to motion along one direction. “Position” refers to position along the x axis, “velocity” means the x component of velocity, and “acceleration” is similarly along the x direction.

After using the test for some time we realized that several potentially interesting conclusions about students' performance were not possible when using the test in its current form. First, we realized that we could not make conclusions

about students' understanding of related objectives (dimensions) of the test (see Table I). For example, we could not make a strict comparison between students' ability to determine the change of position from a velocity graph (objective 3) and the ability to determine the change of velocity from the acceleration graph (objective 4) since the statements of the items of these two related objectives were not similar. Any difference found between objectives 3 and 4 could be due to the differences between the statements of the items of these two objectives.

Second, we also realized we could not make conclusions about comparisons of students' ability to select the corresponding graph from a graph (objective 5), since the original test includes items that require the selection of the velocity graph from the position graph, the acceleration graph from the velocity graph, and the velocity graph from the acceleration graph, but did not include the fourth possible option: the selection of the position graph from the velocity graph. Third, we realized that we could not make complete conclusions about comparisons of students' ability to select a textual description from different graphs (objective 6) since, for example, this objective did not include an item requiring a description of motion with a velocity increasing uniformly from an acceleration graph, an important topic in an introduction course. In the same way, that happened in objective 7 (to select a graph from a textual description). Finally, we also realized that, in some items, the most

*Corresponding author.
genaro.zavala@itesm.mx

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

TABLE I. Objectives and concepts evaluated in the original version of the test. Note that all the graphs are graphed with respect to time.

Objective	Description	Concept
1	Determine the velocity from the position graph	Slope
2	Determine the acceleration from the velocity graph	
3	Determine the change of position in an interval from the velocity graph	Area under the curve
4	Determine the change of velocity in an interval from the acceleration graph	
5	Select the corresponding graph from a graph	
6	Select a textual description from a graph	Slope and/or area under the curve
7	Select a graph from a textual description	

common alternative conceptions were not included as distractors (incorrect options). (More details in Sec. V).

Therefore, we decided to make several modifications to the original version of the TUG- K (adding new items and adding new distractors in some of the original items) that would allow us to establish these types of conclusions about students' performance. The five objectives of this article are to (i) describe in detail the modifications made to the items of the test and the added distractors, (ii) analyze the difficulty, the discriminatory power, and the reliability of the added items, (iii) analyze the effectiveness of the added distractors in terms of their frequency selection and its discriminatory power, (iv) analyze the reliability and discriminatory power of the new version of the test as a whole and compare the values of the new version to the original test, and (v) illustrate the use of the new version of the test, presenting new analysis of students' understanding that was not possible with the original version of the test, specifically, in the objectives and items that in the new version meet parallelism. Sections VI and VII cover these objectives.

II. PREVIOUS RESEARCH ON THE TUG-K

Studies that have used the TUG-K can be classified into four main groups:

- (i) Studies that have used the test as a basis to design new, related tests: for example, a test of understanding of graphs in the context of calculus [6] and a test of understanding graphs in rotational kinematics [10].
- (ii) Studies that have used the test to evaluate the effectiveness of new curricular material. These

new materials incorporate a variety of approaches, including analysis of videos [7], a tutorial activity [8], and an open and interactive multimedia *e*-learning module [11].

- (iii) Studies that have used the test to evaluate the relationship between the ability to interpret kinematics graphs and other variables of the population: A study analyzing the relationship between critical thinking and gender [12].
- (iv) Studies that have used the test to investigate physics instructors' pedagogical content knowledge in the subject of graphs of kinematics; for example, there is a study analyzing this knowledge of teaching assistants [9].

In addition, it is important to mention three points about the studies that have been conducted in the area that have analyzed university students' understanding of kinematic graphs. The first is that several of these studies have referenced the TUG-K original article without using the test as an assessment tool [4,5,13–15]. The second is that no studies have presented modifications to the original version of the test. Finally, the type of studies that have used the original version of the test in the past might use the modified version in the future.

III. METHODOLOGY FOR DESIGNING THE MODIFIED VERSION

To make the modifications to the test we followed an iterative process of four administrations of different versions of the test over two years in a Mexican private university to obtain the final modified version.

In the first administration, we gave the test in its original version. In the second administration, we modified the test with new added items, along with some changes to the distractors and images of the original items. In the third administration, we made some adjustments to the modified distractors from the analysis of the data obtained in the second administration. For each distractor, we analyzed the percentage obtained and the item response curve (IRC) [16]; and for each item we analyzed the difficulty index, the discriminatory index, and the point-biserial coefficient [17]. Finally, in the fourth administration, we made further adjustments using the same procedure followed in the third administration. In this last phase, we administered the final modified version of the test (available in physport.org); this version will be referred to as “the modified version” from now on.

In all administrations, we administered the tests in Spanish. Physics instructors with high proficiency in both languages translated the original test from English to Spanish, something similarly done in other studies [18], and any differences were discussed and reconciled. One of the authors, a native English speaker, reviewed the English translation of the final modified version presented in physport.org.

All participants in this study were enrolled in the Introduction to Physics course, which is a remedial course taken by students entering the university without having scored a passing grade on a physics selection test. This course covers subjects from a traditional high school physics course: dimensional analysis, significant figures, density, trigonometry, kinematics in one dimension, and vectors. The textbook used by the students was *Introduction to University Physics* by Alarcon and Zavala [19], and in class they also worked on collaborative activities taken from the *Activities Manual* by the same authors [20].

As mention before, we had four administrations of different versions of tests to design the final modified version. To cover the objectives of this study, and in the interest of conciseness, in this contribution we only present the results obtained in the first administration (original test, $N = 248$) and in the fourth administration (final modified version, $N = 471$). In the second administration, we did most of the additions and removals of items. In some original items and all new items, we asked students to write their reasoning when choosing their answer. We also used this administration to compare original versus new items administering two versions to students in a random way. With those results, we prepared the third administration in which we almost had the final version. However, in this

third administration, we reworded some items and we implemented a few items with more than five choices to make sure that the choices represented the main students' conceptions for those items. In all cases, the tests were administered in the last session of classes and students were told that the grade they obtained would not count as part of their course grade.

IV. DESCRIPTION OF THE ORIGINAL TUG-K

The original version of the test assesses the concept of the slope and the concept of area under the curve. The test evaluates seven objectives (dimensions) and has 21 items. Table I in Sec. I describes the objectives and concepts evaluated in the original version of the test. As shown in Table I, objectives 1 and 2 are directly related to each other, since they evaluate the concept of slope. Objective 1 assesses obtaining the velocity from the position versus time graph, and objective 2 assesses obtaining the acceleration from the velocity versus time graph. On the other hand, objectives 3 and 4 are also directly related to each other since they evaluate the concept of area under the curve. Objective 3 assesses obtaining the change of position in a time interval from the velocity versus time graph, and objective 4 assesses obtaining the change of velocity from the acceleration versus time graph. As shown

TABLE II. Description of the 21 items of the original version grouped by each of the 7 objectives.

Objective	Item of the original version	Description
1	5	Determine the positive value of velocity in a time from the position graph
	17	Determine the negative value of velocity in a time from the position graph
	13	Determine the highest instantaneous velocity in an interval from the position graph
2	7	Determine the positive value of acceleration in a time from the velocity graph
	6	Determine the positive value of acceleration in a time from the velocity graph
	2	Determine the interval with the most negative acceleration from the velocity graph
3	18	Establish the procedure to determine the change of position in an interval from the velocity graph
	4	Determine the change of position in an interval from the velocity graph
	20	Determine the change of position in an interval from the velocity graph
4	10	Determine the smallest change in velocity in an interval from the acceleration graph
	16	Determine the change of velocity in an interval from the acceleration graph
	1	Determine the greatest change in velocity in an interval from the acceleration graph
5	11	Select the velocity graph from the position graph
	14	Select the acceleration graph from the velocity graph
	15	Select the velocity graph from the acceleration graph
6	8	Particular case evaluated: From the position graph determine that the movement of an object is as follows: it doesn't move, move backward and then stops
	3	From the position graph determine that the object moves at constant velocity
	21	From the velocity graph determine that the object moves at constant acceleration
7	9	Particular case evaluated: Identify the position graph that corresponds to a positive and constant acceleration
	12	Identify the position and velocity graphs that correspond to a constant velocity
	19	Identify the velocity and acceleration graphs that correspond to a constant non-zero acceleration

also in Table I, objectives 5, 6, and 7 evaluate in different ways the understanding of the concept of slope and/or the concept of area under the curve. Objective 5 assesses the selection of the corresponding graph from another graph. Objective 6 assesses the selection of a textual description from a graph. And, finally, objective 7 assesses the selection of a graph from a textual description. Table II shows a detailed description of the 21 items of the original version grouped by each of the seven objectives described above.

V. DESCRIPTION OF THE MODIFICATIONS

Here we cover the first objective of this article, which consists of describing in detail the modifications made to the items of the test and the added distractors.

A. Overview of the modifications

Table III shows a summary of the modifications made in the test. We made two major changes to the test: we added new items and new distractors in some of the original items. Items with new distractors were organized into three different groups: (i) items in which, besides major and minor changes in distractors, we also made some changes in the graph of the statement, (ii) items with only major and minor changes in the distractors but no change in the graph of the statement, and (iii) items with only major changes in the distractors. In Sec. VB we describe in detail the

TABLE III. Summary of changes made in the modified version of the test. Note that the original version has 21 items, that the modified version has 26 items, and that the correspondence between the items of the two versions is shown in Table IV.

Changes	Items of the modified version
New items	6, 10, 13, 17, 20, 21, 23, 25, 26
Changes in the distractors of original items	
Items with changes in the graph of the statement and major and minor changes in the distractors	5, 14, 16
Items with no changes in the graph of the statement and major and minor changes in the distractors	22
Items with only major changes in the distractors	2, 7, 11, 12, 15, 18, 19, 24
No modification	1, 3, 4, 8, 9

TABLE IV. Correspondence of the items in the two versions of the tests. Note that items 6, 10, 13, and 20 of the original version were removed and exchanged for new items and that in the modified version we added nine new items (items 6, 10, 13, 17, 20, 21, 23, 25, and 26).

Modified version	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Original version	1	2	3	4	5		7	8	9		11	12		14	15	16		17	18			19		21		

changes in the new items and in Sec. VC we describe the changes in the distractors.

As shown in Table III, in the modified version nine items were added. In addition, as will be shown next, four items of the original version were removed. The original version has 21 items (see Table II), while the modified version has 26 items. Table IV shows the correspondence of the items in the two versions. That table shows the nine added items: 6, 10, 13, 17, 20, 21, 23, 25, and 26, and the four removed items of the original version: 6, 10, 13, and 20. Note that these items do not appear in the second row of Table IV.

B. New added items and original removed items

As shown in Table III, the first major change made was to add nine new items. These items were added to cover two improvements from the original test: the first is to improve parallelism between related objectives, and the second is to improve in some objectives the assessment of the four possible steps between kinematics graphs and to improve parallelism among the items within a single objective.

In Table V we describe in detail each of the items that we added and the specific reasons that led us to add them. A very important point to mention here is that the modified version does not evaluate more objectives than those of the original version; it evaluates the same seven objectives (see Table I) but tries to achieve a parallelism between objectives and items more systematically.

As shown in Table V, one item was modified in each of the objectives 1 to 4 and a total of five items were added in objectives 5, 6, and 7 (one, two, and two, respectively). The four items from objectives 1 to 4 were modified to achieve parallelism between the items of the related objectives (objectives 1 and 2 and objectives 3 and 4). The item added in objective 5 evaluates the four possible steps among kinematics graphs and thus, achieves parallelism among the items within this objective, and; finally, the four added items in objectives 6 and 7 were included to achieve parallelism among the items within these objectives.

We also removed four items. The original item 6 (determine the positive value of acceleration in a time from the velocity graph) was removed because, within objective 2, this item evaluated the same concept as item 7. The original item 10 (determine the smallest change in velocity in an interval from the acceleration graph) was removed since we decided that, in order to have a small number of items for objective 4 (and objective 3), the three items that evaluated objective 4 (objective 3) were the following: establish the procedure to determine the change

TABLE V. Description of the new added items and the original items removed of each objective, and the motivation behind these changes. Note that the modified version assesses the same seven objectives as the original version (see Table I).

Objective	Added item in the modified version	Description of the new added items and the original items removed	Motivation behind these changes
1	13	The original item 13 was replaced by an item parallel to item 2 of objective 2.	Achieve parallelism between the items of these two related objectives
2	6	The original item 6 was replaced by an item parallel to item 17 of objective 1.	
3	23	The original item 20 was replaced by an item parallel to item 1 of objective 4.	Achieve parallelism between the items of these two related objectives
4	10	The original item 10 was replaced by an item parallel to item 18 of objective 3.	
5	21	An item that requests the selection of the graph of position from the graph of velocity was added.	Evaluate the four possible steps among kinematics graphs and thus achieve parallelism among the items within this objective.
6	17	An item that asks for the description of motion with position increasing uniformly from the graph of velocity was added.	The two original items ask: (i) for the description of motion with constant velocity from the graph of position, and (ii) for the description of motion with constant acceleration from the graph of velocity. Both items were added to achieve parallelism among the items contained within this objective.
	25	An item that asks for the description of motion with velocity increasing uniformly from the graph of acceleration was added.	
7	26	An item that asks for the graphs that correspond to a motion with velocity increasing uniformly was added.	The two original items ask for the graphs that correspond to a motion with: (i) constant velocity, and (ii) constant acceleration. Both items were added to achieve parallelism among the items contained within this objective.
	20	An item that asks for the graphs that correspond to a motion with acceleration increasing uniformly was added.	

of velocity (the change in position) in an interval from the acceleration graph (velocity graph), determine the change of velocity (change in position) in an interval from the acceleration graph (velocity graph), and determine the greatest change in velocity (the change in position) in an interval from the acceleration graph (velocity graph). The original item 13 was removed since it asked for an interval in which the object velocity was the greatest from a position vs time graph. We decided to have two parallel items in objective 1 (new item 13) and objective 2 (item 2) in which both items asked the most negative derivative (of position or of velocity, respectively) instead of having one in which the result was positive (in the original item 13) and the other in which the result was negative (item 2). Finally, item 20 was removed since the item asked for the change in position from a velocity graph when the velocity was uniform. In the same objective, there was item 4 that evaluated the same concept but with a uniformly changing velocity graph. We decided that the concept of calculating the change in position from a velocity graph was better evaluated when the velocity changes.

Table VI shows a complete description of the 26 items of the modified version of the test grouped in each of the 7 objectives. Note the parallelism between the items of the related objectives (objectives 1 and 2 and objectives 3 and

4). Note also the twelve pairs of parallel items in the modified version of the test: 5 and 7, 18 and 6, 13 and 2 in objectives 1 and 2; 19 and 10, 4 and 16, 23 and 1 in objectives 3 and 4; 11 and 14, 21 and 15 in objective 5; 3 and 24, 17 and 25 in objective 6; 12 and 22, 26 and 20 in objective 7.

C. Distractor changes in some of the original items

In Table III we classify the four groups of items in which we made changes in the distractors. As mentioned in Sec. I, these modifications were made to improve the original test, since in some items the most common alternative conceptions were not distractors.

In Tables VII–IX, we describe in detail the changes made to each of these four groups of items and the specific reasons that led us to make these alterations. Table VII shows the modifications in items in which, besides major and minor changes in distractors, there are changes in the graph of the statement. In that table, the major changes in the distractors were made because we wanted to modify the distractor to fulfill parallelism with another item and/or we found a frequent error (with a percentage of selection higher than 10%) in the second or third administrations that made us decide to include that distractor. Table VIII shows

TABLE VI. Description of the items of the modified version grouped in each of the seven objectives. The descriptions in italics correspond to the nine new added items.

Objective	Item of the modified version	Description
1	5	Determine the positive value of velocity in a time from the position graph
	18	Determine the negative value of velocity in a time from the position graph
	13	<i>New item: Determine the interval with most negative velocity from the position graph</i>
2	7	Determine the positive value of acceleration in a time from the velocity graph
	6	<i>New item: Determine the negative value of acceleration in a time from the velocity graph</i>
	2	Determine the interval with most negative acceleration from the velocity graph
3	19	Establish the procedure to determine the change of position in an interval from the velocity graph
	4	Determine the change of position in an interval from the velocity graph
	23	<i>New item: Determine the greatest change in position in an interval from the velocity graph</i>
4	10	<i>New item: Establish the procedure to determine the change of velocity in an interval from the acceleration graph</i>
	16	Determine the change of velocity in an interval from the acceleration graph
	1	Determine the greatest change in velocity in an interval from the acceleration graph
5	11	Select the velocity graph from the position graph
	14	Select the acceleration graph from the velocity graph
	15	Select the velocity graph from the acceleration graph
	21	<i>New item: Select the position graph from the velocity graph</i>
6	8	Particular case evaluated: From the position graph determine that the movement of an object is as follows: it doesn't move, move backward and then stops
	3	From the position graph determine that the object moves at constant velocity
	24	From the velocity graph determine that the object moves at constant acceleration
	17	<i>New item: From the velocity graph determine that the object increases its position uniformly</i>
	25	<i>New item: From the acceleration graph determine that the object increases its velocity uniformly</i>
7	9	Particular case evaluated: Identify the position graph that corresponds to a positive and constant acceleration
	12	Identify the position and velocity graphs that correspond to a constant velocity
	22	Identify the velocity and acceleration graphs that correspond to a constant nonzero acceleration
	26	<i>New item: Identify the velocity and acceleration graphs that correspond to a velocity that increases uniformly</i>
	20	<i>New item: Identify the acceleration graph that corresponds to an acceleration that increases uniformly</i>

the modifications in the items with only major and minor changes in the distractors but no change in the graph of the statement. Table IX shows the modification in items with only major changes in the distractors. Note that, for this article, a major change in a distractor is a change in the alternative conception evaluated in the distractor. Instead, a minor change does not correspond to a different evaluated alternative conception but slight modifications such as a variation in the graph of the statement, rewording, or the order of the distractors.

VI. ANALYSIS OF THE MODIFIED TEST

A. Difficulty, discriminatory power, and reliability of the added items

Here we cover the second objective of this article, that is, to analyze the difficulty, the discriminatory power, and the reliability of the added items. As noted above, nine items were added in the new version of the test: 6, 10, 13, 17, 20,

21, 23, 25 and 26. In Sec. VB, we established the reasons why these items were added. In this section, we perform the three statistical evaluations of the added items recommended by Ding *et al.* [17]: (i) item difficulty index, (ii) item discriminatory index, and (iii) item point-biserial coefficient.

The item difficulty index (P) is a measure of the difficulty of a single test question, the item discriminatory index (D) is a measure of the discriminatory power of each item on a test, and the item point-biserial coefficient (r_{pbs}) (sometimes referred to as the reliability index for each item) is a measure of consistency of a single test item with the whole test [17]. Widely adopted criteria, used by Ding *et al.* [17], suggest that the difficulty index should be between 0.3 and 0.9, the discriminatory index should be above 0.3, and the point-biserial coefficient should be above 0.2.

Table X shows the indexes and coefficients obtained for all items for the new version of the test. Here we discuss

TABLE VII. Description of the modifications and the reasons behind these modifications to the items with changes in the graph of the statement and major and minor changes in the distractors (items 5, 14, and 16; see Table III). Note that in this group of items five major changes in the distractors were made.

Item of the modified version	Description of the item	Change in the graph	Major changes in the distractors	Minor changes in the distractors
5	Determine the positive value of velocity in a time from the position graph	The new graph does not pass through the origin (unlike the original version) in order to discriminate between the correct answer and an error that occurs by computing the position divided by time	Option D: Position divided by time (a frequent error). Option A: Calculation of a “slope” by counting squares (parallelism to option B of item 7).	Option E: The value of the position at the time. Option B: Random value (Both modifications due to change in graph).
14	Select the acceleration graph from the velocity graph	The new graph has sections with different slope absolute values (not the same as the original) to discriminate between frequent errors	Option D: Parallel distractor to distractor C of the original item 11. Option E: Parallel distractor to distractor B of the original item 11.	Option A: Same model as the original option A. Option C: Same model as the original option D. (Both modifications due to change in graph).
16	Determine the change of velocity in an interval from the acceleration Graph	The new graph has different values of acceleration and time at the end of the interval (not the same as the original) to discriminate between frequent errors	Option E: Multiplication of acceleration by time (a frequent error and parallelism with option E of the original item 4).	Option A: Time divided by acceleration. Option B: Acceleration divided by time. Option C: The value of the slope. Option D: The value of time. (The four modifications due to change in graph).

only the nine added items (items 6, 10, 13, 17, 20, 21, 23, 25, and 26). In Sec. VI C, we compare the results to those of the original test and we analyze the rest of the items. Note that these values were obtained with data of the fourth administration’s population who answered the last modified version of the test (see Sec. III). As seen in Table X, all the indexes and coefficients of the nine added items satisfy the recommended values. We can conclude that the newly added items satisfied the statistical tests evaluations recommended by physics education researchers and have satisfactory difficulty level and discriminatory power, and are reliable (in the technical sense) and consistent.

B. Effectiveness of the modified distractors

In Sec. V C we described the distractors added, clustered into three different groups of items: (i) items with changes

in the graph of the statement and major and minor changes in the distractors (Table VII), (ii) items with only major and minor changes in the distractors (Table VIII), and (iii) items with only major changes in the distractors (Table IX). Here, we cover the third objective of this article, which is to analyze the effectiveness of the distractors with major changes. The effectiveness of these distractors is measured in two ways: in terms of their frequency selection, by analyzing the percentage of students selecting the distractor as recommended by Suen [21], and in terms of its discriminatory power, by analyzing the IRC of the distractor as recommended by Morris *et al.* [16].

As shown in Tables VII, VIII, and IX, we added a total of 18 distractors as a major change. In the first group we added five distractors (5D, 5A, 14D, 14E, 16E; Table VII), in the second group one distractor (22E; Table VIII), and in the third group 12 distractors (2A, 7C, 11E, 12C, 12D, 12E,

TABLE VIII. Description of the modifications and the reasons behind these modifications in the items with major and minor changes in the distractors (item 22; see Table III). Note that in the group of items one major change in the distractors was made.

Item of the modified version	Description of the item	Major changes in the distractors	Minor changes in the distractors
22	Identify the velocity and acceleration graphs that correspond to a constant non-zero acceleration	Option E: To choose the correct graph of acceleration but to choose an incorrect graph of velocity that corresponds to a constant velocity (parallelism to option D of added item 12 in the modified version).	Option B: The original option E changes to the position of option B.

TABLE IX. Description of the modifications and the reasons behind these modifications in the items with major changes in the distractors (items 2, 7, 11, 12, 15, 18, 19 & 24; see Table III). Note that in this group of items, twelve major changes in the distractors were made.

Item of the modified version	Description of the item	Major changes in the distractors
2	Determine the interval with the most negative acceleration from the velocity graph	Option A: Most positive acceleration (parallelism to option C of added item 13 in the modified version).
7	Determine the positive value of acceleration in a time from the velocity graph	Option C: Velocity divided by time (a frequent error).
11	Select the velocity graph from the position graph	Option E: Change in the graph of the distractor. The general characteristics of the graph are kept but the relationship between the absolute values of the slopes in specific sections are opposite to the original distractor (two reasons behind this change: to have a distractor with a correct relationship between the absolute values of the slopes in specific sections and because of parallelism to the new option C of item 14).
12	Identify the position and velocity graphs that correspond to a constant velocity	Option C: Incomplete answer (only the graph of velocity) Option D: To choose the correct graph of velocity and include a graph of acceleration that corresponds to constant acceleration. Option E: To choose the two correct graphs but including a graph of acceleration that corresponds to constant acceleration. (The three options are frequent errors).
15	Select the velocity graph from the acceleration graph	Option C: Graph following the rule “if the acceleration is equal to zero, then the velocity is also equal to zero” Option D: A graph that is opposite to the correct one: when speed should be constant, it is variable and when it should be variable, it is constant. (Both options are frequent errors).
18	Determine the negative value of velocity in a time from the position graph	Option D: Position divided by time with negative sign (a frequent error)
19	Establish the procedure to determine the change of position in an interval from the velocity graph	Option D: Multiplication of velocity by time (a frequent error). Option E: Time divided by velocity (to avoid the original answer: “Not enough information to answer”)
24	Determine that the object moves at constant acceleration from the velocity graph	Option C: “The position of the object decreases uniformly” (a frequent error).

TABLE X. Difficulty index (P), discriminatory index (D), and point-biserial coefficient (r_{pbs}) for each of the 26 items of the modified version.

	1	2	3	4	5	6	7	8	9	10	11	12	13
P	0.11	0.62	0.58	0.39	0.39	0.46	0.40	0.51	0.46	0.55	0.45	0.48	0.51
D	0.37	0.51	0.63	0.61	0.52	0.58	0.48	0.72	0.73	0.58	0.66	0.80	0.76
r_{pbs}	0.47	0.42	0.51	0.48	0.43	0.46	0.42	0.58	0.57	0.40	0.51	0.63	0.60
	14	15	16	17	18	19	20	21	22	23	24	25	26
P	0.54	0.46	0.33	0.70	0.41	0.67	0.52	0.50	0.48	0.33	0.32	0.54	0.51
D	0.71	0.64	0.51	0.71	0.54	0.56	0.43	0.60	0.73	0.52	0.71	0.78	0.83
r_{pbs}	0.55	0.50	0.42	0.59	0.43	0.45	0.32	0.44	0.58	0.43	0.60	0.61	0.65

TABLE XI. Results obtained with items of the modified version organized by the seven objectives. Correct answers in bold. Note that the 18 added distractors with important changes are underlined.

Objective	Item of the modified version	Short description of the item	%					
			A	B	C	D	E	N
1	5	Positive velocity in a time	<u>5%</u>	1%	39%	<u>42%</u>	12%	1%
	18	Negative velocity in a time	41%	13%	11%	<u>23%</u>	11%	1%
	13	Most negative velocity in an interval	51%	14%	2%	<u>31%</u>	1%	0%
2	7	Positive acceleration in a time	40%	19%	<u>24%</u>	13%	3%	1%
	6	Negative acceleration in a time	18%	46%	11%	12%	11%	2%
	2	Most negative acceleration in an interval	<u>2%</u>	14%	20%	1%	62%	0%
3	19	Procedure: Change of position in an interval	2%	67%	13%	<u>15%</u>	<u>3%</u>	0%
	4	Change of position in an interval	4%	11%	15%	39%	28%	2%
	23	Greatest change of position in an interval	3%	33%	18%	19%	26%	1%
4	10	Procedure: Change of velocity in an interval	55%	28%	4%	6%	7%	1%
	16	Change of velocity in an interval	12%	27%	12%	33%	<u>15%</u>	0%
	1	Greatest change of velocity in an interval	11%	10%	1%	54%	<u>23%</u>	1%
5	11	Velocity graph from the position graph	15%	24%	13%	45%	<u>3%</u>	0%
	14	Acceleration graph from the velocity graph	10%	54%	7%	<u>16%</u>	<u>12%</u>	1%
	21	Position graph from the velocity graph	8%	50%	6%	19%	17%	0%
	15	Velocity graph from the acceleration graph	46%	6%	<u>17%</u>	<u>11%</u>	21%	0%
6	8	Position graph: Determine particular movement	10%	14%	14%	51%	11%	0%
	3	Position graph: Determine movement at constant velocity	9%	2%	29%	58%	3%	0%
	24	Velocity graph: Determine movement at constant acceleration	32%	42%	<u>15%</u>	8%	1%	3%
	17	Velocity graph: Determine movement with its position increasing uniformly	70%	10%	3%	14%	4%	0%
	25	Acceleration graph: Determine movement with its velocity increasing uniformly	5%	8%	54%	23%	6%	4%
7	9	Positive and constant acceleration: Identify position graph	6%	32%	10%	6%	46%	0%
	12	Constant velocity: Identify position and velocity graphs	7%	48%	<u>18%</u>	<u>15%</u>	<u>11%</u>	0%
	22	Constant acceleration: Identify velocity and acceleration graphs	12%	20%	48%	11%	<u>8%</u>	0%
	26	Velocity that increases uniformly: Identify velocity and acceleration graphs	6%	51%	3%	19%	15%	5%
	20	Acceleration that increases uniformly: Identify acceleration graph	2%	15%	52%	21%	9%	0%

15C, 15D, 18D, 19D, 19E, 24C; Table IX). Table XI shows the results.

Analyzing the obtained percentages in the 18 added distractors in Table XI, we observe two trends that show the effectiveness of these distractors in terms of their selection frequency. The first is that 6 out of 18 (i.e., one-third) were the distractors with the highest percentage within each of its items. These distractors are 5D, 7C, 12C, 14D, 18D, and 19D. The second trend is that 13 out of 18 (i.e., more than two-thirds) had a percentage of selection equal or greater than 10%. These distractors are the six distractors mentioned above and the following seven distractors: 12D, 12E, 14E, 15C, 15D, 16E, and 24C. Conversely, five out of 18 distractors (5A, 2A, 11E, 19E and 22E) did not have a percentage within this range. As mentioned above in Tables VII, VIII, and IX, four of these distractors were added due to parallelism reasons (5A, 2A, 11E, 22E) and one of these distractors was included to avoid the original answer “Not enough information to answer” (19E). These

two trends show that the great majority of the added distractors as major changes are effective in terms of their frequency selection and those that are not that attractive as alternative conceptions are important to have in a test in which comparisons can be made.

C. Comparison with the original test

This section covers the fourth objective of this article: to analyze the reliability and discriminatory power of the new version of the test as a whole and compare the values of the new version with the original test. Therefore, we calculated for both versions the five statistical tests suggested by Ding *et al.* [17]. The three first measures focus on individual test items: the item difficulty index, the item discriminatory index, and the item point biserial. The other two measures focus on the test as a whole, the Kuder-Richardson reliability test, and Ferguson’s delta test. We present a summary of the five statistical tests for both versions in Table XII. Note that these values were obtained with data

TABLE XII. Summary of the results of the five statistical tests suggested by Ding *et al.* [17] for the original and modified versions of the test.

Test statistics	Desired values	Original version	Modified version
Difficulty index	[0.3, 0.9]	Average: 0.51	Average: 0.47
Discriminatory index (25%-25% method)	≥ 0.3	Average: 0.56	Average: 0.62
Point-biserial coefficient	≥ 0.2	Average: 0.74	Average: 0.50
Kuder-Richardson reliability index	≥ 0.7 for group measures	0.81	0.88
Ferguson's delta	> 0.9	0.98	0.99

from the first and fourth administration's populations (see Sec. III).

We can point out two important conclusions that Table XII shows. The first is that the modified version satisfied all the criteria suggested by Ding *et al.* [17]. We can, therefore, conclude that the modified version is a reliable test with satisfactory discriminatory power. The second is that even though the tests were administered to different students with a very comparable background, the results of the original test compared to the modified version are similar, with slightly better averages for the modified version, except for the difficulty index and the point-biserial coefficient. A smaller difficulty index for the modified version probably indicates that the distractors are doing a better job than those of the original version. The drop in the average point-biserial coefficient is harder to explain, but might indicate that the various difficulties students have with interpreting kinematics graphs (which this new test version is better at picking out) are not due to a single set of interrelated misunderstandings. (Recall that the point-biserial coefficient is the correlation between an individual item's correctness and the whole test score.) We hesitate to speculate further on this, but note that the finding clearly indicates the need for further research, perhaps through a combination of factor analysis and interviewing.

In addition, we also observe similar good values in the individual indexes of the items of both versions. As we know, the average difficulty and discriminatory indexes, and the average point-biserial coefficient are calculated by averaging the indexes and coefficients of all the items. In both versions, we observed that most of the items met the difficulty index values recommended by Ding *et al.* [17]: values in the range [0.3, 0.9]. In the original version only two items (items 1 and 16) had values below 0.3, and in the last modified version, only one item (item 1, see Table X) had a value below 0.3. In addition, we observed that all the items of the original and modified version met the discrimination index criterion (≥ 0.3) and the coefficient point biserial criterion (≥ 0.2) recommended by Ding *et al.* [17].

The important thing to conclude, after the above analysis, is that the modified version of the test covers the desired improvements of the original version (which

were described above) and satisfied the reliability and discriminatory power criteria as effectively as the original test.

VII. USE OF THE MODIFIED VERSION: ANALYSIS OF STUDENTS' UNDERSTANDING

This section covers the fifth objective of this study: to illustrate the use of the new version of the test, presenting a new analysis of students' understanding that is possible to perform with the modified version of the test, specifically with the objectives and items that meet parallelism. As mentioned above, in the modified version of the test we added items to achieve parallelism between related objectives (objectives 1 and 2 and objectives 3 and 4) and to achieve parallelism among the items within some objectives (objectives 5, 6, and 7). Table XIII shows the percentages of the correct answer of the items of the modified version.

We present an analysis of (i) the overall performance of students in the test, (ii) degree of items' difficulty, (iii) trends in students' performance in the items of the related objectives, and (iv) differences in students' performance on related items of all objectives.

A. Overall performance of students

The average of the scores of the modified version, from the sample of 471 students of the fourth administration (see Sec. III), is 12.25 out of 26 possible points. This average, expressed in percentage of the total possible points, is 47%, which corresponds to the average difficulty index value shown in Table XII (0.47). The distribution of scores was significantly non-normal [Kolmogorov-Smirnov, $D(471) = 0.094$, $p < 0.01$; Shapiro-Wilk test, $W(471) = 0.963$, $p < 0.01$]. The skewness of the distribution of scores is 0.251 (SE = 0.113), indicating a pile-up to the right, and the kurtosis of the distribution is -0.969 (SE = 0.225), indicating a flatter than normal distribution. The positive skew indicates that the test was difficult for the students. For this type of distribution, it is more useful to use quartiles as measures of spread. The median of the distribution is 12, the bottom quartile (Q1) is 7, and the top quartile (Q3) is 17, so the interquartile range is 10. It is interesting to note that the students at the median (12) had difficulty answering 14 questions (out of 26) correctly.

TABLE XIII. Percentage of the correct answer of the items and the average of each objective organized by objectives. The items of the four first objectives are grouped by the related objectives 1 and 2 and related objectives 3 and 4.

Description	Item	Objective	Item	Objective
Description of the items that evaluate the slope	Item	Objective 1	Item	Objective 2
Determine positive value	5	39%	7	40%
Determine negative value	18	41%	6	46%
Identify the interval in which the slope is most negative	13	51%	2	62%
Average percentage of the objective		44%		49%
Description of the items that evaluate the area under the curve	Item	Objective 3	Item	Objective 4
Establish the procedure to determine the change of a variable	19	67%	10	55%
Determine the change of a variable	4	39%	16	33%
Identify the greatest change in a variable	23	33%	1	11%
Average percentage of the objective		46%		33%
Description of the items requesting to select the corresponding graph from a graph	Item	Objective 5	Related to	
Determine the corresponding graph that implies a step forward	11	45%	Objective 1	
	14	54%	Objective 2	
Determine the corresponding graph that implies a step backward	21	50%	Objective 3	
	15	46%	Objective 4	
Average percentage of the objective			49%	
Description of the items requesting to select a textual description from a graph	Item	Objective 6		
Graph of position: Determine particular movement	8	51%		
Graph of position: Determine movement at constant velocity	3	58%		
Graph of velocity: Determine movement at constant acceleration	24	32%		
Graph of velocity: Determine movement with its position increasing uniformly	17	70%		
Graph of acceleration: Determine velocity with its velocity increasing uniformly	25	54%		
Average percentage of the objective		53%		
Description of the items requesting to select a graph from a textual description	Item	Objective 7		
Positive and constant acceleration: Identify graph of position	9	46%		
Constant velocity: Identify position and velocity graphs	12	48%		
Constant acceleration: Identify velocity and acceleration graphs	22	48%		
Velocity that increases uniformly: Identify velocity and acceleration graphs	26	51%		
Acceleration that increases uniformly: Identify acceleration graphs	20	52%		
Average percentage of the objective		49%		

B. Items

We classified items as “high” difficulty level if they had a proportion of correct answers that was equal to or less than 35%, as “medium” difficulty level if they had a correct proportion of 35% to 55%, and as “low” difficulty level if their proportion of correct answers was equal to or greater than 55%.

As shown in Table XIII, the four items considered to have a high difficulty level in order of decreasing difficulty are: Item 1 (11%) requesting the determination of the largest change of velocity in an interval of the acceleration graph (objective 4); item 23 (33%) selecting the largest change of position in an interval of the velocity graph (objective 3); item 16 (33%) asking for the change of velocity over an interval of the acceleration graph (objective 4); and item 24 (32%) requesting the selection of an object moves at constant acceleration from the velocity graph (objective 6). Two points are interesting to note regarding the most difficult items. The first is that three of the four items are from the related objectives 3 and 4 and that two of these items belong

to the objectives evaluating the identification of the largest change of a variable (items 1 and 23).

The four items with a low difficulty level are item 17 (70%) asking for the object that increases its position uniformly from the velocity graph (objective 6), item 19 (67%) asking for the procedure to determine the change of position in an interval of the velocity graph (objective 3), item 2 (62%) determining the interval with the most negative acceleration of the velocity graph (objective 2), and item 3 (58%) determining which object moves at constant velocity from the position graph (objective 6). An important point to note in these items with low difficulty level is that two of these four items are from objective 6. Finally, the rest of the items (i.e., the great majority of items) have a medium difficulty level with a correct answer proportion between 35% and 55%.

C. Related objectives and related items

The modified test allows us to compare better the students’ performances in the parallel items of the related

objectives 1 and 2 and related objectives 3 and 4. Qualitatively comparing the average percentage of objective 1 (44%) to the average percentage of objective 2 (49%), we observe that these percentages are similar. This shows that, in general, students have similar difficulties with items requesting the determination of the velocity at an instant of time from the position graph (objective 1) as in items asking for the acceleration at a time from the velocity graph (objective 2).

On the other hand, qualitatively comparing the average percentages of the related objectives 3 and 4, we observe a greater difference between the average percentages of the related objectives 3 and 4 (33% in objective 4 vs 46% in objective 3). This seems to show that students have more difficulties with items requesting the change of velocity over an interval from the acceleration graph (objective 4) than with items asking for the change of position during an interval from the velocity graph (objective 3).

Also, in the related objectives 1 and 2 and related objectives 3 and 4, we observe interesting trends in students' performance when we analyze the correct answer of the items (Table XIII). In the related objectives 1 and 2 that evaluate the understanding of the slope concept either as velocity or acceleration, we observe that there are no differences on the items asking for a positive value of a slope (items 5 and 7) and one item involving a negative value of a slope (item 18). The other item involving a negative value of a slope (item 6), the results indicate that it is less difficult, something that might be due to the perfect interval showed; i.e., exact interval from 60 to 120 sec, easiness of reading the vertical axis on those times, and the fact that the question refers to the middle of the interval (90 sec). In the same objectives 1 and 2, items 13 and 2 seem to be less difficult. Those items ask students to identify the interval in which the slope is the most negative, which are conceptual questions.

On the other hand, in the related objectives that evaluate the understanding of the area under the curve concept as either change of position or change of velocity, we note that the items that are the most difficult for the students are the items requesting identification of the greatest change in a variable (items 23 and 1); that the second most difficult are the items seeking the value of the change of a variable (items 4 and 16); and that the least difficult items are the items asking for the procedure needed to determine the change of a variable (items 19 and 10).

The information presented in Table XIII also allows us to analyze the differences in students' performance on related items that assess the same way the same mathematical concept but in different kinematic variables. For example, in objectives 1 and 2, items 5 and 7 evaluate finding the positive value of a slope. The difference is that item 5 requests the finding of the velocity in a position graph while item 7 asks for the acceleration in a velocity graph.

To compare students' answers and detect significant differences, we used the chi-square test following the

procedure described by Sheskin [22]. Observing Table XIII we can identify the twelve pairs of parallel items mentioned above in the modified version of the test: 5 and 7, 18, and 6, 13, and 2 in objectives 1 and 2; 19 and 10, 4 and 16, 23 and 1 in objectives 3 and 4; 11 and 14, 21 and 15 in objective 5; 3 and 24, 17 and 25 in objective 6; 12 and 22, 26 and 20 in objective 7.

Following the procedure described by Sheskin [22], we found that in six of the 12 pairs of the related items there was a significant difference in the selection of the correct answer (with $p < 0.01$ because of the Bonferroni correction). Next, we note these six pairs of items where we found a significant difference:

- In objectives 1 and 2, items 13 and 2, respectively.
- In objectives 3 and 4, items 19 and 10, respectively.
- In objectives 3 and 4, items 23 and 1, respectively.
- In objective 5, items 11 and 14.
- In objective 6, items 3 and 24.
- In objective 6, items 17 and 25.

In an overall analysis of these differences, it is noteworthy that we observe them in six of the seven objectives of the test (objectives 1–6). These significant differences suggest that the kinematic variable requested in the items has an effect on the students' selection of the correct answer, which has a great instructional importance. In future studies focused on the analysis of students' understanding, we will explore these differences and investigate if they persist in students that finish a more advanced course, such as a mechanics course based on calculus.

VIII. DISCUSSION

The TUG-K is one of the earliest and most reliable tests in use for physics education research studies [2]. The test was structured with seven objectives clearly defined. Using the test for some time, we realized from the analysis of results that some conclusions could not be reached with the test in its original form. Two examples are shown here to encourage reader discussion. One example is item 5 of the original test, which we modified. Figure 1 shows the original version and the modified version of the question.

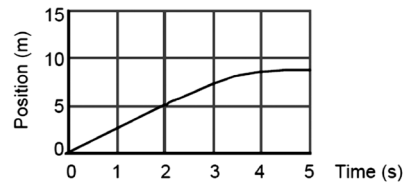
The change we want to focus on in this discussion is the graph modification. In the original version, students calculating the right answer by dividing the change in position by the time interval would obtain answer (c) 2.5 m/s. However, students having the alternative conception that the velocity is the distance (or position) divided by time would obtain the same answer (c) 2.5 m/s. In contrast, in item 5 of the modified version, these two different calculations would obtain different answers (2.5 and 5.0 m/s). Figure 2 shows students' results for these two items in the second administration with the same population, half of the students taking the original version and the other half taking the modified version, described in Sec. III.

Comparing the two graphs in Fig. 2, the results show evidence that in the original version there were both

(a) Original item

5. The velocity at the 2 second point is:

- (A) 0.4 m/s
- (B) 2.0 m/s
- (C) 2.5 m/s
- (D) 5.0 m/s
- (E) 10.0 m/s



(b) Modified item

5. The following figure shows the position versus time graph of an object. The velocity of the object at $t = 2$ s is:

- A) 0.5 m/s
- B) 8.5 m/s
- C) 2.5 m/s
- D) 5.0 m/s
- E) 10.0 m/s

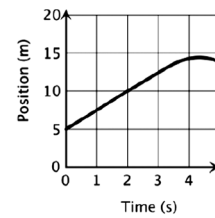


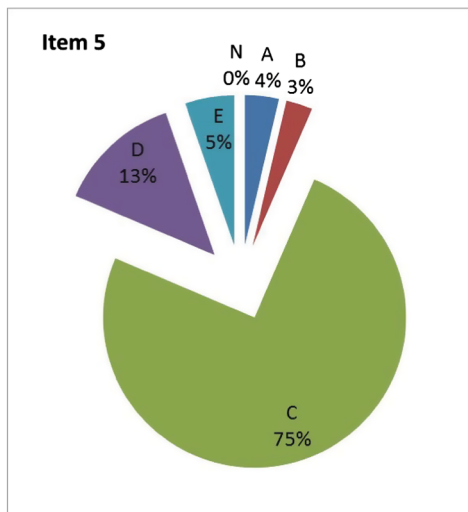
FIG. 1. Original version and modified version of item 5.

students answering option (c) 2.5 m/s because they were calculating the velocity in the right way [option (c) in the modified version’s results] and those who were calculating the velocity by dividing the position by time [option (d) in the modified version’s results].

The second example is related to the parallelism of objectives. We wanted to have a test that would tell us whether in kinematics graphs the understanding of the relation between position and velocity is the same or different from the understanding of the relation between velocity and acceleration. We know that, mathematically, the relations are the same, but students’ conceptual understanding might be different. With great effort, we evaluated

objectives 1 and 2 in the same way, and objectives 3 and 4 also in the same way. As shown in Table VI, objectives 1 and 2 have three questions each. Item 5 from objective 1 is parallel to item 7 of objective 2: both assess the understanding of calculating the slope (velocity or acceleration) in a given time from a graph (position and velocity). Item 18 from objective 1 is parallel to item 6 of objective 2: both assess the understanding of calculating the negative slope (velocity or acceleration) in a given time from a graph (position and velocity). Item 13 from objective 1 is parallel to item 2 of objective 2: both assess the understanding of finding the largest slope (velocity or acceleration) in a given time interval of a graph (position and velocity). The

(a) Original version



(b) Modified version

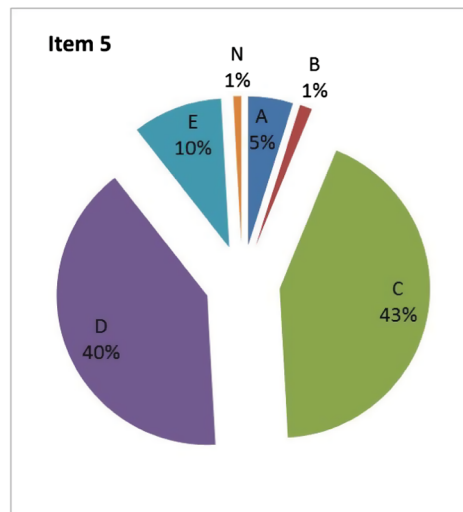


FIG. 2. Students’ results with item 5 of the original and modified version.

parallelisms in these items are the basis of the parallelism of objectives. However, we realize that there might be some differences in the parallel items that may hinder seeing actual differences in the results. We believe, nevertheless, that the results with this test are more comparable and that we can make conclusions that might help us conduct further research or to make adjustments in our teaching. Therefore, the changes to graphs, distractors, and items (some of them in addition to what the original test had) were probed in terms of the difficulty, the discriminatory power, and the reliability for individual items, but also with the whole test. We believe that the modified version of the test is a better instrument for studying student understanding of kinematics graphs.

In Sec. VII, we showed the results and some analysis that could be done with this test. Here, we discuss these in more general terms, since the objective of the paper is to present the test and show possible uses. We, on the other hand, suggest the new test be used in future research with a more extensive focus on students' understanding.

Objective 4 is more difficult for students than objective 3. Items 10 and 1 are significantly more difficult than items 19 and 23, respectively. This result seems to be consistent with what is explained in the Introduction to Physics class, the course in which this study was carried out and a typical course in introductory physics. The class, and the typical classes of this kind, usually emphasizes the relationship of the change in position in a velocity graph and less emphasis on the change in velocity in an acceleration graph. Most of the time devoted to these subjects is mainly dedicated to the first part (what objective 3 assesses). These results might indicate that understanding of the relation of the change in position in a velocity graph does not necessarily transfer to the change in velocity in an acceleration graph.

It is interesting to compare the difference that we reported between the average percentages of objectives 3 and 4 to the results reported in the article in which the TUG-K is presented. Beichner [2] mentions that objective 4 "is by far the most difficult objective." However, analyzing the original test, one can argue that the difference reported between objectives 3 and 4 could be due to the evaluated concept or to the differences between the statements of the items of these two objectives. In this version of the test, the comparison is possible since the objectives are more parallel.

The results in objectives 1 and 2 are similar. In this case, there is not much more time dedicated to the velocity in terms of the slope in a position graph than that to the acceleration in terms of the slope in a velocity graph. There is an interesting difference in one pair of related items. Item 13 from objective 1 is more difficult than item 2 from objective 2. This difference might be due to the differences in the time intervals. In the case of item 2, the time intervals in which the acceleration is negative are of the same amount and those in item 13 in which the velocity is negative are

not. It might be easier for students to compare the slope when the time intervals are the same than to do it when the time intervals are different. This result needs to be addressed in future research of students' understanding.

Another interesting result was found in objective 4 (and also in objective 3). If one poses a question asking for a procedure to calculate the change in velocity (item 10), there are many students who answer correctly (55%). However, if one poses a question requesting a calculation of the change in velocity (item 16), only 33% of students chose the right answer. It seems that providing some help to students, similar to scaffolding, offers benefits; however, removing that help (item 16) might leave only the stronger students to answer this question correctly. One would think that all of those students (33%) understand that the change in velocity is the area under the curve in an acceleration graph; however, only 11% of students answer item 1 correctly, in which the understanding of the area under the curve is crucial. Item 1 has an additional difficulty (and that is why the percentage of correct answers dropped to 11%), the question makes explicit the word maximum change in velocity. The way the question is posed could trigger evoking a derivative (rate of change) instead of area under the curve.

IX. CONCLUSIONS

The original version of the Test of Understanding of Graphs in Kinematics [2] has been a well-received assessment. However, when analyzing this test, we detected several potential improvements, especially regarding the parallelism between related objectives, the parallelism between the items of some objectives, and the representation of the most common alternative conceptions as distractors. To make those improvements, we decided to modify the test, adding new items and modifying some distractors in some of the original items that remained. When analyzing the final modified version of the test, which we are designating TUG-K 4.0, it was found that the added items satisfied the statistical tests of difficulty, discriminatory power, and reliability; that the great majority of the modified distractors were effective in terms of their frequency selection and its discriminatory power; and that the final modified version of the test satisfied the reliability and discriminatory power criteria just as effectively as the original test. We also showed here the use of the new version of the test, presenting a new analysis of students' understanding that was not possible with the original version of the test, specifically in the objectives and items that meet parallelisms in the new version.

We realize that to have a complete parallel test between objectives and within objectives, completely isomorphic questions should be designed. However, we decided not to do it since the test, on one hand would have had some more questions (which is not practical), and on the other hand, the same graphs and wording would have made the test

tedious for students and probably confuse them. We believe that this is a limiting issue; however, we also believe that the test can be used with confidence that it evaluates kinematics concepts in a structured and reliable way with the adequate distractors and a high level of parallelism.

Finally, the test and its analysis have instructional value, as they can help teachers or researchers who wish to increase students' understanding in the topic of graphs in kinematics to plan their instructional methodologies [23]. In this contribution, we presented items with low results; therefore, one of the instructional recommendations is to focus specifically on teaching the skills to solve the high difficulty level items. In particular, a recommendation is to focus specifically on teaching the area under the curve concept in the kinematic variable of change of velocity, since, at least with these students, understanding the same concept as the change in position does not necessarily transfer to understanding the concept of change in velocity.

In physport.org, we present the final modified version of the test. It can be used by teachers and researchers to assess students' understanding of, and learning about, graphs in kinematics. We request that students not be allowed to keep copies of the test or its items, since these can easily be uploaded to the web and then searched for by other students. We also remind readers that this instrument was not intended for use in high stakes testing that could impact student grades.

ACKNOWLEDGMENTS

We would like to acknowledge Juan Velarde for insightful discussions, the coordinators (in different semesters) of the course in which the test was administered, for their great help; all the lecturers who gave up class time for those test administrations, and all the students who participated in this study.

-
- [1] L. C. McDermott, M. L. Rosenquist, and E. H. van Zee, Student difficulties in connecting graphs and physics: Examples from kinematics, *Am. J. Phys.* **55**, 503 (1987).
- [2] R. Beichner, Testing student interpretation of kinematic graphs, *Am. J. Phys.* **62**, 750 (1994).
- [3] D. H. Nguyen and N. S. Rebello, Students' understanding and application of the area under the curve concept in physics problems, *Phys. Rev. ST Phys. Educ. Res.* **7**, 010112 (2011).
- [4] L. Ivanjek, A. Susac, M. Planinic, A. Andrasevic, and Z. Milin-Sipus, Student reasoning about graphs in different contexts, *Phys. Rev. Phys. Educ. Res.* **12**, 010106 (2016).
- [5] L. Bollen, M. De Cock, K. Zuza, J. Guisasola, and P. van Kampen, Generalizing a categorization of students' interpretations of linear kinematics graphs, *Phys. Rev. Phys. Educ. Res.* **12**, 010108 (2016).
- [6] N. Perez-Goytia, A. Dominguez, and G. Zavala, Understanding and interpreting calculus graphs: Refining an instrument, *AIP Conf. Proc.* **1289**, 249 (2010).
- [7] N. Chanpichai and P. Wattanakasiwich, Teaching Physics with Basketball, *AIP Conf. Proc.* **1263**, 212 (2010).
- [8] S. E. Tejada Torres and H. Alarcon, A tutorial-type activity to overcome learning difficulties in understanding graphics in kinematics, *Lat. Am. J. Phys. Educ.* **6**, 285 (2012).
- [9] A. Maries and C. Singh, Exploring one aspect of pedagogical content knowledge of teaching assistants using the test of understanding graphs in kinematics, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020120 (2013).
- [10] W. A. Turner, G. W. Ellis, and R. J. Beichner, A comparison of student misconceptions in rotational and rectilinear motion, *Proceeding at 2014 ASEE Annual Conference, Indianapolis, Indiana, 2014* (American Society for Engineering Education, Washington, DC, 2014), <https://peer.asee.org/19926>.
- [11] C. Watson and V. Brathwaite, An open and interactive multimedia e-learning module for graphing kinematics, *Proceedings of the International Conference on e-Learning* (Academic Conferences and Publishing International Limited, Reading, 2013), p. 409–415.
- [12] B. Bektasli and A. L. White, The relationships between logical thinking, gender, and kinematics graph interpretation skills, *Egitim Arastirmalari - Eurasian Journal of Educational Research* **12**, 1 (2012).
- [13] W. M. Christensen and J. R. Thompson, Investigating graphical representations of slope and derivative without a physics context, *Phys. Rev. ST Phys. Educ. Res.* **8**, 023101 (2012).
- [14] M. Planinic, Z. Milin-Sipus, H. Katic, A. Susac, and L. Ivanjek, Comparison of student understanding of line graph slope in physics and mathematics, *Int. J. Sci. Math. Educ.* **10**, 1393 (2012).
- [15] M. Planinic, L. Ivanjek, A. Susac, and Z. Milin-Sipus, Comparison of university students understanding of graphs in different contexts, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020103 (2013).
- [16] G. A. Morris, L. Branum-Martin, N. Harshman, S. D. Baker, E. Mazur, S. Dutta, T. Mzoughi, and V. McCauley, Testing the test: Item response curves and test quality, *Am. J. Phys.* **74**, 449 (2006).
- [17] L. Ding, R. Chabay, B. Sherwood, and R. Beichner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010105 (2006).
- [18] P. Wattanakasiwich, P. Taleab, M. D. Sharma, and I. D. Johnston, Development and implementation of a conceptual survey in thermodynamics, *Int. J. Innovation in Sci. Math. Educ.* **21**, 29 (2013).

- [19] H. Alarcon and G. Zavala, *Introduction to University Physics* (Trillas, Mexico, 2005).
- [20] H. Alarcon and G. Zavala, *Introduction to University Physics—Activities manual* (Trillas, Mexico, 2006).
- [21] H. K. Suen, *Principles of Test Theories* (Lawrence Erlbaum Associates, Hillsdale, NJ, 1990).
- [22] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, 4th ed. (Chapman and Hall/CRC, Boca Raton, Florida, 2007).
- [23] L. C. McDermott, Oersted Medal Lecture 2001: Physics education research—The key to student learning, *Am. J. Phys.* **69**, 1127 (2001).