



Test of Understanding Graphs in Calculus: Test of Students' Interpretation of Calculus Graphs

Angeles Dominguez

Tecnologico de Monterrey, MEXICO
Universidad Andres Bello, CHILE

Pablo Barniol

Tecnologico de Monterrey, MEXICO

Genaro Zavala

Tecnologico de Monterrey, MEXICO
Universidad Andres Bello, CHILE

Received 3 May 2017 • Revised 26 July 2017 • Accepted 13 September 2017

ABSTRACT

Studies show that students, within the context of mathematics and science, have difficulties understanding the concepts of the derivative as the slope and the concept of the antiderivative as the area under the curve. In this article, we present the Test of Understanding Graphs in Calculus (TUG-C), an assessment tool that will help to evaluate students' understanding of these two concepts by a graphical representation. Data from 144 students of introductory courses of physics and mathematics at a university was collected and analyzed. To evaluate the reliability and discriminatory power of this test, we used statistical techniques for individual items and the test as a whole, and proved that the test's results are satisfactory within the standard requirements. We present the design process in this paper and the test in the appendix. We discuss the findings of our research, students' understanding of the relations between these two concepts, using this new multiple-choice test. Finally, we outline specific recommendations. The analysis and recommendations can be used by mathematics or science education researchers, and by teachers that teach these concepts.

Keywords: concept of the derivative, concept of the antiderivative, context of calculus, reliable evaluation instrument, students' graph understanding

INTRODUCTION

The comprehension of various concepts used in science requires students to have an adequate understanding of a function, its first derivative and its second derivative in their graphical representations. For example, a complete comprehension of the kinematics concepts requires students to have an adequate understanding of the graphs of position (function), velocity (first derivative) and acceleration (second derivative). It is important for students to be able to understand, in the context of kinematics, the concept of the derivative as the slope in the relationships between position and velocity, and between velocity and acceleration. Similarly, students should be able to understand, in the context of calculus, the concept of the derivative as the slope in the relationships between a function and the first derivative ($f(x)$ to $f'(x)$), and between the derivative and the second derivative ($f'(x)$ to $f''(x)$). In the same way, it would be important for students to be able to understand, in the context of kinematics, the

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

Correspondence: Genaro Zavala, *Tecnologico de Monterrey, Mexico.*

✉ genaro.zavala@itesm.mx

Contribution of this paper to the literature

- In this article, we present a test that evaluates students' understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve in the context of calculus, and show that it is a content-valid and reliable evaluation instrument with satisfactory discriminatory power.
- We analyze students' understanding of the concepts evaluated in the test and outline recommendations, based on this analysis, for the instruction of these concepts.
- The analyses, recommendations and the test presented in the Appendix can be used by mathematics or science education researchers, and by instructors teaching these concepts.

concept of the antiderivative as the area under the curve in the relationships between the acceleration and the change velocity and between the velocity and the change in position. Correspondingly, in the context of calculus, the concept of the antiderivative as the area under the curve in the relationships between the second derivative and the change of the first derivative ($f''(x)$ to $\Delta f'(x)$) and between the first derivative and the change of the function ($f'(x)$ to $\Delta f(x)$). In this article, we study university students' understanding of these two concepts (slope and area under the curve) in the context of calculus, using a new multiple-choice test. Tests with this feature are highly valued in the area of mathematics and science education research since they allow the evaluation of conceptual learning of large populations (Redish, 1999; Gurel, Eryilmaz & McDermott, 2015).

Many researchers have analyzed students' understanding of the concepts of slope and area under the curve in the context of science, specifically in physics (McDermott et al., 1987; Beichner, 1994; Woolnough, 2000; Meltzer, 2004; Pollock, Thompson & Mountcastle, 2007; Nguyen and Rebello, 2011), while others have studied this understanding in the context of mathematics (Orton, 1983; Leinhardt et al. 1990; Hadjidemetriou & Williams, 2002; Bajracharya et al. 2012; Christensen & Thompson, 2012; Epstein, 2013). However, to date, no study has presented a multiple-choice test that evaluates students' understanding of these concepts in the context of calculus, with a design that follows the steps recommended by mathematics and science education researchers (Beichner, 1994; Ding et al. 2006, Engelhardt, 2009).

To address this need, we conducted a research study with four objectives: (1) to present a multiple-choice test that evaluates students' graph understanding of the concepts of the derivative and the antiderivative (as the slope of the tangent line to the curve at a certain point and as the area under the curve for a given subinterval, respectively) in the context of calculus and its design process; (2) to show that it is a content-valid and reliable evaluation instrument with satisfactory discriminatory power according to the analysis recommended by science education researchers (Beichner, 1994; Ding et al. 2006, Engelhardt, 2009); (3) to conduct a detailed analysis of students' understanding of the concepts evaluated in the test; and (4) to outline specific recommendations, based on the previous analysis, for the instruction of these concepts. It is important to mention that in previous short articles (Pérez, Domínguez & Zavala, 2010; Perez-Goytia, Dominguez & Zavala, 2010), we have presented results of preliminary versions of the test.

PREVIOUS RESEARCH

This section is divided into three subsections. In the first and second sections, we present the most important findings of the studies that have analyzed students' understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve, respectively. In the third section, we describe the tests designed to evaluate these concepts in the context of mathematics, discussing the differences between those tests and our own. The two first subsections are related to the incorrect options that we established in our test, and the third subsection presents a detailed justification for the need of our study and our test.

Students' Understanding of the Concept of the Derivative as the Slope

Several studies have analyzed students' understanding of the concept of the derivative as the slope in the context of physics: the majority of them use the context of kinematics (McDermott et al., 1987; Beichner, 1994;

Planinic, Milin-Sipus, Katic, Susac, & Ivanjek, 2012; Planinic, Ivanjek & Susac, 2013), while some of them use other contexts (Woolnough, 2000; Planinic, Ivanjek & Susac, 2013). Several studies also analyze this understanding in the context of mathematics (Leinhardt et al. 1990; Hadjidemetriou & Williams, 2002; Christensen & Thompson, 2012; Planinic, Milin-Sipus, Katic, Susac, & Ivanjek, 2012; Planinic, Ivanjek & Susac, 2013; Epstein, 2013).

In this subsection, we focus on the two studies that present an overall classification of students' difficulties with the understanding of the concept of the derivative as the slope (Leinhardt et al. 1990; Beichner, 1994). Leinhardt et al. (1990) classified students' difficulties into three categories: (1) interval/point confusions, in which students focus on a single point instead of on an interval; (2) slope/height confusions, in which students confuse the height of the graph with the slope; and (3) iconic confusions, in which students incorrectly interpret graphs as pictures. Beichner (1994) designed the "Test of Understanding Graphs in Kinematics (TUG-K)" and applied it to 895 high school and college students. He pointed out the most frequent errors that students make regarding understanding the slope concept and notes that these errors are directly related to the three categories, classified by Leinhardt et al. For instance, regarding the first category of Leinhardt et al., Beichner found that students often compute the slope at a point by simply dividing a single ordinate value by a single abscissa value, essentially forcing the line through the origin.

Students' Understanding of the Concept of the Antiderivative as the Area under the Curve

Several studies have analyzed students' understanding of the concept of the antiderivative as the area under the curve in the context of physics: the majority of them use the context of kinematics (McDermott et al. 1987; Beichner 1994; Planinic, Ivanjek & Susac, 2013), although some of them use other contexts (Meltzer, 2004; Pollock, Thompson & Mountcastle, 2007; Nguyen and Rebello, 2011; Planinic, Ivanjek & Susac, 2013). In addition, several studies analyze this understanding in the context of mathematics (Orton, 1983; Bajracharya et al. 2012; Planinic, Ivanjek & Susac, 2013).

Beichner (1994) presents an overall analysis of students' difficulties with the understanding of the concept of the antiderivative as the area under the curve and classifies them into three categories: (1) not recognizing the meaning of areas under the graph, (2) calculating the slope rather than the area, and (3) area/height confusions in which students confuse the height of the graph in the last point of the interval with the area. It is noteworthy that Nguyen and Rebello (2011) found that, when presented with several graphs, students had difficulties in selecting the graph in which the area under the graph corresponded to a given integral, although all of them could state, "the integral equals to the area under the curve."

Related Tests

Our test evaluates students' graph understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve in the context of calculus, each concept in two different steps using the function, the first derivative, and the second derivative. In the literature, there are two tests previously designed that relate to the present study. The first is the "Calculus Concept Inventory (CCI)" designed by Epstein (2013). The second is the mathematics version of a test designed by Planinic, Ivanjek & Susac (2013). We will briefly describe these tests and identify the differences between those and our test.

The "Calculus Concept Inventory (CCI)" designed by Epstein (2013) is a 22-item multiple-choice test of conceptual understanding of the most basic principles of differential calculus. The test has three dimensions: (1) functions, (2) derivatives, and (3) limits, ratios, and the continuum. Although the test is for calculus students, this inventory does not focus on evaluating students' understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve.

The mathematics version of the test, designed by Planinic, Ivanjek & Susac (2013), evaluates students' understanding of graphs and focuses on the same concepts as our test. This test has eight questions: five of them refer to the concept of the slope and three to the concept of the area under the graph. However, there are three major differences between this test and ours. The first difference is that not all the questions designed by Planinic

et al. are multiple-choice questions: only four of them have this format, and the other four are open-ended questions. We believe that instruments in which there are open-ended questions are important for research; however, our goal is to obtain an instrument that not only can be used for research but can also be used to assess large student populations and be as easy to analyze as many other multiple-choice instruments available in the literature (i.e., Epstein, 2013). The second difference is related to the objective of the study and the design of the mathematics version of the test. Their study focuses on comparing the graphical understanding of the slope and the area under the curve in mathematics with two other contexts. The third difference, and the most important one, is that the context of Planinic et al.'s test is mathematics, while our test belongs to the context of calculus specifically. Planinic et al. use the context of mathematics and ask directly to find "the slope in a point" or "the area under a curve in an interval" in graphs plotted in the x and y axes. In contrast, we use the context of calculus in our study, and we ask to find the *derivative* of a function at a point or the change of the *antiderivative* of a function in an interval. As mentioned before, we evaluate these concepts in two steps using the function, the first derivative and the second derivative. This assessment in two steps is not possible in the context of mathematics used by Planinic et al. We believe that the differences between our test and the two related published tests justify the need for our study and our test.

METHODS AND TEST DEVELOPMENT

In this section, we cover the first objective of this study: to present a multiple-choice test that evaluates students' graph understanding of the concepts of derivative and antiderivative (as the slope and as the area under the curve, respectively) in the context of calculus, and its design process.

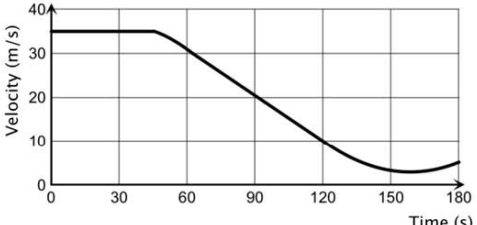
Test Development

We decided to base our new test on the "Test of Understanding of Graphs in Kinematics (TUG-K)" by Beichner (1994), since it is a content-valid and reliable evaluation instrument with satisfactory discriminatory power widely used in the area of science education (see, for example: Chanpichai & Wattanakasiwich, 2010; Bektasli & White, 2012; Tejada Torres & Alarcon, 2012; Maries & Singh, 2013; Mesic, Dervic, Gazibegovic-Busuladzic, Salibasic, & Erceg, 2015; Hill & Sharma, 2015), and also in our modified version of the TUG-K (Zavala et al., 2017). The original version has been a well-received assessment. However, when analyzing this test, we detected several potential improvements, mainly regarding the parallelism between related objectives and the parallelism between the items of some objectives, but also, the representation of the most common alternative conceptions as distractors. To generate those improvements, we decided to modify the test, adding new items and modifying some distractors in some of the original items that remained. That process is described in another study (Zavala et al., 2017). Note that the original version of the TUG-K has 21 items and our modified version of the TUG-K has 26 items. The general idea to design the test presented in this article was to rewrite the items of the TUG-K, removing the context of kinematics and replacing it with the context of calculus. [Figure 1](#) shows an example of this translation.

To create the test described in this article we designed two preliminary versions of the tests and the final version of the test, which we present here. To design the first preliminary version, we rewrote the 26 items of our modified version of the TUG-K removing the context of kinematics and replacing it with the context of calculus. This version was reviewed by physics and mathematics professors, and special care was put into preserving the original structure of the items. This version was administered to university students of introductory courses in physics and mathematics. The results of this administration (Pérez, Domínguez & Zavala, 2010) showed that, while most of the problems of the test had an almost perfect "translation" from kinematics to calculus, there were some items that lost their meaning or were too difficult for the students to answer. Those items corresponded to objectives 6 and 7 of the TUG-K, which focused on the relationship between a kinematics graph and a textual description. Based on this analysis, it was decided that the second preliminary version of the test would have only 16 items from the remaining objectives of the original test. The results of this second version, which was our pilot study, were analyzed briefly in a previous short article (Perez-Goytia, Dominguez & Zavala, 2010). In that work, we proved that the 16 items in the context of calculus behaved satisfactorily. That is, the results indicated that the TUG-C had

Item of the modified Test of Understanding of Graphs in Kinematics (TUG-K): The following figure shows the velocity versus time graph of an object. At $t = 90$ s, the acceleration of the object is:

(A) -0.22 m/s^2
 (B) -0.33 m/s^2
 (C) -1.0 m/s^2
 (D) -2.0 m/s^2
 (E) 20 m/s^2



Item 11 of our Test of Understanding Graphs in Calculus (TUG-C): The following figure shows the graph of $f'(x)$, the first derivative of a function. At $x = 90$, the second derivative of the function is:

(A) -0.22
 (B) -0.33
 (C) -1.0
 (D) -2.0
 (E) 20

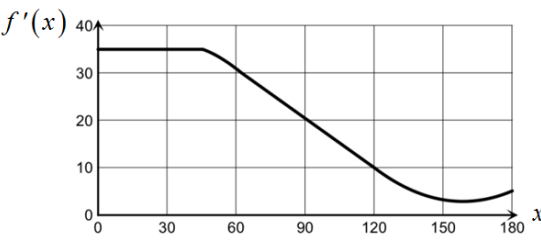


Figure 1. Example of the translation of item 11 of our “Test of Understanding Graphs in Calculus (TUG-C)”

potential to become an appropriate instrument to measure conceptual understanding and graphical interpretation of a function and its derivative.

After this last analysis, we decided to design the final version of the test with the same 16 items, adding some modifications to improve the parallelism of the items. As we will see in the next section, there are several items in the test that are directly related to each other. In this version, we performed several modifications with the distractors and graphs of some items so that the items directly related to each other had the same type of distractors and graphs. This allows us to make direct comparisons between these items (as we will do in the analysis section). In the [Appendix](#) of this article, we show this last version of the test, which is referred to as the “Test of Understanding Graphs in Calculus (TUG-C).” Note that the order of items in this version is different from the previous versions, since we decided to establish a random item order.

Characteristics of the Test

Table 1 shows a description of the TUG-C (the complete test can be found in the [Appendix](#)). The table presents a description of the five dimensions of the test, the items included in each dimension, the concept evaluated (the derivative as the slope, the antiderivative as the area under the curve, or either of them) and the specific step evaluated. Moreover, **Table 2** shows a detailed description of the test’s 16 items grouped in each of the five dimensions.

Table 1. Description of the Test of Understanding Graphs in Calculus (TUG-C)

Dimension	Description	Items	Concept	Steps evaluated
1	Determine $f'(x)$ from the graph of $f(x)$	1, 6, 13	The derivative as the slope	$f(x) \rightarrow f'(x)$
2	Determine $f''(x)$ from the graph of $f'(x)$	7, 11, 3		$f'(x) \rightarrow f''(x)$
3	Determine $\Delta f(x)$ from the graph of $f'(x)$	5, 14, 15	The antiderivative as the area under the curve	$f'(x) \rightarrow \Delta f(x)$
4	Determine $\Delta f'(x)$ from the graph of $f''(x)$	12, 8, 10		$f''(x) \rightarrow \Delta f'(x)$
5	Determine the corresponding graph from a graph	16, 9, 2, 4	Either of the two concepts can be used	The four steps are evaluated in each of the items: $f(x) \rightarrow f'(x); f'(x) \rightarrow f''(x);$ $f'(x) \rightarrow f(x); f''(x) \rightarrow f'(x)$

Table 2. Description of the items of the Test of Understanding Graphs in Calculus (TUG-C)

Dimension	Items (Description)
1	1. Determine the positive value of $f'(x)$ from the graph of $f(x)$
	6. Determine the negative value of $f'(x)$ from the graph of $f(x)$
	13. Identify the interval in which $f'(x)$ is the most negative in the graph of $f(x)$
2	7. Determine the positive value of $f''(x)$ from the graph of $f'(x)$
	11. Determine the negative value of $f''(x)$ from the graph of $f'(x)$
	3. Identify the interval in which $f''(x)$ is the most negative in the graph of $f'(x)$
3	5. Establish the procedure to determine the $\Delta f(x)$ from the graph of $f'(x)$
	14. Determine the value of the change the $\Delta f(x)$ from the graph of $f'(x)$
	15. Identify the $f(x)$ with the greatest change from several graphs of $f'(x)$
4	12. Establish the procedure to determine the $\Delta f'(x)$ from the graph of $f''(x)$
	8. Determine the value of the change of the $\Delta f'(x)$ from the graph of $f''(x)$
	10. Identify the $f'(x)$ with the greatest change from several graphs of $f''(x)$
5	16. Determine the corresponding graph of $f'(x)$ from the graph of $f(x)$
	9. Determine the corresponding graph of $f''(x)$ from the graph of $f'(x)$
	2. Determine the corresponding graph of $f(x)$ from the graph of $f'(x)$
	4. Determine the corresponding graph of $f'(x)$ from the graph of $f''(x)$

As shown in **Table 1**, the first four dimensions contain three items, and the fifth dimension contains four items. Dimensions 1 & 2 of both tests are directly related, since both evaluate the understanding of the concept of the derivative as the slope, and dimensions 3 & 4 are also directly related, since both evaluate the understanding of the concept of the antiderivative as the area under the curve. The difference in these related dimensions lies in which step is evaluated. Dimension 1 evaluates the step from $f(x)$ to $f'(x)$, while dimension 2 evaluates the step from $f'(x)$ to $f''(x)$. On the other hand, dimension 3 evaluates the step from $f'(x)$ to $\Delta f(x)$, while dimension 4 evaluates the step from $f''(x)$ to $\Delta f'(x)$.

Table 2 shows that the related dimensions (dimensions 1 & 2 and dimensions 3 & 4) have related items that evaluate the same concept in the same way, with the only difference being the step evaluated. For example, item 1 of dimension 1 evaluates the determination of the positive value of $f'(x)$ from the graph of $f(x)$, while the related item 7 of dimension 2 evaluates the determination of the positive value of $f''(x)$ from the graph of $f'(x)$. The three items of the related dimensions 1 & 2 ask: (1) to determine the positive value of a derivative, (2) to determine the negative value of the derivative, and (3) to identify the interval in which the derivative is the most negative. On the other hand, the three items of the related dimensions 3 & 4 ask: (1) to establish the procedure to determine the

change of an antiderivative, (2) to determine the value of the change of an antiderivative, and (3) to identify the variable whose antiderivative has the greatest change in a specific interval. Furthermore, in an overview of the test it is possible to observe relations between the three items of dimensions 1 & 2 and the three items of dimensions 3 & 4. The first two items of each of the dimensions focus on obtaining a value of a variable, and the third item focuses on finding a maximum of this variable.

As shown in **Tables 1** and **2**, dimension 5 evaluates selecting, among different graphs, the correct graph according to the relationships that each item requests. The items in this dimension evaluate each of the steps evaluated in the other four dimensions: (1) from $f(x)$ to $f'(x)$; (2) from $f'(x)$ to $f''(x)$; (3) from $f'(x)$ to $f(x)$; and (4) from $f''(x)$ to $f'(x)$. Dimension 5 also has related items that evaluate the same concept in the same way, with the only difference being the step evaluated. Items 16 and 9 evaluate selecting the corresponding graph of the derivative from a graph, while items 2 and 4 evaluate selecting the corresponding graph of the antiderivative from a graph. The main difference, and the reason for it being a dimension in itself, is that dimension 5 is a process from understanding relationships from graph to graph. Dimensions 1-4 are processes from understanding the relationships from an operation of the graph (the slope of the area under the curve). In summary, the eight related items in the test are: 1 and 7, 6 and 11, 13 and 3 in the related dimensions 1 & 2; 5 and 12, 14 and 8, 15 and 10 in the related dimensions 3 & 4; and 16 and 9, 2 and 4 in dimension 5.

Participants

The research was conducted at a large private university in Mexico. The participants in this study were engineering students finishing their introductory calculus-based mechanics course and their first calculus course. The textbook used in the mechanics course was “Physics for Scientists and Engineers” by Serway and Jewett (2008). Students also used the “Tutorials in Introductory Physics” by McDermott, Shaffer, and the Physics Education Research Group (2001). The textbooks used in the calculus course were by Salinas et al. (2000; 2012). This course covers the following main topics: linear function, qualitative analysis of a function and its first and second derivative, quadratic function and Euler’s method (interpretation of the area under the curve), analysis of the characteristics, the derivative and applications of different models (polynomial, exponential, sine), and basic integral with a change of variables. The test was administered as a diagnostic test to 144 students who were completing the courses mentioned above, and it did not count towards the final course grades.

ANALYSIS OF THE TEST

In this section, we cover the second objective of this study: to show that the TUG-C is a content-valid and reliable evaluation instrument with adequate discriminatory power according to the analysis recommended by mathematics and science education researchers (Beichner, 1994; Ding et al. 2006, Engelhardt, 2009). We divide this section into two subsections: (1) content validity, and (2) reliability and discriminatory power.

Content Validity

We checked the content validity of the items of the TUG-C. Content validity measures how well the test items cover the content domain they intend to test (Engelhardt, 2009). In evaluating the TUG-C, we asked eight experts (four mathematics faculty members and four physics faculty members) to rate each item with its corresponding objective (1 being the lowest and 5 the highest), in accordance with the procedure established by Engelhardt (2009). Each of the items on the TUG-C was rated with a high score regarding the match between the test item itself and its stated objective. The lowest average score for any item was 4.25 and the highest was 4.88. Moreover, the overall average score was 4.76. These results are evidence of the high content validity of the TUG-C.

Reliability and Discriminatory Power

We also evaluated the reliability and discriminatory power of the TUG-C, performing the five statistical tests suggested by Ding et al. (2006). The first three measures focus on individual test items: the item difficulty index, the item discrimination index, and the item point-biserial. **Table 3** shows these values for each item on the

Table 3. Item difficulty index (P), item discriminatory index (D), and point-biserial coefficient (r_{pbs}) for each item of the TUG-C

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
P	0.37	0.52	0.53	0.50	0.58	0.47	0.43	0.53	0.61	0.28	0.54	0.66	0.61	0.49	0.26	0.49
D	0.42	0.56	0.72	0.33	0.50	0.83	0.86	0.72	0.64	0.61	0.61	0.64	0.64	0.72	0.75	0.69
r_{pbs}	0.41	0.41	0.57	0.22	0.42	0.63	0.67	0.55	0.46	0.56	0.47	0.49	0.51	0.60	0.63	0.57

TUG-C. The other two measures focus on the test as a whole: the Kuder-Richardson reliability test and Ferguson’s delta test. We discuss the results of these five statistical tests below.

Item difficulty index

The item difficulty index (P) is a measure of the difficulty of a single test question. A widely-adopted criterion, used by Ding et al. (2006), indicates that the difficulty index should be between 0.3 and 0.9. **Table 3** shows the difficulty index P values for each item on the TUG-C. Only two items, items 10 (0.28) and 15 (0.26), have item difficulty indexes slightly lower than desired. Ding et al. also recommended the calculation of the average difficulty value. The criterion range for the average difficulty value is also [0.3–0.9]. For the TUG-C, the average difficulty value is 0.49, which also falls within the suggested range.

Item discrimination index

The item discriminatory index (D) is a measure of the discriminatory power of each item on a test. Ding et al. (2006) established two criteria for this index: (1) eliminate items with negative indexes, and (2) the majority of the test items should have a good discrimination index ($D \geq 0.3$). Table 3 shows the discrimination index D values for each item of the TUG-C (using the 25%–25% method). We observe that the TUG-C satisfies these two criteria, since there are no negative items, and all of the items have a discrimination index over 0.3. Ding et al. also recommended the calculation of the average discrimination index, suggesting a value of ≥ 0.3 . For the TUG-C the average discriminatory value is 0.64 (using the 25%–25% method), which meets this criterion.

Point-biserial coefficient

The point-biserial coefficient (r_{pbs}) is a measure of the consistency of a single item in relation to the whole test, reflecting the correlation between students’ scores on an individual item and their scores on the entire test. A widely-adopted criterion, followed by Ding et al. (2006), is that an item with a satisfactory point-biserial coefficient must be $r_{pbs} \geq 0.2$. **Table 3** shows the point-biserial coefficient for each item on the TUG-C. We can see that all of the TUG-C’s items satisfy this condition. Ding et al. also recommended the calculation of the average point-biserial coefficient, with a criterion range of ≥ 0.2 . The average coefficient of the TUV is 0.51, which also satisfies this criterion.

Kuder-Richardson reliability index and Ferguson’s delta test

The Kuder-Richardson reliability index is a measure of the self-consistency of a whole test. Ding et al. (2006) state that a test with a reliability index that is higher or equal to 0.7 is reliable for group measures. The index for the TUG-C is 0.81, which meets this criterion. Ferguson’s delta test measures the discriminatory power of an entire test by investigating how broadly the total scores of a sample are distributed over the possible range. A widely-adopted criterion, followed by Ding et al., is that a test with a Ferguson’s delta of higher than 0.9 offers a good discrimination. Ferguson’s delta test for the TUG-C is 0.99, which satisfies this requirement.

Summary of the five statistical tests

We present a summary of the five statistical tests in **Table 4**. From the analysis, we can conclude that the TUG-C is a reliable test with satisfactory discriminatory power.

Table 4. Summary of the results of the five statistical tests suggested by Ding et al. (2006) for the TUG-C

Test statistic	Desired values	TUG-C
Difficulty index	[0.3, 0.9]	Average: 0.49
Discrimination index	≥ 0.3	Average: 0.64
Point-biserial coefficient	≥ 0.2	Average: 0.51
Kuder-Richardson reliability index	≥ 0.7 for group measures	0.81
Ferguson's delta test	> 0.9	0.99

ANALYSIS OF STUDENTS' UNDERSTANDING OF THE CONCEPTS OF DERIVATIVE AND ANTIDERIVATIVE

In this section, we cover the third objective of this study: to conduct a detailed analysis of students' understanding of the concepts evaluated by the TUG-C. Specifically, we studied the results of 144 students who had completed their introductory calculus-based mechanics course and their first calculus course.

Overall Performance

The average of the scores of the TUG-C, from the sample of 144 students, is 7.88 of 16 possible points (each test item is worth 1 point). This average, expressed in percentage of the total possible points, is 49%, which corresponds to the average difficulty index value (0.49) shown in the previous section. The distribution of scores was significantly non-normal (Kolmogorov-Smirnov, $D(144) = 0.093$, $p < 0.01$; Shapiro-Wilk test, $W(144) = 0.965$, $p < 0.01$). The skewness of the distribution of scores is 0.152 (SE=0.202), indicating a pile-up to the right, and the kurtosis of the distribution is -0.991 (SE=0.401), indicating a flatter than normal distribution. The positive skew indicates that the test is 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 8, the bottom quartile (Q1) is 4.25, and the top quartile (Q3) is 11, so the interquartile range is 6.75. In this overall analysis, it is noteworthy that the students at the median (8) had difficulty answering correctly eight questions (out of 16) on the TUG-C.

The overall results show that this is not an easy test. Students struggle with questions they may not familiar with. However, the concepts included in the tests are taught in their courses, but probably not in the same way, the test presents them. That the test's statistical tests are satisfactory means that students answer the questions engaged and with interest, even if the questions are not presented in the way they are used to.

Performance on Three Representative Items of the Test

In this subsection, we conduct a qualitative analysis regarding students' performance on three representative items of the test: 1, 14 and 16 (see [Table 5](#) and the [Appendix](#)). As shown in [Table 1](#), item 1 evaluates the concept of the derivative as the slope, item 14 assesses the concept of the antiderivative as the area under the curve, and item 16 evaluates the use of either of the two concepts to determine the corresponding graph from a specific graph.

Table 5. The five dimensions evaluated in the TUG-C, the description of the items, and the percentages selecting a particular choice for each item. (Note that the correct answer is in boldface.)

Dim.	Items description	A	B	C	D	E
1	1. Determine the positive value of $f'(x)$ from the graph of $f(x)$	12%	9%	37%	26%	15%
	6. Determine the negative value of $f'(x)$ from the graph of $f(x)$	47%	11%	8%	24%	10%
	13. Identify the interval in which $f'(x)$ is the most negative in the graph of $f(x)$	13%	61%	10%	11%	5%
2	7. Determine the positive value of $f''(x)$ from the graph of $f'(x)$	43%	6%	25%	17%	10%
	11. Determine the negative value of $f''(x)$ from the graph of $f'(x)$	15%	54%	9%	13%	8%
	3. Identify the interval in which $f''(x)$ is the most negative in the graph of $f'(x)$	6%	22%	10%	9%	53%
3	5. Establish the procedure to determine the $\Delta f(x)$ from the graph of $f'(x)$	3%	58%	33%	3%	3%
	14. Determine the value of the $\Delta f(x)$ from the graph of $f'(x)$	7%	49%	13%	24%	7%
	15. Identify the $f(x)$ with the greatest change from several graphs of $f'(x)$	26%	29%	12%	31%	2%
4	12. Establish the procedure to determine the $\Delta f'(x)$ from the graph of $f''(x)$	66%	18%	7%	6%	3%
	8. Determine the value of the $\Delta f'(x)$ from the graph of $f''(x)$	15%	11%	10%	53%	10%
	10. Identify the $f'(x)$ with the greatest change from several graphs of $f''(x)$	8%	28%	4%	31%	28%
5	16. Determine the corresponding graph of $f'(x)$ from the graph of $f(x)$	3%	27%	14%	49%	7%
	9. Determine the corresponding graph of $f''(x)$ from the graph of $f'(x)$	3%	13%	19%	61%	5%
	2. Determine the corresponding graph of $f(x)$ from the graph of $f'(x)$	9%	52%	2%	24%	12%
	4. Determine the corresponding graph of $f'(x)$ from the graph of $f''(x)$	50%	3%	13%	6%	28%

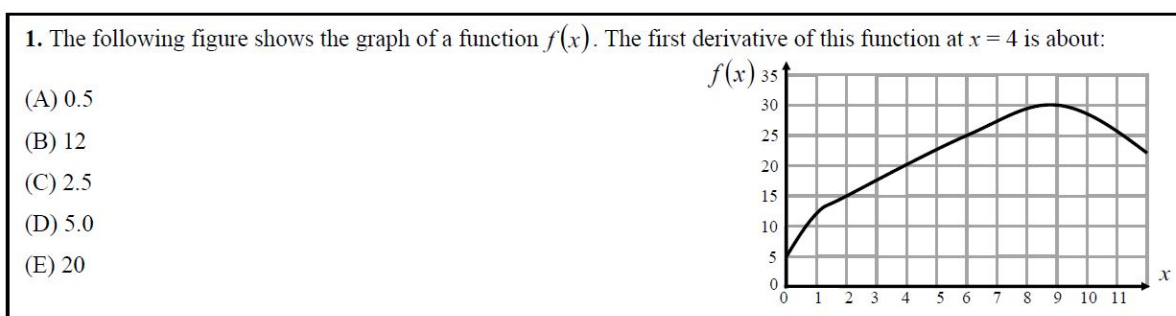


Figure 2. Item 1 of our “Test of Understanding Graphs in Calculus (TUG-C)”

Figure 2 presents item 1 that evaluates the concept of the derivative as the slope asking to determine the positive value of $f'(x)$ at a point from the graph of $f(x)$. Only 37% of students select the correct option C. The most frequent error is to obtain this value dividing the ordinate by the abscissa of the point on the graph that is not valid in this situation (option D, 26%). Moreover, two other incorrect options are selected in similar proportions, above 10%: selecting the ordinate of the point (option E, 15%), and calculating the value of a “slope” by counting squares (option A, 12%).

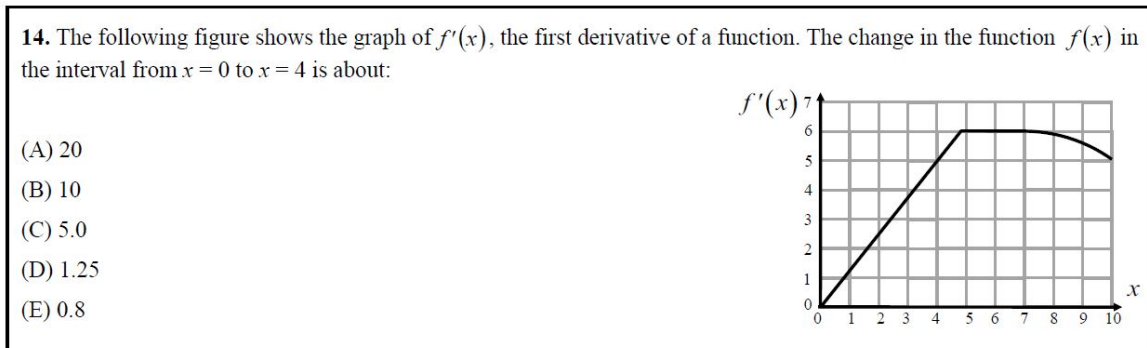


Figure 3. Item 14 of our “Test of Understanding Graphs in Calculus (TUG-C)”

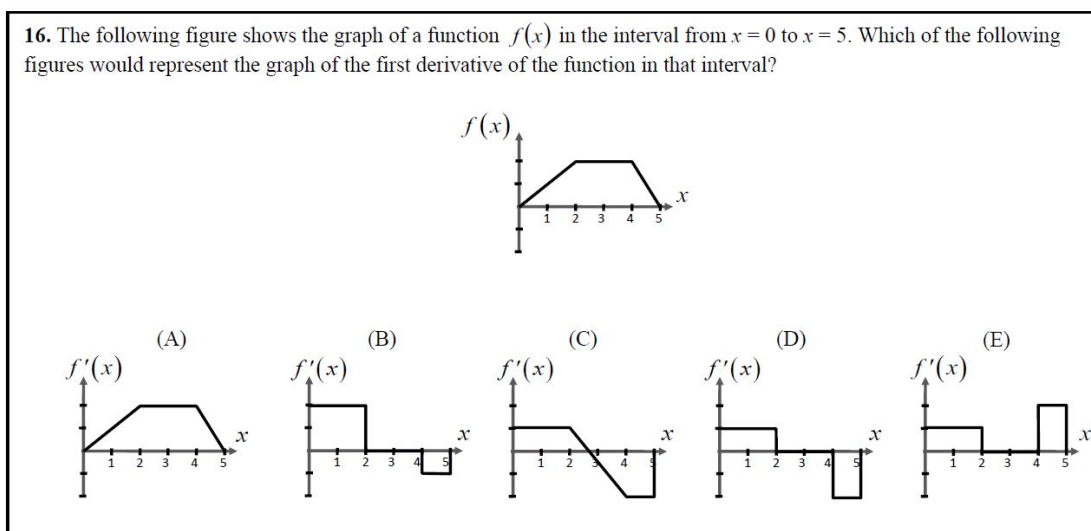


Figure 4. Item 16 of our “Test of Understanding Graphs in Calculus (TUG-C)”

Figure 3 presents item 14 that evaluates the concept of the antiderivative as the area under the curve asking to determine the value of the change $\Delta f(x)$ in an interval from the graph of $f'(x)$. In this item, 49% of students select B, the correct option. The most frequent error is to use the correct procedure to calculate the slope in the interval instead of the area under the curve (option D, 24%). The other three incorrect options are selected in similar proportions. In one of them, students select the ordinate value of the point on the right in the interval, $x = 4$ (option C, 13%). In another, students use an incorrect procedure to calculate the slope of the curve in the interval, dividing the abscissa by the ordinate of the point on the right in the interval (option E, 7%). On the other, students multiply the abscissa by the ordinate of the point on the right in the interval (option A, 7%). It is interesting to notice that the latter multiplication is part of the correct procedure to calculate the area under the curve, but students do not divide this multiplication by two.

Figure 4 presents item 16 that evaluates the determination of the corresponding graph of $f'(x)$ from the graph of $f(x)$. In this item, 49% of students select D, the correct option. In the most frequent error students seem to understand the shape the graph should have but have difficulties relating the relative values of the slopes of the graph, opting for a relationship opposite to the correct one (option B, 27%). In the second most frequent error, students make a mistake only in the section of the graph in which the derivative is zero. Instead of setting a step-type graph with a value of zero in this interval, students choose option C (14%), in which the derivative value decreases uniformly in that interval. Finally, in the third most frequent error (option E, 7%), students only make a mistake with the sign of the value of the slope in the last section of the graph.

Table 6. Correct answer percentages of the related items, correct answer percentages of students selecting the correct choice in both of the related items, and correct answer averages percentages of each dimension

Description of items that evaluate the concept of the derivative	Item	Dim. 1	Item	Dim. 2	Both correct
Determine the positive value of the derivative	1	37%	7	43%	24%
Determine the negative value of the derivative	6	47%	11	54%	38%
Identify the interval in which the derivative is the most negative	13	61%	3	53%	48%
Average of the dimension		48%		50%	

Description of items that evaluate the concept of the antiderivative	Item	Dim. 3	Item	Dim. 4	Both correct
Establish the procedure to determine the change of the antiderivative	5	58%	12	66%	48%
Determine the value of the change of the antiderivative	14	49%	8	53%	40%
Identify the variable whose antiderivative has greatest change in a specific interval	15	26%	10	28%	21%
Average of the dimension		44%		49%	

Description of items that can be solve using the concepts of derivative and/or the antiderivative	Item	Dim. 5	Related to	Both correct
Determine the corresponding graph of the derivative from a graph	16	49%	Dim.1	35%
	9	61%	Dim. 2	
Determine the corresponding graph of the antiderivative from a graph	2	52%	Dim. 3	30%
	4	50%	Dim. 4	
Average of the dimension		53%		

Items and Dimensions

Table 6 shows the proportion of students selecting the correct choice of the related items, the proportion of students selecting the correct choice in both of the related items, and the average of the correct choice of each dimension.

From Table 6, we can note three issues regarding these results. The first is that the five dimensions have very close average values, ranging from 44% to 53%. The second is that the value of these averages is relatively low, around 50%. These results show that students have similar difficulties with the concepts evaluated in the test. Moreover, the third is that individual results for the items range from 26% for item 15 to 66% for item 12. It shows that the concepts evaluated in all items of the test are difficult for students, since in the item with the highest percentage (item 12), a third of the students showed difficulties to answer the question correctly.

In the following subsection, we present two analyses. In the first, a comparison of the related items of the test; and in the second, we cluster the items of the test according to levels of difficulty.

Related items in the test

The related items evaluate the same concept in the same way, with the only difference being the step evaluated. Therefore, it is relevant for instructional reasons to perform a comparison of the correct answers to these related items. When we qualitatively compare the proportion of students answering the related items correctly, we observe that they are very similar. Moreover, comparing students' correct answers in these related items using the chi-square test following the procedure described by Sheskin (2007), we found no significant differences in choosing the correct answer in any of the related items.

When we observe that there is no significant difference in the selection of the correct answer in the related items, we could think that there is consistency in students' answers, that is, the majority of students who correctly answer the item that evaluates the first step, also correctly answer the item that evaluates the second step. However, when we perform a cross-analysis showing the proportion of students answering both related items correctly (see [Table 6](#)), we observe that in several related items this proportion is considerably lower than that for each of the items. Therefore, a considerable number of students correctly answer one of the items but incorrectly answer the other.

For example, for items 1 and 7 of dimensions 1 and 2 (37% and 43% of students answering correctly, respectively) we notice that 24% of the total students answered both items correctly, that is, 65% of students answering item 1 correctly, answer item 7 correctly. We also notice that 13% of students answered item 1 correctly, which evaluates the first step, but answered item 7 incorrectly, which evaluates the second step, and that 19% of students answered item 7 correctly but answered item 1 incorrectly.

For items 5 and 12 of dimensions 3 and 4 (58% and 66% of students answering correctly, respectively), we observe that 48% of students answered both items correctly, that is, 83% of students who answered item 5 correctly, answered item 12 correctly. We also observe that 10% of students answered item 5 correctly, which evaluates the first step, but answered item 12 incorrectly, which evaluates the second step, and that 18% of students answered item 12 correctly but answered item 5 incorrectly.

Finally, for items 16 and 9 of dimension 5 (49% and 61% respectively), we observe that 35% of students answered both items correctly, that is, 71% of students who answered item 16 correctly, also answered item 9 correctly. We also observe that 14% of students answered item 16 correctly, which evaluates the first step, but answered item 9 incorrectly, which evaluates the second step, and that 26% of students answered item 9 correctly but answered item 16 incorrectly.

We can hypothesize that only when a student answers the two related items correctly, he or she may have a complete understanding of the concept. From [Table 6](#), we observe that the proportion of students having a complete understanding of the concept is quite low and range from 21% for items 15 and 10 to 48% for items 13 and 3, and 5 and 12. There is a considerable proportion of students showing only a partial understanding of the concept since they answer one related item correctly but the other incorrectly.

Cluster of items according to difficulty level

According to [Table 6](#), the most difficult items are those from dimension 3 & 4, which evaluate the identification of the variable whose antiderivative has the greatest change in a specific interval. Only 26% of students answered item 15 correctly, which evaluates the first step (from $f'(x)$ to $f(x)$) and only 28% of students answered item 10, which evaluates the second step (from $f''(x)$ to $f'(x)$). These two items have in common that they assess the maximum value of an antiderivative in an interval.

These two questions ask students to choose a graph that has the greatest change of a function (the function for item 15 and the first derivative of the function for item 10) given the graph of the derivative (the first derivative for item 15 and the second derivative for item 10). As we have seen in other items, some students confuse the concept of the slope with the concept of the area under the curve. Therefore, these two questions have some items that might be attractive to those students since in option D for item 15 and option D for item 10 the slope changes continuously. Other students might be confused by the word change by thinking about the change of the function in the graphs. In that case, options C and E for item 15 and options A and C for item 10 change; moreover, in option B for item 15 and option E for item 10, the function changes more since it goes from zero to the maximum and then back to zero. What is not attractive for all those students is option A for item 15 and option B for item 10, which are the correct answers, since in those options neither the slope of the function nor the function change in the interval. These items could represent good discriminatory items for those who understand the concept of the antiderivative as the area under the curve. [Table 3](#) shows the item discrimination index, which is a measure of the discriminatory power of each item on a test. Item 15 is considerably above average (0.75 vs average = 0.64) and item 10 is slightly

below average (0.61). This index is the discriminatory power concerning the test as a whole; it would probably be better for item 10 if we take only the items that correspond to the concept of the antiderivative. On the other hand, the two items have an above average point-biserial coefficient, which is a measure of the correlation between students' scores on an individual item and their scores on the entire test (0.63 for item 15 and 0.56 for item 10 vs average = 0.51). Actually, item 15 is the second highest in the table.

Table 6 also shows the easiest items, which are two groups of related items. The first group is from the items of dimensions 3 & 4 (items 5 and 12 respectively, which evaluate the account of the procedure to determine the change of an antiderivative. Item 5, which evaluates the first step (from $f'(x)$ to $\Delta f(x)$), was answered correctly by 58% of students, and 66% of students answered correctly item 12, which evaluates the second step (from $f''(x)$ to $\Delta f'(x)$). The second group of related items comes from dimensions 1 & 2 (items 13 and 3 respectively), which evaluate the identification of the interval in which the derivative is the most negative. Item 13, which evaluates the first step (from $f(x)$ to $f'(x)$), was answered correctly by 61% of students, and 53% of students answered correctly item 3, which evaluates the second step (from $f'(x)$ to $f''(x)$). The items of these groups have in common that to solve them it is not necessary to make accurate calculations.

Items 5 and 12 (dimensions 3 & 4, respectively) correspond to items in which students have to choose, among different descriptions, the one that represents the concept of the antiderivative as the area under the curve. Items 14 and 8 are items that evaluate the same concept, but in these two cases, students are asked to calculate the change instead of choosing a procedure. The results of students in these two items are considerably better than those of items 5 and 12. It seems that the correct answer to items 5 and 12 attract not only those students able to do the procedure without saying what the procedure is but also, those students who, while presenting the question, would not be able to do it by themselves.

Finally, from **Table 6**, we observe that the other five groups of related items have a medium difficulty level. These groups of items evaluate the determination of the positive and negative value of the derivative (two groups, dimensions 1 & 2), the determination of the change of the antiderivative (one group, dimensions 3 & 4), and the determination of the corresponding graph of the derivative or the antiderivative from a graph (two groups, dimension 5). The items from these five groups have in common that, to solve them, it is necessary to make accurate calculations, unlike the items of the groups that were the easiest and the most difficult for students. This type of calculations are necessary in all of the items of dimension 5 while choosing the correct graphs, since in all items there are incorrect graphs very similar to the correct choice but with slight differences (e.g., the incorrect option B in item 16), and students need to do quality calculations to choose the correct answer.

Most Frequent Errors

In this subsection, we present an overall analysis of the most frequent errors in the items (a) from the related dimensions 1 & 2, (b) from the related dimensions 3 & 4, and (c) from dimension 5. **Table 5** shows the five dimensions evaluated in the TUG-C, the items' descriptions, the results for each option of the items. Note that the percentages of the correct answers correspond to the difficulty indices shown in **Table 3**.

Items of the related dimensions 1 & 2

The items of these dimensions evaluate students' understanding of the concept of the derivative as the slope. Dimension 1 evaluates the determination of $f'(x)$ from the graph of $f(x)$, and dimension 2 evaluates the determination of $f''(x)$ from the graph of $f'(x)$.

Dimensions 1 & 2 have two items that evaluate the determination of a positive and a negative value of a derivative at a point of a curve (dimension 1: items 1 and 6; dimension 2: items 7 and 11). **Table 5** shows that, for all the items, the most frequent error is obtaining this value by dividing the ordinate by the abscissa of the point in the graph (item 1: option D, 26%; item 6: option D, 24%; item 7: option C: 25%; item 11: option A, 15%). It is important to note that in the items in which the derivative is negative (items 6 and 11), students add a negative sign to the

obtained value. An interesting tendency is that the proportion of students answering correctly is higher for items with a negative derivative than it is for items with a positive derivative.

This error is also rather common in the context of kinematics (Beichner, 1994) but in that case, the misunderstanding comes from the conception that velocity (or acceleration) is distance divided by time (velocity divided by time). In this case, the error could come from the way in which students are interpreting the derivative, df/dx , which could be, as in kinematics, a ratio of two quantities, the function and x .

Dimensions 1 & 2 have a third item which evaluates the identification of the interval in which the derivative is the most negative. Dimension 1 evaluates the identification of the interval in which $f'(x)$ is the most negative in the graph of $f(x)$ (item 13), and dimension 2 evaluates the identification of the interval in which $f''(x)$ is the most negative in the graph of $f'(x)$ (item 3). In these two items, we observe two frequent errors: the choosing of an interval in which the derivative is negative but not the most negative (item 13: option D, 11%; item 3: option B, 22%), and the choosing of the point in which the graph has a minimum value (item 13: option A, 13%; item 3: option C, 10%).

An interesting result is that these two errors, an interval in which the derivative is negative and has the most negative value, are connected. Some students choose the former because, in that interval, not only the slope is negative, but the function also becomes negative. The last point of the interval is the most negative value of the function in the graph for both items.

Items of the related dimensions 3 & 4

The items of these dimensions evaluate students' understanding of the concept of the antiderivative as the area under the curve. Dimension 3 evaluates the determination of $\Delta f(x)$ from the graph of $f'(x)$, and dimension 4 evaluates the determination of $\Delta f'(x)$ from the graph of $f''(x)$.

Dimensions 3 & 4 have an item that evaluates the account of the procedure to determine the change of an antiderivative in an interval from a graph (dimension 3: item 5; dimension 4: item 12). Note that the slope of the curves is constant in the interval. As shown in [Table 5](#), the most frequent error in these two items is to account for the procedure to calculate the slope of the curve instead of the area under the curve (item 5: option C, 33%; item 12: option B, 18%).

These dimensions also have an item that evaluates the determination of the value of the change of an antiderivative. Item 14 evaluates the determination of the value of $\Delta f(x)$ from the graph of $f'(x)$ (dimension 3), and item 8 evaluates the determination of the value of $\Delta f'(x)$ from the graph of $f''(x)$ (dimension 4). We observe a pattern: the sum of the percentages of the two answers in which students use correct or incorrect procedures to calculate the slope of the curve, instead of the area under the curve, are similar in the two items, and these two answers are the most frequent errors in both items. The first incorrect choice is to use the correct procedure to calculate the slope in the interval instead of the area under the curve (item 14, option D: 24%; item 8, option B: 11%). The second incorrect choice is to use an incorrect procedure to calculate the slope of the curve in the interval (item 14, option E: 7%; item 8, option A: 15%). The sum of these percentages is 31% for item 14 and 26% for item 8. The difference between these sums is minimal (only 5%), and these two choices together are the most frequent errors in both items.

It seems that the most important challenge for instruction is that the concept of the antiderivative as the area under the curve is misunderstood for that of the derivative as the slope. For students, and probably more commonly for first-year students, the slope is the concept they learn. Thus, they resort to it even in cases in which it does not apply.

Dimensions 3 & 4 have a third item that evaluates the identification of the variable whose antiderivative has the greatest change in a specific interval (dimension 3, item 15; dimension 4, item 10). In these two items, we found the same two most frequent errors. In the first most frequent error, students do not choose the graph of the curve with the greatest area under the curve in the interval, but rather the graph of a curve whose slopes in the

interval are always increasing (item 15 option D, 31%; item 10: option D, 31%). In these items, students seem to be thinking in terms of the slope as in the previous items of these dimensions. This item asks for a variable with the greatest change in the interval and students choose the curve that has the greatest change in positive slopes. The second most frequent error in these items is choosing a graph of a curve that increases in the middle of the interval and decreases in the other half (item 15: option B, 29%; item 10: option E, 28%). (The curve is like an inverted parabola. The left point of the interval is (0, 0) and the right point of the interval is (constant, 0); therefore, the vertical change of the curve in the interval is zero).

In these errors, students seem to be thinking in terms of two resources, which are different ways of thinking about a situation (Hammer, 2000). The first resource is to think in terms of the slope: this item asks for a variable with the greatest change in the interval, and students choose the curve that has the greatest change in slope as it begins in a high positive value and ends at the same value, but negative. The second resource is to think in terms of the vertical value. Although the vertical change is zero in the curve of the graph, students seem to think in terms of the vertical value, and conclude that this curve has the greatest vertical change since it “rises and falls” (as interviews with some students have revealed).

Items of dimension 5

The four items of dimension 5 evaluate the determination of the corresponding graph from a graph. Students can solve these items using either of the two concepts evaluated in the first four dimensions: the concept of the derivative as the slope and/or the concept of the antiderivative as the area under the curve.

As shown in [Table 6](#), two items of this dimension evaluate the determination of the corresponding derivative graph from a graph (items 16 & 9). Item 16 evaluates the determination of the corresponding $f'(x)$ graph from the $f(x)$ graph, and item 9 evaluates the determination of the corresponding $f''(x)$ graph from the $f'(x)$ graph. The two most frequent errors in these two related items are very similar (see [Table 5](#)). In item 16, the most frequent error corresponds to a graph in which students seem to understand the shape the graph should have, but have difficulties relating the values of the slopes of the graph, opting for a relationship opposite to the correct one (option B, 27%). In item 9, the most frequent error corresponds to a graph in which students also seem to understand the shape the graph should have, but have difficulties relating the values of the slopes of the graph, choosing absolute values for the slopes (option C: 19%; see item 9 in the [Appendix](#)). In these two items, the second most frequent errors are the same (item 16: option C, 14%; item 9: option B, 13%). In these choices, students make mistakes only in the sections of the graph in which the derivative is zero. Instead of setting a step type graph with a value of zero in this section, students choose the option in which the constant values are connected by a straight line.

Interestingly, the differences in the first most common error in both items appear to be due to slight differences in the graph shown. The first section of the graph of item 9 goes from negative values to positive values, which is different from what happens in the first section of item 16, which has only positive values. This subtle difference between the graphs (which is not important to the expert) seems to have a certain effect on the errors that are triggered in students. This is consistent with studies of science education that mention that superficial features of problems are very important for novices (Leonard, Gerace & Dufresne, 1999).

There are two other items of dimension 5, which evaluate the determination of the corresponding graph of the antiderivative from a graph (items 2 & 4). Item 2 evaluates the determination of the corresponding $f(x)$ graph from the $f'(x)$ graph, and item 4 evaluates the determination of the corresponding $f'(x)$ graph from the $f''(x)$ graph. These two related items have the same most frequent error (item 2: option D, 24%; item 4: option E, 28%). In this choice, students seem to understand the shape of the antiderivative graph, but have difficulties relating the absolute values of the slopes of the sections with slopes different from zero, opting for a relationship opposite to the correct one (see items 2 & 4 in the [Appendix](#)).

RECOMMENDATIONS FOR INSTRUCTION

This section addresses the fourth objective of the study: to establish recommendations for instruction of the concepts of derivative and antiderivative based on the results obtained from the TUG-C. McDermott (2001) suggests that every curricular change should originate from the research on students' understanding. The previous analysis of student performance presented in this article is part of such research on students' understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve in the context of calculus. Also, it allows us to establish specific recommendations for instruction on these concepts. Next, we summarize the most important findings derived from our analysis of the students' performance, and then we make some recommendations for instruction.

Since the distribution of the students' scores in the test shows a positive skew, we can state that the test presents numerous challenges for students. We noticed that students who are at the median of the distribution (8) had difficulty answering correctly 8 out of 16 items on the test. Since the topics covered on the test are concepts that the students should have learned in early mathematics and science courses at the university level, this result shows the need to modify instruction in order to increase students' conceptual understanding of the concepts of the derivative and the antiderivative.

Moreover, we observe that the value of the average of correct answers for every dimension is relatively low, tending to 50%, and that students have similar difficulties in these five dimensions. This shows that the need to modify the instruction should be done in the instruction of the skills and concepts evaluated in all of the five dimensions of the test.

Interestingly, in our analysis we found no significant differences when choosing the correct answer in any of the related items of the test. From this we notice: (a) that students' performance in the items of the dimension 1, that evaluate the determination of $f'(x)$ from the graph of $f(x)$, is similar to students' performance in the items of the dimension 2, that evaluate the determination of $f''(x)$ from the graph of $f'(x)$; (b) that students' performance in items of the dimension 3, that evaluate the determination of $\Delta f(x)$ from the graph of $f'(x)$, is similar to the students' performance in the items of dimension 4, that evaluate the determination of $\Delta f'(x)$ from the graph of $f''(x)$; (c) that students' performance in the items that evaluate the determination of the corresponding graph of the derivative from a graph in the two steps evaluated in the test is similar, and (d) that students' performance in the items that evaluate the determination of the corresponding graph of the antiderivative from a graph in the two steps evaluated in the test is also similar. These results could be positive, since we can infer that students are learning no matter whether we talk about the function, its first derivative or the second derivative. This could mean that students have a level of understanding of the relationships of derivatives and antiderivatives no matter the derivative order, which is encouraging. However, taking into account the low performance of students in the test we can also think that they are similarly lacking in understanding of the relationships in calculus, particularly in graphs. Also, if we take into account that only a student who answered the two related items correctly, has a complete understanding of the concept, from [Table 6](#), we observe that the proportion of students having a complete understanding of the concept is quite low and range from 21% for items 15 & 10 to 48% for items 13 & 3 and 5 & 12. There are a considerable proportion of students showing only a partial understanding of the concept, since they answer one related item correctly but the other incorrectly.

According to the classification of items by difficulty level, the most difficult items for students are the items of dimensions 3 & 4 that evaluate the identification of the variable whose antiderivative has the greatest change in a specific interval. These items have in common that they assess the maximum value of the antiderivative in an interval. Therefore, a general instructional recommendation is to specifically focus on teaching the skills to solve this type of items.

McDermott (2001) proposes also that persistent conceptual errors must be explicitly addressed during instruction. We identified the most frequent error for the related items of the test. Mathematics and Science teachers can use this catalog of errors when planning their instruction for the concepts of the derivative and the

antiderivative. Moreover, analyzing the most frequent errors identified in the previous section, we noticed that there are four errors that have a percentage of selection higher than 20% in both related items. The first error is to obtain the value of a derivative at a point of a curve (that is positive derivative) by dividing the abscissa by the ordinate of the point in the graph (item 1: option D, 26%; item 7: option C: 25%). The second and third errors are in the two items that evaluate the identification of the variable whose antiderivative has the greatest change in a specific interval. In these two items, we found that students don't choose the graph of the curve with the greatest area under it, but a graph of a curve whose slopes in the interval are always increasing (item 15: option D, 31%; item 10: option D, 31%), and a graph of a curve that increases in the middle of the interval and decreases in the other half (item 15: option B, 29%; item 10: option E, 28%). Finally, the fourth error can be found in the items that evaluate the determination of the corresponding graph of the antiderivative from a graph. In this error, students seem to understand the general shape of the antiderivative graph, but have difficulties relating the absolute values of the slopes of the sections with slopes different from zero, choosing a relationship opposite to the correct one (item 2: option D, 24%; item 4: option E, 28%). We recommend that mathematics and science teachers focus on these errors in particular when planning their instruction. The instructional materials that help to teach these topics should include sections in which students reflect on their own learning to realize that the concept and procedure of calculating a first derivative from a function are the same as that of calculating a second derivative from the first derivative of the function. The materials should also include sections in which students reflect to realize that the concept and procedure of calculating a change in a function from the graph of the first derivative are the same as calculating the change of the first derivative from the graph of the second derivative.

CONCLUSION

In this article, we first presented the "Test of Understanding Graphs in Calculus (TUG-C)" and its design process. Later, we showed that the TUG-C is a content-valid and reliable evaluation instrument with satisfactory discriminatory power according to the analysis recommended by mathematics and science education researchers (Beichner, 1994; Ding et al. 2006, Engelhardt, 2009). Then, we conducted a detailed analysis of students' understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve evaluated in the TUG-C. Finally, we outlined specific recommendations, based on the previous analysis, for the instruction of these concepts.

This article has two main implications. The first is that the test presented in the [Appendix](#) can be used by mathematics or science education researchers, and by teachers covering these concepts. It is important to note that the TUG-C is the first test for evaluating students' understanding of the concept of the derivative as the slope and the concept of the antiderivative as the area under the curve in the context of calculus that satisfies all the criteria recommended by mathematics and science education researchers. The test could be used to analyze students' understanding of these concepts in different institutions, to investigate students' learning performance, and to test the effectiveness of new instructional material based on research designed to increase student knowledge and understanding (Hake, 1998; Redish, 1999). The second implication is that the instructional recommendations established in this article could also be taken into account by researchers and teachers, and could guide the design of new instructional material intended to increase students' understanding of these concepts.

REFERENCES

- Bajracharya, R. R., Wemyss, T. M., & Thompson, J. R. (2012). Student interpretation of the signs of definite integrals using graphical representations. *AIP Conference Proceedings*, 1413, 111-114.
- Beichner, R. (1994). Testing student interpretation of kinematic graphs. *American Journal of Physics*, 62(8), 750-762.
- Bektasli, B., & White, A. L. (2012). The relationships between logical thinking, gender, and kinematics graph interpretation skills. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 12(48), 1-19.
- Chanpichai, N., & Wattanakasiwich, P. (2010). Teaching Physics with Basketball, *AIP Conference Proceedings*, 1263, 212-215.

- Christensen, W. M., & Thompson, J. R. (2012). Investigating graphical representations of slope and derivative without a physics context. *Physical Review Special Topics - Physics Education Research*, 8, 023101.
- Ding, L., Chabay, R., Sherwood, B., & Beichner R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical Review Special Topics - Physics Education Research*, 2(1), 010105.
- Engelhardt, P. (2009). *An Introduction to Classical Test Theory as Applied to Conceptual Multiple-Choice Tests in Getting Started in PER*, Vol. 2. Retrieved from <http://www.compadre.org/Repository/document/ServeFile.cfm?ID=8807&DocID=1148>.
- Epstein, J. (2013). The calculus concept inventory-measurement of the effect of teaching methodology in mathematics. *Notices of the American Mathematical Society*, 60(8), 1018-1026.
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(5), 989-1008.
- Hadjidemetriou, C., & Williams, J. (2002). Children's graphical conceptions, *Research in Mathematics Education*, 4(1), 69-87.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *American Journal of Physics*, 66, 64.
- Hammer, D. (2000). Student resources for learning introductory physics, *American Journal of Physics*, 68, S52.
- Hill, M., & Sharma, M. D. (2015). Students' representational fluency at university: A cross-sectional measure of how multiple representations are used by physics students using the representational fluency survey. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), 1633-1655.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60, 1-64.
- Leonard, W., Gerace, W., & Dufresne, R. (1999). Concept-based problem solving: Making concepts the language of physics. Retrieved from <http://www.srri.umass.edu/publications/gerace-1999cbp>
- Maries, A., & Singh, C. (2013). Exploring one aspect of pedagogical content knowledge of teaching assistants using the test of understanding graphs in kinematics. *Physical Review Special Topics - Physics Education Research*, 9, 020120.
- McDermott, L. C. (2001). Oersted medal lecture 2001: "Physics education research - The key to student learning". *American Journal of Physics*, 69(11), 1127-1137.
- McDermott, L. C., & Shaffer, P. (2001). *Tutoriales para Física Introductoria*. México, D. F., México: Pearson.
- McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55, 503-513.
- Meltzer, D. E. (2004). Investigation of students' reasoning regarding heat, work, and the first law of thermodynamics in an introductory calculus-based general physics course. *American Journal of Physics*, 72(11), 1432-1446.
- Mesic, V., Dervic, D., Gazibegovic-Busuladzic, A., Salibasic, D., & Erceg, N. (2015). Comparing the impact of dynamic and static media on students' learning of one-dimensional kinematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(5), 1119-1140.
- Nguyen, D., & Rebello, N. S. (2011). Students' understanding and application of the area under the curve concept in physics problems. *Physical Review Special Topics - Physics Education Research*, 7, 010112.
- Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics*, 14(1), 1-18.
- Pérez, N., Domínguez, A., & Zavala, G. (2010). Diseño de un instrumento para evaluar entendimiento de gráficas en cálculo. *24ALME*.
- Perez-Goytia, N., Dominguez, A., & Zavala, G. (2010). Understanding and interpreting calculus graphs: Refining an instrument. *AIP Conference Proceedings*, 1289, 249-252.
- Planinic, M., Ivanjek, L., Susac, A., & Milin-Sipus, Z. (2013). Comparison of university students' understanding of graphs in different contexts. *Physical Review Special Topics - Physics Education Research*, 9.

- Planinic, M., Milin-Sipus, Z., Katic, H., Susac, A., & Ivanjek, L. (2012). Comparison of student understanding of line graph slope in physics and mathematics. *International Journal of Science and Mathematics Education*, 10(6), 1393-1414.
- Pollock, E. B., Thompson, J. R., & Mountcastle, D. B. (2007). Student understanding of the physics and mathematics of process variables in P-V diagrams. *AIP Conference Proceedings*, 951, 168-171.
- Redish, E. (1999). Millikan Lecture 1998: Building a science of teaching physics. *American Journal of Physics*, 67, 562.
- Salinas, N. P., Alanís, J. A., Garza, J. L., Pulido, R., Santos, F. X., & Escobedo, J. C. (2000). *Elementos del Cálculo: Reconstrucción Conceptual para el Aprendizaje y la Enseñanza*, 1. México: Trillas.
- Salinas, N. P., Alanís, J. A., Garza, J. L., Pulido, R., Santos, F. X., & Escobedo, J. C. (2012). *Cálculo Aplicado: Competencias matemáticas a través de contextos*, 1. México: Cengage Learning.
- Serway, R. A., & Jewett, J. W. (2008). *Física para ciencias e ingeniería* (Vol. 1). México, D.F.: Cengage Learning.
- Sheskin, D. J. (2007). *Handbook of Parametric and Nonparametric Statistical Procedures* (4th ed.). Florida, Boca Raton: Chapman and Hall/CRC.
- Tejada Torres, S. E., & Alarcon, H. (2012). A tutorial-type activity to overcome learning difficulties in understanding graphics in kinematics. *Latin American Journal of Physics Education*, 6, Suppl. I, 285-289.
- Woolnough, J. (2000). How do students learn to apply their mathematical knowledge to interpret graphs in physics? *Research in Science Education*, 30, 259-267.
- Zavala, G., Tejada, S., Barniol, P., & Beichner, R. J. (2017). Modifying the test of understanding graphs in kinematics. *Physical Review Physics Education Research*, 13, 020111.

APPENDIX

DIAGNOSTIC QUESTIONNAIRE

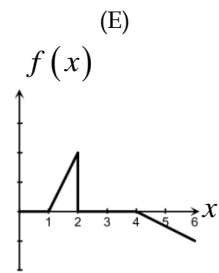
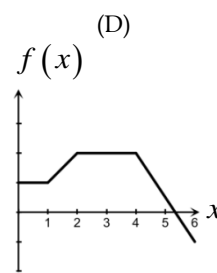
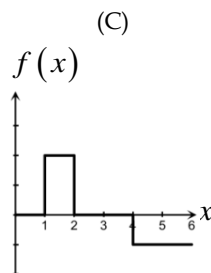
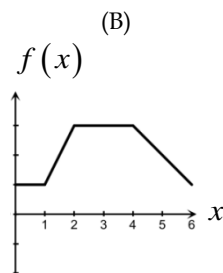
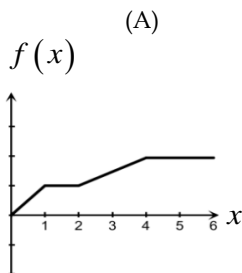
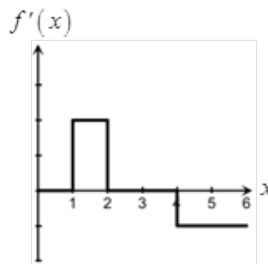
Notice that all the graphs in this questionnaire refer to a function f which depends on x , that is $f(x)$. Also, note that $f'(x)$ corresponds to the first derivative of the function with respect to x , and $f''(x)$ is the second derivative of the function with respect to x . Moreover, $f(x)$, $f'(x)$ and $f''(x)$ are graphed with respect to x , this means x is the variable on the horizontal axis.

1. The following figure shows the graph of a function $f(x)$. The first derivative of this function at $x = 4$ is about:

- (A) 0.5
- (B) 12
- (C) 2.5
- (D) 5.0
- (E) 20

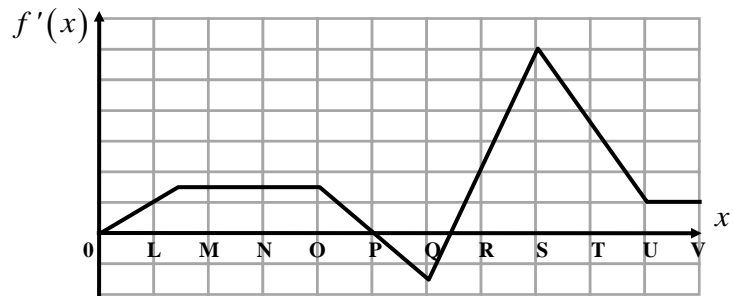


2. The following figure shows the graph of $f'(x)$, the first derivative of a function, in the interval from $x = 0$ to $x = 6$. Which of the following figures would represent the graph of the function $f(x)$ in that interval?

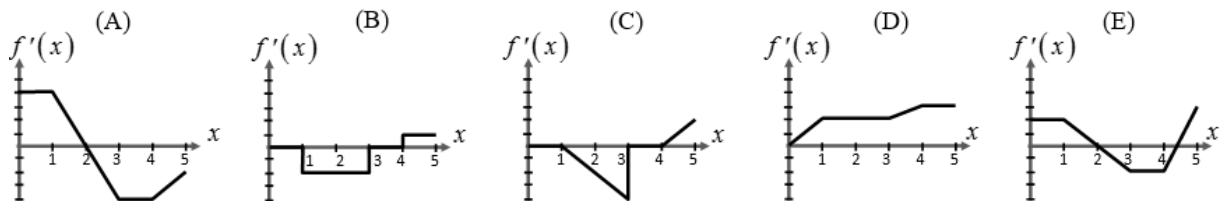
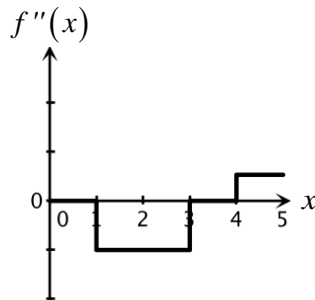


3. The following figure shows the graph of $f'(x)$, the first derivative of a function. Which of the following options corresponds to the case when the second derivative of the function is the most negative?

- (A) Q to S
- (B) O to Q
- (C) at Q
- (D) at S
- (E) S to U

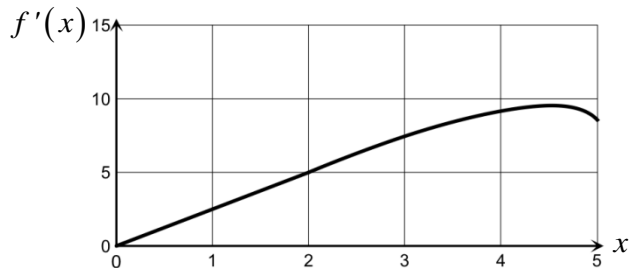


4. The following figure shows the graph of $f''(x)$, the second derivative of a function, in the interval from $x = 0$ to $x = 5$. Which of the following figures would represent the graph of the first derivative of the function in that interval?



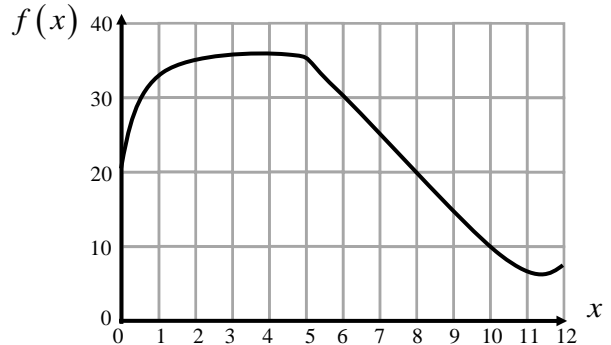
5. The following figure shows the graph of $f'(x)$, the first derivative of a function. If you wanted to know the change in the function $f(x)$ in the interval from $x = 0$ to $x = 2$, from the graph you would:

- (A) Read 5 directly off the vertical axis.
- (B) Find the area between the line and the horizontal axis by calculating $\frac{5 \cdot 2}{2}$.
- (C) Find the slope of that line segment by dividing 5 by 2.
- (D) Find the value multiplying 5 by 2.
- (E) Find the value by dividing 2 by 5.



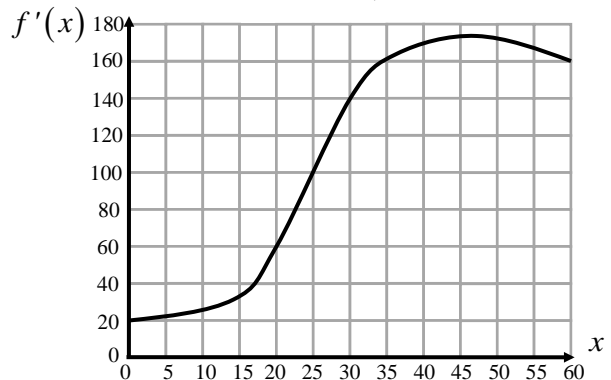
6. The following figure shows the graph of a function $f(x)$. The first derivative of this function at $x = 8$ is about:

- (A) -5.0
- (B) -2.0
- (C) -0.5
- (D) -2.5
- (E) 20



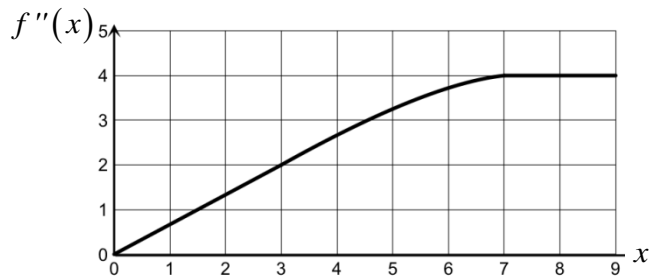
7. The following figure shows the graph of $f'(x)$, the first derivative of a function. At $x = 25$, the second derivative of the function is about:

- (A) 8.0
- (B) 2.0
- (C) 4.0
- (D) 100
- (E) 60

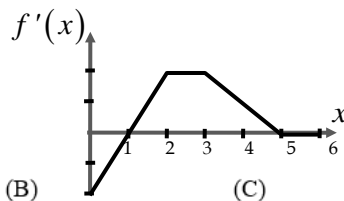


8. The following figure shows the graph of $f''(x)$, the second derivative of a function. The change in the first derivative of the function in the interval from $x = 0$ to $x = 3$ is about:

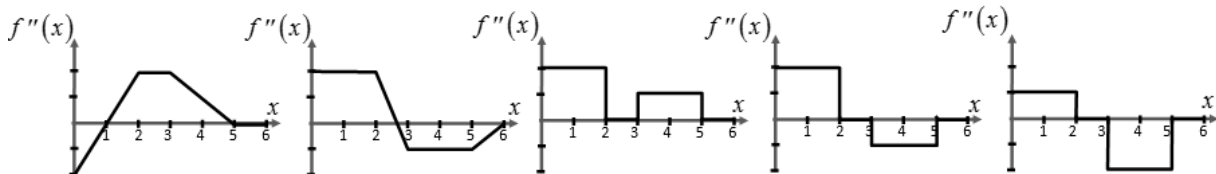
- (A) 1.5
- (B) 0.67
- (C) 2.0
- (D) 3.0
- (E) 6.0



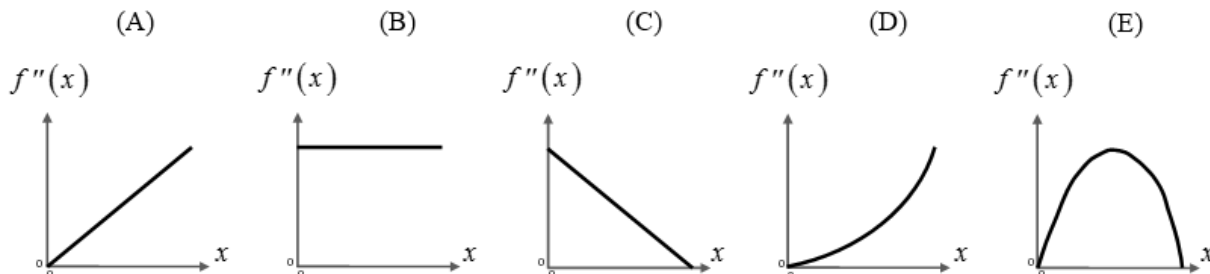
9. The following figure shows the graph of $f'(x)$, the first derivative of a function, in the interval from $x = 0$ to $x = 6$. Which of the following figures would represent the graph of the second derivative of the function in that interval?



- (A) (B) (C) (D) (E)

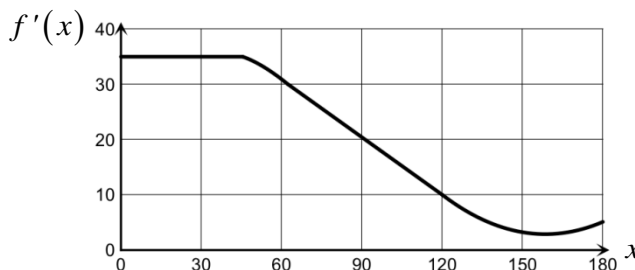


10. Graphs of the second derivative for five functions are shown below. All axes have the same scale. Which of the following graphs corresponds to the function with the greatest change in its first derivative in the interval?



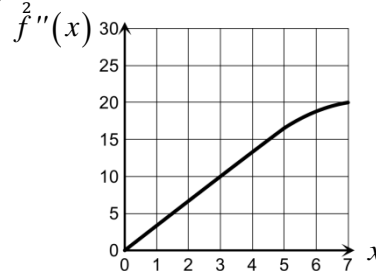
11. The following figure shows the graph of $f'(x)$, the first derivative of a function. At $x = 90$, the second derivative of the function is about:

- (A) -0.22
- (B) -0.33
- (C) -1.0
- (D) -2.0
- (E) 20



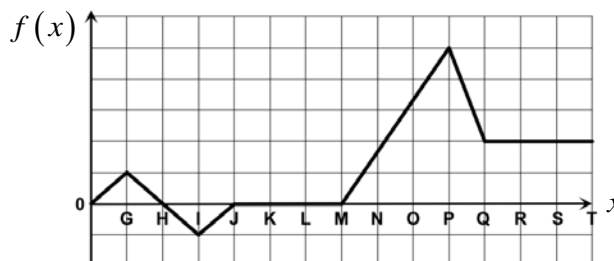
12. The following figure shows the graph of $f''(x)$, the second derivative of a function. If you wanted to know the change in the first derivative of the function in the interval from $x = 0$ to $x = 3$, from the graph you would:

- (A) Find the area between the line and the horizontal axis by calculating $\frac{10 \cdot 3}{2} \cdot \frac{10 \cdot 3}{2}$
- (B) Find the slope of that line segment by dividing 10 by 3.
- (C) Read 10 directly off the vertical axis.
- (D) Find the value by dividing 3 by 10.
- (E) Find the value by multiplying 10 by 3.



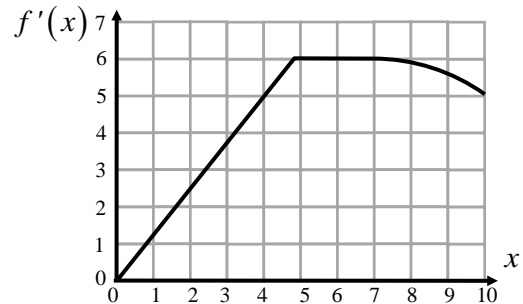
13. The following figure shows the graph of a function $f(x)$. Which of the following options corresponds to the case when the first derivative of the function is the most negative?

- (A) at I
- (B) P to Q
- (C) M to P
- (D) G to I
- (E) at P

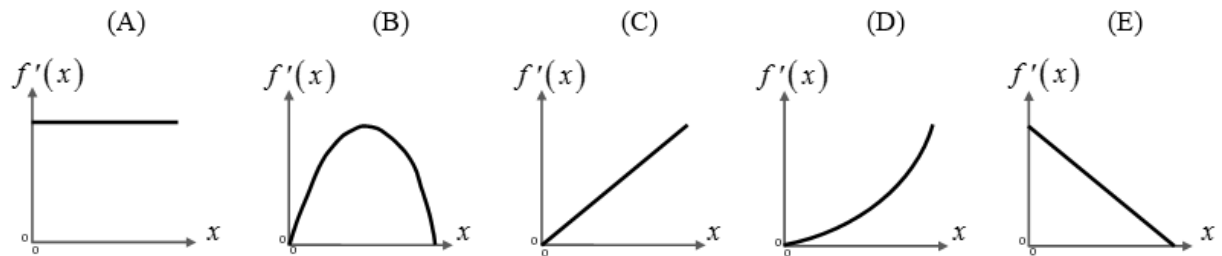


14. The following figure shows the graph of $f'(x)$, the first derivative of a function. The change in the function $f(x)$ in the interval from $x = 0$ to $x = 4$ is about:

- (A) 20
- (B) 10
- (C) 5.0
- (D) 1.25
- (E) 0.8



15. Graphs of the first derivative for five functions are shown below. All axes have the same scale. Which of the following graphs corresponds to the function $f(x)$ with the greatest change in the interval?



16. The following figure shows the graph of a function $f(x)$ in the interval from $x = 0$ to $x = 5$. Which of the following figures would represent the graph of the first derivative of the function in that interval?

