# Historical Reflections on Teaching the Fundamental Theorem of Integral Calculus

**David M. Bressoud**

**Abstract.** This article explores the history of the Fundamental Theorem of Integral Calculus, from its origins in the 17th century through its formalization in the 19th century to its presentation in 20th century textbooks, and draws conclusions about what this historical development tells us about how to teach this fundamental insight of calculus.

**1. INTRODUCTION.** Nothing is considered more basic to calculus than the Fundamental Theorem of Integral Calculus, which is commonly presented as follows:

**Fundamental Theorem of Integral Calculus (FTIC).** *For any function $f$ that is continuous on the interval* $[a, b]$,

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x), \quad \text{for } a < x < b, \tag{1}$$

*and, if $F'(x) = f(x)$ for all $x$ in* $[a, b]$, *then*

$$\int_a^b f(t)\,dt = F(b) - F(a). \tag{2}$$

Equation (1) is known as the *antiderivative part* of the FTIC for it shows how to use the definite integral to construct an antiderivative. Equation (2) is known as the *evaluation part* of the FTIC for it shows how to use an antiderivative to evaluate the definite integral.

There is a fundamental problem with this statement of this fundamental theorem: few students understand it. The common interpretation is that integration and differentiation are inverse processes. That is fine as far as it goes. The problem arises from the fact that this theorem assumes that the definite integral has been defined as a limit of Riemann sums. For most students, the working definition of the definite integral is the difference of the values of "the" antiderivative. When this interpretation of the theorem is combined with the common definition of integration, this theorem ceases to have any meaning.

The difficulty arises from the fact that the statement of the FTIC given in the first paragraph is a highly refined product of 19th century mathematics. Its language was created by Cauchy in 1823 and its packaging in this form as *the* fundamental theorem of integral calculus comes to us from du Bois-Reymond in 1876. In the past few decades, the adjective "integral" has been dropped, compounding the tendency to interpret it as equivalent to a statement of the inverse nature of differentiation and integration.

It is correct to claim the FTIC as a—perhaps even the—basic result of calculus. But it incorporates many conceptually difficult ideas. The historical development of this theorem can help us understand the conceptual difficulties along this road, difficulties we should expect our own students to encounter.

doi:10.4169/amer.math.monthly.118.02.099

The purpose of this article is to explore how this theorem has been understood and explained over the centuries, how it came to appear in the form we now find in our calculus texts, and what this tells us about the difficulties our students are likely to encounter as they try to understand it. The first two sections (Sections 2 and 3) look at the earliest development of this theorem during the 17th century as the connection between integration and differentiation came to be recognized. There are two very distinct streams of understanding: the geometric view as exemplified in the work of Leibniz and a more primitive dynamic understanding as exemplified by Newton. Sections 4 and 5 follow the development of this theorem in the 19th century, its incorporation into English-language textbooks in the years around 1900, and its evolution into the form we know today. The last section (Section 6) deals with the pedagogical implications of this story.

**2. LEIBNIZ AND THE GEOMETRIC FTIC.** Although Gottfried Leibniz and Isaac Newton are commonly credited with discovery of the FTIC, the development of this theorem did not end with them, nor did it begin with them. As Carl Boyer points out in the conclusion to *The History of the Calculus and Its Conceptual Development* [**3**], we need to recognize the contributions of Newton and Leibniz as steps along a continuum, a continuum that stretches back at least as far as Eudoxus of Cnidus and has continued through Euler, Cauchy, and Weierstrass.

When we turn to the 17th century, integrals and derivatives cannot be found as operators in their modern sense. Not even the objects of these operators, functions, existed in anything like our modern understanding. Instead of functions, the objects to which calculus is applied are curves. Integrals are understood to be areas and derivatives are defined by the ratio of sides of the *characteristic triangle*, the triangle formed by the horizontal axis, the line segment from this axis to the corresponding point on the curve (usually vertical but, in the 17th century, the ordinate was not always measured perpendicular to the axis), and the line tangent to the curve at that point. Rather than limiting himself to the derivative, Leibniz works with the *differential triangle*, a triangle similar to the characteristic triangle, but with sides representing the differentials of the variables and the slope of the hypotenuse representing the derivative.

Behind this geometric view of calculus, there is also a dynamic understanding that sees integration as an accumulation of a quantity described by its rate of change. This is the viewpoint favored by Newton when he speaks of fluents (variable quantities) and fluxions (the rates of change of such quantities). Even more than recognizing the inverse nature of integration and differentiation, the genius of Newton and Leibniz lies in their ability to move easily between the geometric and dynamic instantiations of calculus. Liebniz recognizes integration as an accumulation or summation of infinitesimal areas. Newton uses geometric models to reason about the relationships of acceleration, velocity, and distance

Today we tend to conflate the geometric and the dynamic aspects of calculus. After all, the only real difference between them is whether we view the independent variable as marking distance or elapsed time. But conceptually they are quite distinct. Calculus emerged because the geometric and dynamic conceptions of the integral and derivative came to be seen as manifestations of common general principles, but it took time and genius to extract those general principles. It should not surprise us when our students fail to grasp these principles immediately and intuitively.

We are also often too glib in simply asserting that the problems of areas and tangents are inverse to each other. It is a useful corrective to look back at how Leibniz himself wrestled with this complementarity.

D. J. Struik labels one particular passage from Leibniz, published in 1693, as "The Fundamental Theorem of Calculus":

> I shall now show *that the general problem of quadratures* [areas] *can be reduced to the finding of a line that has a given law of tangency* (*declivitas*), that is, for which the sides of the characteristic triangle have a given mutual relation. Then I shall show how this line can be described by a motion that I have invented. [**35**, pp. 282–283]

Leibniz seeks to find the area beneath a curve. He demonstrates how to construct an auxiliary curve for which the slope (the ratio of the sides of the characteristic triangle) is proportional to the vertical height of the original curve (the vertical height divided by a constant, $a$). See Figure 1. This sets up the opportunity to employ a summation argument to demonstrate that the area in question is proportional to the ordinate of the auxiliary curve. If an explicit formula for the auxiliary curve is known, that provides a formula for the area. In modern notation, if we can find a function $F$ for which the original curve is described by $y = F'(x)$, then the area is expressed in terms of $F$, yielding the evaluation part of the FTIC.

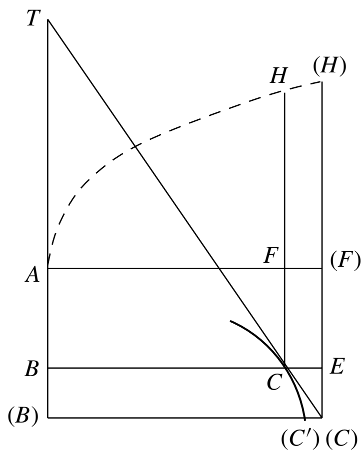



**Figure 1.** Leibniz's illustration for his proof of the FTIC.

A few explanatory words about Leibniz's illustration are in order. The line $A(F)$ is what we would call the horizontal axis and the line $T(B)$ is almost the vertical axis. It differs from our modern representation in that both above and below the horizontal axis it records the positive distance from the horizontal axis. Thus, $AH(H)$ is the curve beneath which we wish to find the area, and $C(C')$ is the curve whose derivative at $C$ is equal to $FH$, the ordinate of the curve $AH(H)$. Rather than superimposing these two positive curves, Leibniz has taken the graph of the antiderivative and reflected it across the horizontal axis. The values taken by the curve $C(C')$ are *increasing* because their distance from the horizontal axis is increasing.

The curve $C(C')$ is constructed so that the ratio of the sides of the characteristic triangle, $TB : BC$, is equal to the ratio of $FH$ to a constant length $a$:

$$TB : BC = FH : a. \tag{3}$$

To our modern eye, the introduction of this constant $a$ seems strange. We would simply construct the curve so that $FH = TB/BC$. But to Leibniz and his contemporaries, a ratio of lengths can only equal another ratio of comparable quantities. Since $FH$ is a length, it must be in ratio to another length. There is a useful pedagogical lesson in this. The historical importance of avoiding ratios of heterogeneous quantities hints at the difficulty we should expect students to have with such ratios.

Leibniz now assumes that the curve $AH(H)$ is specified by a relationship between $z$ (upward vertical displacement) and $y$ (horizontal displacement), so that $FH = z$ and $AF = y$, while the curve $C(C')$ is specified by a relationship between $x$ (downward vertical displacement) and $y$. He assumes, though never explicitly states, that this second curve passes through the point $y = 0$, $x = 0$. The differential triangle is given by $CE(C)$ with $dy = CE = F(F)$ and $dx = E(C)$. Their ratio is equal to the ratio of sides of the characteristic triangle,

$$TB : BC = dx : dy. \tag{4}$$

If we combine this with equation (3), we see that

$$dx : dy = z : a, \quad \text{or} \quad a\,dx = z\,dy. \tag{5}$$

The product $z\,dy$ is essentially the area $FH(H)(F)$, while the product $a\,dx$ is the area of the rectangle of height $E(C)$ and width $a$. Summing these incremental additions to the area, we get that $ax = \int z\,dy$. In other words, the area of $AHF$ is equal to the area of the rectangle whose height is $FC$ and whose width is the constant $a$. Leibniz has established that the area defined by $\int z\,dy$ is equal to the change in the value of the antiderivative over the interval $AF$.

As Victor Katz explains [**21**, p. 572], what enabled Leibniz to find this proof was his understanding of areas and tangents as manifestations of sums and differences. As such, their inverse nature becomes obvious.

The geometric version of the FTIC did not begin with Leibniz. We can find a very similar geometric proof in Isaac Barrow's *Lectiones geometricae* (*Geometric Lectures*) of 1670, where Barrow also uses the same technique of reflecting one of the curves across the horizontal axis. Struik also refers to this result as "The Fundamental Theorem of Calculus." Barrow actually proves the antiderivative part of the FTIC, and, appropriately, it is the curve that we identify with the derivative that he places below the horizontal axis. Barrow constructs what we would call an antiderivative, a curve above the axis whose height is proportional to the area under the original curve from some starting point up to the point above which the auxiliary curve is plotted, and he shows that the ratio of the sides of the characteristic triangle of this antiderivative is equal to the ratio of the height of the original curve to an appropriate constant.

What Barrow is doing is constructing an accumulator function, and Katz [**21**, p. 537] effectively argues that lying behind his geometric argument is a dynamic understanding, with the curve below the horizontal axis understood as the velocity and the curve above as the accumulated distance.

Barrow was not the first to give a proof of the geometric form of the FTIC. As Adolf Prag explained on the occasion of the tercentenary of James Gregory's birth [**32**] and Andrew Leahy has since expounded [**24**], the FTIC is contained within Gregory's *Geometriae pars universalis* (*Universal Part of Geometry*) published in 1668 [**15**]. Gregory showed how to find the length of a curve by finding the area under a related curve. In modern notation, he states and proves that for a suitable constant $c$, chosen so that ratios are equated to comparable ratios, the length of the curve defined by $y = f(x)$ from $x = a$ to $x = b$ is equal to the area under the curve defined by $y = c\sqrt{1 + (f'(x))^2}$ between these same limits. He then poses and solves the inverse problem: Given a curve defined by $y = g(x)$ over the interval $[a, b]$, find a related curve defined by $y = u(x)$ so that area under the first curve is equal to the length of the second curve. He knows that the slope of the tangent of $u$ is given by $c^{-1}\sqrt{g^2(t) - c^2}$. He now proves a statement equivalent to the antiderivative part of the FTIC, that if we define $u(x)$ in terms of the area beneath a curve whose ordinate is specified by $z/c$, in modern notation

$$u(x) = \frac{1}{c}\int_a^x z(t)\,dt,$$

then $z/c$ describes the slope of the tangent to $u$. It follows that

$$u(x) = \frac{1}{c}\int_a^x \sqrt{g^2(t) - c^2}\,dt.$$

Beyond the statement of the FTIC buried within his proof, something else important happens in this work: the recognition that area can be used to represent something else, in this case arc length. This theme of using area to describe the accumulation of some other quantity will be pursued in the next section.

It is worth mentioning that Gregory was not the first to discover the general formula for calculating arc length in terms of the area under a related curve. This honor goes to Hendrick van Heureat, whose work, which does not include the inverse problem, was published in 1659 in van Schooten's Latin edition of Descartes' *Geometry*.

The roots of the geometric form of the FTIC go even earlier. As Struik explains [**35**, p. 253], many people had been studying problems that involved the characteristic triangle. The geometric form of the FTIC does not require an algebraic representation of the curves. The earliest results on the problem of constructing a curve from knowledge about its tangent were obtained by Simon Stevin and Pedro Nuñez in the 16th century, and by Descartes and Torricelli in the early 17th.

Why then does Leibniz get the credit for the discovery of calculus that is denied Gregory or Barrow? As Katz points out [**21**, p. 539], Barrow never recognized the importance of his discovery. In "Patterns of Mathematical Thought in the later Seventeenth Century," D. T. Whiteside admits the impossibility of pinpointing the moment at which (or even the person with whom) calculus came into existence, but he does offer the following:

> It is very tempting, nevertheless, to admit two criteria into a working definition [of calculus] (without excluding others); first, that differentiation and integration be seen as inverse procedures; and, secondly, that both be defined with respect to an adequate algorithmic technique. [**40**, p. 365]

While both Gregory and Barrow possessed much of this "adequate algorithmic technique," they never married this technique to the general application of this theorem. This is perhaps the most important pedagogical lesson we can draw from this episode in the history of calculus: it is precisely in the combination of algorithmic technique with a grasp of the full meaning of the FTIC that calculus becomes a useful tool.

**3. NEWTON AND THE DYNAMIC FTIC.** The dynamic understanding of the FTIC views the function to be integrated as a rate of change and the definite integral as an accumulator of this change. This idea sits behind Leibniz's integral as a sum of infinitesimals, but it is most explicit in the calculus as exposited by Newton. The prime example of the dynamic approach views the derivative as a velocity, with distance traveled as the accumulation of small increments of distance that are proportional to the velocity at a given time.

Newton had worked out his ideas of calculus in Lincolnshire during the plague years of 1665–1666. His result on the rate of change of area can be found in his manuscript, "The October 1666 Tract on Fluxions," written two years before Gregory published his *Geometriae pars universalis*. In Problem 5, he asks "To find the nature of the crooked line whose area is expressed by any given equation" [**28**, p. 427]. In other words, if we have an expression for the area under a curve as a function, $y$, of the abscissa (distance $ab$ in Figure 2), then how do we find the equation of the bounding curve? His solution is to observe that "the motion by which $y$ increaseth will bee $bc = q$," the ordinate of the curve. This is the antiderivative part of the FTIC: the rate of change of the area is given by the ordinate of the bounding curve.

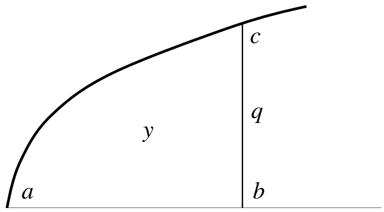



**Figure 2.** Newton's illustration for his statement of the FTIC.

The evaluation part of the FTIC appears in Problem 7: "The nature of any crooked line being given to find its area, when it may bee" [**28**, p. 430]. From the solution to Problem 5, the known ordinate $bc$ is the rate of change of the area. If an antiderivative is known, it provides the formula for the area. Newton then illustrates the power of this method by handling a nontrivial example: the area under the curve $y = ax/\sqrt{a^2 - x^2}$, a problem that is easily solved since he earlier showed that the fluxion or derivative of $(2c/nb)\sqrt{a + bx^n}$ is $cx^n/x\sqrt{a + bx^n}$.

The key to the dynamic understanding of the FTIC is the ability to recognize the ordinate as the rate of change of the area. This is intimately connected to an older problem, that of graphing velocity as a function of time and recognizing that accumulated distance is the area beneath the curve of this graph. This insight has a pedigree that takes us back at least to the 14th century.

The scholars of Merton College, Oxford explored the relationship between velocity and distance. In 1335, William Heytesbury published his *Regule solvendi sophismata* (*Rules for Solving Sophisms*) which included a remarkably modern definition of instantaneous velocity: "In nonuniform motion, however, the velocity at any given instant will be measured by the path which would be described by the most rapidly moving point if, in a period of time, it were moved uniformly at the same degree of velocity with which it is moved in that given instant, whatever [instant] be assigned" [**8**, p. 236]. Heytesbury went on to observe the Mertonian rule, that the distance travelled by an object under uniformly increasing velocity, starting at $V_o$ and ending at $V_t$, is equal to the distance travelled over the same interval of time with constant velocity $V_m = (V_o + V_t)/2$.

Other scholars at Merton College began to explore the representation of velocity by line segments. As Katz points out [**21**, p. 356], they probably were inspired by the Aristotelian view of time, distance, and length as *magnitudes*, infinitely divisible quantities in distinction to *numbers* that are based on indivisible units. Once they included velocity as a magnitude, it was a short step to representing it as distance. This approach was brought to fruition around 1350 in the *Tractatus de configurationibus qualitatum et motuum* (*Treatise on the Configuration of Qualities and Motions*) by Nicole Oresme. He represented velocity by vertical distances so that Heytesbury's rule became a geometric observation (see Figure 3). Significantly, Oresme explains that, however the velocity might vary, the total area of the figure expresses the total distance [**8**, p. 364].

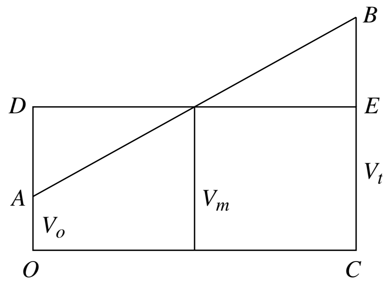



**Figure 3.** Oresme's demonstration of the Mertonian rule.

In the journal of the Dutch scientist Isaac Beeckman from the year 1618, we find the first explanation of the Mertonian rule in terms of a limit. In studying the distance traveled by a point whose velocity increases uniformly from zero (line $AB$ in Figure 4), Beeckman approximates the distance by replacing the triangle with circumscribed rectangles obtained by adding the areas of the small triangles: $k, l, \ldots, q, r$. The width of each rectangle represents one moment or interval of time. Since the velocity is constant in that moment, the area of the rectangle represents the distance traveled in that time. He then observes that as the moments of time become smaller, the additional area to the right of the line representing the velocity also becomes smaller until the sum of these additional areas "would be of no quantity when a moment of no quantity is taken" [**8**, p. 418]. It follows that the area of the triangle *ABC* represents the total distance.

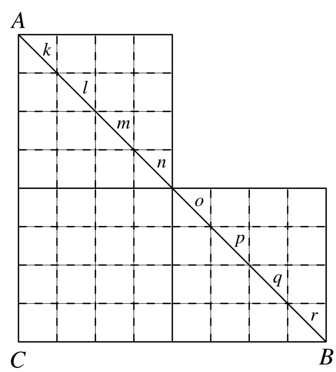



**Figure 4.** Beeckman's illustration. Time axis is vertical from $A$ to $C$.

The use of limit arguments applied to areas and volumes is ancient, but this is the first time it is explicitly applied to accumulating distance given velocity of known variability. This is the essence of the evaluation part of the FTIC, that if the rate of accumulation of any quantity can be represented as a distance above an axis representing time and we consider the curve that records these distances, then the area beneath this curve describes the accumulation of that quantity.

Beeckman was a friend and teacher of René Descartes. It is not clear how influential he was on the development of calculus, but these ideas were in the air. One finds a similar but less precise argument in Galileo's last great work, *Dialogues and mathematical demonstrations concerning the Two New Sciences* [**13**, Book III, Theorem I]. A comparison of the approaches of Oresme and Galileo can be found in [**29**].

The pedagogical lesson to be drawn from this discussion is that there is a considerable conceptual step that must be made from recognizing area as a limit of approximating rectangles to seeing area as a vehicle for expressing accumulation of any quantity for which the rate of accumulation is known. But once this step has been made, one is really quite close to the FTIC.

**4. THE FTIC IN THE 19TH CENTURY.** While calculus as understood in the late 17th and early 18th centuries was recognized to have broad applications to variable phenomena, it was always presented as a tool for analyzing curves. Euler was the person chiefly responsible for the shift from an analysis of curves to an analysis of functions, and this was cemented in his *Introductio in analysin infinitorum* (*Introduction to Analysis of the Infinite*) of 1748. Yet, however they might formally be defined, functions were still understood to be limited to those relationships that could be expressed in terms of a standard repertoire of basic functions, a repertoire that gradually expanded to include all functions expressible as power series.

Although they relied heavily upon it, the mathematicians of the 18th century paid little explicit attention to the FTIC, and no one in this century described this theorem as "fundamental." We can see this neglect reflected in the texts of Sylvestre-François Lacroix. Lacroix authored the most widely read calculus books of the first half of the 19th century. His massive *Traité du calcul différentiel et du calcul intégral* of 1797–1798 was written to consolidate all knowledge of calculus. The shorter *Traité Élémentaire*, first published in 1802, became *the* standard calculus textbook in both French and English. Lacroix's definition of integration was common for his time:

> Integral calculus is the inverse of differential calculus. Its goal is to restore the functions from their differential coefficients. [**22**, vol. 2, p. 1]

Perhaps the most surprising aspect of his books on calculus is that—right through the sixth edition of *Traité Élémentaire* published in 1837, the last edition to be printed during Lacroix's lifetime—nowhere does Lacroix explicitly justify any result that is equivalent to either form of the FTIC. His first chapter on integration in the *Traité Élémentaire* explores the basic methods of reversing differentiation. The next chapter is on applications to plane areas, curve lengths, surface areas, and volumes. He begins by simply stating that integration can be used to find areas under curves and then launches into the use of the rules of integration to calculate these areas.

A seventh edition of *Traité Élémentaire* was issued in 1867 with additional notes by Charles Hermite and Joseph Alfred Serret. Serret adds the observation that $\int_{x_0}^{X} f(x)\,dx$, which is defined to be $F(X) - F(x_0)$ where $F$ is any function whose derivative is $f$, is equal to

$$\lim \sum_{x=x_0}^{x=X} f(x)\,dx,$$

"which shows that the definite integral is the limit, for $dx = 0$, of the sum of values of the differential $f(x)\,dx$ as $x$ varies from $x_0$ to $X$ over intervals of length $dx$" [**23**, vol. 2, p. 174].

The limit definition of the definite integral was introduced by Cauchy in 1823. It is instructive to compare Cauchy's treatment, which is entirely modern, with the traditional approach taken by his colleague, Siméon-Denis Poisson, only three years earlier. Poisson seeks to show that the difference of the antiderivatives is equal to the integral:

$$\int f(x)\,dx = F(b) - F(a). \tag{6}$$

> This quantity $F(b) - F(a)$ is what is called the *definite* integral, taken from $x = a$ to $x = b$, while $F(x) + c$ is the general or indefinite integral. [**31**, p. 319]

He subsequently refers to equation (6) as "the fundamental proposition of the theory of definite integrals," the first time that I am aware of that anyone referred to this theorem as "fundamental." His proof relies on Lagrangian error analysis and concludes with an appeal to the Leibnizian view of $\int f(x)\,dx$ as an infinite sum of products in which $dx$ is understood to be an infinitesimal.

Poisson assumes that $F$, an antiderivative of $f$, possesses a power series expansion valid over the interval $[a, b]$. It follows that for each pair of positive integers $j \leq n$, there is a $k \geq 1$ and a function $R_j$ such that

$$F(a + jt) = F\big(a + (j-1)t\big) + tf\big(a + (j-1)t\big) + t^{1+k} R_j(t),$$

where $t = t(n) = (b - a)/n$. It is then possible to express the difference of the antiderivatives as

$$\begin{aligned} F(b) - F(a) &= \sum_{j=1}^{n} \big[F(a + jt) - F\big(a + (j-1)t\big)\big] \\ &= \sum_{j=1}^{n} tf\big(a + (j-1)t\big) + t^{1+k} \sum_{j=1}^{n} R_j(t). \end{aligned}$$

Poisson now asserts that the functions $R_j(t)$ stay bounded. In fact, we know by the Lagrange remainder theorem that we can take $k = 1$ and these functions are bounded by the supremum of $|f'(x)|/2$ over all $x$ in $[a, b]$. It follows that as $n$ increases, the second summation can be made as small as we wish. In Poisson's own words, the limit of the first summation "is the sum of the values of $f(x)\,dx$ as $x$ increases by infinitesimal amounts from $x = a$ to $x = b$."

Whereas Poisson implicitly assumes that all functions are analytic, Cauchy begins his study of integration in Lecture 21 of *Résumé des leçons ... sur le calcul infinitésimal* (*Summary of lectures ... on infinitesimal calculus*) with the explicit statement that the only restriction on the functions he will study is that they are continuous. He forms what today we would call a left-hand Riemann sum,

$$(x_1 - x_0)f(x_0) + (x_2 - x_1)f(x_1) + \cdots + (X - x_{n-1})f(x_{n-1}),$$

uses the intermediate value theorem to show that each of these sums can be expressed as $(X - x_0)f(x_0 + \theta(X - x_0))$ for some $\theta$ strictly between 0 and 1, and justifies that this approaches a limit as "the values of the intervals become very small and the number $n$ very large" [**7**, p. 81].

As an interesting aside, Cauchy considers the question of how to denote the definite integral, presenting three options then in use:

$$\int_a^b f(x)\,dx, \quad \int f(x)\,dx \begin{bmatrix} a \\ b \end{bmatrix}, \quad \int f(x)\,dx \begin{bmatrix} x = a \\ x = b \end{bmatrix}.$$

Fortunately for us, he chose the first of these, the notation invented by Joseph Fourier and popularized in his 1822 *Théorie analytique de la chaleur* (*Analytic Theory of Heat*).

There follow twenty pages of results on definite integrals before he tackles indefinite integrals. He begins Lecture 26, labeled simply *Indefinite integrals*, by defining

$$\mathcal{F}(x) = \int_{x_0}^{x} f(x)\,dx.$$

He uses the previously established mean value theorem for integrals,

$$\mathcal{F}(x+a) - \mathcal{F}(x) = a\, f(x+\theta a), \quad \text{for some } \theta,\ 0 < \theta < 1,$$

and the assumption that $f$ is continuous to prove that $\mathcal{F}'(x) = f(x)$, the antiderivative part of the FTIC. He then proves, from the mean value theorem, that any function whose derivative is 0 must be constant, and therefore any two antiderivatives of $f$ must differ by a constant. Since $\mathcal{F}$ is an antiderivative of $f$, it follows that if $F$ is any antiderivative of $f$, then

$$F(x) - F(x_0) = \mathcal{F}(x) - \mathcal{F}(x_0) = \int_{x_0}^{x} f(x)\,dx,$$

the evaluation part of the FTIC. This is still the standard proof of the FTIC. As the title of Lecture 26 makes clear, Cauchy's intent in proving this result was to establish the connection between his definition of the definite integral and the common conception of integral as antiderivative.

Despite its importance, neither Cauchy nor the later French mathematicians of the 19th century referred to this theorem as "fundamental," with the sole exception of Charles de Freycinet who employed this adjective liberally throughout his calculus text of 1860 [**10**]. The designation FTIC seems to have originated in Berlin. Paul du Bois-Reymond used it in the appendix to his 1876 paper on Fourier series [**11**], where it appears to be a recognized title rather than a simple description. He begins the appendix by describing this as "the most important and useful theorem of integral calculus," states both parts (though he highlights the evaluation part), gives what is essentially Cauchy's proof, and refers to it throughout by the full title, *Fundamentalsatz der Integralrechnung*.

The introduction of Fourier series in the early 19th century had expanded the notion of function, and under the influence of Dirichlet, Weierstrass, and others it came to take on the full generality we know today, including the possibility of being nowhere continuous. In his *Habilitation* of 1854, *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe* (*On the representability of a function by a trigonometric series*), Bernhard Riemann explores necessary and sufficient conditions for a function to give rise to a trigonometric series that converges to the function itself. He sets out to study the most general functions possible, including those that are nowhere continuous. He begins by asking, "What is to be understood by $\int_a^b f(x)\,dx$?" [**2**, p. 22]. His response is what today we know as the Riemann integral, the limit of a sum of terms of the form $f(x_i^*)\,(x_i - x_{i-1})$ where, as with Cauchy, the intervals $[x_{i-1}, x_i]$ can be of unequal length. Riemann introduces the extra freedom that $x_i^*$ can be any point in $[x_{i-1}, x_i]$. The purpose of adding this freedom is to facilitate the exploration of the question: how discontinuous can a function be and still be integrable? As he immediately shows by example, a function can be discontinuous on a dense set of points yet still integrable.

Riemann's thesis was published in 1867. It ushered in the modern age that seeks to minimize the hypotheses of analysis. In 1880, du Bois-Reymond explored what happens to the FTIC in the case that $F'$ is not continuous [**12**]. Removing the assumption of continuity greatly weakens the conclusions. Gaston Darboux had already observed that Riemann's example of an integrable, nowhere continuous function provides an example of a function for which the antiderivative part does not hold. Vito Volterra exploded the hope that continuity might not be needed for the evaluation part in 1881 [**39**] when he constructed a function that is differentiable everywhere and whose derivative stays bounded, yet the derivative is not integrable (see [**4**, pp. 89–94]).

The Lebesgue integral goes a long way toward enabling a more satisfactory statement of the FTIC, though even it does not guarantee that the evaluation part holds for every differentiable function $F$ (see [**4**, pp. 291–298]).

The pedagogical point is that the statement of the FTIC given at the start of this article is a product of centuries of development. It has been refined and polished so that it can be considered in the context of functions of full generality and has been totally divorced from its context-rich origins. This is not to say that it should not be presented in first-year calculus, but rather that, without discussion of its development, we should not be surprised when it fails to be meaningful. For first-year calculus, a less precise and more intuitive statement can be much more useful.

**5. FTIC IN ENGLISH-LANGUAGE TEXTBOOKS.** The earliest appearance of the phrase "Fundamental Theorem of the Integral Calculus" in English that I have been able to track down is in Daniel A. Murray's *An Elementary Course in the Integral Calculus* [**26**], published in 1898. Murray, a Canadian who had studied in Berlin and Paris before earning his Ph.D. at Johns Hopkins, was an instructor at Cornell at the time. It may have been in Berlin that he became aware of this theorem by this designation. He defines the definite integral as a limit of sums and employs the mean value theorem to show that it can be evaluated by finding an antiderivative. His proof, which does not explicitly assume the continuity of the function to be integrated, is problematic. One indication that the designation as the FTIC was not yet common is the fact that five years later, when Murray published a general introduction to calculus, *A First Course in Infinitesimal Calculus* [**27**], the treatment is identical, but he never uses the phrase "fundamental theorem."

It appears that the term was popularized in English in 1907 by Ernest W. Hobson in his influential book *The Theory of Functions of a Real Variable and the Theory of Fourier's Series* [**19**]. Hobson wrote his book to bring the latest developments in real analysis, especially Lebesgue's theory of measure and integration, to English-speaking mathematicians. His text contains full statements and proofs of the FTIC, both for the Riemann integral and for the Lebesgue integral, and eventually for the Denjoy integral.

A possible indication of the impact of Hobson's book can be traced through two influential though very different books on calculus that appeared at about the same time. William A. Granville's *Elements of the Differential and Integral Calculus* [**14**] was published in 1904 as a general introduction to calculus. Over the succeeding decades, it would become the standard American introduction to the subject. G. H. Hardy's *A Course of Pure Mathematics* [**16**] appeared in 1908. It sets out a rigorous and careful introduction to the subject. Though results equivalent to the FTIC appear in both books, neither makes any mention that these are fundamental theorems within calculus.

The rapidly increasing awareness of the FTIC under that name is reflected in Arthur Berry's 1910 review of Hardy's book [**1**]. One of his few complaints is that "the proof of the fundamental theorem of the integral calculus, though contained in §§137–142, is nowhere very clearly enunciated or proved." By the time the second edition of *A Course of Pure Mathematics* appeared in 1914, Hardy includes an explicit statement of the theorem, appropriately labeled.

An explicit reference to the FTIC as such was added to the second (1911) edition of Granville's *Elements of the Differential and Integral Calculus*. While Hardy found this edition to be "on the whole clear, readable, and reasonably accurate" [**17**], he focused particular criticism on Granville's treatment of integration: "Another part of the book which is unsatisfactory, because the foundations have not been properly laid, is that which deals with integration as summation, the Fundamental Theorem of the Integral Calculus, and so on." By 1913, Hardy not only refers to this theorem explicitly, he chooses to capitalize the key words.

Both Hardy and Granville introduce the definite integral in the context of finding the area beneath a (presumably) positive function, and both prove what amounts to the antiderivative part of the FTIC. Granville states this as:

> **Theorem.** The differential of the area bounded by any curve, the axis of $X$, and two ordinates is equal to the product of the ordinate terminating the area and the differential of the corresponding abscissa. [**14**, p. 355]

Hardy's formulation is simpler:

> Thus, the ordinate of the curve is the derivative of the area, and the area is the integral of the ordinate. [**16**, p. 238]

Using this result, both authors show that the area can be computed as the difference of values of an antiderivative (the evaluation part of the FTIC).

In Richard Courant's *Differential and Integral Calculus* [**9**], he devotes a section to "the fundamental theorems of the differential and integral calculus." He singles out the antiderivative part,

$$\frac{d}{dx}\int_a^x f(u)\,du = f(x),$$

when $f$ is a continuous function, as "the root idea of the whole of the differential and integral calculus." The remainder of this section contains a number of theorems that build to the evaluation part of the FTIC.

In the first edition of George Thomas's *Calculus and Analytic Geometry* [**36**], the book that would come to dominate the American calculus textbook market during its author's lifetime, Thomas privileges the evaluation part of the theorem. He begins with a discussion of an area function, shows that the derivative of this area function is the value of the ordinate, uses that to demonstrate that the area can be computed using an antiderivative, and then defines area rigorously as a limit of approximating sums and designates this limit by the notation of the definite integral. He has an interesting statement of the FTIC that avoids the notation of the definite integral until the very end:

> Let $x_k^*$ be any value of $x$ in the $k$th subinterval, $x_k \leq x_k^* \leq x_{k+1}$, and form the sum
>
> $$S_n = \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k.$$
>
> Then as the number $n$ of subdivisions increases indefinitely while the lengths of the individual subintervals approach zero, the sums $S_n$ have a limit given by
>
> $$\lim \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k = F(b) - F(a),$$
>
> where $F(x)$ is any function whose derivative is $f(x)$. [**36**, p. 159]

In a concluding sentence, he recalls that this limit is denoted by the definite integral.

At some point, perhaps as late as the 1950s, authors of calculus texts began to drop the adjective "integral," referring to this result simply as the Fundamental Theorem of Calculus. The earliest such usage I can find is in Gaylord Merriman's *Calculus: An Introduction to Analysis, and a Tool for the Sciences* of 1954 [**25**]. R. Creighton Buck's *Advanced Calculus* of 1956 [**5**] still called it the Fundamental Theorem of Integral Calculus, a label that was used in textbooks at least as late as 1968 when A. Wayne Roberts published his *Introductory Calculus* [**33**].

Tracking the use of the terms "fundamental theorem of integral calculus" and "fundamental theorem of calculus" through the 1950s and '60s, one finds authors who use both within the same book or article. In fact, in the index to Merriman's *Calculus*, it is listed as "Fundamental Theorem of integral calculus" (to identify it among several fundamental theorems), while in Buck's index it is listed simply as "Fundamental theorem of calculus." When both are used, the former is usually the correct, formal name, while the latter is used more informally, often to save space in a title, running head, or index. It appears that by the 1970s, by which time any ambiguity in referring to this result simply as the fundamental theorem of calculus had disappeared, most authors chose to use only the shorter designation.

**6. PEDAGOGICAL IMPLICATIONS.** There are three big lessons that I draw from this overview of the historical development of the FTIC. The first two arise from the observation that, given the long time it took before the mathematical community was able and chose to define integration by means of a definite integral that is a limit of Riemann sums, it should not surprise us that our students have difficulty with this definition.

The first lesson is that it never will be enough simply to introduce integration as a limit of sums. This limit, taken over all pairs consisting of a partition of the interval in question and a set of tags or points from subintervals, is conceptually complex. If we want students to understand integration as a limit, then they need experience working with these sums in contexts that lead them to appreciate the importance of this definition. Working with Riemann sums does not mean forcing students to evaluate the area under a parabola using the sum of squares formula. There is a common belief that students who have struggled through this will then appreciate the computational ease of invoking the FTIC. My own experience is that they are more likely to resent being forced to do this problem the hard way when there was a much easier route, and the resentment is only compounded when many or most of the students arrive in college having already seen enough integration to know there is an easier approach.

We need to keep in mind Guershon Harel's *Necessity Principle*: "Students are most likely to learn when they see a need for what we intend to teach them, where by 'need' is meant intellectual need" [**18**, p. 501]. Students require experience with a variety of problems for which the easiest and most natural approach is in terms of an accumulation that can be modeled with a sum of products, multiplying the rate of accumulation by the small increment in the independent variable, problems for which the evaluation of the definite integral is the last step. The heart of the intellectual activity should be converting a problem into an accumulation that can be calculated via a limit of appropriate sums.

In theory, this could be done with problems of areas and volumes. Unfortunately, it is impossible to find a college calculus course to which no one arrives aware of the connection between integrals and areas. It is even difficult to find a college class in which no one has seen the formula for computing a volume of a solid of revolution in terms of the definite integral. If we want students to see an intellectual need for evaluating limits of Riemann sums, then we need to give students extensive experience with good and unfamiliar problems that involve accumulation, problems that require thought in constructing approximations to something that is accumulating continuously and for which the evaluation of the limit can be accomplished via the FTIC. These problems must be tailored to the students in the class so that they are challenging but not overwhelming, but I know of no other way of getting students to understand and appreciate the core idea of the FTIC.

The second lesson is that, despite our efforts to define integration as a limit of sums, the working definition of integration for most students will continue to be antidifferentiation. This tendency is aggravated by the fact that the majority of students in college calculus arrive with the antidifferentiation definition deeply embedded from their high school experience with calculus. Even those students who arrive as *tabulae rasae* quickly learn from peers and their own experience with the way we assess their knowledge of integration that it is most efficient to think of integration as antidifferentiation. When discussing the FTIC, we must be prepared to deal with the fact that many students will persist in thinking of integration as the inverse process of differentiation. Because of this, when presenting this theorem we must:

- remind students that the definite integral is shorthand for the limit of Riemann sums,
- point out that equation (1) states that we can use this limit to produce an antiderivative for *any* continuous function, and
- point out that equation (2) states that, when an antiderivative is known, this limit can be evaluated in terms of that antiderivative.

I also believe that we should re-introduce the adjective "integral" into the name of this theorem, emphasizing the fact that it tells us about the dual nature of integration, which can be viewed either as a limit of sums or as antidifferentiation.

The third lesson draws on the historical development of the concepts of calculus in the 17th century. Integration and differentiation have two distinct conceptual settings, one geometric and the other dynamic. We mathematicians are sufficiently comfortable moving between them that we often forget how difficult it can be for students to grasp their equivalence. The historical lesson is to focus first on the dynamic understanding, and then to use this to build the geometric realization of the theorem.

With regard to the third lesson, Patrick Thompson, Marilyn Carlson, and others [**6, 20, 30, 37, 38**] have explored the pedagogical obstacles that students must overcome before they can comprehend the dynamic FTIC. The very first step is to understand functions as descriptions of covarying relationships. Most students think of functions as static objects, either algebraic expressions—in Thompson's words, as "a short expression on the left and a long expression on the right, separated by an equal sign" [**37**, p. 164]—or as the geometric object that is the graph of the function. We cannot expect students to comprehend the dynamic view of the FTIC unless they are able to see functions as describing a dynamic relationship between covarying quantities.

Understanding covariation is only the first step toward the FTIC. Next comes accumulation. We would do well to require our students to formulate their own explanations of the Mertonian rule. Accumulation is hard. A recent article in *Science* [**34**] describes a survey of graduate students at MIT, students with undergraduate degrees in economics or the sciences and thus who had studied calculus, who were asked to explain the behavior of an accumulation function when a nonconstant rate of accumulation was specified. Even these students had difficulty with this task.

The next step is to consider the rate of change of the accumulation function. Rate of change is already a difficult concept. Compounding this with an accumulator makes it all the more difficult. But once all of these pieces are in place and are understood, the FTIC is virtually self-evident. Newton never proves it. He simply observes that, of course, the rate of change of the accumulated quantity is the rate at which that quantity is accumulating. Calculus emerges from the awareness of the power of this observation.

**7. CONCLUSION.** Many years ago, my eyes were opened when I dared to probe what students really take away from their experience of calculus. I was particularly horrified by one undergraduate who had completed two full years of calculus yet, despite all my efforts to evoke something more, was able to remember nothing about the meaning of calculus beyond recalling that differentiation is a process of turning functions into "simpler" functions—in the sense that quadratic polynomials become linear and cubic polynomials become quadratic—and that integration is the reverse of this process. As we think about how we should teach the FTIC, we must keep in mind what it is that we want students to remember from this course, and then we must work hard to ensure that they do.

**ACKNOWLEDGEMENTS.** I wish to thank Larry d'Antonio and Ivor Grattan-Guinness who helped me track down the emergence of the adjective "fundamental" with respect to this theorem, the referees who provided much useful criticism and many suggestions for additional sources, and, most especially, David Smith who taught me to probe student understanding.

REFERENCES

1. A. Berry, review of *A Course of Pure Mathematics* by G. H. Hardy, *Math. Gaz.* **5** (1910) 303–305. doi:10.2307/3602606
2. G. Birkhoff, *A Source Book in Classical Analysis*, Harvard University Press, Cambridge, MA, 1973.
3. C. B. Boyer, *The History of the Calculus and its Conceptual Development*, Dover, New York, 1959; reprint of *The Concepts of the Calculus*, Hafner, New York, 1949.
4. D. M. Bressoud, *A Radical Approach to Lebesgue's Theory of Integration*, Cambridge University Press, Cambridge, 2008.
5. R. C. Buck, *Advanced Calculus*, McGraw-Hill, New York, 1956.
6. M. P. Carlson, J. Persson, and N. Smith, Developing and connecting calculus students' notions of rate-of-change and accumulation: The fundamental theorem of calculus, in *Proceedings of the 2003 Meeting of*

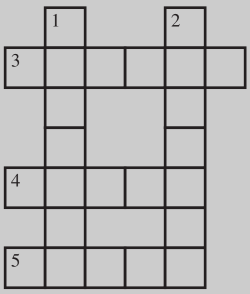
*the International Group for the Psychology of Mathematics Education–North America*, vol. 2, University of Hawaii, Honolulu, HI, 2003, 165–172.
7. A.-L. Cauchy, *Résumé des leçons donnés a L'École Royale Polytechnique sur le Calcul Infinitésimal*, vol. 1, L'Imprimerie Royale, Paris, 1823.
8. M. Clagett, *The Science of Mechanics in the Middle Ages*, University of Wisconsin Press, Madison, WI, 1959.
9. R. Courant, *Differential and Integral Calculus* (trans. E. J. McShane), 2nd ed., Blackie & Son, London, 1937.
10. C. de Freycinet, *De L'analyse infinitésimale. Étude sur la Métaphysique du Haut Calcul*, 2nd ed., Gauthier-Villars, Paris, 1881.
11. P. du Bois-Reymond, Der Beweis des Fundamentalsatz der Integralrechnung: $\int_a^b F'(x)\,dx = F(b) - F(a)$, *Abhandlungen der Mathematisch-Physikalischen Classe der Königlich Bayerischen Akademie der Wissenschaften zu München* **12** (1876) 161–166.
12. ———, Anhang über den Fundamentalsatzes der Integralrechnung, *Mathematische Annalen* **16** (1880) 115–128. doi:10.1007/BF01459233
13. G. Galilei, *Dialogues Concerning Two New Sciences* (trans. H. Crew and A. de Salvio), 2nd ed., Dover, New York, 1954; reprint of Macmillan, New York, 1914.
14. W. A. Granville, *Elements of the Differential and Integral Calculus*, 1st/2nd ed., Ginn, Boston, MA, 1904/1911.
15. J. Gregory, *The Universal Part of Geometry* (trans. A. Leahy), available at http://math.knox.edu/aleahy/gregory/WORKING/gpu.html.
16. G. H. Hardy, *A Course of Pure Mathematics*, 1st/2nd ed., Cambridge University Press, Cambridge, 1908/1913.
17. ———, review of *Elements of the Differential and Integral Calculus* by W. A. Granville, *Math. Gaz.* **7** (1913) 23–24. doi:10.2307/3604382
18. G. Harel, Two dual assertions: The first on learning and the second on teaching (or vice versa), *Amer. Math. Monthly* **105** (1998) 497–507. doi:10.2307/2589401
19. E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Cambridge University Press, Cambridge, 1907.
20. J. J. Kaput, Democratizing access to calculus: New routes to old roots, in *Mathematical Thinking and Problem Solving*, A. H. Schoenfeld, ed., L. Erlbaum, Hillsdale, NJ, 1994, 77–156.
21. V. J. Katz, *A History of Mathematics: An Introduction*, 3rd ed., Addison-Wesley, Boston, 2009.
22. S. F. Lacroix, *Traité du calcul différentiel et du calcul intégral*, 2nd ed., Courcier, Paris, 1810–19.
23. ———, *Traité élémentaire du calcul différentiel et du calcul intégral*, 7th ed., Gauthier-Villars, Paris, 1867.
24. A. Leahy, A Euclidean approach to the FTC, *Loci*, available at http://dx.doi.org/10.4169/loci002156.
25. G. M. Merriman, *Calculus: An Introduction to Analysis, and a Tool for the Sciences*, Henry Holt, New York, 1954.
26. D. A. Murray, *An Elementary Course in the Integral Calculus*, American Book, New York, 1898.
27. ———, *A First Course in the Infinitesimal Calculus*, Longmans, Green, New York, 1903.
28. I. Newton, The October 1666 tract on fluxions, in *The Mathematical Papers of Isaac Newton*, vol. 1, 1664–1666, D. T. Whiteside, ed., Cambridge University Press, Cambridge, 1967, 400–448.
29. O. Nicodemi, Galileo and Oresme: Who is modern? Who is medieval? *Math. Mag.* **83** (2010) 24–32. doi:10.4169/002557010X479965
30. M. Oehrtman, M. Carlson, and P. W. Thomson, Foundational reasoning abilities that promote coherence in students' function understanding, in *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*, M. P. Carlson and C. Rasmussen, eds., Mathematical Association of America, Washington, DC, 2008, 27–41.
31. S.-D. Poisson, Suite du mémoire sur les intégrales définies, *Journal de L'École Royale Polytechnique* **11** (1820) 295–335.
32. A. Prag, On James Gregory's *Geometriae pars universalis*, in *James Gregory Tercentenary Memorial Volume*, H. W. Turnbull, ed., G. Bell, London, 1939, 487–509.
33. A. W. Roberts, *Introductory Calculus*, Academic Press, New York, 1968.
34. J. D. Sterman, Risk communication on climate: Mental models and mass balance, *Science* **322** (2008) 532–533. doi:10.1126/science.1162574
35. D. J. Struik, ed., *A Source Book in Mathematics: 1200–1800*, Princeton University Press, Princeton, NJ, 1986.
36. G. B. Thomas, *Calculus and Analytic Geometry*, Addison-Wesley, Reading, MA, 1951.
37. P. W. Thompson, Images of rate and operational understanding of the Fundamental Theorem of Calculus, *Education Studies in Mathematics* **26** (1994) 229–274. doi:10.1007/BF01273664

38. P. W. Thomson and J. Silverman, The concept of accumulation in calculus, in *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*, M. P. Carlson and C. Rasmussen, eds., Mathematical Association of America, Washington, DC, 2008, 43–52.
39. V. Volterra, Alcune osservazioni sulle funzioni puteggiate discontinue, *Giornale di Matematiche* **19** (1881) 76–86.
40. D. T. Whiteside, Patterns of mathematical thought in the later seventeenth century, *Archive for History of Exact Sciences* **1** (1960–1962) 179–388. doi:10.1007/BF00327940



**DAVID M. BRESSOUD** is DeWitt Wallace Professor of Mathematics at Macalester College and was President of the Mathematical Association of America from 2009 to 2011. He is an avid consumer of the history of mathematics and a firm believer in its importance in helping us understand how to teach this subject.
*Department of Mathematics, Statistics, and Computer Science, Macalester College, St. Paul, MN 55105*
*bressoud@macalester.edu*


**A Pretty Small Crossword Puzzle—with two entirely different answers**




### Across

3 Philosophical logician with contributions to proof theory and modal logic.

4 Mathematical logician who studied the natural numbers, and also the existence of solutions to certain equations.

5 Mathematical logician who proved half the independence of the continuum hypothesis.

### Down

1 Mathematical logician, contributed to the foundations of infinity; middle-European, but the name sounds sort of Italian.

2 Mathematical logician who also wrote children's books.

(Solution on p. 160)

—Submitted by Jim Henle, Smith College