Fractional Differential Calculus
Volume 6, Number 2 (2016), 281–292 doi:10.7153/fdc-06-19

A FRACTIONAL RATE MODEL OF LEARNING

George Dassios, George Fragoyiannis and Konstantia Satrazemi

(*Communicated by J. A. Tenreiro Machado*)

Abstract. A fundamental principle of Cognitive Psychology states that the rate at which the human brain learns a certain amount of knowledge is proportional to the amount of knowledge yet to be learned. This is the so called pure memory or tabula raza law of learning. The mathematical formulation of this principle leads to a simple ordinary differential equation of the first order. Here we expand the existing mathematical model to a fractional differential equation which allows for a more realistic model having a much higher freedom to fit possible experimental data, as well as allowing for memory effects during the learning process. Two different definitions of the fractional derivative are used, one is the standard Riemann-Liouville global definition and the other is a local definition based on the choice of the unit that measures functional variation. A detailed comparison with the conventional model both at the analytical and the numerical level is included.

1. Introduction

The question of how the human mind learns was a major philosophical problem since the days of Plato and Aristotele. In fact, Aristotele introduced the theory of *tabula raza* according to which the human mind starts as a blank slate on which knowledge, coming initially through experience, is imprinted.

Since the Antiquity there have been many theoretical descriptions of how the brain learns. These theories involve different philosophical explanation but also mathematical models both at the deterministic and the stochastic level [2, 4, 5, 6, 12, 15, 17, and the references there]. From the deterministic point of view, it seams that the most widely accepted mathematical models are those that are founded on the principle of *tabular raza*, on the *connectedness* and on *tutoring*.

The tabular raza, or pure memory model, is based on the principle that the human mind learns at a rate that is proportional to the amount of pre-determined knowledge yet to be learned.

The connectedness model is based on the constructivist view that the subject learns by constructing a relation between the new and some prior knowledge. A more sophisticated model is to consider a linear combination between the previous two simple models. In this model, we choose a fraction γ of the learning to be connected and the remaining $(1-\gamma)$ fraction of learning is assumed to be pure memory.

Mathematics subject classification (2010): 26A33, 34A08, 97C30.
Keywords and phrases: learning model, fractional rate.

The third model of learning is the choice of a tutor. This model provides instruction as a tailor-made procedure where the instructor provides exactly what the subject needs to know without wasting time on repetitions.

All these models have been constructed on the basis of experimental results made mainly in schools and universities. These data lead to differential equations for learning quantities as functions of time.

In the present work, we attempt to redefine one of this learning models, that is, the pure memory model, in terms of fractional order derivative and compare the solution with the corresponding model with conventional differentiation. In fact, we analyze the learning model which is based on the standard law that *we learn at a rate which is proportional to the amount of knowledge we need to learn*. In mathematical terms, this law is expressed as an analogy between the derivative of the acquired amount of knowledge and the amount of knowledge yet to be learned at any particular time. The proportionality constant of this analogy is a positive number which specifies the intellectual ability of the subject. The higher this constant the higher the derivative and therefore the faster the brain acquires a pre-assigned material.

On the other hand, the intellectual ability for learning can also be calibrated, even in a finer way, by using a fractional derivative of the acquired knowledge at any particular time. That is, instead of keeping the order of differentiation equal to one and introduce the intelligence as a multiplicative factor, we keep the intelligence factor to be equal to one and introduce a fractional derivative of the instantaneous knowledge obtained. This is exactly what we have done in the present work, were we have solved the problem using the conventional, the Riemann-Liouville, as well as a local differentiation models and compare the obtained results. This comparison involves the connecting function between the intelligence factor and the intelligence rate of absorbing knowledge.

This paper is organized as follows. Section 2 analyzes the convectional model where the intelligence enters as a multiplicative constant. A short discussion of the fractional derivative and its meaning is included in Section 3, where we also introduce a local fractional derivative based on the unit used to calculate the rate of change. In Section 4 we solve the fractional model of learning both with the Riemann-Liouville as well as with our local definition of the fractional derivative. We observed that no significant difference between the two definitions of the derivative is identified. In Section 5 we provide graphical interpretations of the obtained results. Finally Section 6 re-states all the results.

2. The conventional model

Let M be the total amount of knowledge that has to be learned. This knowledge could be interpreted, for example, as a course material for an advanced educational program. Let $y_1(t)$ be the amount of knowledge that has been acquired at time$t$, assuming that

$$y_1(0) = \varepsilon M, \quad 0 \leqslant \varepsilon \leqslant 1 \tag{1}$$

that is, ε determines the percentage of the material M that is initially known. The value $\varepsilon = 0$ corresponds to complete initial ignorance for the subject. Then the mathematical model is stated as

$$\frac{dy_1(t)}{dt} = A(M - y_1(t)), \quad t > 0 \tag{2}$$

where A is the *intelligence factor* of the particular person. The rate at which the material is absorbed is proportional to the amount of material left to be learned [1]. The solution of this simple conventional model is given by the function

$$y_1(t) = M[1 - (1 - \varepsilon)e^{-At}], \quad t > 0 \tag{3}$$

where, obviously,

$$\lim_{\varepsilon \to 0} y_1(t) = M[1 - e^{-At}] \tag{4}$$

and

$$\lim_{\varepsilon \to 1} y_1(t) = M \tag{5}$$

in which case all the material is known at the first place. As we observe in Figure 1, due to the relaxation that appears as we approach the acquisition of the whole material, we can not expect to finish our task in finite time, and this is true for every initial ε. In fact, if we denote by $t_{h,\varepsilon}$ the time needed for learning half of the material, then

$$t_{h,\varepsilon} = \frac{\ln[2(1 - \varepsilon)]}{A} \tag{6}$$

which shows that the half acquisition time $t_{h,\varepsilon}$ is inversely proportional to the intelligence factor A.

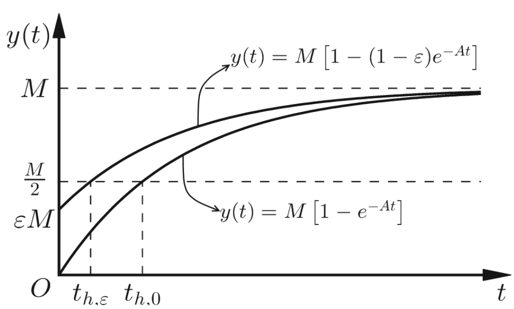

Figure 1.

A consequence of the above learning process is the fact that the only way to achieve a complete knowledge of the material at finite time is to raise the expected level of knowledge and attempt to learn an amount of material $M_1 > M$. Then, the material M can be learnt at the finite time

$$t_0 = \frac{1}{A} \ln \frac{M_1(1 - \varepsilon)}{M_1 - M} \tag{7}$$

where, the larger the distance between M_1 and M, the shorter the time t_0. This is depicted in Figure 2.

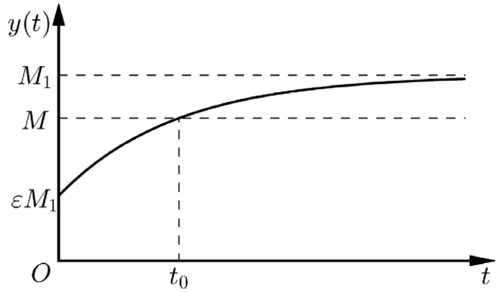

Figure 2.

A more realistic model is the one that takes into consideration a loss of the acquired knowledge due to time passage. In this case, the rate of learning is diminished by a term that is proportional to the part of M that is already acquired. Then the learning equation becomes

$$\frac{dy_2(t)}{dt} = A(M - y_2(t)) - Ly_2(t), \quad t > 0 \tag{8}$$

where the *loss factor* L is a coefficient depending on the ability of a person to keep up with the acquired knowledge as time passes by. A person with a good memory has a low L. The solution of equation (8) under the initial condition (1) is given by

$$y_2(t) = M\left[\frac{A}{A+L} + \left(\varepsilon - \frac{A}{A+L}\right)e^{-(A+L)t}\right] \tag{9}$$

We observe that in the presence of memory loss, even as $t \longrightarrow \infty$, the total amount of knowledge does not exceeds the $A/(A+L)$ percentage of M. Hence, in this case it is more important to raise the value of M to a higher level. The half time learning in this case is equal to

$$t_{h,\varepsilon} = \frac{1}{A+L}\ln\frac{\frac{A}{A+L} - \varepsilon}{\frac{A}{A+L} - \frac{1}{2}} \tag{10}$$

where, since we are interest in the half time, ε should be less than $1/2$, so that the logarithm assumes positive values. Furthermore, A should be larger than L, since in any other case no learning process occurs. If L is larger than A, then the subject forgets faster than he learns. In particular, for $\varepsilon = 0$, we obtain

$$t_{h,o} = \frac{1}{A+L}\ln\frac{2A}{A-L} \tag{11}$$

3. The meaning of fractional differentiation

The problem of Fractional Calculus is as old as the Conventional Calculus. However, the field of Fractional Calculus and its applications expanded very rapidly only within the last two decades [7, 9, 10, 11, 13]. Many authors attempted to provide some meaning to fractional integrals and, from a point of view, this is enough, since the fractional derivative is expressed via ordinary differentiation of fractional integration [13, 14].

Considering the well known definition of the fractional integral of order $a > 0$

$$ {}_0D_t^{-a} f(t) = \frac{1}{\Gamma(a)} \int_0^t (t-\tau)^{a-1} f(\tau) d\tau \tag{12} $$

we can interpret it as the area of a Stieltjes integral

$$ {}_0D_t^{-a} f(t) = \int_0^t f(\tau) d\mu_t(\tau) \tag{13} $$

with respect to the measure

$$ \mu_t(\tau) = \frac{1}{\Gamma(a+1)} [t^a - (t-\tau)^\alpha] \tag{14} $$

which depends both on a and t.

Besides this geometrical interpretation, it is possible to give a physical meaning to the integral (12), or equivalently to the integral (13), in terms of the distance traveled by a moving body when the time is measured with respect to a clock which clicks at the points $\mu_t(1)$, $\mu_t(2)$, $\mu_t(3), \ldots$ specifying non-equidistant time intervals [11].

A more direct interpretation of the integral (12) is to consider it as a continuous linear combination of the values of f in the interval $[0,t]$, taken with a weight given by the function $(t-\tau)^{\alpha-1}$. That is, ${}_0D_t^{-a} f(t)$ is a weighted average of the function f in the interval $[0,t]$. Note that for $a \in (0,1)$ the weight function is an increasing function of τ which ends with an integrable singularity at $\tau = t$, while for $a > 1$ the weight function is a continuous decreasing function of $\tau \in [0,t]$.

The fact that the ordinary integral represents a very special case, in fact a peculiarity, of integration becomes obvious if we observe that the integral in (12) becomes an ordinary average with constant weight if and only if a is a positive integer. For $a = n \in \mathbb{N}$, ${}_0D_t^{-n} f(t)$ denotes the $(n-1)$-th repeated ordinary integrals, as we can see from the Cauchy formula for multiple integration.

There exist in the literature some direct definitions of fractional derivatives which have a local character [1, 3, 8, 10, 16]. We offer here another such definition which is based on the physical interpretation of measuring the rate of change of a function with respect to a non-uniform unit of functional variation.

Starting from the fact that, in all theories of fractional differentiation the a-derivative of the power function t^l reduces the exponent l by the order of differentiation a, we reexamine the physical meaning of the ordinary derivative

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{15}$$

The ratio

$$\frac{f(x) - f(x_0)}{x - x_0},$$

measures the variation of the function f, using as a unit of the variation the identity function $idx = x$. The main characteristic of the identity function is that it has a constant variation, equal to one, at every point of its domain. In other words, the identity function provides a homogeneous unit for measuring functional variations.

However, nothing prevents us to use an inhomogeneous unit of measurement, especially if this unit is compatible with the variations of the quantities entering a specific model. For example, if we use as a unit of measuring functional variations the function $g(x) = \sqrt{x}$, then

$$\lim_{x \to x_0} \frac{x - x_0}{\sqrt{x} - \sqrt{x_0}} = \lim_{x \to x_0} \left(\sqrt{x} + \sqrt{x_0}\right) = 2\sqrt{x_0} \tag{16}$$

that is

$$_0D_x^{1/2} x = 2\sqrt{x} \tag{17}$$

and for a function f

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{\sqrt{x} - \sqrt{x_0}} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \lim_{x \to x_0} \frac{x - x_0}{\sqrt{x} - \sqrt{x_0}} = 2\sqrt{x_0} f'(x_0) \tag{18}$$

Hence, we arrive at

$$_0D_x^{1/2} f(x) = 2\sqrt{x} f'(x) \tag{19}$$

In the generic case, where $0 < a < 1$, the fractional derivative is defined as

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x^a - x_0^a} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \lim_{x \to x_0} \frac{x - x_0}{x^a - x_0^a} = f'(x_0) \frac{1}{a x_0^{a-1}} \tag{20}$$

that is

$$_0D_x^a f(x) = \frac{1}{a} x^{1-a} f'(x) \tag{21}$$

In the case where $a > 0$, with $m = [a]$, we define the a-order local derivative as

$$_0D_x^a f(x) = {}_0D_x^{a-m} f^{(m)}(x) = \frac{1}{a-m} x^{1-a+m} f^{(m+1)}(x) \tag{22}$$

On the other hand the fractional integral is still defined as in the Riemann-Liouville definition (12).

It seems that the interpretation of the derivative as a measure of functional variation with respect to a non uniform unit, which is specified by the order of differentiation, does not generate the multiplicative constants of the fractional derivatives generated by the Riemann-Liouville or the Caputo derivative [13]. This is due to the fact that the proposed interpretation preserves the local character of the derivative as opposed to the global character incorporated in the Riemann-Liouville or the Caputo derivative.

The next question is associated with the usefulness of such non-uniform units of measuring functional variations. Well, some quantities entering particular models vary at a non-uniform rate and this rate can be represented by such a non-uniform unit in a natural and tailor-made choice for the model at hand. Obviously, this choice simplifies the logistics of the model. For example, if we know that a quantity varies proportionally to the cubic root of the variable x, then by choosing the cubic root of x as the unit of functional variation, it follows that this quantity has a constant derivative of order one third. Actually, this is exactly what Abel did in his celebrated 1823 paper initiating the applications of fractional calculus.

4. The fractional model

In the conventional model the intelligence of the 'student' enters the problem via the intelligence factor A, while the order of the rate of learning is assumed to be equal to one. A more advanced model concerns the case where the ability of the student to learn is measured by a derivative of fractional order $a \in (0,1]$. In this case, we replace the intelligence factor A with the *intelligence rate* a and the corresponding mathematical model reads

$$ {}_0D_t^a y_3(t) = M - y_3(t), \quad t > 0 \tag{23} $$

where we considered the normalized intelligence factor $A = 1$, and for simplicity we assume that $y_3(0) = 0$, so that the Riemann-Liouville derivative ${}_0D_t^a$ coincides with the Caputo derivative ${}_0^cD_t^a$. It is well known that the solution of the fractional differential equation

$$ {}_0D_t^a u(t) + Au(t) = Ag(t) \tag{24} $$

with $u(0) = 0$ and $0 < a < 1$ is given by

$$ u(t) = -g(t) * \frac{d}{dt} E_a(-At^a) \tag{25} $$

where $*$ denotes the convolution integral and

$$ E_a(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(an+1)}, \quad z \in \mathbb{C} \tag{26} $$

is the Mittag-Leffler function of one parameter [10]. The Mittag-Leffler function is a generalization of the exponential function and in particular

$$ E_1(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n+1)} = \sum_{n=0}^{\infty} \frac{z^n}{n!} = e^z \tag{27} $$

Hence, the solution of (22) is given by

$$ y_3(t) = -M * \frac{d}{dt} E_a(-t^a) \tag{28} $$

where

$$\frac{d}{dt}E_a(-t^a) = \sum_{n=1}^{\infty} \frac{(-1)^n t^{an-1}}{\Gamma(an)} \tag{29}$$

and therefore

$$y_3(t) = -M \int_0^t \sum_{n=1}^{\infty} \frac{(-1)^n (t-\tau)^{an-1}}{\Gamma(an)} d\tau = -M \sum_{n=1}^{\infty} \frac{(-1)^n t^{an}}{\Gamma(an+1)} = M[1 - E_a(-t^a)]. \tag{30}$$

Compare solution (30) with solution (4), for $A = 1$, in order to understand the role of the Mittag-Leffler function in connection to the a-fractional derivative.

If we use our local fractional differentiation approach, equation (23) is written as

$$\frac{t^{1-a}}{a} y_5'(t) = M - y_5(t), \quad t > 0 \tag{31}$$

which is also written as

$$(e^{t^a} y_5(t))' = Mat^{a-1} e^{t^a} \tag{32}$$

and in view of the initial condition $y_5(0) = 0$, implies that

$$y_5(t) = M(1 - e^{-t^a}), \quad t > 0 \tag{33}$$

where $0 < a < 1$.

5. Comparison of the models

The conventional model (CM), with zero initial condition, gives

$$y_1(t) = M(1 - e^{-At}) \tag{34}$$

while the Riemann-Liouville fractional model (FM), under the same condition, gives

$$y_3(t) = M[1 - E_a(-t^a)] \tag{35}$$

and the local model (LM) gives

$$y_5(t) = M(1 - e^{-t^a}) \tag{36}$$

Graphs of $y_1(t)$, $y_3(t)$, $t \in [0,4]$ for $(A,a) = (1.08, 0.80)$ and $(A,a) = (1.21, 0.30)$, are depicted in Figures 3 and 4 respectively.

Note that after we pass the half learning time with the CM we learn faster than the FM. This is of course due to the difference between the values of the two parameters.

Again the CM is more effective after the half-learning time, but things are the opposite way before that time.

Since, for all models, the acquisition of M cannot be achieved in finite time, a possible measure of comparison is to find the relation that connects the intelligence

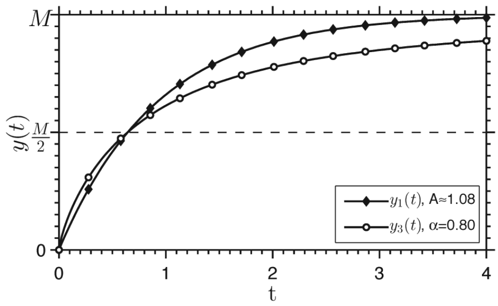

Figure 3.

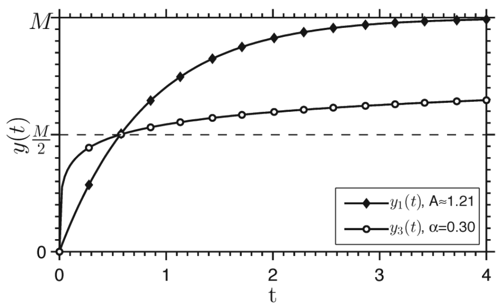

Figure 4.

factor A with the intelligence rate a under the assumption that the half-learning time for the relative models coincide.

Along this line we observe that the half-learning time for y_1 is

$$t_1 = \frac{\ln 2}{A} \tag{37}$$

for y_3 is the solution t_3 of the equation

$$\frac{1}{2} = E_a(-t_3^a) \tag{38}$$

and for y_5 is equal to

$$t_5 = \sqrt[a]{\ln 2} \tag{39}$$

Since, we assumed the same half-learning time, i.e. $t_1 = t_3 = t_5$, it follows that the parameters A and a are connected with the relation

$$E_a\left(-\left(\frac{\ln 2}{A}\right)^a\right) = \frac{1}{2} \tag{40}$$

which is depicted in the next Figure 5, for y_1 and y_3. This curve shows the pairs of intelligence factors and intelligence rates that lead to identical half-learning periods for the conventional and the fractional models.

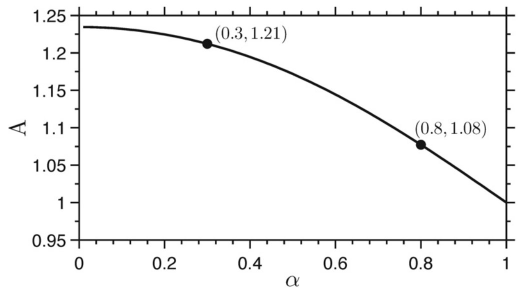

Figure 5.

Also, the relation

$$A = (\ln 2)^{\frac{a-1}{a}} \tag{41}$$

connects the half-learning periods for y_1 (conventional) and y_5 (local derivative).

Figures 6–8 provide some graphs of the functions $E_a(-t^a)$ and e^{-t^a} for comparison.

6. Conclusions

There are three major mathematical models for learning, the tabula raza, the connectedness and the tutorial. These models are formulated via three corresponding differential equations that differ on the expression for the rate of absorbing knowledge. In the work at hand, we have picked one of this models, i.e. the tabula raza model, and we re-defined its basic law terms in fractional derivative of order $a \in (0,1)$. This way, the multiplicative intelligence factor of the conventional model is now replaced by the intelligence rate of change, which is a more fine parameter of the model. The relation that connects these two different parameters of the problem is identified and depicted.

The fractional differential equation is interpreted in two different ways. One uses the standard Riemann-Liouville global definition of the fractional derivative. The other uses a local definition of the fractional derivative which is based on measuring functional variation by an appropriate non-homogeneous unit of fractional variation. As we can see from the relative numerical values, the two kinds of fractional derivatives do not show any significant difference. However, the local definition frees the definition of the derivative from the functional history of the evaluated quantity, which is implied by the global character of the Riemann-Liouville definition.

Acknowledgements. The authors want to thank Dr. Fotini Kariotou for fruitful discussion during the development of this work.

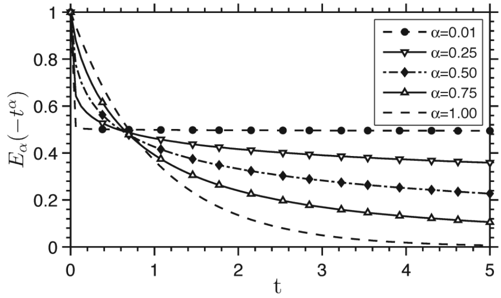

Figure 6.

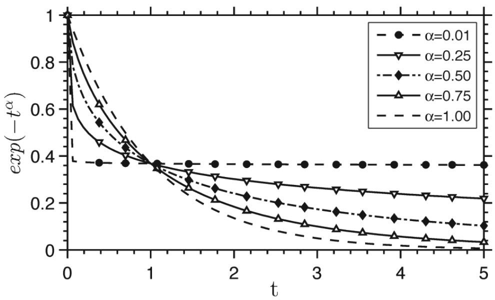

Figure 7.

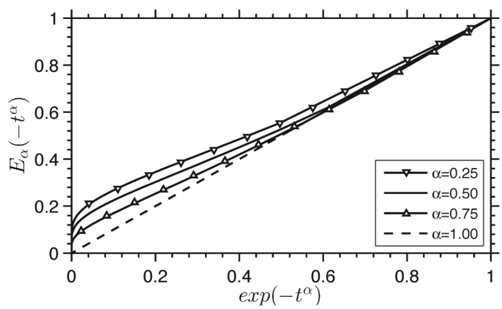

Figure 8.

REFERENCES

[1] T. ABDELJAWAD, *On conformable fractional calculus*, arXiv:1402.6892vl [math.DS] 27 feb 1014.

[2] B. ALDRIDGE, *A mathematical model for mastery learning*, Journal of Research in Science Teaching, vol. **20**, pp. 1–17 (1983).

[3] D. ANDERSON AND D. ULNESS, *Properties of the Katugmpola fractional derivative with potential application in quantum mechanics*, Journal of Mathematical Physics, vol. **56**, 063502 (2015).

[4] I. ANTONIOU AND F. BOSCO, *On the spectral properties of a Markov model for learning processes*, International Journal of Modern Physics, vol. **11**, pp. 213–220 (2000).
[5] R. ATKINSON AND R. SHIFFRIN, *Mathematical Models for Memory and Learning*, Institute for Mathematical Studies in the Social Sciences, Technical Report no. **79**, Stanford University (1965).
[6] R. BUSH AND F. MOSTELLER, *A mathematical model for simple learning*, The Phychological Review, vol. **58**, pp. 313–323 (1951).
[7] R. HERRMANN, *Fractional Calculus. An Introduction for Physicists*, World Scientific, New Jersey (2014).
[8] U. KATUGAMPOLA, *A new approach to generalized fractional derivatives*, Bulletin of Mathematical Analysis and Applications, vol. **4**, pp. 1–15 (2014).
[9] A. KILBAS, H. SRIVASTAVA AND J. TRUJILLO, *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam (2006).
[10] K. KOLWANKAR, *Local fractional calculus: a review*, arXiv:1307.0739v1 [nlin.CD] **2** July 2013.
[11] K. OLDHAM AND J. SPANIER, *The Fractional Culculus*, Academic Press, New York (1974).
[12] P. PREECE, *Mathematical modeling of learning*, Journal of Research in Science Teaching, vol. **21**, pp. 953–955 (1984).
[13] I. PODLUBNEY, *Fractional Differential Equations*, Academic Press, San Diego (1999).
[14] I. PODLUBNEY, *Geometric and physical interpretations of fractional differentiation and fractional integration*, Fractional Calculus and Applied Analysis, vol. **5**, pp. 367–386 (2002).
[15] D. PRITCHARD, Y.-J. LEE AND L. BAO, *Mathematical learning models that depened on prior knowledge and instructional strategies*, Physical Review Special Topics-Physics Education Research, vol. **4**, 010109 (2008).
[16] V. TANASOV, *Local fractional derivatives of differentiable functions and integer-order derivatives or zero*, International Journal of Applied and Computational Mathematics, DOI 10.1007/S40819-015-0054-6.
[17] M. WELLS, D. HESTENES AND G. SWACKHAMER, *A modeling method for high school physics instruction*, American Journal of Physics, vol. **63**, pp. 606–619 (1995).

(Received April 7, 2016)

George Dassios
University of Patras and FORTH/ICE-HT
Greece

George Fragoyiannis
University of Patras
Greece

Konstantia Satrazemi
University of Patras
Greece