The place of education/learning in the hierarchy of Engels' curves of consumption: the quantitative basis of mechanisms finally elucidated.

(Short Title): **Mechanism based metrics and econometrics of learning/retention**

Vetury Sitaramam* and Pallavi Vetury-Iyer

"Vetury",# 18, S.No. 245, Anant Coop. Hsg. Society, Aundh, Pune - 411 067, India,

*Corresponding author, sitaramamv@gmail.com

Affiliations: NIL (*Corresponding author, retired as a Professor from University of Pune, Pune, India. The second author is still under internship for teaching).

Copyright, Pune 2017.

Abstract:

The study focuses quantitatively on ways to assess school education, currently under severe criticism, in terms the first of its three major components, students, teachers and curriculum,. We developed a novel method to assess long term retention as a measure of true learning in tests that do not involve any cramming and found to our dismay that it falls progressively more towards recent years. Our initial work on poverty showed that consumption on commodities is hyperbolic (as in catalysis) and hierarchic and, that the total income is not fractionally distributed among the commodities, as economists assume; the leftover (residual) income after the more essential items is spent on the commodities of lesser preference. Education is also a naturally ordered commodity and parental income quantitatively determined school dropouts/attendance obeying the same hyperbolic relationship, as we could show. We could thereby affirm the first-ever mechanism for Engels' three curves of consumption, based on the time constants associated with the corresponding hyperbolic functions. The disturbing observation was that long term retention as a true measure of learning decreases rapidly with years and seems to be independent of socioeconomic circumstances and gender but depends on subjects. The test also revealed that those who are not good at cramming could exhibit good retention, though the conventional tests fail to identify them. Causes of the progressive loss of retention for the recent years, either as limits to learning or as limitations in imparting learning, need to be assessed by further studies. This would hopefully begin an era of inclusive education for a just society for which methodology beyond pious intentions never existed.

Key words: Social justice, STEM education, examinations, rote learning, cramming, retention, true learning, econometrics, dropouts, income, commodities, hierarchy of consumption, hyperbolic function.

Introduction:

Improvement in school education requires a clear understanding of the economic factors that determine the actual participation of the stakeholders in the society. When education is in trouble due to poor performance of institutions and even a nation, it, in turn, affects the economics of the population in a competitive world [1,2]. Both education and economics of that nation spiral downwards, the mutual interaction between education and economic wellbeing in either direction tending to be autocatalytic [3-6]. The remedial measures currently offered in pedagogy are themselves without a metric, generally limited to a loose headcount approach rather than mechanism/model based. The reasons are not far to seek: the econometric theory itself is still highly empirical in handling commodities and the social rules governing the expenditures beyond what was proposed by Engel [7]. It follows that educationists are also without a clue as to how parents treat education as a commodity. Any educational innovation cannot be assessed for its real impact [8], which needs to distinguish between the socioeconomic influences in learning as opposed to true learning per se.

In order to arrive at clear insights in such sociological (in terms of economics) and socio-biological (in terms of limits to learning in individuals) issues inherent to school education, large scale, well curated data is of critical importance. Collection of data at the national level is hugely expensive and requires a justification for spending the tax payers' money. However data alone does not guarantee understanding without analytical rigour on one hand and conceptualization of the core issues into testable models on the other. Obviously, the problem is to be arrived at by consolidating, stepwise, what is already available and develop probes to obtain critical insights that clearly justify large scale measurements. Probes are just that: small, clear, concise and definitive. History is replete with examples of pioneering observations, small in number and yet definitive, that guided later thinking over centuries. Notable examples are the work of Ernst Engel in econometrics and Gregor Mendel in genetics [7, 9]. Ebbinghaus discovered similarly the exponential decay in memory of random words with himself as the only subject! Much debated for its 'objectivity, bias, and sample size' and so on, his study still represents a major paradigm shift in experimental psychology from qualitative descriptions to rigorous quantitation [10].

Thus, the incisive part in a novel work lies in the design for newer ideas built on intuitive modeling based on what is already available. Our work over the last three decades on the econometrics and biology of poverty and deprivation showed that existing econometric theories of poverty have only led to massive confusion despite having high quality national

data [11-13]. The reason was in partitioning income into expenditures on commodities and the basis of such a partition, an aspect of Engel's original work that remained at an empirical level, never translated into a model or a mechanism. This paper is thus in three parts, each with distinct and yet interlaced methodology. The first part enunciates the first-ever mechanism and explanation for Engel's curves of expenditures, which permits ordering of preferences of various commodities such as food, clothing and shelter, to show that health and education also form part of this spectrum of preferences. In the second part, we apply our analytical methods to the only naturally ordered commodity, viz., education, to validate the methodology of ordering hierarchy of expenditures and demonstrate the logic behind school attendance and the dominant influence of parental income on it [cf.11, 13]. In both instances, we use the published, well curated data from national and international sources [14]. From these studies, we identified the need to develop a unique metric for school education thus far not attempted, viz., long term retention, without which true learning cannot be assessed [15]. Since the usual tests that involve cramming are unduly influenced by socioeconomic factors, the objective is to examine whether such the novel test is independent of the school grades and, hence, socio-economic factors. We report the feasibility of such a test that qualifies to be enlarged for validation at the national level.

Raison d'être of education is in its social context – living in society as a contributor (to 'perform') and sharing its value system (to 'belong') [16-18]. A serious lacuna persists in evolving appropriate frameworks for pedagogy in a just society: the act of learning per se could not be evaluated independent of the socio-economic influences, thus far [19]. Education creates social opportunities, while unequal access as well as an unjust ranking of performance damages the very social fabric in which pedagogy operates. A major rethink on the structuring of pedagogy in terms of teaching and evaluation practices is the need of the hour in the growing aspiration for an effective and universal STEM education [20]. Students, teachers and curriculum form a triad in the context of learning, requiring interdependent methods of evaluation, necessarily quantitative, as it behoves STEM education itself, which enhances rigour and does not deemphasize learning in the humanities as well. The nub of the matter is in that the current methods of evaluation in education reduce the individual performance to a set of scores or grades, much debated for their insight and comprehensiveness [21-22]. The triad itself remains barely visible in current evaluation practices.

Three questions thereby arise when we examine the state of education and assessment in schools, which relate to our assumptions as much as to our practices [23, 24]. Firstly there

is a universal mistrust of examination marks/grades which largely determine the future engagements of the students and their life course. These examinations are primarily time-bound, content-bound (i.e., depending on the grade) and largely age-bound, which ensure cramming as a necessity with little choice. Secondly, there is an aversion to the so-called rote learning, which tends to be confused with cramming, while the former should really reflect memorizing without any association/context in the strict sense of Ebbinghaus [cf.10]. Cramming is a direct sequel to the practices of evaluation and not intrinsic to learning. Both require some degree of comprehension and the ability to reproduce and to use the concepts in the given context and yet learning requires that it be retained, the text and the context, with clarity [25]. The third and yet the most important aspect of education is that an evaluation of teaching and learning should transcend the socioeconomic barriers and aid and assess the child's actual intellectual potential and growth, independent of circumstances[26]. Any method that makes a clear distinction is a significant step forward.

I. Modeling consumption expenditures and the mechanism of Engel's curves of consumption.

I.1. Methodology: The aggregate data was obtained from National Sample Survey Organization reports since inception (usually of the order of 100, 000 households, rural and urban in the ratio of ~60:40) and analyzed and modeled. All the results were independently confirmed using unit data of independent households [27]. The socio-econometric data was obtained as aggregate data from the publicly available reports of NSSO, New Delhi, which has urban, rural, regional, community-wise data over several decades. The data of interest here consists of expenditure of independent households for commodities expressed as monthly per capita expenditures (MPCE). It is important to note that income is defined strictly as the sum of all identified expenditures; each income class is given, not as the midpoint for the class interval but as the sum of actual expenditures for that class interval, while the average expenditure was given for each commodity in the aggregate data. While we illustrated the expenditure patterns from the 50th round (1993-94), the pattern repeats over several decades to date, barring inflation.

Our work differs from all the past econometrics work in one important aspect: we did not trust linear or non-linear regression (the latter with Marquardt-Levenberg algorithm using Sigmaplot versions, currently 13.0) without assessing one fundamental requirement for regression largely ignored by econometricians, i.e., the residuals must be random. Fig. 1

shows a theoretical hyperbolic function as in Eq. 3 fitted after a log-log transformation as in Eq. 1. The runs test rejects the regression due to residual pattern very clearly while the coefficient of correlation remains quite convincing [28]. The presence of errors in real data would make it more difficult to make such a distinction as with theoretical data unless the data is very coherent and well curated as NSSO data is.

Fig 1: *Influence of illegitimate transformation in regression analysis.*

Given the theoretical data for $y=3x/(10+x)$, i.e., without error and with $r = 1$, (Fig1, Inset), Fig1 shows a plot of the log-log transformation of the data and then fitted to linear regression ($r = 0.997$). Clearly the data points deviate from the regression line, runs test showing 2 runs (too few runs) rejecting that the residuals are random ($P < 0.0002$). Conventional analyses would find the deviation negligible and would accept the linear regression as a good model, particularly when the data has relatively large error. Well curated data (e.g., of Costa Ribeiro) and well curated large data bases (e.g., NSSO, which made possible much of poverty work by econometricians) reveal illegitimacy of regression readily and hence better modelling.

1.2. Theoretical considerations and Results:

The theoretical section deals with the use of a hierarchy of consumption all of which is hyperbolic, which was developed and used by us throughout our past work, empirically as well as algorithmically [cf. 11]. Fig. 2 depicts cereals, clothing and accommodation (rent) from the urban and rural households in the 50th round of National Sample Survey Organization (NSSO) data of 1993-94 in comparison with education and health (out of pocket expenditures) as an example since the pattern repeats over all rounds [cf. 27]. A typical functional form often employed by economists to depict mutually independent expenditures (Fig. 3, Top) and considered by them to be a good approximation to the observed data is:

$$C = Ay^\theta \quad \dots \text{Eq. 1.}$$

or $\log C = \log A + \theta \cdot \log y$, where θ is referred to as the elasticity coefficient. Thus, $\theta < 1$ reflects necessities and $\theta > 1$ reflects luxuries. These are represented in Fig. 4 Left with three curves; a convex, a concave and intermediate, monotonic relationships, the essence of Engel's curves.

While the log-log relationships have been adopted for the purpose of defining income dependence (or elasticity) thus far, the methodology was found to be, in reality, highly unsatisfactory (*vide infra*). The foremost requirement in linear regression (on direct or transformed data) is that the residuals should be random, i.e., there should be no residual pattern. By using runs test [28], a useful non-parametric test that detects non-randomness of residuals, the log-log fits by regression were rejected uniformly though the r^2 itself was of the order of ≥ 0.9 in all NSSO data over some 20 odd rounds, both urban and rural [11]! The data on cereals (food) clearly indicated a saturating behaviour and we observed that when fitted to a hyperbolic function (of the kind, $y = ax/(b+x)$), the residuals were invariably random unlike the log-log fit. That is,

$$\text{Exp}_1 = \text{Exp}_{\max 1} \cdot I / (K_1 + I) \quad \dots \text{Eq. 2.}$$

where Exp_1 is the expenditure on the 1st commodity (i.e., with highest preference such as cereals, the staple source of carbohydrate), $\text{Exp}_{\max 1}$ is the maximum value that this expenditure can take, K_1 is the inverse of preference for the commodity and I , the income. Given the consumption data as in Fig. 2, it was shown that the parameters are estimated readily from the NSSO data for items such as cereals, which account for the concave form of Engel's curve (necessities), quite adequately [7, 11]. It occurred to us that if the preferences are ordered or hierarchic, the expenditure on the item of second preference would also be hyperbolic except that the expenditure on the first item needs to be subtracted from the total income to give the residual income I_r , against which is plotted the second expenditure of preference (Fig. 3 Bottom). Given a sequence of commodities, this could be repeated in the order of increasing θ , which approximately indicates K for each commodity. Thus

$$\begin{aligned} \text{Exp}_1 &= \text{Exp}_{\max 1} \cdot I / (K_1 + I) \\ \text{Exp}_2 &= \text{Exp}_{\max 2} \cdot I_r / (K_2 + I_r) & \text{where } I_r = I - \text{Exp}_1 \\ \text{Exp}_3 &= \text{Exp}_{\max 3} \cdot I_r' / (K_3 + I_r') & \text{where } I_r' = I - (\text{Exp}_1 + \text{Exp}_2) \quad \dots \text{Eq. 3.} \end{aligned}$$

...and so on. Fig.4 (Right) gives these theoretical profiles of expenditures constructed as a hierarchy and we readily see that the relationships create all the three types of Engel's curves (cf. Fig. 4 Left). In other words, all consumption is hyperbolic once we strip away the influence of hierarchy. The NSSO data in available rounds, for both urban and rural, could be

subjected to an algorithm that selects for a hierarchy based on higher preference and accepts the models based on i. a better r^2 and ii. random residuals in the runs test [cf. 28]. The models chosen were the linear regression with and without log transformation and the rectangular hyperbola by non-linear regression. The rectangular hyperbola succeeded in every round over some 50 years [11]. This allowed us to define the hierarchically ordered consumer basket for the urban and the rural populations of India for the first time algorithmically using NSSO data without additional assumptions [27].

Fig 2: A typical expenditure data on some commodities in the National data. The data is from 50th Round (1993-94) of NSSO published aggregate data (Rural and Urban). **Top:** Rural (68598 households), **Bottom:** Urban (46148 households). Monthly per capita expenditures (MPCE) (in Indian rupees) of selected items: Filled circles, Cereals (major food component), Unfilled squares, Rent, Filled squares, Health, Unfilled triangles, Education, Unfilled circles, Clothing. Note that income in NSSO data always indicates the sum of all expenditures actually obtained in surveys and the data over years shows an invariant pattern barring inflation.

Fig 3: *Conventional and Hierarchic models of consumption in econometrics.*

Top: Conventional dominant models in economics that partition income into various expenditures, considered at the same level such that each relationship is defined by an elasticity coefficient directly related to income.

Bottom: Our hierarchic model of consumption. Each RI is the *residual income* after considering the expenditure for the commodit(y)ies of the higher preference. The hierarchical model does not exclude some expenditures being at the same level as in top figure but the three curves of Engel must follow the hierarchy wherein the residual income is obtained solving for Eq.3 serially.

Fig 4: *Engel's curves of expenditure and method of their reconstruction.*

Left: The three Engel curves of expenditure, Type I, Necessities, Type III, Luxuries and the intermediate, Type II.

Right: The Engel curves were reconstructed from a hierarchy of expenditures as in Eq.3 replicating the phenomenology associated with the Engel curves, each expenditure being hyperbolic itself as a function of income. The arrow represents increasing θ (cf. Eq. 1)

Any convexity or even linearity associated with the expenditure of a commodity implies presence of higher preferences and both education and health are such expenditures of lower preference, reflected in the rural and the urban with their own order of preferences. Thus, the educational expenditure is also best understood as a commodity in its income dependence, amenable to similar analyses. Two major developments follow this analysis.

Firstly, the value of modelling arises from the successful insights and useful predictions in otherwise obscure phenomenology. In order to achieve that, one must examine the origins of the model developed. The hyperbolic model initially derived for cereal consumption for its obvious relation to food and, thereby, biochemistry of consumption, led to a major departure in defining the poverty of people [11, 13]. The nutritional norm (x calories needed for y Joules of work and sustenance) [29] has been shown to be faulty based on our biochemical work; we could show that the energy intermediate, adenosine triphosphate, ATP, generated for a given amount of glucose is **not** constant, as is taught in all standard textbooks of biochemistry to date [30], but varies with the physical state of the mitochondrion inside the cell, shown specifically *in vitro* based on high quality preparations [31-33]. Indeed, cell biologists have observed for long that the mitochondrion inside the living cell continuously changes its size and shape [34], significance of which was lost on biologists. These were important consequences not visualized earlier in an explanatory world dominated by enthalpic considerations, while the framework needed and driven is by entropic considerations, positive (dissipative) and negative(ordered) [35-38]. The dependence of the stoichiometry between work and calories on the physical state of cells and organelles shows that presumption of constant work-calorie relationship, the nutritional norm for poverty definition, is without substance. The prevailing constant discomfort in the definition of a poverty line based on such fixed norms is justified since our work could define a poverty index without an arbitrary over-all poverty line shifting the paradigm to commodity specific poverty; this permits a continuous evaluation of the deprivation as the population enhances its economic status [11-13, 39-40]. The new commodity specific poverty index arises from Eq. 2 such that the shortfall of Exp_i from Exp_{maxi} integrated over all expenditure (income) classes and normalized over populations yields a commodity specific poverty. This liberates the

policy maker from a constraining overall poverty to a realistic commodity-specific range of poverties, which vary with time as the country progresses. In the absence of a clear delineation of this progress, a blunderbuss approach to poverty with a single index that ‘fits all’ led to massive debates and upheavals in public debate. In recognition of the work by the author, the independent confirmation with unit data of this work was associated with renaming this poverty index as Sitaramam Index (SI) [39,40].

Secondly, the work is still far from complete since the critical question was not addressed to either by us or the others working on SI. It is imperative that we understand that ever since Engel determined in 1857 the three types of expenditure in the area of econometrics, it continued without an insight/mechanism and remained a brute empirical observation. This is now finally corrected here. The essence of the hyperbolic models of the kind, $y = ax/(b+x)$ is commonly used in other disciplines as well, such as enzyme kinetics in biochemistry and predator-prey relationships in ecology as Holling type II model [41-43]. The proof of these models derives from the simple fact that a micro-time constant associated with the material that is being processed exists...such as the processing of each insect by a bird while eating or the making or breaking of a bond in catalysis by an enzyme or a catalyst. Saturation occurs simply because the overall rate cannot even remotely approximate the micro-time constant, much less supersede it. In the case of economics, clearly the temporal distinction lies between perishables and durables which represent a continuous spectrum of commodities rather than two distinct classes and it is surprising that such a formal distinction has not been made. [11-13, 39-40]. Table 1 derives a critical comparison between these three disciplines that use the hyperbolic function.

A major reason for the continuation of the subject matter of econometrics to date is by adopting a language of mathematics, ensuring obfuscation for those who need it most and therefore the comparisons in Table 1, as in rest of this manuscript, are made simple enough to be understood by those not at home with formal mathematics. Finally a mechanistic explanation based on highly curated national data has been achieved here for Engel’s curves of expenditures on commodities for the first time in econometrics [44,45].

Two aspects of education in children, educational expenditures and school dropouts have vast data with regard to their dependence on the income of the parents [5,7,27,46,47]. Educational expenditures vary directly with the income of the parents (Fig. 2), well within the ambit of Engel’s curves of the concave necessities and the convex luxuries and the in-betweens (Fig. 4 Left) [7]. The dropout rates (or its converse, attendance), which largely

depend on parental income, would be the right choice to analyze the socioeconomic dependence of education.

Table 1: *A comparison of the three uses of the hyperbolic function in different domains of biochemistry, ecology in comparison with the present use in econometrics/education. Stepwise derivation of the proof.*

	Enzymology (catalysis)	Predator-prey relationships*	Econometrics
1. What saturates?	Total time for catalysis, T_t	Total time spent in predation, T_t	Total time for occupancy for a commodity, T_t
2. What has not	Time not involved in actual catalysis, T_s	Time spent in the search for the prey, T_s	Time to spare, T_s
3. The micro-event	Time to handle one substrate(turnover time) $T_h = 1/T_n$, T_n = turnover number (or V_{max}/E_t , as conventional)	Time to handle one prey, T_h	Time to handle one unit of commodity, T_h
4. Number of events	No. of substrate molecules catalyzed per unit time, N_a	No. of prey captured, N_a	Number of units of commodity purchased, N_a
5. Density	Concentration of substrate, S	Prey density, C	Expenditure total (=income), Exp_t
	$N_a \propto S.T_s$ i.e., $N_a = a .T_s.S$	$N_a \propto C.T_s$ i.e., $N_a = a.C.T_s$	$N_a \propto Exp_t.T_s$ $= a.Exp_t.T_s$
6. Balance equation	$T_t = T_s + N_a T_h$	$T_t = T_s + N_a .T_h$	$T_t = T_s + N_a T_h$

7. Saturable function	Rate of catalysis $= N_a/T_t$ $= aS/(1+a.S.T_h)$ $= (1/T_h).S/((1/a.T_h) + S)$ $= T_n.S/(T_n.a)+S$	Rate of consumption $= N_a /T_t$ $= a.C/(1+a.C.T_h)$ $= [(1/T_h).C] / [(1/a.T_h)+C]$	Rate of consumption $= N_a /T_t$ $= a.Exp_t / (1+a.Exp_t.T_h)$ $= (1/T_h).Exp_t / ((1/(a.T_h))+ Exp_t)$, where $(1/T_h)$ is the maximum rate of consumption.
-----------------------	--	---	---

*Holling Type 2

II: Hierarchy of school attendance/dropouts based on parental income.

II.1. Methodology: We examined available data/reports from United Nations, World Bank and several assorted collections including PISA and Pratham from India, dominated by head count reporting and % changes and could not find data that closely matches the quality of the NSSO data for econometric analyses [27,46-49]. Where does one find socioeconomic influences on education besides the usual school tests which are difficult to model with any level of accuracy or insight? An examination of UNESCO or World Bank data [46,47] on completion of primary schooling as a function of gross domestic product or per capita income has enormous scatter wherein it is difficult to discern any meaningful relationships. We came across an important Brazilian report, published decades earlier, with data that was amenable to precise modelling.

This unique 1985 publication from Brazil by Costa Ribeiro[14] had a systematic evaluation of school attendance in a 3D graph that was promptly digitized (Fig. 5) for its components. Data was manually digitized from the published work with less than 1% error and used. The data when plotted in the usual 2D mode had a close resemblance to Engel's curves and the parental income and the cumulative fractional attendance added up to a nearly linear function of income. This prompted us to restate the educational expenditures in the framework of consumption curves based on hierarchical, hyperbolic analysis with the added advantage that the grades are naturally ordinal and we needed no additional effort to find their preferences as we had to do with commodities as with NSSO data. Once again, when analysed for the hyperbolic nature of the data, the data yielded unique insight into the nature of societal choices.

Fig 5: *Similarity of income versus participation in schooling with Engel's curves.*

Top: Income (in units of minimal wages of 113 R\$) versus participation upto 12 years of schooling in Brazil 1983, corrected for census in the 3D format from digitized data as published. **Bottom:** Same data plotted in 2D format to indicate the similarity with Engel's curves. The arrow indicates increasing grades from 1 to 12.

II.2. Theoretical considerations and Results:

II.2.a. The education model:

The educational model is an extension of our earlier econometric work, recognizing the dominant hyperbolic nature of income-dependent societal preferences including education. This analysis, completely new in econometrics as well as in education, also rests on a hierarchical analysis of hyperbolic consumption behaviours initiated from us [11]. The distinction between existing analytical practices and our approach is summarized in Figs. 1

and 3 and Eqs. 1 and 3. Though it is common knowledge that a regression analysis is considered proper if and only if the residuals are random, the high coefficients of correlation (usually > 0.9) lulled decades of consumption and poverty research into a pursuit of logarithmic transformations and elasticity coefficients, deeply flawed for two reasons. Firstly, correlation cannot be confused with causality regardless of how good the correlation coefficients are, without compelling reasons *extraneous* to the analysis. Secondly, there was no compulsion to look into the mechanism subserving Engel's classification of commodities, since the assumption of existence of θ , really a token of gross analytical simplification, became frozen in the econometric realm without even an attempt at substantiation. This precluded any possibility of searching for a mechanism and the subject matter remained without insights for several decades on what constituted the dependence of a society on specific commodities.

A derived model for education became feasible due to the availability of published empirical data from Brazil [cf.14], tailored, post facto, to the construction of a serial model similar to Engel's curves of consumption inductively (Fig 5). The measurement relates to fractional attendance in each grade corrected for census, plotted as a function of income class of the parents. The overall relationship, f_{ij} , between educational involvement over all the grades and the parental income is nearly linear, (Fig 6, Left):

$$A_{ij} = f_{ij}(C_i, I_j \dots) \quad \dots \text{Eq. 4.}$$

where A_{ij} is the fractional attendance in the i^{th} grade that also depends on the j^{th} income class, C_i refers to the grade (class of study) i , I_j refers to the *income* for j^{th} class, measured in multiples (1-10) of minimum wages (in Brazilian Reals, 113 R\$ then relevant in Brazil), such that,

$$\sum_{i=1}^{i=12} A_{ij} = f_j(I) = mI_j + c_j \quad \dots \text{Eq. 5.}$$

where m appears to be rather independent of j .

The fractional attendance follows a natural order such that higher the income, higher the fractional attendance and shows concavity for the lower grades and convexity for the higher grades (classes) (Fig.5, Bottom). Therefore A_{ij} should be of the form

$$A_{ij} = \frac{a_1 I_j}{b_1 + I_j} \quad \dots \text{Eq. 6.}$$

which is a concave function. Since $\{A_{1j}\}$ is but a subset of the entire data set $\{A_{ij}\}$, the hierarchy of preference renders, at $i = 2$, *in a hypothetical set where there are only two grades of study considered*,

$$A_{2j} = mI_j + c_j - \frac{a_1 \cdot I_j}{b_1 + I_j} \quad \dots \text{Eq. 7a.}$$

$$\text{or, } A_{2j} = mI_j + c_j - A_{1j} \quad \dots \text{Eq. 7b.}$$

where grade 1 is concave and thereby grade 2 is convex such that the sum is linear, consistent with reality when actually summed up over 12 grades. It simply means that the concave function becomes progressively more convex with each consecutive grade of study, from 1-12, as a function of income. Here A_i and A_j are data given while a_i and b_i are estimated. It is worth keeping in mind that this is a unique and valuable national data (of Brazil) corrected for census several decades ago, matching in its curating prowess the NSSO data! It also follows that, except for the last grade that remains linear when the influence of grades i_{-1} are removed, i_{+1} becomes concave, as we have analyzed earlier with commodities (Eq. 3) [11].

Despite huge efforts worldwide, mass literacy programs are limited by parental incomes in developing countries [cf.47]. Though primary education is mostly free, even with subsidies tried in many countries, (e.g., cash, food provided for the school going children etc.) the dropout rates continue to be high [cf.14,47].

II.2.b Hierarchical, hyperbolic nature of school attendance

Fig.6 Left shows the cumulative fractional attendance (sum of attendance in each grade as a fraction of the census such that the sum cannot exceed the total number of grades) (cf.14), plotted as a function of income for the Brazilian population. While the relationship is monotonic, clearly the fractional attendance shows a plateau midway in income classes. More importantly, fractional attendance and parental income are reasonable proxies for each other such that we could emulate our earlier analyses of Engel's curves [7].

Based on the justifiable assumption that some expenditure is incurred for each grade, the fractional attendance representing a measure of that cost, we subjected the data to analysis similar to consumption data of NSSO [27]. The results were striking in that while the maximum attendance did not exhibit much change, the K value for each grade increased and subsequently decreased (Fig.6, Right), approximating a parabolic relationship at high level of significance ($P < 0.001$).

These findings were consistent with the following observations:

- i. Education, the world over, is income related;

- ii. This income dependence is quite similar to other consumption data;
- iii. The income dependence of schooling follows the natural hierarchy of grades exceptionally well; and
- iv. The explanation of a biphasic relationship in K values for education is easily understood...at lower incomes and lower grades with the end far in sight, education is seen more as a burden, a situation that changes midway as the end is in sight and the attitudes correspondingly change.

Fig 6: An interpretation of variation in the affinity for school participation as related to parental income.

Left: The cumulative fractional attendance in the Brazilian education data was assessed for dependence of the income class of the parents. The data across grades fitted a linear regression, $y = 2.38 + 0.761x$, $r = 0.99$, $P < 0.001$. However, the residuals were not random such that the non-parametric Wald–Wolfowitz runs test revealed runs consistent with the probability for non-randomness, $P < 0.0005$.

Right: The plot shows $V_{\max}$ (filled circles), K (unfilled circles) in arbitrary units (axis on the left). The values of K at corresponding grades could be approximated to a parabolic equation, $y = a.x (b - c.x)$ representing a shift in K at the critical grade, b at which the direction of the preference (inverse of K) changes: $K = 0.385x(4.34 - 0.324x)$ the coefficient of correlation, $r = 0.804$, $P < 0.001$, where x is the grade, b approximates the 4-5th grade (the 12th grade omitted in the analyses, it being a straight line).

Thus, the hierarchical analyses appear to capture the social nuances in the decisions that parents indulge in, rather well. This needs to be contrasted with heroic attempts thus far made to capture the reason for high drop-out rates world over, wherein the detail attempted lost the essence, viz., parental income [50, 51]. When proxy variables were used, the dependence of dropout rates on parental income was lost [52]. Thus, the decision for dropout rate at the

lower school level, tragic as it is, is parental, while the decision at the collegiate level is an individual adult decision [53], unfortunate in terms of wastage of national resources.

III. A Novel metric of true learning?

There is a large body of metadata in the country and elsewhere that shows how school grades are strongly influenced by the socioeconomic status, how this requires urgent remedy and that the performance is progressively and even alarmingly deteriorating [4, 34]. No solutions are yet in sight. We took school grades as a common metric, keenly reflective and susceptible to socioeconomic status of students as a self-evident assumption. Socioeconomic data would be valid only over large samples of the order of the NSSO data used here, running into several thousand households, urban and rural, far beyond the scope of a prospective exploratory study such as this.

III.1. Theoretical considerations:

III.1.a. Magnitude of the problem to defines the boundaries of the experimental design.

A comprehensive analysis of school education in India, comparable to NSSO data, is prohibitively expensive both in terms of time and money. We have 720 dialects, 22 official languages, over 13 scripts, at least 3-4 boards of educational systems dispersed over 29 states, ~300 million students by 2017 and only a USD 7 billion annual budget for school education nationally. Which population are we to represent, when we are trying to develop a rapid methodology universally applicable (given further testing)? What homogeneity in data acquisition would be acceptable to all in a country known for millennia for its diversity? Even for one city with only central board involving some 20+ schools analysing only a cross section of students of all grades would exceed a million dollars and a couple of years of intensive work. The solution lies in restricting the problem by removing any novelty in the formulation of the examination pattern. The same teachers that teach and conduct the annual tests for the school (since Board examinations are only at 10th and 12th grades), were asked to give the representative questions for each class and subject similar to what they give for each year at the year-end examination. Thus, there is no difference in the content of examination nor the pattern such that no standardization of any kind is required for the test.

III.1.b.Retention of the past learning in a school

Since the conventional examinations in schools actually ensure cramming, we tested the 9th grade students for their retention from grade 1 to even grade 10 (as a control for the grade they were yet to attend) *without prior notice* of the nature of the examination. The

multiple choice tests were based on what they were actually taught and the questions set by the teachers who taught these subjects. Thus the comparison was only between routine cramming- based annual examination scores and retention-oriented examination scores the nature of which remained unknown to students eliminating cramming. Fig. 7 summarizes the possible outcomes...of the ability of the 9th grade students to recall what they learnt in the previous years, i.e., retention. Had the learning been perfect, the ideal representation would be invariant over the years as in AB (~90% as in Model I, actually seen with students with conventional examinations and selected for higher education in national tests [cf.53]). Had the learning been strictly by rote, it would be increasingly more difficult to recall from the previous years and the results would approximate rote learning and memory measured empirically as with Ebbinghaus [cf.10], i.e., Model II. Though his model was based on time dependence of the limiting case of discrete, disconnected words without meaning, various reinforcements for better memorizing such as repetition would tend to enhance recall, i.e., the angle β to 90° approximating Model I. At the population level, precise relationships such as exponential decay etc., would not be discernible beyond just trends, since the data would be strongly influenced by distributions. The third alternative, in principle, would represent a disturbing element in assessing education: the actual retention declines (α tending to zero) over years relentlessly (Model III). The anticipation was that the students should not be able to answer the 10th grade questions, set as a logical control since they are yet to begin to attend the grade.

Fig 7: Models of possible outcomes of testing for long term retention in children.

Possible outcomes of retrospective testing of grade 9 students tested up to grade 10 after completion of their grade 9 year-end examinations as a linear approximation. A-B represents

a perfect score of ~100 for all the years up to 9 and goes down to 0 at grade 10 which the students are yet to begin (Model I). If the education were entirely by disconnected rote learning, the Ebbinghaus model would predict $B(-x)$, where β follows the memorized content which disappears progressively towards earlier grades (Model II). At best, variants of Ebbinghaus model on memory/retention would only predict that β tends to zero at no retention or a right angle at full retention (Model I). However, if retention/learning is compromised in education, the performance would deteriorate (α tending to zero) with successive years (A-C) (Model III). Note that one would expect B or C to drop to near zero for Grade 10 component of the test. If it were to merely follow the regression line of A-C, i.e. $C-(+x)$, it would represent a major anomaly that requires an explanation.

III. 2. Ethical considerations for the novel test for retention in a school:

III. 2. a. Study design:

The study was to administer a **single, standard school test with the only difference that it was primarily to test for retention of what was learnt in all the completed grades**, created by the teachers themselves who actually taught the subjects, without any influence either by the investigators or any models in their mind. These were simply truncated versions of each grade of the year-end tests pooled without any innovations. The authorities of an urban English medium school were approached with a request to arrange for the test following the test requirements (vide infra). There is no Indian law (education being on the concurrent list for both central and state governments) that requires any ethical clearance for the schools to administer an **academic** test. The school stipulated, in a series of meetings, specific conditions to be met with to conduct the test since the suggested procedure by the investigators was **well within the authority** of the administration. These were:

- i. The investigators make a formal request to the management and assure continuation of the testing if the management finds it useful;
- ii. The test would be limited to a small and significant sample of students;
- iii. The school has the provision for conducting tests, internal (preparatory to Board examinations) and external (for scholarships, awards, Olympiads etc.). The school would decide the time, (a day after the year end examinations, as it were) so that there was no interference with either regular studies or their own examinations;
- iv. The school would directly inform the parents for their concurrence with the strict proviso that the novel test was merely exploratory and entirely voluntary and any inability to participate would bear no consequences for the students;
- v. Any information extracted from the existing school records would be handled by the school only and forwarded to the investigators ; no data was collected of any kind from the students or parents for the test , socioeconomic or otherwise;

- vi. The results should be made available to the school for it to decide if these would be of use to the students;
- vii. The investigators would not interact with the students at any time or be known to them;
- viii. The test would not be taxing and the school would provide a break and refreshments for the students during the test.

The investigators, in turn, stipulated that: a. Data on teachers' qualifications, classes and subjects taught with duration, years of experience be made available and satisfied themselves that these match the requirements of the State School Board requirements; b. The teachers would submit 10 questions for each subject and class as they do for the year-end examinations to be evaluated by the investigators only for ambiguity and correctness and the incorrect ones be corrected/replaced by the teachers themselves; c. A minimum of 30 students are required for statistical reasons, obtained from the previous year's school grades of 105 students, obeying a platykurtic distribution (vide infra); d. All questions would be restricted to multiple choice types to remove subjectivity in the assessment the new test.

The investigators were not privy to the internal school procedures (such as correspondence with the parents and children and announcements), meetings and minutes if any for internal meetings and all communications after arriving at the above stipulations were by email with the Principal only, with the exception of the sealed answer sheets with the answer key handed over physically to the investigators.

III. 2. b. Test design:

The teachers initially gave 575 multiple choice questions when requested for 10 questions per subject (that they actually taught) per year (~600 questions). **Validation by the investigators was limited to identification of ambiguous questions and their correction / removal was done by teachers themselves.** These were initially with $\geq 20\%$ errors. The test was reduced by the school to 280 questions with 4-6 questions per subject per year as the teachers felt that even a 5 hours test (of 575 questions) would be too taxing and with inadequate time. The final set of questions still had ~10% errors, where the students were given the benefit. The actual test took 88 ± 14 min (i.e., <20 sec per question) though the time allotted was 240 min (~50 sec/question) scheduled in two sessions of 120 min each. The test includes grade 10 questions which they have not attended since the clear expectation would be that they should score near zero as a control.

III. 2. c. Recruitment of students:

National and international data on education/economics from most sources was unusable as there was marked scatter in the data not amenable a level of scrutiny that we achieved with NSSO data and Ribeiro's Brazilian data. We have attempted to analyze Brazilian data for subsequent years after Ribeiro's work, which was un-curated and not normalized for census, and no sense could be made due to high scatter. The problem in the design was how to overcome the negative influence of errors and distributions in the data and the errors themselves. Transformations, often suggested to reduce the influence of distributions, also affect the distribution of errors difficult to anticipate.

The alternative was to emulate laboratory practices. The interest in linear regression analysis is to obtain a clear relationship,

$$y = mx + c + \varepsilon$$

where, ε is the error (whose distribution is assumed than measured) associated with the measurement. Thus untransformed data is ideal for curve fitting as the estimates of m and c do not suffer due to manipulation of ε . In order to estimate blood glucose, for example, we make a standard curve of a UNIFORM distribution of standard samples (x) such as 1, 2,3,4 and 5 mg glucose, create a standard curve and compute the unknown in an X on Y regression from the slope. Since the 105 students showed a near normal distribution, we chose to obtain a uniform distribution of a subset of students from these 105 students for a critical comparison between the two sets of grades, i.e., from school grades and the new test. When we tested the students, albeit of higher classes up to Master's level, without such a selection, the results were less than satisfactory with large deviations, year-wise [53].

Since no standard remedy exists, we even tested this procedure of uniform sampling in the case of rice seeds by their weight in our earlier studies to predict yield in a single day from the seeds successfully [54]. Clearly, if the procedure is haphazard or without coherence in logic or sampling, the ensuing results would be even worse and it would be the ultimate verdict against the validity of the method. It is the test design and not blind adherence to either an exhaustive sampling or a random sampling that decides the outcomes. Sampling theory would be rather impervious to whether the sample is rice seeds or grades of school children.

The performance of 105 students in their earlier Grade 8 year-end examination was plotted for their scores (theory only). A subset of students falling into a uniform distribution

was selected, exclusively based on a nearly constant interval between scores, to obtain 32 students, out of which 28 actually completed the test in two parts of 140 questions each. 19 boys and 13 girls out of 32 students selected were of the age group of 14-15 years, anthropometric data within the expected range for the height of each child, barring one girl, heavier for her height. The 4 absentees were all boys from the lower half of the performance in their 8th grade.

The answer sheets were sent with a key to the investigators who graded the answer sheets for each class and subject accordingly and obtained the new scores against their previous year's school record. The questions were routinely used in the schools for the tests, and as there could be future use of similar tests, the questions were kept confidential between the school and the investigators. **Since the investigators were given the names of the students, the identities of the school and the authorities as well as the year of the test were also kept confidential.** The school records with any personal information of the students were out of bounds for the investigators. The scores in the new test were fully disclosed to the school if they wished to use them in any manner. Since they made little sense by themselves, nothing of direct use was gained by the school. The answer sheets remained sealed with the investigators to be shredded after the publication of the manuscript.

The teachers firmly declined to take the test themselves on the grounds that they would not be familiar with subjects other than what they teach.

In summary, the study aimed at and succeeded in developing a focused methodology based on a small sample test which can be readily applied at a larger scale for the benefit of evaluation of educational practices as the larger goal. In the process, none of the results/methodology in any way harmed any student, either by interactions or the results of the study and the study was fully compliant with the procedures and permissions as determined by the school itself. Neither of the authors have any financial competing interests in this work.

III. 3. Further methodological aspects: Quantification of subject content:

Since scoring for retention should relate to the content learnt, we scanned the official textbooks from the State Board to obtain a word count per page (though the random sampling tended to be biased to more OCR (optical character recognition) software-friendly, as the books have number of figures and unfilled spaces in tables. For Indian languages (Hindi and Marathi) we extrapolated the number of pages from equal sized books in English and font size subjected to word count, further confirmed by actual word count. The content of each subject was added for each year to get the total content per year in words. The content was

scaled over the years, considering the total content in the first year as 1.0. This permitted us to make a distinction between retention as the performance in questions for each grade retrospectively, and learning, in arbitrary units, could be defined as a product of this performance multiplied by the content for that grade.

III. 4. Results of the Study on Retention

The school tests are performed on all the students, marked by a near normal distribution of the results. A major problem in designing these studies was that the anticipated distribution of the performance of students would tend to confound the statistical relationships due to grouping around the mean. Our design and the consequent results differ from the routine school tests significantly, though the test content year-wise does not differ from the school examinations:

- i. We selected students based on equidistant performance in the latest available 8th grade school scores to yield a near uniform (platykurtic) distribution, kurtosis being of major concern to yield clear, discernible relationships (Fig. 8).
- ii. Since the aim was to test for true retention by the students without prior preparation (notably the last minute cramming), the nature of the test was not disclosed, other than that it is novel, to the students. The test itself was not different from usual school tests in that it included representative questions formulated by the teachers who taught the subjects in the same school. The difference was only in that the questions were multiple choice only from all subjects **for all classes and in a single test**.
- iii. A particularly important device was to include the 10th grade questions, which the students were yet to attend, as a control, not encountered in any standard examination.
- iv. By measuring the content of all subjects over years, a clear distinction could be made between fractional retention and quantum of learning (*vide infra*).

Fig. 9, Left shows the aggregate 8th grade scores plotted against the current test aggregate scores. The relationship is significant even though 4 students among the 32 selected did not take the test. The frequency distributions of the new test scores were considerably more leptokurtic than the sampled uniform distribution based on 8th grade scores (Fig. 8), showing that the new test did not merely replicate but measured something different than the conventional school tests [cf.23].

Fig 8: *Frequency distributions of students actual, selected, and from the retention test.* Frequency distribution of the 8th grade marks as % (filled circles, 67.8 ± 10 , $n= 105$, kurtosis, -0.166), students for the new test selected to achieve a uniform distribution (unfilled circles, 68.9 ± 14 , $n= 28$, kurtosis, -0.846) and the final result of the new test (filled triangles, 67.8 ± 8.5 , $n=28$, kurtosis, 0.648). Note that the mean of 3 sets of results did not significantly differ. Data of 32 students originally selected showed a mean of 67.2 ± 14.3 , kurtosis, -1.084 , frequency distribution omitted in the figure.

The critical observation was that the new test score showed an increase in variation at lower scores i.e., marked heteroscedasticity (Fig.9, Left). This is important because at least 20% of the children who would be labelled as poor performers in the class tests emerged with better scores indicating comparable retentive ability as with the better performers, suggesting an unjust misclassification! Fig.9, Right, reveals what was not anticipated. The score for 10th grade was **not low** and the Model III regression, i.e., past performance, determined even the future 10th grade score (cf. Fig. 1, C- +x trend as opposed to C – 10 trend), raising distinct possibilities. Firstly, if the fall in scores shown by Model III was due to lesser retention, i. it reflects a general tendency in education that the ability to learn per se decreases, grave, if true; ii: they are taught little in 10th grade, (largely spent in cramming, preparative to Board exams), the merit of which itself is questionable. Secondly, acquiring scores higher than anticipated for 10th grade not yet taught could mean that a. they attend coaching classes (which is true with majority of students), or, b. this test per se was very lenient as they were not anyway taught much in 10th grade or even earlier. Reality could well be a combination of all these.

Fig 9: Comparison of school scores and the retention test scores and the anomalous high performance in the 10th grade questions

Left: A comparison of performance by school tests and the current test (normalized to 100) for retention. Linear regression of the 28 students who attended the test was by the method of least squares, $y = 0.224x + 52.30$, $r=0.374$, $p < 0.05$. Note that a subset of 5 students (~20%, highlighted as unfilled circles) scored significantly higher than predicted by the school tests in which they fared ~50% scores or less. Filled triangles, the students who did not appear for the test. **Right:** A comparison of test scores obtained for Std.10 with those predicted by individual regression analyses of 28 students up to 9th grade, the 10th grade scores obtained by extrapolation of the regression equations. The predicted and recorded test scores (both normalized to 1) were compared by linear regression obtained by the method of least squares, $y = 0.629x + 0.249$, $r = 0.523$, $p < 0.01$, indicating an anomaly (*vide text*).

Age and year of study dominate much of our thinking in education while the real important variable is the content. Obtaining a proxy for the content (which is, of course, mostly written in words in text books) was simply that we evaluated it as the sum of the number of words that can be recognized in scans of pages of all textbooks and arrived at an aggregate number for each subject and year. The average score in the current test shows a clear relationship to the year including 10th grade, indicating that the Model III behaviour was consistent with progressively compromised retention of the content (Fig 10, Top Left). It is important to note that the content increased significantly over the years such that the content showed a clear second order relationship to the year of study (Fig 10, Top Right). While the scores obtained per year reflect normalization, the quantum of learning needs to be a product of the normalized score and the content for that year. As expected, we see that the quantum of learning increases with each year while that fraction retained decreases dramatically (Fig. 10, Bottom), plotted as a second order relation for comparison. *How significant is the loss of retention as in Model III?*

Fig 10: *Progressive loss of long term retention vis-a-vis learning.*

Top, Left: Test scores averaged for the relevant subjects normalized to 1.0 (max) for Grade 1-10 among 28 students. Vertical bars represent $\pm$ standard deviation among subjects for each grade/year. Regression line for all subjects, $y = -0.040x + 0.943$, $r = 0.889$, $P < 0.001$, clearly in accordance with Model III (cf. Fig. 1). **Top, Right:** The aggregate content for all subjects was plotted over years/grades. (See materials and methods for calculating content of subjects, total content of Grade 1 taken as 1.0). The relationship between content plotted vs. The corresponding grade shows a highly significant relationship, $y = 0.200x^2 - 0.047x + 1.211$, $r = 0.965$, $P < 0.001$. **Bottom:** Relationship between the average test scores (as a measure of retention, shown as filled squares, $y = 0.002x^2 - 0.069x + 1.003$, $r = 0.915$, $P < 0.001$) and the test scores multiplied by the content for each year (as a measure of learning for that year in arbitrary units, shown as filled circles, $y = 0.057x^2 + 0.555x + 0.293$, $r = 0.901$, $P < 0.001$) as a function of the grade studied, both fitted to a quadratic equation for ease of comparison.

Performance in education is considered to be highly susceptible to social variables and nature of subjects [cf.23,55,56]. Table 2 shows that Model III behaviour is common to Science, Mathematics, History and English (the medium of instruction). The slope as well as its variance varied with the subject (content and teaching) such that the slopes were significantly different between Science and English and Mathematics considered critical for STEM education. The lesser sample size and larger variance did not allow a better

assignment of a model for Geography, Hindi (the national language) and Marathi (the local language).

Table 2: *Applicability of Models for various subjects in the current test.*

Subject	Grades	Slope* (S.E)	r ²	P	Model
Science	1-10	-0.0605 (0.0079)	0.877	<0.001	III
Mathematics	1-10	-0.0289 (0.0079)	0.623	<0.01	III
History	3-10	-0.0515 (0.0169)	0.607	<0.01	III
English	1-10	-0.0259 (0.0122)	0.360	<0.05	III
Geography	5-10	-0.0047 (0.034)	0.004	NS	-
Hindi	5-10	0.0061 (0.0481)	0.004	NS	-
Marathi	5-10	-0.0288 (0.0237)	0.270	NS	-

*Current performance is given for individual subjects over years, including the prospective 10th Grade. *Slope =average score per student (max=1)/no. of grades, obtained by linear regression by the method of least squares. Slope for Science differed significantly (p<0.05) with English and Mathematics only.*

Sign indicates models as in Fig.1. S.E., standard error of the slope. N.S., Not significant.

We ranked the 8th grade scores into groups of three and assessed the slope across years. It is generally presumed, often correctly, that the lower performers in schools are from the lower socio-economic strata, a major concern and source of debate in a pluralistic society that is striving to be more inclusive [57]. While the intercept (i.e., at grade 0) relates to school performance, *the negative slopes are quite independent of the school performance or gender, significance of which requires special attention* (Table 3). Table 4 shows the descriptive statistics for various subjects and lack of gender selectivity for these 28 students in the retention test, to indicate that frequency distributions among various subjects were not identical. Herein lies the limitation of a limited data set, which limits the extent of finer analyses that the data can be further subjected to. Further significant analyses would be possible only if the sample size, questions per subject per year as well as number of students are scaled by an order (or two) in magnitude.

Table 3. School scores for Grade VIII vs current test scores for groups of three students across all subjects, for 1-10 years of study, based on linear regression by the method of least squares.

VIII Scores Mean $\pm$ S.D (n)	Slope	Intercept	r	P
175 $\pm$ 3.0 (3)	-0.0429	1.007	0.865	<0.001
170 $\pm$ 1.5 (3)	-0.0407	0.994	0.887	<0.001
161 $\pm$ 3.2 (3)	-0.0367	0.921	0.799	<0.01
151 $\pm$ 1.7 (3)	-0.0365	0.951	0.794	<0.01
142 $\pm$ 3.5 (3)	-0.0406	0.973	0.860	<0.001
132 $\pm$ 4.0 (3)	-0.048	0.927	0.9	<0.001
121 $\pm$ 3.2 (3)	-0.0427	0.940	0.902	<0.001
110 $\pm$ 6.2 (3)	-0.0439	0.889	0.834	<0.001
88 $\pm$ 11 (4)	-0.0417	0.926	0.950	<0.001
150.46 $\pm$ 24.5 (13 girls)	-0.0410	0.964	0.891	<0.001
126.7 $\pm$ 27.6 (15 boys)	-0.0420	0.932	0.909	<0.001

The slope was not related to the school score by linear regression, $y = 4 \times 10^{-5}x - 0.047$, $r = 0.354$, N.S. On the other hand, the intercept was significantly related to school scores $y = 9.3 \times 10^{-4}x + 0.818$, $r = 0.708$, $P \leq 0.01$. Number of students of each score group, n, is given in parentheses. The slopes for each group of three were indistinguishable from each other, though the slope itself for each group was significant, shown by the r and P values. Bottom two sets of data relate to gender non-selectivity.

Table 4: Descriptive statistics associated with individual subject scores in the retention test, and gender (non-)selectivity.

	Total	English	Maths	Science	History	Geography	Hindi	Marathi	Girls (13)	Boys (15)
Mean	0.6776	0.7479	0.7459	0.6992	0.5846	0.5216	0.6786	0.6592	0.6988	0.6593
Median	0.6821	0.7692	0.7404	0.7212	0.5870	0.5526	0.7083	0.6667	0.6964	0.6714
Mode	0.6714	0.7692	0.7308	0.7308	0.5217	0.5526	0.7500	0.7917	0.6714	0.7320*
Standard deviation	0.0850	0.0806	0.1105	0.1092	0.1137	0.0931	0.1210	0.1559	0.0458	0.1066
Variance	0.0072	0.0065	0.0122	0.0119	0.0129	0.0087	0.0146	0.0243	0.0021	0.0114
Coefficient of Variation	12.5415	10.7807	14.8087	15.6169	19.4495	17.8530	17.8323	23.6527	6.5532	16.1632
Skewness	-0.7840	-0.9612	-0.1952	-0.5373	-0.7023	-0.5866	-1.2243	-0.9775	0.4777	-0.3687
Kurtosis	0.6484	1.9030	-1.1287	-0.1989	1.2258	1.7058	2.5322	1.5732	0.1197	-0.7301

* calculated from grouped data. Numbers in parenthesis indicate n for subsets, while the rest were all 28 students.

The essential difference in this new test compared with the usual school examinations was both in the nature of the examination and the performance in relationship to the year-wise content. Since the test was given by the teachers themselves following the norms of school examinations for each year and subject, there was no influence of any theory of pedagogy or ideology in the test per se. Higher variance in humanities and languages, which also contribute to Model III, particularly need to be traced to problems in teaching methods and content.

That this test, with all its limitations and no external intervention, could give a robust attribute, the slope α , raises two distinct possibilities to ponder. Either it is a measure of the limits to mass learning or it is a measure of the current practices of learning and offers a completely novel and unique metric for the competence of systems of education, i.e., the students, teachers and the curriculum by appropriate emphases in analyses.

IV. Discussion:

There is nothing that prohibits large, complex social systems from exhibiting highly tractable and even simple behaviours [58], while confirming the derived models may not be all that simple. This is also at the heart of physical reasoning as seen with thermodynamics. In the last some 150 years, this is the first time that a clear mechanism for Engel's curves based on micro-time constants has been developed and a hierarchy of consumption thereby established a major departure, as also seen with the other publications from our group, from conventional econometrics. It is definitely surprising that the modified logic of hierarchical hyperbolic consumption of various commodities has taken this long to establish since Engel's pioneering observations and that, despite publications spanning two decades from our own group, significant corrections have not been taken by economists to date. Now that a welfare design for national development based on poverty line is established to be an absurd notion unsupported by scientific evidence such as the nutritional norm and is lost in the polemics of political debate, the commodity specific poverty Index (SI) [11-13,39,40] in an appropriate form needs to be applied at the national level. This permits taking care of individual commodities in the creation and delineation of the needs of people based on their weightages, education being the major concern in this manuscript. We must point out that the existence of independent specialists for each commodity such as education or health or food supplies in isolation without the overall view of consumption has not been helpful in the development of clear-cut policies in defining the needs in the country.

IV. 1: The need for good data and competent handling by modelling:

It is not often realized that statistics is a relatively young discipline (compared to, say, physics). It is ideally suited for small samples, the *raison d'être* of its development, which also limits the classes of problems that can directly be handled. Statistical design and even outside the box thinking therefore becomes of paramount importance, tailoring the data requirements to specific ends. However limited is the possibility of chance associations in the face of excellent regression values, selective studies by design need to be enhanced to the actual population level and verified. Since that would be very expensive to start with, the focussed small studies are critical to the development of the insights into planning large scale data acquisition as we show here.

The plea by Iannidis [59] for statistical veracity in medical research was independently found to be very valid in education and econometrics in our own studies, the basis of much of the novel results reported here. General consumption of commodities requires to be ordered by analysis since its ordinal nature is not given *a priori*. Our results obtained on econometrics and econometrics of education differ from the rest, not because of any preconceived models, but by strict adherence to the best practices, e.g., i. avoiding transformations on data and ii. Directly handling data with non-linear regression methodology, iii. actually and assiduously confirming at every stage that the residuals were indeed random by non-parametric tests such as runs test and, since correlation does not guarantee causality, and iv. arriving at derivable models based on 'mechanisms' in the physical sense to assess the validity of the conclusions. It is amazing that a log-log relationship is still adopted by econometricians without ever validating whether the residuals were acceptably random (Fig.1). Our hierarchical model of consumption could get rid of the arbitrary and contentious poverty line and led to a commodity-specific poverty index without any axiomatic approaches that defy independent proof [cf. 11-13, 39, 40]. The hierarchy inherent to consumption and even education as a commodity has not been visualized in previous studies [60, 61]. Certain credit goes to Piketty for having not just an alternative view in economics but a distinct disenchantment with overworked mathematics that lost a sense of reality among practitioners of economics. Piketty rightly warns of magical mathematical solutions that economists love to solve societal problems with, which they indeed do not, but use abstruse arguments only to scare away the less formally trained scholars from the more needy social disciplines [62]. Pedagogy is no exception and the use of language is an equally potent source of obfuscation in handling complexity as is mathematics. Education being the only commodity that is naturally ordered (in the sense that attendance in grades 1-12, is

obligatorily sequential), it also confirms similar work from our group on the consumption preferences among rural and urban Indian populations [11].

Certain aspects of education are outside the scope of economics, in the sense of the need for intervention by agencies external to the community. One is that total literacy and even primary education can be achieved, despite lack of eradication of poverty, as seen with the example of Kerala state in India [63-65]. Another somewhat taboo topic is the limitations of learning in children, i.e., whether there are biological limits to learning in a particular child to a particular level at a specific age [66]. Acceptance of gifted children as a reality needs to consider the reality of the less endowed as well [67]. Natural variation would justify such a proposition to some level, but the socio-economic determinants would not permit any clear answer to the all-important question: how much should a child be taught [68]? Thus, only when we distinguish true learning (defined as long term retention) from that influenced by the socio-economic ambience, it will be possible to address the question dispassionately and on an even keel.

IV. 2. Education as a commodity

Education is one among many such universal social systems/processes. Is it really a commodity (that is purchased) or an investment (voluntarily deployed by a family towards the future of the wards)? The answer would be that such a distinction would be superfluous if the income dependence of any metric of education manifests hyperbolic behaviour. Its economic dependence is manifest in the early years of schooling wherein it follows the rigorous contours of a hierarchical econometric theory, highly consistent with our previous work [cf. 11-13, 39-40]. The influence of socioeconomic conditions percolates into higher education in conventional schooling and performance which involves cramming. The mistrust in the conventional grading system based on year-end examinations is justified to the extent that it tends to miscategorise those who learn and retain and yet are less capable of cramming. Nowhere in the world is a given grade/score as robust as it is presumed to be and neither has the system has offered a grade along with the confidence interval that must be associated with it. Periodic intensive cramming evidently requires a social setting that allows the student to minimize distractions in the short term for the purpose of examinations. These insights have become possible with concrete data for the first time.

Our studies on education also focus on social justice. An evaluation of learning that is susceptible to social forces as opposed to learning per se is a deep unresolved anomaly in the

practice of pedagogy. This can now be addressed to in concrete terms, i.e. α , which is independent of school grades, ergo, socio-economic status.

IV. 3. School dropouts:

It is clear that the results reported here regarding retention in learning require large scale confirmation in and outside the country, possible only by a strict adherence to the methodological innovations introduced here. It is necessary to establish a quantitative relationship between current school grades and the socioeconomic status, while the existing studies relate more qualitatively. Also one needs to establish in prospective studies the nature and possible invariance of distributions in grades since one cannot presume them to be constant. In fact, the future of pedagogy lies in arriving at and deriving as many performance indices as possible across populations and time. On the other hand, school dropouts and attainment of universal literacy (even if reduced to the simplest form of reading and writing) represent a well-documented and yet, a highly remediable situation, despite the economic constraints [49]. It demands a social momentum as the prime mover as demonstrated in the state of Kerala in this country even without gross economic development [63-65]. Since the major focus on school dropouts shifts to rural areas, it is prudent to consider that the economic upliftment requires equal attention besides a spread of awareness; the huge social cost of dropouts (which far exceeds that of any intervention) requires that it must be minimized by any and all means.

IV. 4. Fall in Retention:

The major result here relates to the revelation of a metric, α , of fall in retention that is **independent** of the socio-economic status and gender, inferred from their independence of routine school grades. What to do to ameliorate the situation is evident from the results already. Since the fall, α , varies with the subject (i.e. content), it defines the focal point for development in the design and delivery of the content and the examination system as well. What makes a subject be readily forgotten? Clearly the learning that fails to relate to our lives, i.e., the knowledge that is ‘bookish’ as opposed to ‘practical’, has little to enforce its retention [69]. We see that the retention remains highest in the first three years of schooling, since these basics are a prerequisite for all the subsequent years, while the subsequent years do not offer such a rigorous continuity.

IV. 5. Learning in ‘silos’:

The irrelevance of the subjects taught has two foci: firstly each subject, as evidenced by the textbooks, is independent of all others, i.e., in silos; secondly, so do teachers restrict

their involvement to little beyond their own teaching in a specified subject [70]. How do we expect the students to correlate, integrate and develop a homogenous view of their learning independent of the content and the teachers? The usability of the subject has two considerations. One relates to extent of usage **across subjects** such as Science in Mathematics or English in Science and the other relates to **across years**, whose continuity can be assessed by concept maps and other tools [71-73]. A major revision of content should aim at, say, more than 30% overlap between subjects while the teachers themselves be proficient in more than 2-3 subjects. The aim, which does not exclude teachers, is obviously rendering α to 90° .

IV. 6. Testing as an adjunct to actual learning:

Evaluation by examinations itself is a practical way to enhance retention, to the extent that we need to consider examinations not as something to be feared but as a useful adjunct to learning [74, 75]. Then the examination system has to change from the year-end torture to self-help *ad libitum* by students themselves, or, as assisted by teachers, largely to appreciate what they have **not** learnt. What if the content itself is measured by the number of non-repetitive class of questions that it can generate, rather than the number of pages or even the weight of the book? That needs to be supplemented by efforts that use and reuse the earlier content in the later years, supported by clearly continuous concept maps creating a demand for recall and utilization. That relates well to the normal practices in professional education, e.g., medicine or engineering. The student has to see the relevance of what is learnt within the learning process itself and/or in his life at school and home.

These were also deeply embedded notions in education in India for millennia as evidenced by Panchatantra [cf.17], which even taught value systems as practicable entities. The ultimate example of the finest mixture of poetry and language and mathematics is embodied in Lilavati that Bhaskara II wrote in chaste Sanskrit for his daughter [76]. One may identify many texts in English and other languages and even creative writing of new interdisciplinary texts is eminently feasible.

All this is possible once we accept, unconditionally, that evaluation is an integral form of education. We had on several occasions discussed the relevant matters with several teachers, parents and students as to their perceptions, though not measured. Dilution in the methods of evaluation has become indispensable; the school administration is answerable to the performance statistics rather than performance per se [77]. The only way is to promote most, if not all, students as the schools need to cope with large number of failing students

[78]. Justifications include the ‘psychological stress’ [79] imposed on the students by examinations, a perennial source of anxiety for parents! This argues strongly for a fresh look at examinations as a learning device and using modern technology to reduce the tedium for both students and teachers. Laptops and automation are better suited for testing and not as a substitute to teaching.

IV. 7. Universality of problems of pedagogy: the general case

Our results are unlikely to be specific for the country, India or Brazil, and share the universal concern for the poor state of education, with each country visualizing and battling its own problems. This in itself points out the inadequacy of the existing theories of pedagogy in visualizing a common metric. Barring war-torn, oppressive and disaster-struck areas [80], if the decline in retention is independent of peoples and intrinsic to the current process of education itself, a major effort to examine the alternatives will be worthwhile. The teachers who formulate examinations may not themselves fare well in the same examinations across the subjects. For instance, if the teachers take the same tests as the students and the performance of the two mutually match and outrank the other pairs of teachers and students between schools... that would be a realistic assessment of the teaching process. Learning necessarily implies retention. The current results do not even follow the direction predicted by models (e.g. Model II) of the much maligned rote learning, which should really be viewed as **pre-principle learning**, mandatory if learning has to be effective. While rote learning has a definitive place in schooling, rote teaching has none. The resulting reorientation in educational methodologies and reforms will not be small in eliminating learning/teaching of subjects in silos (the content mostly suffers from this ‘isolation’ of subject matter), which appears to be a major source of disorder in achieving coherence in teaching!

The decline in learning could indicate limits to learning itself (i.e., of principles) at the mass level, a worrisome but a real prospect, since it also sets practical limits to what can be achieved. Any approach that makes education prohibitively expensive (e.g., due to unregulated privatization and market forces) can be evaluated for its true effectiveness and curtailed if not justified. We need to extend the current focused and yet limited study, primarily empirical rather than theory based, to wider national and international levels that include students, teachers and curriculum(and possibly value systems), a feat within easy reach in these days of automation and big data handling capabilities, given the will[81].

Lastly, the studies yielded totally unforeseen dividends in that the ‘affinity’ (1/K) for investing on education in a family goes downhill initially but actually begins to increase midway through schooling as the end is in sight. It is correct to reject a regression based on

there being runs, not acceptable to the requirement of residuals being random. Even better is to find out what that residual pattern exactly is, as we have seen with Brazilian education. Even if primary education is considered 'free', the hidden costs of that education that the family incurs which includes the attention paid to the children determines the success of continuing education towards completion of schooling. Value systems play a major role which includes the decision processes at home. Unfortunately, each commodity, education, health etc., represent domains of expenditure and are analysed independent of each other, while a family makes a decision on each expenditure, representing the sum total of the needs and aspirations. We see that in rural India, education plays a secondary role to health expenditures while the urban India sees them with equal emphasis.

Clearly, the triad of learning, teaching and retention should be assessed independent of socioeconomic circumstances. We also see that socioeconomic circumstances could be assessed by different parameters such as dropout rates quite independently. More importantly, the strongly empirical testing approaches developed here can be duplicated anywhere in the world and can be assessed for the veracity of such a conclusion, which it should be. Viewing testing, as in conventional sense, as a necessary evil should cease and testing which does not require cramming but helps identify for the student, what is not known, should become an integral part in learning. The importance of a measure of learning per se, i.e. α , independent of socioeconomic circumstances, is novel and much desired and to be further validated elsewhere. The results from large data bases are emphatic enough and have been proven to be valid in all the years from NSSO data. The models so arrived at speak volubly for their validation in other landscapes of learning since pedagogy is one area of human social interest undimmed over millennia and yet fraught with much conflict, theoretical, value based, social and religious and carries its own baggage.

IV.8. The need for a change in set paradigms.

The difficulty in reforms lies in their execution in a community that does not have the clarity of what the prospects are, as with STEM education [cf.20]. The isolationist policy of cherry-picking students after school, based on conventional (both routine and specially crafted) evaluations, into extremely selective institutes (like IITs etc., in India, prevalent in rest of the world as well) cannot substitute for a national strategy for STEM education [82] since it needs to be mass-based, aiming at the creators as well as users of that knowledge. At the same time, the income-dependent [cf. 26,46], high dropout rates (> 30% in India) and grossly inadequate learning at the mass level do not indicate any success of public policies

(cf. studies by Pratham [49] and elsewhere [83], followed by preliminary reports of baseline tests by the several Indian states [84-86] as with the rest of the world [48]). Insights that facilitate causally relevant intervention for students are few. Evaluation of teachers is equally dismal [87]. Percentage of teachers that pass the National teacher eligibility test (TET) at the middle school level in India varies from 0.6% (2012) to 12% (2015) [88,89]. The third critical component, the curriculum design itself, is with hardly any metric or feedback [cf. 74, 90]. The ability to formulate varied and yet unambiguous questions offers a good method to assess both the teachers and the curriculum, when quantified.

Through the social pressures in the community, the relatively less aware may find an isolationist approach to subjects and careers preferable; it is so not because of any virtue associated with compartmentalized learning but due to the intellectual myopia characteristic of ingrained ignorance of preceding generations that learnt mostly in silos. It appears very plausible that reforms in teaching, testing and curricula that aim at a progressive erasure of learning in silos are most likely to achieve an improvement in α towards 90° . The practice of elimination of students not capable of mere cramming alone would amply justify these remedial efforts.

Educational expenditures, after validation of the adequacy and cost-efficacy of the education that the institutions brought about, will contribute to the future assessment of battle against economic deprivation of populations. The future course, besides a large scale confirmation of the results on retention by students, clearly requires that the teachers and curriculum also be assessed by similar means. Being a social process, it essentially depends on the political will generated by the people. Introduction of automation for the entire testing process in schools would ultimately help eradicate malpractices, substandard teaching, and multiple entrance tests for higher studies and would also help ultimately answer the central question: how much should a child be taught.

Acknowledgements:

The experiment was **not financially supported** by any external agency and was primarily executed by the authors themselves, supported by teachers and students in an English medium school in an urban Indian setting.

Individual contributions:

Sitaramam (VS) formulated the problem and designed the experiment and completed the analysis, modeling and finalizing the paper. PVI (VS's daughter) initially raised the question of how the teacher training in Bachelor of Education courses match the teaching requirements and subsequent performance of both teachers and students. She helped in screening the question banks and preparing the school data for analyses, discussions and preparation of the manuscript. There are no competing interests. The authors had or have no association with the school nor was the school associated with analyses and preparation of the manuscript.

References

1. Hanushek, E. A., Woessmann L, The Role of Education Quality for Economic Growth (February 2007). World Bank Policy Research Working Paper No. 4122. Available at SSRN: <https://ssrn.com/abstract=960379>
2. Barro, R.J., 2001. Education and economic growth, The contribution of human and social capital to sustained economic growth and well-being, pp.13-41 <http://www.oecd.org/edu/innovation-education/1825455.pdf>
3. McMahon, W.W., 2009. *Higher learning, greater good: The private and social benefits of higher education*. JHU Press.
4. Aghion, P., Boustan, L., Hoxby, C. and Vandenbussche, J., 2009. The causal impact of education on economic growth: evidence from US. *Unpublished paper*.
5. White KR (1982) The relation between socio-economic status and academic achievement, *Psychological Bulletin* 91: 461-481.
6. Hanushek, E.A. and Woessmann, L., 2012. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of economic growth*, 17(4), pp.267-321.
7. Engel E (1857), "Die Productions- und Consumptionsverhaeltnisse des Koenigsreichs Sachsen," *Zeitschrift des Statistischen Bureaus des Koniglich Sachsischen Ministeriums des Inneren*, No. 8 und 9. Reprinted in the Appendix of Engel (1895). "Die Lebenskosten Belgischer Arbeiter-Familien Fruher and jetzt", *International Statistical Institute Bulletin*, 9: 1-74.
8. Glatthorn, A. A., & Jailall, J. (2000). Curriculum for the new millenium. In R. S. Brandt (Ed.), *Education in a new era* (pp. 97–121). Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).
9. Bateson, W. and Mendel, G., 2013. *Mendel's principles of heredity*. Courier Corporation.
10. Ebbinghaus H (1885), *Memory: A Contribution to Experimental Psychology*, (Translated by H. A. Ruger & C. E. Bussenius, 1913); Originally published in New York by Teachers College, Columbia University. <http://psychclassics.yorku.ca/Ebbinghaus/index.htm>
11. Sitaramam V, Paranjpe SA, Krishnakumar T, Gore AP, Sastry JG (1996). Minimum needs of the poor and priorities attached to them, *Economic and Political Weekly*, Special Number, September, 2499-2505.
12. Gore AP, Kumar TK, Paranjpe SA, Sitaramam V (1996), Measurement of Poverty through Commodity specific Consumption Deprivation. *Proceedings of the Thirty second Annual Conference of the Indian Econometric Society* March 2: 21-23
13. Kumar TK, Gore AP, Sitaramam V (1996), Some conceptual and statistical issues on measurement of poverty. *J. Stat Plan Inference* 49: 53-71.
14. Costa Rebeiro SA (1992), Educação e a Inserção do Brasana Modernidade, In: Série Documentos. Instituto de Estudos Avançados AJSP. In: Cadernos de Pesquisa no 84, Cortez Editora, Fundação Carlos Chagas, fevereiro de, 63-82 (1993).
15. Kember D (1996), The intention to both memorize and understand: Another approach to learning? *Higher Education*, 31: 341-354.
16. Plato (1991), *The Republic*, 2nd Edition (English Translation, Allan Bloom), Basic Books, New York.

17. VisnuSarma, Panchatantra (1993), (Translated by Chandra Rajan), Penguin Books India (P) Ltd.
18. Gandhi MK (1931), Collected works of Mahatma Gandhi, Speech at Chatham House Meeting, 48: 193-206, Retrieved from
https://www.gandhiheritageportal.org/cwmg_volume_thumbview/NDg=#page/224/mode/2up
19. Battle J, Lewis M (2002), The increasing significance of class: The relative effects of race and socioeconomic status on academic achievement. *Journal of Poverty* 6:21-35. Published online: (2008)
http://dx.doi.org/10.1300/J134v06n02_02
20. STEM Education: To build a scientist, (2015) *Nature*, 523: 371-373,
21. Saxton E, Burns R, Holveck S, Kelley S, Prince D, Rigelman N, Skinner EA (2014), A Common Measurement System for K-12 STEM education: Adopting an educational methodology that evaluates theoretical foundations and systems thinking, *Studies in Educational Evaluation* 40: 18-35.
22. Schinske J, Tanner K (2014) Teaching More by Grading Less (or Differently). *CBE Life Sciences Education*. 13: 159-166. doi:10.1187/cbe.CBE-14-03-0054.
23. Black P (1998), Testing: friend or foe? Theory and practice of Assessment and Testing, London, Falmer Press.
24. Pellegrino JW, Chudowsky N, & Glaser R, (Eds.) (2001). Knowing What Students Know – The Science and Design of Educational Assessment. National Academy Press, Washington, DC:
25. Kember D (1996). The intention to both memorize and understand: Another approach to learning? *Higher Education*, 31: 341-354.
26. Educating for a Sustainable Future, A Transdisciplinary vision for Concerted Action (1997), UNESCO, EPD-97/CONF.401/CLD.1
http://www.unesco.org/education/tlsf/mods/theme_a/popups/mod01t05s01.html
27. NSSO, (National Sample Survey Organization), 50th round (June 1993 – July 1994). Reports and Publications. Ministry of Statistics and Programme implementation, Govt. of India. Differences in level of consumption among socio-economic groups. (NSSO, MOSPI),
http://mospiold.nic.in/national_data_bank/ndb-rpts-50.htm
28. Gibbons JD (1976). Nonparametric methods for quantitative analysis. International series in decision process. In: Massey H, (Ed.), American Science Press, Inc., Ohio, USA
29. Dandekar VM, Rath N, (1971): Poverty in India, Indian School of Political Economy
30. Lehninger AL, (2004), *Principles of Biochemistry*. 4th Edition. Ed. , D.L.Nelson, M.M.Cox, W.H.Freeman. New York
31. Sitaramam V, Sambasivarao D, Mathai JC, (1989) Differential effects of osmotic pressure on mitochondrial respiratory chain and indices of oxidative phosphorylation. *Biochim. Biophys. Acta* 975, 252 - 266.
32. Mathai JC, Sitaramam V, (1989) Variable porosity of the mitochondrial inner membrane induced by energization. *Biochim. Biophys. Acta* 976, 214-221.
33. Mathai JC, Sauna ZE, John O, Sitaramam V, (1993) Rate limiting step in electron transport: osmotically sensitive diffusion of quinones through voids in the bilayer. *J. Biol. Chem.*, 268, 15442-15454.
34. Wainio W, (1970) *The Mammalian Mitochondrial Respiratory chain*, Academic Press, New York
35. Rajagopal K and Sitaramam V, (1998) Biology of the inner space: voids in biopolymers. *J. theoret. Biol.* 195, 245-271.

36. Sitaramam V, Janardana Sarma MK, (1981) Gravitational field enhances permeability of biological membranes to sucrose: An experimental refutation of sucrose- space hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 78, 3441-3445.
37. Sambasivarao D, Sitaramam V, (1983) Studies on the non-linear osmotic pressure-volume relationship in mitochondria and entry of sucrose into the matrix space during centrifugation. *Biochem. Biophys. Acta.* 722, 256-270.
38. Pradhan GR, Pandit SA, Gangal AD, Sitaramam V (2003) Shape anisotropy of lipid molecules and voids *J.theoret. Biol.* 220, 189-199.
39. Kumar TK, Mallick S, Holla J (2009), Estimating Consumption Deprivation in India Using Survey Data: A State-Level Rural-Urban Analysis Before and During Reform Period, *Journal of Development Studies*, 45: 441- 470
40. Kumar TK, Holla J, Guha P (2008), Engel curve method for measuring poverty, *Economic and Political Weekly*, July: 115-123; <http://www.mse.ac.in/wp-content/uploads/2016/10/Engel-Curve-Method-for-Measuring-Poverty.pdf>
41. Holling CS (1959a), The components of predation as revealed by a study of small-mammal predation of the European sawfly. *Can. Entomol.* 91: 293–320
42. Holling CS (1959b), Some characteristics of simple types of predation and parasitism. *Can. Entomol.* 91: 385– 398
43. Dawes JHP, Souza MO (2013). A derivation of Holling's type I, II and III functional responses in predator–prey systems. *Journal of theoretical biology* 327: 11-22.
44. Banks, J.; Blundell, R.; Lewbel, A. (1997). "Quadratic Engel Curves and Consumer Demand". *Review of Economics and Statistics*. **79** (4): 527–539. doi:[10.1162/003465397557015](https://doi.org/10.1162/003465397557015)
45. https://en.wikipedia.org/wiki/Engel_curve
46. UNESCO Statistical Year Book (1994), UNESCO, Paris.
47. World Bank (2013), World Development Indicators 2013, Washington, DC: World Bank. doi: 10.1596/978-0-8213-9824-1. License: Creative Commons Attribution CC BY 3.0, <http://wdi.worldbank.org/tables>
48. OECD iLibrary (2016), *PISA 2015 Results (Vols I and II), Excellence and Equity in Education, Policies and Practices for Successful Schools*, OECD Publishing, Paris. DOI: <http://dx.doi.org/10.1787/9789264266490-en>; <http://dx.doi.org/10.1787/9789264267510-en>
49. Annual Status of Education Report, Pratham, Maharashtra, India, Accessed on 6th Nov, (2016). http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202015:%20Maharashtra/asr2015_maharashtra.pdf
http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/asr2016_nationalpressrelease.pdf
50. Vos R (1996), *Educational Indicators: What's To Be Measured?*, Integration and Regional Programs Department, Inter-American Institute for Social Development, Inter-American Development Bank,.
51. Samuel M, Silva A (1994). Problems and performance in primary education: Why do systems differ? Social Agenda Policy Group, Inter-American Development Bank, Washington D.C. (mimeo).

52. Filmer D, Pritchett LH (2001), Estimating Wealth Effects without Expenditure Data-Or Tears: An Application to Educational Enrolments in States of India. *Demography* 38: 115-132.
53. Sitaramam V (1998), On recommending students: the curious case of all students being in the top ten per cent. *Current Science*, 75: 647-652
54. Sitaramam V, Bhate R, Kamalraj, P. *Physiol. Mol. Biol. Plants* (2008), Respiration hastens maturation and lowers yield in rice., 14:253-271
55. Entwisle DR, Hayduk LA (1981), Academic expectations and the school attainment of young children. *Sociology of Education* 54: 34-50.
56. Coleman JS, Campbell EQ, Hobson CJ, McPartland J, Mood AM, Weinfeld FD, York RL (1966), Equality of Educational Opportunity survey, (A publication of the National Center for Educational Statistics), Washington, D.C., U.S. Government Printing Office, <http://doi.org/10.3886/ICPSR06389.v3>
57. National Council for Educational Research and Training, National Curriculum Framework (2005), India <http://www.ncert.nic.in/rightside/links/pdf/framework/english/nf2005.pdf>, <http://www.ncert.nic.in/departments/nie/desm/publication/pdf/teacher'shandbook.pdf>
58. Simon HA (1962), The Architecture of Complexity, *Proceedings of the American Philosophical Society*, 106: 467-482
59. Ioannidis JPA (2005) Why most published research findings are false. *Plos Medicine* 2: 0696-0701, e124.
60. Deaton A, (2003) How to monitor poverty for the millennium development goals, *J. Human Development* 4: 353-378.
61. Sen AK, (1981) *Poverty and Famines*. Oxford University Press, Oxford
62. Piketty, T., (2013) *Capital in the twenty-first century*, Harvard University Press, translated by Arthur Goldhammer
63. Shalom SR, (1999). Lessons and hope from Kerala. <https://zcomm.org/zcommentary/lessons-and-hope-from-kerala-by-stephenl-shalom/>.
64. Kerala State Literacy Mission Authority (1991) <http://www.literacymissionkerala.org/>
65. George KK, Sunaina P, (2005), Dynamics of Change in Kerala's Education System: The Socio-economic and Political Dimensions, Working paper No.12, Centre for Socioeconomic and Environmental Studies, <http://cseindia.org/admin/modules/cms/docs/publication/12.pdf>
66. Chi, M.T., 1976. Short-term memory limitations in children: Capacity or processing deficits?. *Memory & Cognition*, 4(5), pp.559-572.
67. Allan, S.D., 1991. Ability grouping research reviews: What do they say about grouping and the gifted. *Educational Leadership*, 48(6), pp.60-65.
68. Jarvis, M., 2005. *The psychology of effective learning and teaching*. Nelson Thornes.
69. Dewey J, *Experience and Education*, Free Press (reprint edition 1997)
70. Thorp H, Goldstein B, (2010), *Engines of Innovation: The Entrepreneurial University in the Twenty-First Century*, Chapel Hill: UNC Press
71. Ausubel DP (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston.
72. Novak JD, Cañas AJ (2006), *The Theory Underlying Concept Maps and How to Construct and Use Them*, Technical Report IHMC CmapTools, Florida Institute for Human and Machine Cognition, 2008, available at: <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>

73. Starr ML, Krajcik JS (1990), Concept maps as a heuristic for science curriculum development: Toward improvement in process and product. *J. Res. Sci. Teach.*, 27: 987–1000. doi:10.1002/tea.3660271007
74. Duncan RG, Hmelo-Silver C (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal for Research in Science Teaching*, 46: 606–609.
75. Shepard AL (2016), The Role of Assessment in a Learning Culture, *Educational Researcher*, 29: 4-14
76. Lilavati of Bhaskaracarya (set of 2 volumes) (2015); Translated and edited by A.B. Padmanabha Rao, Chinmaya International Foundation Shodha Sansthan, Ernakulam, Kerala
77. Connor-Greene P, (2000), Assessing and Promoting Student Learning: Blurring the Line between Teaching and Testing, *Teaching of Psychology*, 27: 84-88.
78. Leonard B (1997), *Passing on Failure: District Promotion Practices and Policies*, Diane Publishing Company, Darby, PA
79. McDonald AS (2001), The Prevalence and Effects of Test Anxiety in School Children, *Educational Psychology* 21: 89-101, <http://dx.doi.org/10.1080/01443410020019867>
80. Gindeel R, Etim JS, (2014) Education in War Affected Areas (Internally Displaced Persons communities) in Eastern Sudan, *Journal of Studies in Education*, 4: 108-121.
81. Kohn A, (1999b). *The schools our children deserve: Moving beyond traditional classrooms and “tougher standards.”* Houghton Mifflin. Boston.
82. Mehta MG, Sharan R, (2016), IITs and the Project of Indian Democracy; *Economic and Political Weekly*, LI: 12-14.
83. Nanda P, *Livemint newspaper*, Jan 14, (2015), Accessed on 6th Nov, 2016.
<http://www.livemint.com/Politics/ZNRkJBrnB4NG9bJzLx5IaO/India-disappoints-in-educational-outcome-test.html>
84. Maharashtra State Board of Secondary and Higher Secondary Education reports, Pune division
<https://mahahsscboard.maharashtra.gov.in/statisticssc.htm>, accessed on 6th Nov. 2016.
85. Gunotsav Results of Gujarat: <http://www.gunotsav.org/gunotsavaheval.aspx>, accessed on 6th Nov. 2016.
86. Sharma R, <http://indianexpress.com/article/cities/ahmedabad/six-lakh-class-vi-students-cant-write-in-gujarati-5-lakh-cant-read-2821226/> accessed on 29 May 2016.
87. Shankar VK, Saini R (2015), What does an MA know? Postgraduate Learning Deficits and the Diploma Disease in Social Sciences, *Economic and Political Weekly*; L: 40-48.
88. Central Teacher Eligibility Test, Central Board of Secondary Education, Delhi
<http://cbse.nic.in/attach/Press%20Note%20reg%20Successfully%20conduct%20CTET.pdf>
89. Summary of analysis of results, <http://www.entrancecorner.com/results/ctet-result/> retrieved on 6th Nov 2016.
90. Porter AC, Smithson, John L, (2001) Defining, Developing, and Using Curriculum Indicators, CPRE Research Report Series, RR-048, Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education, December. <https://secure.wceruw.org/seconline/Reference/rr48.pdf>