Economy of thought: a neglected principle of mathematics education

Alexandre Borovik

Introduction

The aim of science is to seek the simplest explanations of complex facts.
We are apt to fall into the error of thinking that the facts
are simple because simplicity is the goal of our quest.
The guiding motto in the life of every natural philosopher should be,
"Seek simplicity and distrust it."

Alfred North Whitehead

I contribute this paper to a volume on the fascinating topic of simplicity in mathematics; my paper is about the role of simplicity and "economy of thought" in mathematics education; it focuses on the early age, elementary level mathematics education. Originally I was planning to extend the narrative at least up to Bourbaki's project (it is worth remembering that the latter started as a pedagogical exercise[1]), but I soon discovered that elementary school mathematics already provided more material than I could fit in a paper. So I mention Bourbaki only briefly, see Section 7.

This paper is not supposed to be a kind of theoretical musing; indeed many of its passages come from my letters to education professionals and civil servants written in 2011–14, mostly in the context of discussions around the National Curriculum reform in England. The paper is written for adults, not for children – please do not

Department of Mathematics, The University of Manchester, Manchester, UK
web: www.borovik.net
email: alexandre ≫at≪ borovik.net

[1] Bourbaki's original aim was a compact textbook of functional analysis "where every theorem is proved only once" – and they succeeded in turning a few of their books or chapters from books – say, the celebrated *Topologie générale* – into true masterpieces of pedagogical exposition and simplicity in mathematics. See Corry (2011) for more detail.

see it as a source of learning materials for primary school, even if most problems are very accessible. The selection principle for problems was the potential depth of didactic analysis that they allowed, not possibility of the immediate use in the class.

An emphasis on old Russian sources is easy to explain: I am frequently asked to comment on the Russian tradition of mathematics education. The latter might appear to be outdated (it suffices to say that the state and the social system where it has flourished no longer exist), but, I believe, it continues to be relevant. After all, as Stanislas Dehaene (2000) quipped in his book *The Number Sense*,

> *We have to do mathematics using the brain which evolved 30 000 years ago for survival in the African savanna.*

For that reason, I believe, a discourse on mathematics education should involve historic retrospection on a timescale longer than a few years or even a few decades.[2]

I focus on examples from arithmetic and from elementary set theory, mostly for lack of space for anything else in a short paper, and I wish to warn those readers who are not very familiar with mathematics:

> *Arithmetic is not the whole of mathematics, it is only one of its beginnings. Mathematics competence is more than "numeracy" because even competence in arithmetic is much more than "numeracy".*

I hope that the present paper proves this thesis. I can claim more:

> *Restricting mathematics education to teaching "numeracy", "practical mathematics", "mathematics for life", "functional mathematics" and other* ersatz *products is a crime equivalent to feeding children with processed food made of mechanically reconstituted meat, hydrogenated fats, starch, sugar and salt.*

Following this culinary simile, real simplicity in mathematics education is not fish nuggets made from "seafood paste" of unknown provenance, it is sashimi of wild Alaskan salmon or Wagyu beef. Unlike supermarkets, huge Internet resources provide ingredients for a simple, healthy, tasty, exciting, even exotic gourmet cuisine for mathematics education *for free*. But we have no cooks.

1 What can be simpler than $3-1=2$?

What follows is a translation of a fragment from Igor Arnold's (1900–1948) paper Arnold (1946). It goes to the heart of the role of simplicity in mathematics education. For research mathematicians, it may be interesting that I. V. Arnold was V. I. Arnold's father.

[2] See more about mathematics education in Soviet Russia in my forthcoming paper Borovik (2017).

> *Existing attempts to classify arithmetic problems by their themes or by their algebraic structures (we mention relatively successful schemes by Aleksandrov* (1887)*, Voronov* (1939) *and Polak* (1944) *are not sufficient [...] We need to embrace the full scope of the question, without restricting ourselves to the mere algebraic structure of the problem: that is, to characterise those operations which need to be carried out for a solution. The same operations can also be used in completely different concrete situations, and a student may draw a false conclusion as to why these particular operations are used.*
>
> *Let us use as an example several problems which can be solved by the operation*
>
> $$3 - 1 = 2.$$

Igor Arnold then gives a list of 20 problems of which we quote only a few.

(a) *I was given three apples, and then ate one of them. How many were left?*
(b) *A barge-pole three metres long stands upright on the bottom of the canal, with one metre protruding above the surface. How deep is the water in the canal?*
(c) *Tanya said: "I have three more brothers than sisters". How many more boys are there in Tanya's family than girls?*
(d) *How many cuts do you have to make to saw a log into three pieces?*
(e) *A train was due to arrive one hour ago. We are told that it is three hours late. When can we expect it to arrive?*
(f) *A brick and a spade weigh the same as three bricks. What is the weight of the spade?*

⋮ ⋮

(s) *It takes 1 minute for a train 1km long to completely pass a telegraph pole by the track side. At the same speed the train passes right through a tunnel in* 3 *minutes. What is the length of the tunnel?*

These 20 completely different arithmetic problems, all solvable by the operation $3 - 1 = 2$, make it abundantly clear that the so called "word problems" of arithmetic involve identification of mathematical structures and relations of the real world and mapping them onto better formalised structures and relations of arithmetic, or, in Igor Arnold's words,

> *These examples clearly show that teaching arithmetic involves, as a key component, the development of an* ability to negotiate situations whose concrete natures represent very different relations between magnitudes and quantities.

And many of the 20 problems are deep – they are concerned with combinatorial properties of sets of objects in the world, with geometry of space and time, and even with what some adults call *simplicial homology*: Problem (d) is a one-dimensional version of the Euler Formula and a seed (well, maybe a spore) of the entire algebraic topology.

Even more important is Igor Arnold's characterisation of arithmetic:

> *The difference between the "arithmetic" approach to solving problems and the algebraic one is, primarily the need to make a concrete and sensible interpretation of all the values which are used and/or which appear at any stage of the discourse.*

I suggest that Igor Arnold's observation deserves to be raised into one of the characteristic aspects of simplicity in mathematics:

> *An important, and, at the early stages of mathematics education, predominantly important class of "simple" definitions, arguments, or calculations in mathematics is the one where all intermediate structures and values have an immediate interpretation in some lower level and better understood mathematical theory, or in the "real world" of physics, economics, etc.*

I suggest calling it *Arnold's Principle*, intentionally blurring the line between Igor Arnold and Vladimir Arnold; the famously controversial writings by the latter about mathematics education made it obvious that he was much influenced by his father's ideas Arnold (1998). Importantly, Vladimir Arnold republished his father's paper as Arnold (2008) and endorsed it in his touching foreword.

For the needs of elementary mathematics teaching, Arnold's Principle can be reformulated in shorter form:

> ***A "simple" mathematical calculation or argument is the one where all intermediate values and statements have a concrete, immediate, and sensible interpretation in the "real world".***

Among other, and much more advanced sources of simplicity in mathematics we find *abstraction by irrelevance*: removal of all irrelevant details from a concept or a statement and subsequent re-wording of the essence of the matter in a most general form. The classical examples here are Bourbaki's definition of uniform structure and uniform continuity and Kolmogorov's definition of a random variable as a measurable function. Remarkably, both of these celebrated definitions have elementary facets which allow them to be compliant with Arnold's Principle. I briefly discuss them later in the paper.

2 Encapsulation and de-encapsulation

Arnold's Principle fits into the all-important dynamics of encapsulation and de-encapsulation in learning mathematics with precision so remarkable that it deserves some analysis.

The terms "encapsulation" and "de-encapsulation" are not frequently used, and a few words of explanation will be useful; I quote Weller et al. (2004):

> The encapsulation and de-encapsulation of processes in order to perform actions is a common experience in mathematical thinking. For example, one might wish to add two functions f and g to obtain a new function $f+g$. Thinking about doing this requires that the two original functions and the resulting function are conceived as objects. The transformation is imagined by de-encapsulating back to the two underlying processes and coordinating them by thinking about all of the elements x of the domain and all of the individual transformations $f(x)$ and $g(x)$ at one time so as to obtain, by adding, the new process, which consists of transforming each x to $f(x)+g(x)$. This new process is then encapsulated to obtain the new function $f+g$.

Mathematical concepts are shaped and developed in a child's mind in a recurrent process of encapsulation and de-encapsulation, assembly and disassembly of mathematical concepts. It helps if building blocks are simple and easy to handle.

My computer science colleague commented on the quote above that the importance of encapsulation goes beyond mathematics education: it is an important concept in practical computer programming, where it also helps if building blocks are simple.[3]

Anna Sfard (1991, 2016) uses a similar but subtly different concept of *reification*. I write this paper with two hats on me, of a mathematics teacher and a mathematics researcher. For me as a teacher, encapsulation is the use and re-use of a ready-to-use capsule received from a teacher or learned from a book. For me as a researcher, reification is crystallisation, in the mind of a particular problem-solver, and subsequent explicit formulation, of a *previously unknown and non-existing* mathematical concept, object, etc.[4] At the level of school mathematics, reification, as I understand it, happens only in solving nonstandard, olympiad class problems – and almost never in a mathematics class in a mainstream school. Reification suggests high level of autonomy, it is used in an open-ended high risk work-flow. In this paper, I stick to the terms encapsulation / de-encapsulation.

3 De-encapsulation in action: "questions" method

Here is an example of encapsulation and de-encapsulation in action[5].

In 2011 I was asked by my American colleagues to give my assessment of mathematical material on the Khan Academy website[6]. Among other things I looked for the so-called "word problems" and clicked on a link leading to what was called there an "average word problem" but happened to be a "word problem about averages".

> *Gulnar has an average score of* 87 *after* 6 *tests. What does Gulnar need to get on the next test to finish with an average of* 78 *on all* 7 *tests?*

Solution I. What follows are hints provided, one after another, by the Khan Academy website:

Hint 1 Since the average score of the first 6 tests is 87, the sum of the scores of the first 6 tests is $6 \times 87 = 522$.

Hint 2 If Gulnar gets a score of x on the 7th test, then the average score on all 7 tests will be:

---

[3] Chris Stephenson, Private communication.

[4] The term "reification" is even used as a description of a specific computational procedure in my hard core research paper Borovik and Yalçınkaya (2015).

[5] I re-use some material from my paper (actually, a blog post in the pdf format) Borovik (2013a).

[6] Khan Academy. `http://www.khanacademy.org/about`. Last Accessed 14 Apr 2011.

$$\frac{522+x}{7}.$$

Hint 3 This average needs to be equal to 78 so:

$$\frac{522+x}{7}=78.$$

Hint 4 $x=24$.

Solution II. And here is how the same problem would be solved by the "steps" or "questions" method as it was taught in Russian schools in 1950–60s.

Question 1 How many points in total did Gulnar get in 6 tests?
Answer: $6\times 87=522$.
Question 2 How many points in total does Gulnar need to get in 7 tests?
Answer: $7\times 78=546$.
Question 3 How many points does Gulnar need to get in the 7th test?
Answer: $546-522=24$.

Questions 1 and 2 *represent the de-encapsulation of the concept of average.* And this disassembly, de-encapsulation, makes the solution very simple.

Solution III. There is a quicker solution[7] which requires understanding of averages beyond straightforward de-encapsulation:

Question 1 How many "extra" – that is, above the requirement – points did Gulnar get, on average, in 6 tests?
Answer: $87-78=9$.
Question 2 How many "extra" points does Gulnar have?
Answer: $9\times 6=54$.
Question 3 How many points does Gulnar need to get in the last test?
Answer: $78-54=24$.

Finding this solution is next to impossible without mastering some higher level thinking – I will return to this issue in the next Section.

4 Self-directing questions

Crucially, the whole point of the "questions" method is that questions are not supposed to be asked by a teacher: students are taught to formulate these questions *themselves*.

Teaching the "questions method" focuses on the development of each student's ability to start his/her "questions" attempt at a word problem asking *himself or herself* appropriate *self-directing questions* (they are called *auxiliary questions* in the

[7] Suggested by John Baldwin.

Russian pedagogical literature, but in England, the words "an auxiliary question" are loaded with expectation that the question is asked by a teacher to help a struggling pupil).

In the case of Gulnar's problem, these self-directing questions are likely to be something like

Solution II, Question A "Gulnar has an average score of 87 after 6 tests." *What questions can be asked about these data?*

Solution II, Question B "Gulnar needs to get an average of 78 on all 7 tests." *What questions can be asked about these data?*

Solution II, Question C "Gulnar has 522 points. She needs 546 points." *What questions can be asked about these data?*

Therefore the use of the "questions" method in mathematics education involves gently nudging a child towards *reflection* and analysis of his/her own thought process. This should be done, it needs to be emphasised, at a level actually accessible to the child—and this can be done, as it was confirmed by mathematics education practice of dozens of countries around the world. I prefer the term "questions method" to the more commonly used, in British education literature, name "steps method" because the word "'questions" emphasises the pro-active and reflective components of thinking, while the word "steps" might inadvertently imply a passive procedural approach.

And what is even more important, self-directing questions are *meta-questions*, that is, questions aimed at finding the optimal way of reasoning.

From a basic pedagogical point of view, if the didactic aim of the problem is to reinforce the understanding of particular concept (say, averages – as in Gulnar's problem) then the "questions" method appears to be more useful; it gives a student a joint and cohesive vision of the concept.

For a teacher, self-directing questions give a useful tool for assessment of didactic aspects of a problem and its potential solutions. Let us look at possible self-directing questions for Solution III:

Solution III, Question A "Gulnar has an average score of 87 after 6 tests. She needs 78 points on average" *What questions can be asked about these data?*

We immediately see that, unlike in Solution II, a child has to handle two chunks of information simultaneously, not one. Even more: some of this information – namely, the number of previously taken tests – is unnecessary for the first step:

Solution III, Question A* "In previous tests, Gulnar had an average score of 87. She needs 78 points on average." *What questions can be asked about these data?*

But initially discarded information re-appears at the second step:

Solution III, Question B "In each of the 6 tests, Gulnar got, on average, 9 "extra" – that is, above the requirement – points." *What questions can be asked about these data?*

Solution III, Question C "Gulnar has 54 extra points. She needs an average of 78 and has one more test to take." *What questions can be asked about these data?*

I think a teacher may see that this approach requires from a student confident handling of "*structural arithmetic*", in terminology of Tony Gardiner (Gardiner, 2014, Section 2.1.1.2). Indeed, a mental shortcut of Question A is meaningless if a student cannot see an arithmetic equivalent of an algebraic identity

$$\frac{a_1 + a_2 + \cdots + a_n}{n} - b = \frac{(a_1 - b) + (a_2 - b) + \cdots + (a_n - b)}{n}$$

hidden deep in the problem.

Tony Gardiner defines structural arithmetic as

> *an awareness of the algebraic structure lurking just beneath the surface of so many numerical or symbolical expressions, as in*
>
> $$3 \times 17 + 7 \times 17 = \ldots$$
>
> *or [. . .]*
>
> $$16 \times 13 - 3 \times 34 = \ldots$$

He adds:

> *This habit of looking for, and then exploiting,* algebraic *structure in* numerical *work is what we call* structural arithmetic.

And I hope that it is obvious to the reader that a self-directing question is an application of Arnold's Principle, a pro-active response to

> *the need to make a concrete and sensible interpretation of all the values which are used and/or which appear in the discourse*

as formulated by Igor Arnold.

Julia Brodsky, one of the leaders of American mathematics homeschooling and mathematical circles movement (see her book Brodsky (2015)), wrote to me:

> *Self-directed questions is probably the most important skill the students need to learn – not only in math, but for their future life as well (and as a basis for critical thinking). In my experience, the best way to teach that is to model the self-questioning in front of the students by the teachers, as well as playing the "questions game" for novice students, where the students first ask all types of questions about the problem, and then select the most useful ones. This is a skill that takes time and nurturing, and should be taught to the teachers as one of the basic tools.*

In my opinion, as far as didactics of mathematics is concerned, mathematics homeschoolers and mathematics circles volunteers are ahead of the game in comparison with the mainstream mathematics educators – this is a very symptomatic development which I discuss elsewhere in Borovik (2016a).

5 Distributed quantities

It is time to take a closer look both at the differences and at the deep connections between arithmetic and algebra (and other chapters of more advanced mathematics) as emphasised by Igor Arnold:

> *The difference between the "arithmetic" approach to solving problems and the algebraic one is, primarily the need to make a concrete and sensible interpretation of all the values, relations and operations which are used and/or which appear at any stage of the discourse.*

Obviously, not every algebraic problem can be solved by arithmetic means. Still, the power of arithmetic should not be underestimated.

My favorite example is Markov's Inequality:

> *If X is any nonnegative random variable and $a > 0$, then*
> $$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

It is the first fundamental result of the theory of random variables – and a basis of entire statistics.

In its essence, Markov's Inequality is a primary school level observation about inequalities and can be formulated as an arithmetic "word problem" about anglers and fish.

What I formulate now is a result of a very straightforward didactic transformation of Markov's Inequality: de-encapsulation

- of mathematical expectation (or average – recall Gulnar's problem), and
- of probability in its frequentist interpretation,

followed by substitution of concrete values:

> *If* 10 *anglers caught on average* 4 *fish each, then the number of anglers who caught* 5 *or more fish each is at most* 8*.*

A proof of this statement is simple. Together, anglers caught $10 \times 4 = 40$ fish. Assume that there were *more than* 8 anglers who caught *at least* 5 fish each; then these 8 anglers caught together *more than* $8 \times 5 = 40$ fish – a contradiction.

Unfortunately, we cannot expect that all students entering English universities are able to produce this argument. There are three reasons for that:

- This is a proof.
- Even worse, this is a proof from contradiction.
- The argument requires simultaneous handling of two types of inequalities, "*x is more than y*", denoted $x > y$, and "*x is at least y*", denoted $x \geq y$. (I say more on difficulties caused by relations "x is at least y" and "x is at most y" in Section 8.)

But the statement still belongs to the realm of arithmetic and we can continue its *didactic transformation* (Borovik (2016b)) by replacing "proof" by "solving" and converting the statement into a proper "word problem".

> *If* 10 *anglers caught on average* 4 *fish each, what is the maximal possible number of anglers who caught* 5 *or more fish each?*

And here is a solution.

> Anglers caught $10 \times 4 = 40$ fish. So we have to distribute 40 fish between anglers in a way ensuring that as many anglers as possible get 5 or more fish. To achieve that, we should not give more than 5 fish to an angler – that way more fish are left to other anglers, and more of them get their 5 fish. Hence we give 5 fish to an angler. How many of them will get their share? $40 \div 5 = 8$ anglers.

Fish caught by anglers is a classical example of a *random variable*. In the context of arithmetic, I would prefer to use the words "distributed quantity": it is a quantity attributed to, or distributed among, objects in some class: fish to anglers, test marks to students in the class, and, the last but not least, pigeons to pigeonholes, as in the "Pigeonhole Principle".[8] Crucially, we are not interested in its specific values, but only in how frequently particular values appear and how frequently they exceed (or not) particular thresholds. Notice that, in the solution above, we manipulate fish as a distributed quantity, limiting its dispensation to five fish per angler.

In short, what we have is a toy frequentist version of Kolmogorov's definition of a random variable as a measurable function Kolmogoroff (1933). As simple as that.

6 The Arithmetic/Algebra boundary

Everything should be made as simple
as possible, but not simpler.
Apocryphal, attributed to A. Einstein

In previous sections we have explored implications of Arnold's Principle, now we turn our attention to is limitations.

I will be using a beautiful example promoted by Vladimir Arnold.

Vladimir Arnold once said in an interview to Lui (1997):

> *The first real mathematical experience I had was when our schoolteacher I.V. Morotzkin gave us the following problem.*
>
> > Two old women set out at sunrise and each walked with a constant speed. One went from A to B, and the other went from B to A. They met at noon, and continuing without a stop, they arrived respectively at B at 4pm and at A at 9pm. At what time was sunrise on that day?

---

[8] Indeed, I believe that the famous Pigeonhole Principle (it is traditionally formulated in one of the simplest special cases, rather than in a "general" form):

"*if you put* 6 *pigeons in 5 holes than at least one hole contains more than one pigeon*"

should be part of the standard arithmetic curriculum. In the world of adult science, it is one of the most basic concepts of Computer Science and Programming; mathematically, it belongs to Ergodic Theory.

> *I spent a whole day thinking on this oldie, and the solution* (*based on what are now called scaling arguments, dimensional analysis, or toric variety theory, depending on your taste*) *came as a revelation.*
>
> *The feeling of discovery that I had then* (*1949*) *was exactly the same as in all the subsequent much more serious problems – be it the discovery of the relation between algebraic geometry of real plane curves and four-dimensional topology* (*1970*)*, or between singularities of caustics and of wave fronts and simple Lie algebra and Coxeter groups* (*1972*)*. It is the greed to experience such a wonderful feeling more and more times that was, and still is, my main motivation in mathematics.*

This is a very strong statement, and deserves some analysis.

A classical solution makes use of a chapter of arithmetic almost completely forgotten nowadays: theory of proportions. This solution is given below, and it demonstrates a boundary between Arithmetic and Algebra: we see "intermediate values", in terminology of Igor Arnold, which have no obvious real world interpretation.

Assume that the two old women walked from A to B and from B to A, respectively, and that they met at point M. Then the first lady covered the distance from A to M in from sunrise to noon and then distance from M to B in 4 hours. Since she walked at constant speed,

$$\frac{\text{distance from } A \text{ to } M}{\text{distance from } M \text{ to } B} = \frac{\text{time from sunrise to noon}}{4 \text{ hours}}.$$

Similarly, for the second woman

$$\frac{\text{distance from } M \text{ to } A}{\text{distance from } B \text{ to } M} = \frac{9 \text{ hours}}{\text{time from sunrise to noon}}.$$

Since it does not matter in which direction we measure distance, from A to M or from M to A, etc.,

$$\frac{\text{distance from } A \text{ to } M}{\text{distance from } M \text{ to } B} = \frac{\text{distance from } M \text{ to } A}{\text{distance from } B \text{ to } M}$$

and consequently we get a proportion

$$\frac{\text{time from sunrise to noon}}{4 \text{ hours}} = \frac{9 \text{ hours}}{\text{time from sunrise to noon}}.$$

Solving it, we have

$$\text{time from sunrise to noon } = \sqrt{4 \times 9} = \sqrt{36} = 6 \text{ hours}.$$

Therefore the sunrise was 6 hours before the noon, that it, at $12 - 6 = 6$am hours.

What is remarkable, if we trace the world lines of the two ladies on the time-distance plane, we immediately discover that the proportions have immediate geometric meaning and are related to similarity of triangles, Figure 1.

As we can see from this solution in its two shapes, arithmetic and geometric, we have an uncomfortable operation of multiplying time by time, and, even worse, extracting square root from the result.

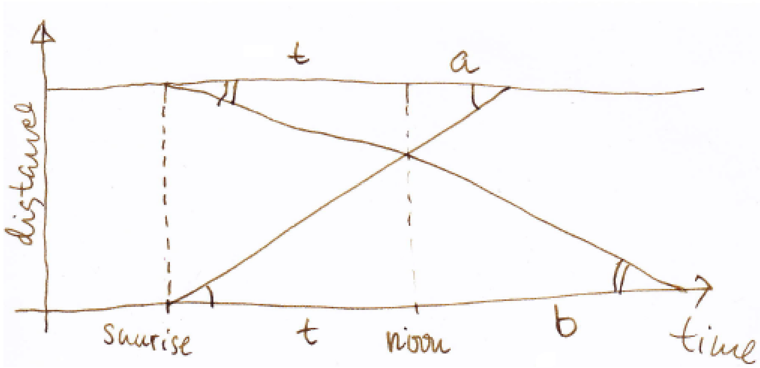

Fig. 1 We have from similarity of triangles $\frac{a}{t} = \frac{t}{b}$ and $t = \sqrt{ab}$.

Even worse, the diagram uses angles. I marked equal angles at the diagram due to a kind of a Pavlovian dog reflex, because I was conditioned to behave that way at school and retained the reflex for the rest of my life. But angles have no meaning on the time-distance plane unless we are on the Minkowski plane of special relativity theory with a fixed quadratic form relating time to space. So, if we use angles in the solution, we are in a modernised version of the problem:

> Two old women flew, on photon spaceships at speeds close to the speed of light, one from galaxy A to galaxy B, and the other from B to A. They set out at … *whoops!* What does it mean "they set out at the same time" if we are in the relativistic context? There is no absolute time in the world of special relativity.

Luckily, angles can be removed from the geometric solution: instead of similarity of triangles, we can stay within affine geometry and use, in the proof of the proportion

$$\frac{a}{t} = \frac{t}{b},$$

properties of central projection from a line to a parallel line.

Still, this example shows that an attempt to look for an "immediate real world interpretation" of intermediate values in a solution of a relatively elementary problem can open Pandora's box of difficult questions about relations between mathematics, mathematical models of reality, and reality itself.

7 Uniform convergence and "likeness"

One day I will find the right words,
and they will be simple.
Jack Kerouac

Now I wish to discuss a beautiful application of the both Arnold's Principle and abstraction by irrelevance in "advanced" mahtematics

The concept of uniform continuity of a function, after a long and torturous development (see Sinkevich (2016)) was transformed, in André Weil's paper Weil (1937) into a strikingly abstract definition of *uniform structure* which uses only basic concepts of elementary set theory: sets and binary relations. Uniform structures had been immediately adopted by Bourbaki; the concept became one of the crown jewels of his *Éléments de mathématique*.

A definition of a *uniform structure* (and its developments, uniform space and uniformly continuous function) is remarkably simple and uses only intuitive elementary set theory; it is a classical example of abstraction by irrelevance: all the details and features of uniform continuity are stripped to the bare logical skeleton.

We start by defining a *tolerance* T on a set X as a *reflexive* (that is, for all $x \in X$, xTx holds) and *symmetric* (that is, for all $x, y \in X$, xTy implies yTx) binary relation on X. Tolerance is a mathematical formalisation of similarity or resemblance relations between objects of the real world Schrader (1971).

A *uniformity basis* on X is a non-empty family $\mathscr{T}$ of tolerances on X which is

- closed under taking intersections (or conjunctions, which is an equivalent way of saying): if $T, S \in \mathscr{T}$ then $T \wedge S \in \mathscr{T}$, and
- allows decomposition: for a tolerance $T \in \mathscr{T}$ there exists a tolerance $S \in \mathscr{T}$ such that $S \circ S \subseteq T$.

Notice that the *inclusion* relation $\subseteq$ and operations of *intersection* "$\cap$" (which is the same as conjunction "$\wedge$") and *composition* "$\circ$" of binary relations have very intuitive meaning.

For example, the relation on the set of people

$$xSy \Leftrightarrow s \text{ is a sibling of } y$$

includes the relation

$$xBy \Leftrightarrow s \text{ is a brother of } y$$

and therefore $B \subseteq S$.

Conjunction/intersection is also easy: let

$$xTy \Leftrightarrow \text{ Tom thinks that } x \text{ and } y \text{ are alike}$$

and

$$xSy \Leftrightarrow \text{ Sarah thinks that } x \text{ and } y \text{ are alike}$$

then

$$x(T \wedge S)y \Leftrightarrow \text{ Tom and Sarah both think that } x \text{ and } y \text{ are alike.}$$

And here is an example of composition: if xCy means that a person x is a child of a person y and xGz means that x is a grandchild of z then $G = C \circ C$. (One more example of composition is given a few lines below.)

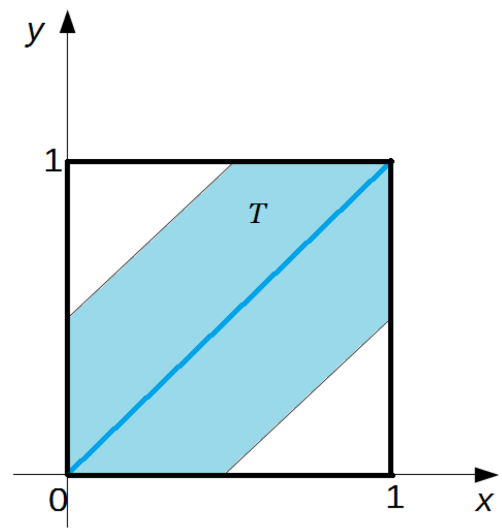

Fig. 2 Tolerance relation $x\,T\,y \Leftrightarrow |x-y| < \frac{1}{2}$ on the segment $X = [0,1]$. It is reflexive because it contains the diagonal of the square $X \times X$ and symmetric because it is symmetric with respect to the diagonal.

A canonical example of a uniformity basis is the one responsible for the uniform continuity of real valued functions on the real segment $[0,1]$ (Figure 2):

$$\begin{aligned} \mathscr{T} &= \{\, T_n \; n = 1,2,3,\ldots \},\ \text{where} \\ T_n &= \left\{ (x,y) \in [0,1] \times [0,1] : |x-y| < \frac{1}{2^n} \right\} \end{aligned}$$

or, if you prefer predicates to sets,

$$xT_ny \Leftrightarrow |x-y| < \frac{1}{2^n}.$$

The operation of composition is especially clear in that example: indeed,

$$T_{n+1} \circ T_{n+1} \subseteq T_n$$

because if

$$|x-y| < \frac{1}{2^{n+1}} \text{ and } |y-z| < \frac{1}{2^{n+1}}$$

then

$$\begin{aligned} |x-z| &= |(x-y)+(y-z)| \\ &\leq |x-y|+|y-z| \\ &= \frac{1}{2^{n+1}}+\frac{1}{2^{n+1}} \\ &= \frac{1}{2^n}. \end{aligned}$$

(Actually, a *uniform structure* on X generated by a uniformity basis $\mathscr{T}$ is the *filter* $\mathscr{F}$ on $X \times X$ generated by $\mathscr{T}$, that is, the set of all binary relations on X which include some tolerance from $\mathscr{T}$.)

"Similarity", "resemblance", "likeness" – all that stuff formalised in the mathematical concept of tolerance are real life concepts, sophisticated but intuitively understood by young children. It looks as if kids can be really excited by real life problems about choosing, identifying and sorting built around "resemblance" and "likeness"; such problems make an excellent propaedeutic for more abstract mathematics.

I will show to the reader examples taken from a little gem of early mathematics education, the book *Socks are like Pants, Cats are like Dogs*[9]. As the title suggests, most problems are about resemblance. Some of them require construction of a tolerance relation (but the term is of course not used in the book) on a finite set, see Figure 3.

Problems on sorting, if described in technical terms, are about constructing equivalence classes containing given objects when the equivalence is given as an intersection of several tolerances (perhaps with subsequent taking the transitive closure).

The book says:

> *Too often the sorting jobs we give our children are not very challenging. Their young brains are capable of differentiating complex patterns like those of identifying beetle families. Let them flex these sorting muscles!*

And children are asked to sort beetles, see examples on the page preceding this chapter with the first and the last sorting problem. Children are given instruction "Follow the directions on the left side and collect only the beetles that are indicated."

I would not mention these problems in my paper if I had not had a chance to watch how 7 and 8 years old boys were sorting beetles with unbelievable enthusiasm; the youngest was even more impressive, his attention was totally focused on minute details of antennae, mandibles, legs, hairs, segmentation of bodies. Beetle Sort works! Pedagogical advice given in the book is realistic and sound:

> *Encourage children to discuss why they think a beetle should be collected. Ask children to explain their reasoning. Accept all answers with explanations as possibilities. Mistakes should be expected. When working on the book, one of the authors (Dr. Gordon Hamilton) solved two of the puzzles wrong, at least according to the current scientific classification*

---

[9] M. Rosenfeld and G. Hamilton, Socks Are Like Pants, Cats Are Like Dogs: Games, Puzzles & Activities for Choosing, Identifying & Sorting Math! Natural Math, vol. 4. Delta Stream Media, 2016. Print ISBN: 978-0-9776939-0-0.

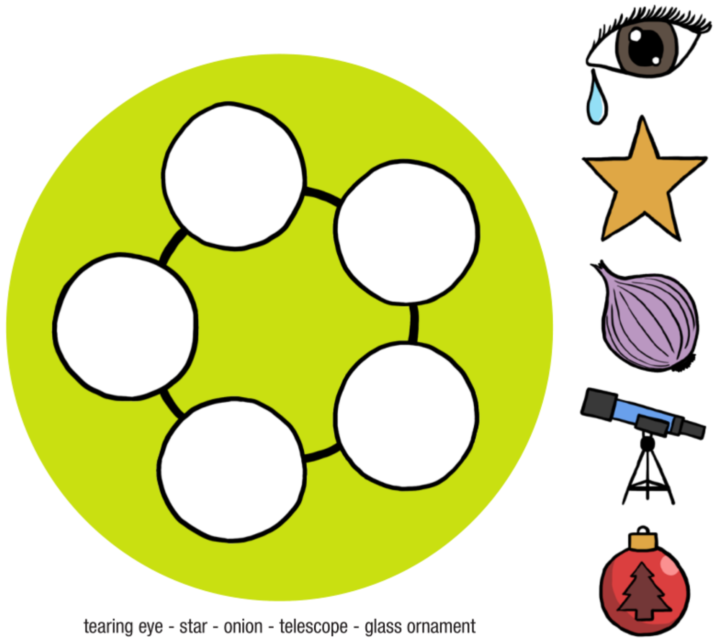

Fig. 3 "This game creates a chain of association between seemingly unrelated objects. Look at each object in the puzzle and place them in the circles so that objects in connected circles share a common trait."

of beetles in the answer keys. Free play on their own terms helps children feel good about math. Toward that goal, children can arrange beautiful beetles in their own ways. On the other hand, tenacity in the face of failure also protects against math anxiety. To build up tenacity, help kids to figure out how the scientific classification works.

I really love the last piece of advice:

If your child is getting frustrated, blame the beetles! It's their fault the puzzle is so difficult!

8 Order and equivalence – and abstraction by irrelevance

There is a class of binary relations which is simpler than tolerance and is more intuitive than any other kind of binary relations: *strict order* $<$. It is characterised by axioms of *transitivity*:

it $x < y$ *and* $y < z$ *than* $x < z$

and *anti-symmetricity*:

$x < y$ and $y < x$ cannot be true simultaneously.

Notice that the anti-symmetricity implies the *anti-reflexivity*:

$x < x$ is never true.

A classical "real life" example of a strict order is the relation on the set of people

x is a descendant of y.

A strict order relation is *linear* if

for every distinct x and y either $x < y$ or $y < x$.

The descendence relation is not linear. But the *counting order* (well-known to most children of age 4), that is, the strict ordering of natural numbers

$$1 < 2 < 3 < 4 < \cdots$$

is the mother of all strict orders. The counting order is easy for children; but the (non-strict) order $x \leq y$, that is, "x is less or equal y" might cause serious difficulties, and not only at the primary school level. Every year, I meet a freshman in my class (at a university!) who asks something like "How can we claim that $2 \leq 3$ if we already know that $2 < 3$?".

Various kinship relations are remarkably self-evident to children, and for very deep cognitive and evolutionary reasons – already apes and even monkeys have sophisticated kinship systems.

The remarkable book *Baboon Metaphysics* Cheney and Seyfarth (2008) provides some astonishing evidence – and please notice that the book contains a formal definition of transitivity of a relation:

> The number of adult males in a baboon group at any given time ranges widely, from as few as 3 to as many as 12. Regardless of their number, however, the males invariably form a linear, transitive dominance hierarchy based on the outcome of aggressive interaction (a linear, transitive hierarchy is one in which individuals A, B, C and D can be arranged in linear order with no reversal that violate the the rule 'if A dominates B and B dominates C, then A dominates C'). Although the male dominance hierarchy is linear, transitive, and unambiguous over short periods of time, rank changes occur often (Kitchen et al. 2003b), and a male's tenure in the alpha position seldom lasts for more than a year. [p. 51]

Human boys in less humane places such as various kinds of borstals, reformatories, and juvenile prisons form a similar strict linear order hierarchy recalculated every day as a result of fights.

The hierarchy of female Baboons is more sophisticated and interesting. As emphasised in the book, it is "Jane Austen's world".

> Like males, female baboons form linear, transitive dominance hierarchies. There, however, the similarity ends. Whereas male dominance ranks are acquired through aggressive challenges and change often, female ranks are inherited from their mothers and remain stable for years at time. Furthermore, most female dominance interactions are very subtle. Although threats and fights do occur, they are far less common and violent than fights among males. Instead, most female dominance interactions take the form of supplants: one female simply

> approaches another and the latter cedes her sitting position, grooming partner, or food. The direction of supplants and aggression—and the resulting female dominance hierarchy—is highly predictable and invariant. The alpha female supplants all others, the second-ranking supplants all but the alpha, and so on down the line to the 24th- or 25th-ranking female, who supplants no one. [p. 65]

Therefore it is not surprising that the concept of linear strict order is so self-evident to humans. But anyone who taught freshmen knows that the concept of *equivalence relation* is incomparably harder. The reason is that the transitivity of dominance is obvious at the level of the monkey bits of our brains. But an equivalence relation is, by definition, a transitive tolerance relation. Therefore

- an equivalence relation is a transitive "likeness";
- a strict order is a transitive "unlikeness".

This makes all the difference. If we understand "equality" in its common sense, as in "all people are equal", not in the sense of "identity" or "sameness", then it appears that the transitivity of equality is a much later, in evolutionary and historic terms, social construct than the transitivity of "dominance" or "superiority".

In a powerful scene in the film *Lincoln*[10], Abraham Lincoln says to his astonished aids:

> Euclid's first common notion is this: Things which are equal to the same thing are equal to each other. That's a rule of mathematical reasoning. It's true because it works. Has done and always will do. In his book, Euclid says this is 'self-evident.' You see, there it is, even in that 2,000-year-old book of mechanical law. It is a self-evident truth that things which are equal to the same thing are equal to each other.

The scene is a fiction, but a brilliant and very convincing fiction expressed in simplest possible terms accessible to all cinema-goers. It is very true in its spirit to a number of well-documented quotes from Lincoln where he uses references to Euclid as a logical and rhetoric device:

> One would start with confidence that he could convince any sane child that the simpler propositions of Euclid are true; but, nevertheless, he would fail, utterly, with one who should deny the definitions and axioms. The principles of Jefferson are the definitions and axioms of free society. And yet they are denied, and evaded, with no small show of success. One dashingly calls them 'glittering generalities'; another bluntly calls them 'self-evident lies'; and still others insidiously argue that they apply only 'to superior races.'[11]

I write this paper from the position of a remedial teacher at the school/university interface, this is why I am keen to have a *holistic* view of mathematics education at all levels, especially interconnections between various parts of mathematics as they are presented to students starting from pre-school.

Unfortunately too many students reach mathematics courses at the university level with ability for abstract thinking suppressed; even after three years in the university, some of them still cannot make usable mental picture of abstract equivalence

[10] *Lincoln*, `http://www.thelincolnmovie.com/`. Director: Steven Spielberg; in the title role: Daniel Day-Lewis; screenplay: Tony Kushner.

[11] A. Lincoln, *Collected Works*, 3:375, quoted in Morrissey (2012), who in his turn quoted (Havers, 2009, p. 72).

relations – you may find more on that in Borovik (2013b). I wholeheartedly agree with one of the commentators on an earlier version of my paper, Wes Raikowski, who wrote to me

> "the series of abstractions and generalisations must, in my view, be rooted in one's own sensory experiences of bodily interactions with the physical world."

Indeed, abstraction is negation, in Hegelian terms; it can start only when concrete real mathematics (of Igor Arnold's $3-1=2$ kind) is sufficiently interiorised by a child in all its richness.

This explains why

> *efficient abstraction by irrelevance is Arnold's Principle in its dialectically negated form:*

in his/her first encounters with abstraction, a student has to have a clear understanding of what he or she discards, treats as irrelevant, abstracts away from.

When writing the paper, I was thinking about, and would love to be useful to, a homeschooling parent or a leader of a mathematical circle, someone who was engaged in a direct Socratic dialogue with children. Abstraction by irrelevance can start by a casual remark from a teacher: "ah, it does not matter", especially if this remark is prepared in advance and strategically positioned. For example, write on a board a problem

> Mary has some cats and some chicken, and together her pets have 5 heads and 14 legs. How many cats does Mary have?

and in the process of collective solving the problem start talking about dogs instead of cats, triggering, with some luck, kids' protests, and then lead children to recognising that, in this problem, dogs and cats are interchangeable because they have the same number of legs / paws.

One of the first examples of abstraction accessible to very young children is the use of numbers as *classifiers* – dogs, cats, rabbits are quadrupeds, they have four legs /paws. And what about kangaroo?

In Beetle Sort problems of Section 7, the number of legs is constant (six) but the number of body segments varies from one family to another, and acts as a classifier (not always sufficient – two different families of beetles might have the same number of body segments – but still useful).

My university colleagues widely accept that the fundamental theorem:

> equivalence classes of an equivalence relation E on a set X form a partition of X

is the *Pons Asinorum* of elementary set theory. In my classes, I do some propaedeutics by preceding the introduction of this theorem by explaining, to my students, that an equivalence relation E on a set X can in many cases be usefully thought about in terms of a *classifying function* $f : X \longrightarrow A$ to some simpler set A, with the characteristic property that

$$xEy \Leftrightarrow f(x) = f(y).$$

In practical classification problems, it frequently happens that one number valued function does not suffice, but even one function can make a decent approximation, like a number of petals in a flower in Linnaeus' celebrated classification of plant species.

To summarise Sections 7 and 8: they provide an example of an advanced concept of mathematics – uniformity – reducible, within bounds of Arnold's Principle, to much simpler and much more intuitive concept of elementary set theory – tolerance relations, while the latter are further reducible, again within bounds of Arnold's principle, to simple intuitive real life concepts of likeness and resemblance. But we have also had a chance to see that two concepts closely related to tolerance relations: equivalence and order – behave very differently when we try to find for them simple, convenient, and intuitive "real life" interpretations.

9 Arnold's Principle and the "questions method" in the historic and social context

Igor Arnold's paper of 1946 reflected *Zeitgeist* of Russian culture in the aftermath of WWII: a quest for simple, reproducible, *scalable* solutions to technological – and educational – problems.

Scalability (that is, feasibility of a wide, unlimited dissemination and implementation) is very difficult to achieve without simplicity, and a few words about scalability need to be said.

As a child, I learnt the "questions" method in my primary school in a direct face-to-face communication with a live teacher and with my peers, not from a video recording on the Internet – as Khan Academy's students learn mathematics – and I describe it here as it was widely and routinely used in all primary schools in Russia in the 1960s. A colleague, responding to an earlier version of my notes on the "questions method", indicated that I was lucky to have an "excellent mentor" who was using "the richness of the Socratic questioning". I loved my teacher – but it needs to be explained that she was a village school teacher in Siberia and was educated (up to the age of 16) in the same village school and then for two years (up to the age of 18) in a pedagogical college in the town of Kyakhta – look it up on the GOOGLE map! Even now its location can be best described as being in the middle of nowhere – imagine what it was half a century ago!

If "policymakers" will ever read my paper, this is my message to them:

> *My teacher's skills in arithmetic were a guaranteed and enforced* minimum *compulsory for every teacher.*

Arnold's Principle was just one example of didactics generated by an approach to education based on scalable solutions at every level: in general education policy, in curriculum development, in methodology of mathematics education, in direct recommendation to teachers on classroom practice.

But it should not be lost from view that the mathematics education policy of Russia at that time was concerned not only with achieving a "guaranteed minimum" outcome, but also with educating an engineering and scientific elite.

Only recently I learned how my *alma mater*, FMSh (the Physics and Mathematics Preparatory Boarding School of Novosibirsk University; I describe it in Borovik (2012). [12]) , was born. It was one of four specialist mathematics boarding schools (the famous Kolmogorov School in Moscow was one of them) opened in 1963. What was not widely known for decades that the decree of the Council of Ministers of the USSR was signed by an immensely powerful man, Dmitry Ustinov, at that time the First Deputy Prime Minister. Ustinov cared about mathematics – including elite, research level mathematics – this was part of *Zeitgeist*. Actually, the FMSh came into existence before the decree was formally signed – and no-one knew where the funds for its upkeep were coming from. And the last but not least: the school was temporarily housed in a building built for a new military academy for training Red Army's political officers – and the opening of the academy was for that reason postponed. This is what I call policy priorities.

10 The economy of thought

A child of five could understand this.
Send someone to fetch a child of five.
Groucho Marx

Mentioning the FMSh, an academically selective establishment (to the extremes of selectivity – the school had the catchment area with population of 40 million people) allows me to move to discussion of a characteristic trait of many of so-called "mathematically able" children[13]: "economy of thought", a (mostly subconscious) inclination to seek clarity and simplicity in a solution.

In relation to arithmetic, Arnold's Principle shows that the "economy of thought" means, first of all, ability to see relations, structures, symmetries of the "real world" and use them to simplify arithmetic reasoning.

Vadim Krutetskii's classical study of psychology of mathematical abilities in children Krutetskii (1976) was a serious work based on hundreds of interviews, numerous tests and statistical analysis (Kruteskii even received advice on the use of

[12] It is instructive to compare my paper with an insider's description of Lycée Louis-le-Grand in Paris, Lemme (2012).

[13] All children have mathematical abilities but not all of them are given a chance to develop them in full.

statistics from Andrei Kolmogorov). The tendency for "economy of thought" is emphasised as one of the most important traits of the so-called "mathematically able" children. This is what he writes about 8 years old Sonya L.

> *Sonya is notable for a striving to find the most economical ways to solve problems, a striving for clarity and simplicity in a solution. Although she does not always succeed in finding the most rational solution to a problem, she usually selects the way that leads to the goal most quickly and easily. Therefore many of her solutions are "elegant." What has been said does not apply to calculations (as was stated above, Sonya is unfamiliar with calculation techniques). Consider a few examples.*
>
> *Problem:*
>
> > *"How much does a fish weigh if its tail weighs 4 kg, its head weighs as much as its tail and half its body, and its body weighs as much as its head and tail together?"*
>
> *Solution:*
>
> > *"Its body is equal in weight to its head and tail. But its head is equal in weight to its tail and half its body, and the tail weighs* 4 *kg. Then the body weighs as much as* 2 *tails and half the body – that is,* 8 *kg and half the body. Then* 8 *kg is another half of the body, and the whole body is* 16 *kg."*
>
> *(We omit the subsequent course of the solution. The problem is actually already solved.)*

This remark:

> *Sonya is unfamiliar with calculation techniques*

is very interesting: Sonya goes directly to what Igor Arnold describes as

> *concrete and sensible interpretation of all the values which are used and/or which appear in the discourse*.

Sonya identifies mathematical structures and relations of the real world and maps them onto better formalised structures and relations of arithmetic, as it is obvious in another episode from Krutetskii's book:

> Problem. *A father and his son are workers, and they walk from home to the plant. The father covers the distance in* 40 *minutes, the son in* 30 *minutes. In how many minutes will the son overtake the father if the latter leaves home* 5 *minutes earlier than the son?*
>
> Usual method of solution [by 12-13 year old children]: *In 1 minute the father covers* $1/40$ *of the way, the son* $1/30$*. The difference in their speed is* $1/120$*. In 5 minutes the father covers* $1/8$ *of the distance. The son will overtake him in*
>
> $$\frac{1}{8} : \frac{1}{120} = 15 \text{ minutes.}$$
>
> Sonya's solution: *"The father left* 5 *minutes earlier than the son; therefore he will arrive* 5 *minutes later. Then the son will overtake him at exactly halfway, that is, in* 15 *minutes."*

Sonya sees symmetries of the world – and not only in space, but in time, too (the latter is more impressive) – and links the symmetry of time with symmetry of space. For an adult, this relation is best expressed by the world lines of the father and the son in the time-distance plane, Figure 4. We will never know what kind of picture

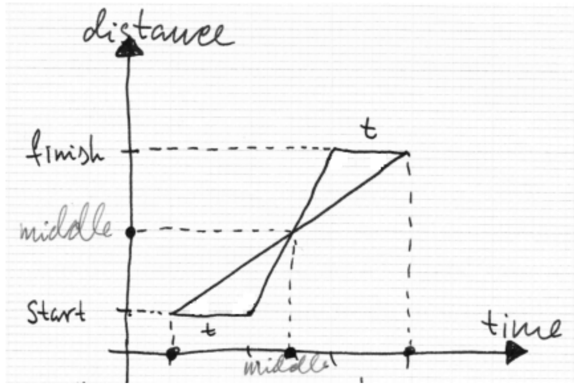

Fig. 4 This is what happens when two bodies move along the same distance from start to finish with constant speeds, but one of them starts t minutes earlier and finishes t minutes later than another: the faster body overtakes the slower one at mid-time and mid-distance. I leave it as an exercise to the reader to check that this follows from a well-known theorem of affine geometry: the two diagonals and the two midlines of a parallelogram intersect at the same point.

(if any) Sonya had in her mind, but she had a feeling of some essential properties of this relation.

In Krutetsii's words:

> *To a certain extent she is characterized by a distinct inclination to find a logical and mathematical meaning in many phenomena of life, to be aware of them within the framework of logical and mathematical categories. In other words, her tendency to perceive many phenomena through a prism of logical and mathematical relationships was marked at an early age* (*7 or 8*).

I wish to emphasise these words:

> *find mathematical meaning in phenomena of life.*

Mathematics is simpler than life and for that reason helps to understand "phenomena of life". This is what mathematics education, especially early stages of mathematics education, should be about: teaching students

- to see "phenomena of life" and use at least some basic mathematics vocabulary and technique for their description and analysis,

and, conversely,

- use their understanding of "mathematical meaning in phenomena of life" for simplifying their mathematics.

These skills should not remain just an exotic trait of a small number of children who somehow attained them by absorbtion from their cultural environment and, as a result, are classified as "able" or "gifted".

I firmly believe that every child should be given a chance to see "phenomena of life" through mathematical lenses.

And the last but not the least: the anonymous referee suggested, in her/his most helpful and enlightening comments, to move the mathematical material of this Section to Section 6, and, as a consequence, put the world line diagrams of Figures 1 and 4 next to each other. I have some difficulty with that: the development of elementary mathematics can not be linear because we have *mathematics for pupils* and *mathematics for teachers* which inevitably live in different dimensions. The time-space plane of world lines is too hard for children, but, in my opinion, it should be part of mathematics education of every primary school teacher.

11 A bonus of "economy of thought": "reduced fatigability"

And this is another quote from Krutetskii:

> *The reduced fatiguability in mathematics lessons that characterizes Sonya should also be noted. Not only is she very hard-working and fond of solving problems "on reasoning": she tires comparatively seldom during these lessons (excluding long, involved calculations, which she does in her head). Neither the lessons at home nor those with the experimenter were ever ended on her own initiative. Even prolonged lessons (for her age) did not lead to marked fatigue. For experimental purposes we set up a few lessons with her of an hour and a half, without interruption (a* 45*-minute lesson doubled!). Only at the very end of this period did the little girl of* 8 *show signs of fatigue (mistakes, slackening of memory). When occupied with other types of work (music, reading, writing), Sonya tires normally.*

In Krutetskii's voluminous book, Sonya is not the only subject. In particular, he interviewed a number of school teachers. This is Krutetskii's quote from one of them:

> *"The mathematically able are distinguished by a striking ability not to tire even after extended lessons in mathematics. I have constantly noticed this. And for some of them mathematics lessons are relaxing. This is probably related to the fact that a capable pupil spends very little energy on what incapable pupils work to exhaustion doing" (Ya. D.,* 18 *years of service).*

Millions of parents could only dream of their children attaining "reduced fatigability" in mathematics work.

I understand that I commit the mortal sin of using introspection as a source of empirical evidence, but, as a research mathematician and teacher of mathematics with 40 years of experience, I suggest that "spending very little energy" is directly related to "economy of thought" and, in its turn, the "economy of thought", at least at earlier stages of learning mathematics, is determined by the right balance of encapsulation and de-encapsulation in mathematical thinking.

These are skills and traits that can be developed in children. They are well-known in pedagogical literature, and they are sometimes called "style".

I quote from *The Aims of Education* by Alfred North Whitehead (1929) (of the *Principia Mathematica*, Russell and Whitehead (1913) fame):

> *The most austere of all mental qualities; I mean the sense for style. It is an aesthetic sense, based on admiration for the direct attainment of a foreseen end, simply and without waste.*

Style in art, style in literature, style in science, style in logic, style in practical execution have fundamentally the same aesthetic qualities, namely, attainment and restraint. The love of a subject in itself and for itself, where it is not the sleepy pleasure of pacing a mental quarter-deck, is the love of style as manifested in that study.

. . . Style, in its finest sense, is the last acquirement of the educated mind; it is also the most useful. It pervades the whole being. The administrator with a sense for style hates waste; the engineer with a sense for style economises his material; the artisan with a sense for style prefers good work. Style is the ultimate morality of mind.

. . . Style is the fashioning of power, the restraining of power with style the end is attained without side issues, without raising undesirable inflammations. With style you attain your end and nothing but your end. With style the effect of your activity is calculable, and foresight is the last gift of gods to men. With style your power is increased, for your mind is not distracted with irrelevancies, and you are more likely to attain your object. Now style is the exclusive privilege of the expert. Whoever heard of the style of an amateur painter, of the style of an amateur poet? Style is always the product of specialist study, the peculiar contribution of specialism to culture.

In children, "economy of thought" is still not a skill to work "without waste", but an instinctive striving to think economically and choose among possible approaches to a problem the one which promises the most streamlined and elegant solution.

"Economy of thought" in young children can be compared with style in sport. There were times, say, in swimming, when young boys and girls frequently won over adults – and it was before steroids came into use. Interestingly, this more frequently happened in long distance swimming, where economy of effort was paramount, and not in short distances, where sheer power and force of adults prevailed.

As a boy in Siberia, I did a bit of cross-country skiing – without much success, I have to say, but with great enjoyment. Aged 15, I knew a 12 years old girl from my school who could beat me at any distance. She liked to tease 18 year old male army conscripts on their compulsory 5 km skiing test, by flying past them, effortlessly like a snowflake in the wind. For sweating and short-breath guys, it was the ultimate humilation. The little girl had style.

But the correct technique, efficient style of swimming or skiing can be taught – if training starts at a right age and done properly. The same can be achieved in mathematics. It is simply a bit more expensive than the standard mass education because it requires investment in proper mathematics and pedagogic education of teachers, smaller classes, etc. – I do not wish to expand on the obvious. More generally, the entire socio-economic environment of Western industrial democracies is becoming increasingly unfavourable, even hostile, to mathematics education – I am writing about that in Borovik (2016a).

Still, "style" in the sense of Whitehead is not something which can be attained at a very young age in a fully developed form. But it is something which (in case of mathematics) can be irreversibly compromised at early stages of education if a student accumulates bad habits and mannerisms.

Acknowledgements I thank my mathematics homeschooler colleagues Maria Droujkova (see her book McManaman and Droujkova (2014)) and Julia Brodsky – their heroism is a source of inspiration for me, and this paper is written with them as readers in mind. Maria Droujkova's new

project, *Avoid Hard Work!*, Droujkova et al. (2017), focuses on simplicity as a guiding principle of mathematics education.

I am grateful to Roman Kossak and Philip Ording, the initiators and editors of this volume, for their patience and persistence.

Ron Aharoni, Natasha Artemeva, John Baldwin, Rick Booth, Julia Brodsky, Ronald Brown, David Corfield, Yagmur Denizhan, Maria Droujkova, Julian Gilbey, Daniel S. Helman, Roman Kossak, Brendan Larvor, Josef Lauri, Azadeh Neman, David Pierce, Wes Raikowski, Peter Ransom, Carlos Santacruz, Seb Schmoller, Anna Sfard, Chris Stephenson, Natasa Strabic, Adam Sulich, Yusuf Ünlü, and the anonymous referee provided most useful feedback and corrections. Tony Gardiner kindly corrected my translation of a fragment from Igor Arnold's paper.

Some parts of the paper were written thanks to my involvement with CMEP, the Cambridge Mathematics Education Project. I thank my CMEP colleagues, and especially Martin Hyland and Tony Gardiner, for many useful discussions—but neither they nor CMEP are responsible for my views expressed here.

This paper was completed when I enjoyed the warm hospitality offered to me in the Nesin Mathematics Village[14] in Şirince, Izmir Province, Turkey; my thanks go to Ali Nesin and to all volunteers and staff who have made the Village a mathematics research paradise, an oasis of proper mathematics education, and a garden of philosophy and arts Alladi and Nesin (2015), Karaali (2014). *Simplicity is the cultural dominant of the Village.*

Disclaimer

I write in my personal capacity; my views do not necessarily represent the position of my employer or any other person, corporation, organisation, or institution.

References

Alladi, K. and G. A. R. Nesin (2015). The Nesin Mathematics Village in Turkey. *Notices Amer. Math. Soc. 62*(6), 652–658.

Arnold, I. V. (1946). Principles of selection and composition of arithmetic problems. *Izvesiya APN RSFSR 6*, 8–28. In Russian.

Arnold, I. V. (2008). *Principles of selection and composition of arithmetic problems*. MCNMO. In Russian.

Arnold, V. I. (1998). On teaching mathematics. *Russ. Math. Surv. 53*(1), 229–236.

Borovik, A. (2012). "Free Maths Schools": some international parallels. *The De Morgan Journal 2*(2), 23–35.

Borovik, A. (2013a). Relationality of teaching, the Khan Academy, and word problems. *Selected Passages From Correspondence With Friends 1*(6), 45–50.

Borovik, A. (2013b). The strange fate of abstract thinking. *Selected Passages From Correspondence With Friends 1*(3), 9–12.

Borovik, A. (2016a). Calling a spade a spade: Mathematics in the new pattern of division of labour. In B. Larvor (Ed.), *Mathematical Cultures: The London*

[14] `http://nesinkoyleri.org/eng/`.

Meetings 2012–14, pp. 347–374. Springer. A pre-publication version: `goo.gl/TT6ncO`.

Borovik, A. (2016b). Didactic transformation. *Selected Passages From Correspondence With Friends 4*(1), 1–8.

Borovik, A. (2017). The Golden Age of Soviet mathematics education: The Ponzi scheme of social mobility. in preparation.

Borovik, A. V. and Ş. Yalçınkaya (2015). Adjoint representations of black box groups $\mathrm{psl}_2(\mathbb{F}_q)$.

Brodsky, J. (2015). *Bright, Brave, Open minds—encouraging young children in math inquiry*. Natural Math.

Cheney, D. and R. Seyfarth (2008). *Baboon Metaphysics: The Evolution of a Social Mind*. University of Chicago Press.

Corry, L. (2011). Writing the ultimate mathematical textbook: Nicolas Bourbaki's *éléments de mathématique*. In E. Robson and J. Stedall (Eds.), *The Oxford Handbook of the History of Mathematics*, pp. 565–588. Oxford University Press.

Dehaene, S. (2000). *The Number Sense: How the Mind Creates Mathematics*. Oxford University Press.

Droujkova, M., J. Tanton, and Y. McManaman (2017). in preparation.

Gardiner, A. D. (2014). Teaching mathematics at secondary level. *The De Morgan Gazette 6*(1), 1–215.

Havers, G. (2009). *Lincoln and the Politics of Christian Love*. University of Missouri Press.

Karaali, G. (2014). Nesin Math Village: Mathematics as a revolutionary act. *The Mathematical Intelligencer 36*(2), 45–49.

Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer.

Krutetskii, V. A. (1976). *The Psychology of Mathematical Abilities in Schoolchildren*. University of Chicago Press. Translated from Russian by J. Teller, edited by J. Kilpatrick and I. Wirszup; the Russian original published in Moscow in 1968.

Lemme, M. (2012). Utter elitism: French mathematics and the system of classes prépas. *The De Morgan Journal 2*(2), 5–22.

Lui, S. H. (1997). An interview with Vladimir Arnold. *Notices Amer. Math. Soc. 44*(4), 432–438.

McManaman, Y. and M. Droujkova (2014). *Moebius Noodles: Adventurous Math for the Playground Crowd*. Delta Stream Media. ASIN: B00IHPL1NU.

Morrissey, C. S. (December 19, 2012). Spielberg's *Lincoln*: Politics as mathematics. *The Catholic World Report*.

Russell, B. and A. N. Whitehead (1910, 1912, 1913). *Principia Mathematica, Volume I-III*. Cambridge University Press.

Schrader, Y. (1971). *Equality, Similarity, Order*. Nauka. In Russian.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects a different sides of the same coin. *Educational Studies in Mathematics 22*(1), 1–36.

Sfard, A. (2016). Objectifying symbols: The uneasy journey of the mathematical "signified" from the world, to mind, to discourse. Preprint, 23 pp.

Sinkevich, G. I. (2016). History of the concept of uniform continuity and the idea of coverings of a segment. *Hist. Sci. Tech. 4*, 3–17.

Weil, A. (1937). Sur les espaces à structure uniforme et sur la topologie générale. *Act. Sci. Ind.* (551), 3–40.

Weller, K., A. Brown, E. Dubinsky, M. McDonald, and C. Stenger (2004). Intimations of infinity. *Notices Amer. Math. Soc. 51*(7), 741–750.

Whitehead, A. N. (1929). *The Aims of Education and Other Essays*. Macmillan Company.