Educ Stud Math (2009) 71:263–278
DOI 10.1007/s10649-008-9175-8

Intuitive vs analytical thinking: four perspectives

Uri Leron · Orit Hazzan

Published online: 10 January 2009

Abstract This article is an attempt to place mathematical thinking in the context of more general theories of human cognition. We describe and compare four perspectives—mathematics, mathematics education, cognitive psychology, and evolutionary psychology—each offering a different view on mathematical thinking and learning and, in particular, on the source of mathematical errors and on ways of dealing with them in the classroom. The four perspectives represent four levels of explanation, and we see them not as competing but as complementing each other. In the classroom or in research data, all four perspectives may be observed. They may differentially account for the behavior of different students on the same task, the same student in different stages of development, or even the same student in different stages of working on a complex task. We first introduce each of the perspectives by reviewing its basic ideas and research base. We then show each perspective at work, by applying it to the analysis of typical mathematical misconceptions. Our illustrations are based on two tasks: one from statistics (taken from the psychological research literature) and one from abstract algebra (based on our own research).

Keywords Mathematics education · Cognitive psychology · Evolutionary psychology · Mathematical errors · Rationality debate · Dual process theory · Medical diagnosis problem · Bayesian thinking · Base-rate neglect · Group theory · Lagrange's theorem

I think, therefore I err.[1]

[1]Gigerenzer (2005), p. 1.

U. Leron · O. Hazzan (✉)
Department of Education in Technology and Science, Technion–Israel Institute of Technology, Haifa, Israel
e-mail: oritha@techunix.technion.ac.il

1 Introduction

This article is an attempt to place mathematical thinking in the context of more general theories of human cognition. We describe and compare four perspectives, each routed in a different research community, and each offering a different view on mathematical thinking and learning and, in particular, on the sources of mathematical errors.

1. Mathematics: where the focus is on the rules and norms of the mathematical research community.
2. Mathematics education: where the focus is on the individual and social processes in a community of learners, in and out of the classroom.
3. Cognitive psychology: where the focus is on the universal characteristics of the human mind and behavior, which are shared across individuals, cultures, and different content areas.
4. Evolutionary psychology: where the focus is on the evolutionary origins of human cognition and behavior and their expression in "universal human nature."

The four perspectives represent four levels of explanation, and we see them not as competing but as complementing each other, much as descriptions at different levels of the same phenomena in other sciences are necessary for fuller understanding of the observed phenomena. For example, here is a typical quote on the usefulness of multilevel view in biological psychology.[2] The "explanatory pluralism" discussed in that article seems very relevant to the multilevel character of mathematics education.

> [Theories] at different levels of description, like psychology and neuroscience, can co-evolve, and mutually influence each other, without the higher-level theory being replaced by, or reduced to, the lower-level one. Such ideas seem to tally with the pluralistic character of biological explanation. In biological psychology, explanatory pluralism would lead us to expect many local and non-reductive interactions between biological, neurophysiological, psychological and evolutionary explanations of mind and behavior. (Looren de Jong, 2002, p. 441)

In the classroom or in research data, all four perspectives may be observed. They may differentially account for the behavior of different students on the same task, the same student in different stages of development, or even the same student in different stages of working on a complex task.

The article can also be considered as an attempt to deal with the (admittedly vague) question, "Are mathematical errors good or bad?" (We are referring here not to accidental errors, but to those of the universal recurring kind.) The mathematical perspective typically views errors as bugs, something that went wrong due to faulty knowledge, and needs to be corrected.[3] The mathematics educational perspective typically views errors as partial knowledge, still undesirable, but a necessary intermediate stage on the way towards attaining professional norms, and a base on which new or refined knowledge can be constructed. Cognitive psychologists[4] typically view errors as an undesirable but unavoidable feature of the human mind, analogical to optical illusions, which originate at the interface between

[2] See also Nisan and Schocken (2005) for such a multilevel view of computer science.

[3] This is meant to describe a typical view in the community. Some individuals or subcommunities may, of course, hold different views.

[4] We refer here mainly to the reasoning and decision-making subcommunities.

intuitive and analytical thinking (Evans, 2008; Kahneman & Frederick, 2005; Stanovich & West, 2000). Evolutionary psychologists, in contrast, view errors as stemming from useful and adaptive features of human cognition (Barkow, Cosmides & Tooby, 1992; Buss, 2005; Cosmides & Tooby, 1996; Gigerenzer, & Todd, the ABC research group, 1999). In this view, mistakes are not really errors but, rather, the expression of an intelligent system that had been adapted by natural selection to Stone-Age ecology, and now has been "tricked" by unexpected conditions of modern civilization or in the psychologist's lab. According to this perspective, people make mistakes (at least of the universal recurring kind) not because of deficiencies in their intelligence or their knowledge but because the requirements of modern mathematics, logic, or statistics clash with their "natural" intelligence. In addition, by the very nature of an intelligent system, it must calculate and make inferences beyond the information given; hence, it must, to a certain extent, be prone to errors in the face of atypical conditions. Hence, the opening quotation, "I think, therefore I err." (Gigerenzer, 2005, p. 1).

The main part of the article consists of the application of the four perspectives to two mathematical tasks, one from statistics (taken from the psychological research literature) and one from abstract algebra (taken from our own research). The two tasks, introduced in Section 2, are known to elicit high rates of errors even among students in top universities. In Section 3, we introduce each of the perspectives by reviewing its basic ideas and research base. In Sections 4 and 5, we put the four perspectives to work by applying them to the analysis of the two tasks and the errors they engender.

2 The two tasks

The following two tasks will be used later in the paper to illustrate how each of the four perspectives can be used to explain students' misconceptions. The tasks were selected for the prevalence and magnitude of the misconceptions they engender.

2.1 Statistical thinking: "Are humans good intuitive statisticians after all?"[5]

Many "biases" concerning statistical thinking are discussed in the psychological research literature. One of the most famous is the "medical diagnosis problem" and the related phenomenon of "base-rate neglect." Here is a standard formulation of the task and data, taken from Samuels, Stitch and Tremoulet (1999):

> Casscells, Schoenberger, and Grayboys (1978) presented [the following problem] to a group of faculty, staff and fourth-year students at Harvard Medical School.
>
> > If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%,[6] what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs? ___%
>
> Under the most plausible interpretation of the problem, the correct Bayesian answer is 2%. But only eighteen percent of the Harvard audience gave an answer close to 2%. Forty-five percent of this distinguished group completely ignored the base-rate information and said that the answer was 95%. (p. 79–80)

[5] The quote is from Cosmides and Tooby (1996), p. 2.

[6] This means that 5% of the people who test positive do not have the disease.

This task is intended to test what is usually called *Bayesian thinking*: how people update their initial statistical estimates (the *base rate*) in the face of new evidence. In this case, the initial estimate (the prevalence of the disease in the general population) is 1/1,000, the new evidence is that the patient has tested positive (together with the 5% false-positive rate), and the task is intended to uncover how the subjects will update their estimate of the chance that the patient actually has the disease. *Base-rate neglect* reflects the widespread fallacy of ignoring the base rate, instead simply subtracting the false-positive rate (5%) from 100%.

A formal solution to the task is based on Bayes' theorem, but there are many complications and controversies surrounding the interpretation and application of that theorem.[7] This is a deep and fascinating issue involving mathematics, psychology, and philosophy, which is beyond the scope of this paper. See Cosmides and Tooby (1996) and Barbey and Sloman (2007) for extensive discussion.

Interestingly, it is possible to arrive at the Bayesian solution intuitively without using the formal Bayes' theorem: Assume that the population consists of 1,000 people and that all have taken the test. We know that one person will have the disease (because of the base rate) and will test positive (because no false-negative rate has been indicated). In addition, 5% of the remaining healthy 999 people (approximately 50) will test false-positive—a total of 51 positive results. Thus, the probability that a person that tests positive actually has the disease is 1/51, which is a little less than 2%.

We digress for a classroom-oriented comment. Confronting people who gave the 95% answer with the correct mathematical solution is not, in our view, a satisfactory way to deal with the conflict between their primary intuition and the mathematical solution (in either its intuitive or formal versions). In our experience, students will reluctantly accept the analytical solution but will remain confused and unhappy. This is because, we submit, they still hold in their mind the two solutions: one immediate and intuitive but now declared illegitimate; the other mathematically respectable but clashing with the intuitive one. And we as mathematics educators have not really done much to help them alleviate the tension between the two incompatible entities that simultaneously inhabit their mind.

A good way of helping students deal with this clash is to look at some variations of the original problem, in particular extreme cases. For example, assume there is a country where the disease has been totally eliminated (that is, its base rate is zero). Now if you ask your students what is the probability that someone from that country who tested positive actually has the disease, they will likely say, "you must be kidding, you just told us that no one has the illness, so of course the probability is zero"—no base-rate neglect here! But now assume that the disease in that country is *almost* totally eliminated, say only 1 out of 1,000 inhabitants has it, would you not expect that the answer to the same question would be, "the probability is *nearly* zero"? Similarly, we might consider the same question in an unfortunate country where *everybody* has the disease. In general, we could imagine a

[7] Here is Cosmides and Tooby's (1996, p. 4) explanation:

> In science, what everyone really wants to know is the probability of a hypothesis given data—p(H|D). That is, *given these observations, how likely is this theory to be true?* This is known as an *inverse probability* or a *posterior probability*. The strong appeal of Bayes' theorem arises from the fact that it allows one to calculate this probability:
>
> $p(H|D) = p(H)p(D|H)/p(D)$, where $p(D) = p(H)p(D|H) + p(\sim H)p(D|\sim H)$.
>
> Bayes' theorem also has another advantage: it lets one calculate the probability of a single event, for example, the probability that a particular person, Mrs. X, has breast cancer, given that she has tested positive for it.

computer microworld (or even let the student develop one) that models this problem, with a slider that assigns variable base rates and demonstrates the resulting probabilities.

2.2 Mathematical reasoning: "Students' use and misuse of Lagrange's theorem"[8]

The data are drawn from the performance of university students on a group theory task. We first lay out the necessary mathematical background.

The entire task takes place within the group Z_6, consisting of the set {0,1,2,3,4,5} and the operation of addition modulo 6, denoted by $+_6$. For example, $2 +_6 3 = 5$, $3 +_6 3 = 0$, $3 +_6 4 = 1$, and, in general, $a +_6 b$ is defined as the remainder of the usual sum $a+b$ on division by 6. Z_6 is a *group* in the sense that it contains 0 and is *closed* under addition mod 6: if a and b are in Z_6, then so is $a +_6 b$.[9] Similarly, we define Z_3 to be the group consisting of the set {0,1,2} and the operation $+_3$ of addition modulo 3. A *subgroup* of Z_6 is a subset of {0,1,2,3,4,5}, which is, in itself, a group under the operation defined in Z_6. For example, it can be checked that the subset {0,2,4} is a subgroup of Z_6, since it contains 0 and is closed under $+_6$.

All the groups in this discussion are *finite*, in the sense that they have a finite number of elements; this number is called the *order* of the group. Thus, the order of Z_6 is 6 and the order of Z_3 is 3. Finally, an important theorem of group theory, called *Lagrange's theorem*, states that *if H is a subgroup of G, then the order of H divides the order of G*; for example, if H is a subgroup of Z_6, then the order H divides 6. Thus, the order of H cannot be 4 or 5 but 3 is possible, and indeed, we have seen above an example of a subgroup of Z_6 with three elements. For what follows, it is relevant to mention that the *converse* of Lagrange's theorem is *not* true in general. For example, it is possible to give an example of a group G of order 12 that does not contain a subgroup of order 6 (see, e.g., Gallian, 1990, example 13, p. 151).

2.3 The task and data (Hazzan & Leron, 1996)

The following task was given to 113 computer science majors in a leading Israeli university, who had previously completed courses in calculus and in linear algebra (an abstract approach), and were now in the midst of an abstract algebra course:

A student wrote in an exam, "Z_3 is a subgroup of Z_6."
In your opinion, is this statement true, partially true, or false?
Please explain your answer.

An incorrect answer was given by 73 students, 20 of whom invoked Lagrange's theorem, in essentially the following manner:

Z_3 is a subgroup of Z_6 by Lagrange's theorem, because 3 divides 6.

Mathematical remark 1 The correct answer is that Z_3 is *not* a subgroup of Z_6. The reason is that Z_3 is not closed under the operation $+_6$ (for example, $2 +_6 2 = 4$, and 4 is not in Z_3). The question is tricky because Z_3 is a sub-*set* of Z_6 and is a group (relative to $+_3$), but it is not a sub-*group* of Z_6 (since it is not a group relative to $+_6$).[10] There is a sophisticated sense

[8] Adapted from Hazzan and Leron (1996).

[9] In the present context, these conditions are equivalent to the standard definition of a group.

[10] The elements of the groups Z_n are often taken to be equivalence classes, not numbers as in our definition, which would lead to a different mathematical analysis of the task. The present analysis, however, is the one relevant for the version that our students have learned.

in which the statement "Z_3 is a subgroup of Z_6" is partially true, namely, that Z_3 is *isomorphic* to the subgroup {0, 2, 4} of Z_6. We would, of course, have been thrilled to receive this answer, but none of our 113 students gave it.

Mathematical remark 2 As can be seen from the previous remark, the correct solution does not use Lagrange's theorem. It is relevant to mention that, in spite of superficial resemblance, there is no way Lagrange's theorem could even *help* on this task, since "*H* is a subgroup" is the hypothesis of Lagrange's theorem, not its conclusion. What the students seem to be using is an incorrect version of a nontheorem (namely, the *converse* of Lagrange's theorem).[11]

3 Reviewing the four perspectives

We now introduce each of the four perspectives outlined in the introduction. Obviously, not all members of the relevant communities hold the same views, and sometimes, not even similar views. What we are portraying is what we believe is a typical view held in the relevant community, or at least in a significant subcommunity. Where possible, we support our profiles with appropriate references to the literature.

Since the first two perspectives (mathematics and mathematics education) are well known to readers of this journal, we dwell on them only briefly, as a basis for comparison with the other perspectives (cognitive and evolutionary psychology), which are introduced and discussed more thoroughly.

3.1 A mathematical perspective: errors as bugs

This perspective looks at mismatches between students' work and the norms of the professional mathematician. The source of students' errors is typically assumed to be their faulty mathematical knowledge, and the way to address them contains teaching suggestions, such as explaining a difficult point, giving more examples or exercises, etc.

3.2 A mathematics education perspective: errors as expressing intermediate state in knowledge construction

In the mathematics education literature, especially in research on misconceptions, knowledge is typically viewed as being actively constructed by the learner, rather than being transmitted from teacher to student. The process of knowledge construction may be lengthy, and in the intermediate stages, the learner will have partial knowledge, which may result in errors. However, these errors are viewed as a normal and acceptable part of the learning process rather than expressing faulty performance. Moreover, errors are valued as offering a window into the learner's states of mind in the process of learning. Thus, researchers document recurring errors and develop theoretical interpretations to account for the process that yields these errors (e.g., Davis, Maher, & Noddings, 1990).

[11] Hazzan and Leron (1996) discuss data on two more tasks, which show that this misuse of Lagrange's theorem is deeper and more prevalent than might appear merely from the data presented here.

3.3 A cognitive psychology perspective: errors as a clash between two systems of thinking

Here, we will mainly discuss the *dual-process theory* (introduced below) and the heuristics-and-biases research program in cognitive psychology (led by Kahneman & Tversky over the last 30 years; cf., e.g., Evans 2003, 2008; Gilovich, Griffin, & Kahneman, 2002; Kahneman, 2002; Stanovich & West, 2000, 2003).

Dual-process theory The ancient distinction between intuitive and analytical modes of thinking has achieved a new level of specificity and rigor in what cognitive psychologists call *dual-process theory*. The first application of this theory to mathematics education research, to the best of our knowledge, appears in Leron and Hazzan (2006); the present exposition and analysis is an abridged version of the one given in that paper.

According to dual-process theory, our cognition and behavior operate in parallel in two quite different modes, called *System 1* (S1) and *System 2* (S2), roughly corresponding to our commonsense notions of intuitive and analytical modes of thinking. These modes operate in different ways, are activated by different parts of the brain, and have different evolutionary history (S2 being evolutionarily more recent and, in fact, largely reflecting *cultural* evolution). The distinction between perception and cognition is ancient and well known, but the introduction of S1, which sits midway between perception and (analytical) cognition, is relatively new and has important consequences for how empirical findings in cognitive psychology are interpreted, including the wide-ranging *rationality debate* (Samuels et al., 1999; Stanovich & West, 2000, 2003; Stein, 1996), and the application to mathematics education research.

Like perception, S1 processes are characterized as being fast, automatic, effortless, unconscious and inflexible (hard to change or overcome); unlike perceptions, S1 processes can be language-mediated and relate to events not in the here-and-now. S2 processes are slow, conscious, effortful, computationally expensive (drawing heavily on working memory resources), and relatively flexible. The two systems differ mainly on the dimension of *accessibility*: how fast and how easily things come to mind. In most situations, S1 and S2 work in concert to produce adaptive responses, but in some cases (such as the ones concocted in the Heuristics-and-biases and in the reasoning research), S1 generates quick automatic *nonnormative* responses, while S2 may or may not intervene in its role as monitor and critic to correct or override S1's response.

Many of the nonnormative answers people give in psychological experiments—and in mathematics education tasks, for that matter—can be explained by the quick and automatic responses of S1, and the frequent failure of S2 to intervene in its role as critic of S1.

Here is a striking example (Kahneman, 2002) for a combined failure of both systems:

> *A baseball bat and ball cost together one dollar and 10 cents. The bat costs one dollar more than the ball. How much does the ball cost?*

> Almost everyone reports an initial tendency to answer '10 cents' because the sum \$1.10 separates naturally into \$1 and 10 cents, and 10 cents is about the right magnitude. […] many intelligent people yield to this immediate impulse: 50% (47/93) of Princeton students, and 56% (164/293) of students at the University of Michigan gave the wrong answer. (p. 451)

According to dual process theory, the fast-reacting S1 "hijacks" the subject's attention and jumps automatically and immediately with the answer of 10 cents, since the numbers one dollar and 10 cents are salient, and since the orders of magnitude are roughly appropriate. For many people, S1's conclusions are accepted uncritically; thus, in a sense, they "behave irrationally." For others, S1 had also immediately jumped with this answer, but in the next stage, their S2 interfered critically and made the necessary adjustments to give the correct answer (5 cents). Significantly, the way S1 worked here, namely, coming up with a very quick decision based on salient features of the problem and of rough sense of what is appropriate in the given situation, usually gives good results under natural conditions, such as searching for food or avoiding predators.

Researchers in cognitive psychology do not usually consider educational implications of their research, so we offer instead our own thoughts on the implications of dual-process theory. Because many of the misconceptions come from combined failure of both S1 and S2, we propose that the most important educational implication is the need to train people *to be aware of the way S1 and S2 operate, and to include this awareness in their problem-solving toolbox*. This suggestion has an interesting (almost paradoxical) recursive nature: It in effect implies that S2 should monitor not only the operation of S1 (its standard role), but the S1-S2 interaction as well; thus, S2 has to monitor its own functioning in monitoring S1. In a way, we might say that an operation of a "System 3" is needed here (to monitor S2), but in practice, this function is recursively assumed by S2 itself.[12] While monitoring and critiquing S1 is one of the reasons S2 has evolved in the first place, monitoring the S1-S2 interaction seems to be what Geary (2002; see next section) has called *biologically secondary skills*, one which will not normally develop without explicit instruction.

When cognitive psychologists do mention education, it may sound a bit naïve to practitioners in the educational field:

> The characteristics that determine analytic system intervention, other than cognitive ability, are *dispositional*. People may *choose* to engage in effortful analytic thinking because they are inclined to do so by strong deductive reasoning instructions […] or, perhaps, because they have personal motivation.
>
> […]
>
> This dispositional aspect of the analytic system also provides encouragement to those who believe that our educational systems can and should encourage people to think in a more "rational," abstract, and decontextualized manner[…]. (Evans, 2006, p. 383)

While this advice is backed up by empirical data from the psychologist's lab, experienced mathematics educators facing the realities of the mathematics classroom usually prefer to deal with misconceptions through appropriate activities and discussions rather than instructions.

3.4 An evolutionary psychology perspective: errors as a clash between human nature and modern civilization

We take from the young discipline of evolutionary psychology (EP) the scientific view of *human nature* as a collection of universal, reliably developing, cognitive and behavioral abilities—such as walking on two feet, face recognition, and the use of language—that are spontaneously acquired and effortlessly used by all people under normal development (Cosmides & Tooby, 1992, 1997; Pinker, 1997, 2002; Ridley, 2003; Tooby & Cosmides, 2005). We also take from EP the evolutionary origins of human nature; hence, the frequent

[12] See Stanovich (2008) for a recent attempt to formulate a tri-process theory.

mismatch between the ancient ecology to which it is adapted and the demands of modern civilization. We all do manage, however, to learn many modern skills (such as writing or driving, or some mathematics) because of our mind's plasticity and its ability to "co-opt" ancient cognitive mechanisms for new purposes (Bjorklund & Pellegrini, 2002; Geary, 2002). However, this is easier for some skills than for others, and nowhere are these differences more manifested than in the learning of mathematics. The ease of learning in such cases is determined by the *accessibility* of the co-opted cognitive mechanisms rather than the complexity of the task.

We emphasize that what is part of human nature need not be innate: we are not born walking or talking. What seem to be innate are the motivation and the ability to engage the species-typical physical and social environment in such a way that the required skill will develop (Geary, 2002). This is the ubiquitous mechanism that Ridley (2003) has called "nature *via* nurture." We also emphasize that what is *not* part of human nature, or even what goes *against* human nature, need not be unlearnable. Individuals in all cultures have always accomplished prodigious feats such as juggling 10 balls while riding a bicycle, playing a Beethoven piano sonata, or proving an abstract mathematical theorem in a formal language. However, research on people's reasoning, and on mathematical thinking in particular, usually deals with what most people are able to accomplish under normal conditions. Under such conditions, many people will produce nonnormative answers if the task requires reasoning that goes against human nature. In terms of mathematical education, this means that learning successfully such skills will require a particularly high level of motivation and perseverance—conditions that are hard to achieve for a long time and for many people in the standard classroom. Finally, it is in order to note here that EP is a hotly debated discipline. Much of the criticism leveled at EP is ideologically or emotionally motivated, but see, e.g., Stanovich and West (2003) and Evans (2003) for a sample of scientifically respectable alternative views.

We take Geary (2002) as an example of the EP community's effort to create a discipline of "evolutionary educational psychology." Geary (2002) distinguishes between *biologically primary skills*, which emerge spontaneously and naturally in all people under normal development, i.e., they are part of human nature as conceived by evolutionary psychologists, and *biologically secondary skills*, which require more effort and more motivation to learn, usually within teaching-oriented situations.

> The principles of evolutionary educational psychology will provide a much needed anchor for guiding instructional research and practice. An evolutionarily informed science of academic development is in fact the only perspective that readily accommodates basic observations that elude explanation by other theoretical perspectives, such as constructivism […]
>
> For instance, it follows logically from the principles of evolutionary educational psychology that children will learn language easily and without formal instruction, and years later many of these children will have difficulty learning to read and write even with formal instruction. (p. 340)

Recently, the EP framework has been applied to explain a curious phenomenon concerning functions (Paz & Leron, 2008), which may offer some insight into the relationship between mathematical thinking and human nature. Paz & Leron study this relationship in the case of what they call the action-on-objects scheme (AOS: the Piagetian notion that *when you act on an object it changes its properties, but still remains the same object*) and the function concept. Based on their empirical findings, they propose that the same parts of human nature (the AOS in this case) that may initially support the learning of a mathematical concept (functions in this case) can later clash with more advanced versions of that same concept. This offers an intriguing view of the (cultural) evolution of mathematical concepts, as being initially anchored in human nature, but then developing in ways that clash with the very intuitive foundations that gave rise to them in the first place.

We conclude this concise review of EP with a word about the relation of human nature to dual-process theory (Section 3.3). Human nature consists, by definition, of a more-or-less fixed collection of traits and behaviors that all human beings in all cultures acquire spontaneously and automatically under normal developmental conditions. System 1, in our view, contains all the traits and behaviors that comprise human nature but, on top of that, also all the traits and behaviors that have become S1 for a particular culture or a particular person because of specific (nonuniversal) developmental conditions. For example, learning language is part of human nature and, thus, part of S1 for all human beings under normal developmental conditions; in contrast, speaking English is not part of human nature but is an S1 skill for people whose mother tongue is English. Similarly, driving a car is not part of human nature but has become an S1 skill for experienced drivers, as evidenced by their ability to hold an intellectual conversation (an S2 task, fully engaging the working memory resources) while driving.

An interesting corollary of an EP perspective is that the persistent errors of the kind discussed here stem not from weakness of the human cognitive system but from its *strength*. Indeed, human nature is a collection of all the skills people are naturally good at, and the errors stem from the clash between the requirements of modern society with these mechanisms that are adapted to the ecology of our Stone-Age ancestors.

4 Putting the four perspectives to work: statistical thinking

We now apply the four perspectives for the analysis of the medical diagnosis problem and the data presented in Section 2.1. As noted before, the two familiar perspectives (mathematics and mathematics education) will be only mentioned briefly, as a basis for comparison with the two relatively novel perspectives (cognitive and EP).

A mathematical point of view would point to an error in statistical thinking: students neglect to take into account the base rate, or, alternatively, fail to correctly apply Bayes' theorem.

Since we did not find in the *mathematics education* literature an analysis of the medical diagnosis problem, we will make up a possible interpretation with the hope of capturing some of the spirit of that community (or one of its subcommunities). One interpretation (in the spirit of Nesher & Teubal, 1975, or Leron & Hazzan, 1997) is that people do not look deeply into the problem but, instead, do some routine calculations based on verbal cues. For example, because of the meaning of "5% false positive," they may classify this problem as a "subtraction problem," and just do the subtraction 100%−5%, which leads to the observed base-rate neglect. Another interpretation is that people actually do take into account *some* base rate, but not necessarily the tiny one (0.001) postulated in the problem. In our experience, some people tend to personalize the problem ("if *I* tested positive…"), bringing in a whole baggage of realistic conditions that are abstracted away in the original formulation. For example, they would not normally take the test if they did not have an a priori *serious* worry that they might have the disease, that is, if they did not assume a very high base rate. In this case, a positive test would indeed mean a high probability of having the disease. Of course one could devise empirical tests to decide between these interpretations (Barby & Sloman, 2007).

We now proceed to analyze the medical diagnosis problem from a *cognitive psychology* perspective, more specifically, from a dual-process theory perspective. An initial analysis may look rather similar to the bat-and-ball analysis in Section 3.3. S1 quickly and effortlessly generates the 95% response because of its accessibility (subtracting the 5% error rate from 100%), and, because of the automatic, intuitive interpretation of "false positive," the base rate of 1/1,000 is completely ignored. As in the bat-and-ball analysis, too, "the dormant S2" failed to catch the error in its role as critic and monitor of S1's output. The difference between the two problems lies in the complexity of what S2 is required to notice and correct: adjusting the *difference* between the costs of the bat and ball in the former vs attending to the multiple *nested-set relationships* in the latter (the network of subsets among the whole population, the people who have the disease, the ones who were tested and the ones who tested positive). Thus, for S2 to notice and correct S1's response needs only a simple alert in the first case but a much greater effort and skill in the second, which may account for the difference in the percentage of correct responses.

More specifically, according to Evans (2006), in order to solve the medical diagnosis problem correctly, subjects must integrate all the information in a *single* mental model, and this is facilitated in formulations that make the nested-set structure salient (including, e.g., the frequentist formulation explained next).[13]

> It appears that heuristic processes cannot lead to correct integration of diagnostic and base rate information, and so Bayesian problems can only be solved analytically. This being the case, problem formats that cue construction of a single mental model that integrates the information in the form of nested sets appears to be critical. (p. 391)

Researchers with *evolutionary and ecological orientation* (Cosmides & Tooby, 1996; Gigerenzer, & Todd, the ABC research group, 1999) claim that people are "good statisticians after all" if only the input and output are given in "natural frequencies" (integers instead of fractions or percentages).

> In this article, we will explore what we will call the "frequentist hypothesis"—the hypothesis that some of our inductive reasoning mechanisms do embody aspects of a calculus of probability, but they are designed to take frequency information as input and produce frequencies as output. (Cosmides & Tooby, 1996, p. 3)

The EP explanation is that the brains of our hunter–gatherers ancestors developed such a module because it was vital for survival and reproduction and because this is the format that people would naturally use under those conditions. The statistical formats of today, however, are the result of the huge amount of information that is collected, processed, and shared by modern societies with modern technologies. To demonstrate a typical EP theorizing, it is worth quoting at some length from Cosmides and Tooby (1996):

> In our natural environment, the only database available from which one could inductively reason was one's own observations, and possibly those communicated by the handful of other individuals one lived with. More critically, the "probability" of a single event is intrinsically unobservable.
>
> No sense organ can discern that if we go to the north canyon, there is a .25 probability that today's hunt will be successful. Either it will or it won't; that is all one can

[13] In the following quotation, Evans uses *heuristic processes* instead of Stanovich & Kahneman's System 1 and *analytic processes* instead of System 2.

> observe. As useful as a sense organ for detecting single-event probabilities might be, it is theoretically impossible to build one. No organism can evolve cognitive mechanisms designed to reason about, or receive as input, information in a format that did not regularly exist.
>
> What *was* available in the environment in which we evolved was the encountered frequencies of actual events—for example, that we were successful 5 out of the last 20 times we hunted in the north canyon. Our hominid ancestors were immersed in a rich flow of observable frequencies that could be used to improve decision-making, given procedures that could take advantage of them. So if we have adaptations for inductive reasoning, they should take frequency information as input.
>
> Once frequency information has been picked up, why not convert it into a single-event probability? Why not store the encountered frequency—"5 out of the last 20 hunts in the north canyon were successful"—as a single-event probability—"there is a .25 chance that a hunt in the north canyon will be successful"? There are advantages to storing and operating on frequentist representations because they preserve important information that would be lost by conversion to a single-event probability. (pp. 15–17)

Because of such considerations, EP researchers have sometimes been accused of telling "just-so stories." However, this accusation is misconceived. As usual in EP methodology, such evolutionary theorizing is not taken as evidence but only as a theoretical framework for generating and explaining hypotheses. The test of the hypotheses is done under the standard psychological methodologies. Indeed, Cosmides and Tooby (1996) have replicated the experiment of Casscells et al. (1978), but with natural frequencies replacing the original fractional formats, and the base-rate neglect has all but disappeared:

> Although the original, non-frequentist version of Casscells et al.'s medical diagnosis problem elicited the correct bayesian answer of "2%" from only 12% of subjects tested, pure frequentist versions of the same problem elicited very high levels of bayesian performance: an average of 76% correct for purely verbal frequentist problems and 92% correct for a problem that requires subjects to construct a concrete, visual frequentist representation. (ibid, p. 58)

These data, and the evolutionary claims accompanying them, have been consequently challenged by other researchers (see Samuels et al., 1999; Evans, 2008). In particular, Evans (2008) claims that what makes the subjects in these experiments achieve such a high success rate is not the frequency format per se, but that "there is now much evidence that what facilitates Bayesian reasoning is a problem structure that cues explicit mental models of nested-set relationships" (p. 6.13). However, the fresh perspective offered by EP has been seminal in reinvigorating the discussion of statistical thinking in particular, and of cognitive biases in general. The very idea of the frequentist hypothesis, and the exciting and fertile experiments that it has engendered by supporters and opponents alike, would not have been possible without the novel evolutionary framework. Here is how Samuels et al. (1999) summarize the debate:

> But despite the polemical fireworks, there is actually a fair amount of agreement between the evolutionary psychologists and their critics. Both sides agree that people do have mental mechanisms which can do a good job at bayesian reasoning, and that presenting problems in a way that makes frequency information salient can play an important role in activating these mechanisms. (p. 101)

5 Putting the four perspectives to work: mathematical reasoning

We now apply the four perspectives for the analysis of students' use and misuse of Lagrange's theorem and the data presented in Section 2.2.

A *mathematical analysis* of the students' errors in the above application of Lagrange's theorem would include, for example: Students do not understand how to use the theorem, fail to check the initial conditions, or confuse between the theorem and its converse. From *the mathematical education perspective*, Hazzan and Leron (1996) and Leron and Hazzan (1997) give detailed analyses of the Lagrange's theorem data, both from a "misconceptions" perspective, and—using the novel tool of virtual monologue—from a "coping" perspective.

From the *cognitive psychology perspective*, Leron and Hazzan (2006), applying dual-process theory, proposed that some cases of students' misuse of Lagrange's theorem reflect a combined S1-S2 failure. The analysis closely resembles Kahneman's analysis of the bat-and-ball data, except for the somewhat surprising demonstration that S1 can "hijack" cognitive behavior even in advanced mathematical settings, where the name of the game is explicitly reasoning and analytical thinking (i.e., S2 mode).

As usual, the S1 response is invoked by what is most immediately accessible to the students in the situation, which also looks roughly appropriate to the task at hand. Specifically, the students know that using a theorem in such situations is expected; they also know more-or-less immediately and effortlessly that Lagrange's theorem says something about groups and subgroups and the divisibility of their orders (it is the *details and logic* of what the theorem says that requires the effortful and pedantic intervention of S2); finally, the appearance of the two numbers 3 and 6 as orders of the groups Z_3 and Z_6, and the fact that 3 divides 6, immediately and automatically cues Lagrange's theorem, yielding the answer, "Z_3 is a subgroup of Z_6 by Lagrange's theorem, because 3 divides 6." This is a striking example for an answer that looks entirely appropriate by the "logic" of S1, but is extremely inappropriate by the logic of S2.

In addition to S1's inappropriate reaction, S2 too fails in its role as critic of S1, since there is nothing in the task situation to alert the monitoring function of S2. The missing judgment—mainly that Lagrange's theorem cannot be used to establish the existence of a subgroup but only its absence—clearly requires S2 processes. It is important to note that some of the students may well have the knowledge required to produce the right answer, had they only stopped to think more (that is, invoke S2). The problem is, rather, that they have no reason to suspect that the answer is wrong; thus, the "permissive System 2" (Kahneman, 2002) remains dormant:

> [An] evaluation of the heuristic attribute comes immediately to mind, and […] its associative relationship with the target attribute is sufficiently close to pass the monitoring of a permissive System 2. (p. 469)

Just as in the bat-and-ball situation, the final (erroneous) response in this case is a combination of S1's quick and effortless reaction, together with S2's failure to take a corrective action in its role as critic and monitor. Since the operation of S1 is effortless and that of S2 so effortful, students will not make the extra effort unless something in the situation alerts them to such need. It is a feasible (and eminently researchable) hypothesis that, at least for some of the students, a small nonmathematical cue would be enough to set them on the path for a correct answer. They may already have all the necessary (S2) knowledge to solve this problem correctly, but a nudge by the interviewer (even just raising an eyebrow or looking doubtful) is needed to mobilize this knowledge. This shows, incidentally, that the dual-system framework leads not only to new explanations but also (like all good theories) to interesting new research *questions*.

We now turn to look at students' misuse of Lagrange's theorem from an *evolutionary psychology perspective* by referring to the *logic of social exchange*. Cosmides and Tooby (1992, 1997) have used the *Wason card selection task* (Wason & Johnson-Laird, 1972) to uncover what they refer to as people's evolved reasoning "algorithms." In a typical example of the Wason task, subjects are shown four cards, say [A], [6], [T], and [3], and are told that each card has a letter on one side and a number on the other. The subjects are then presented with the rule, "if a card has a vowel on one side, then it has an even number on the other side," and are asked the following question: *What card(s) do you need to turn over to see if any of them violate this rule?* The notorious result is that roughly 90% of the subjects, including science majors in college, give an incorrect answer. Many similar experiments have been carried out, using rules of the same logical form "if P then Q," but varying the content of P and Q. The error rate has varied depending on the particular context, but mostly remained high (over 50%).

The motivation behind the original Wason experiment was partly to see if people naturally behave in accordance with the Popperian paradigm that science advances through *refutation* of held beliefs (rather than their confirmation). The normative response to the Wason task depends on the question: What will *refute* the given rule? The answer is that the rule is violated if and only if a card has a vowel on one side but an *odd* number on the other. Thus, according to mathematical logic, the cards you need to turn are [A] (to see if it has an odd number on the other side) and [3] (to see if it has a vowel on the other side). Most subjects, however, choose the [A] card and sometimes also [6], but rarely [3].

Cosmides and Tooby (1992, 1997) have also presented their subjects with many versions of the task, all having the usual logical form "if P then Q," but varying widely in the contents of P and Q and in the background story. While the classical results of the Wason task show that most people perform poorly on it, Cosmides & Tooby demonstrated that their subjects performed significantly better on tasks involving conditions of *social exchange*. In social exchange situations, the individual receives some benefit and is expected to pay some cost. On theoretical grounds, and from what is known about the evolution of cooperation, certain kinds of social skills are expected to have conferred evolutionary advantages on those who excelled in them and, thus, would be naturally selected during evolutionary history. In the Wason task, social exchange situations are represented by statements of the form "if you get the benefit, then you pay the cost" (e.g., if you give me your watch, then I give you $20). A *cheater* is someone who takes the benefit but do not pay the cost. Cosmides & Tooby argue that when the Wason task concerns social exchange, a correct answer amounts to detecting a cheater. Since subjects performed correctly and effortlessly in such situations, and since evolutionary theory clearly shows that cooperation cannot evolve in a community if cheaters are not detected and punished, Cosmides & Tooby have concluded that our mind contains evolved "cheater detection algorithms."

Significantly for the Lagrange's theorem task discussed here, Cosmides & Tooby also tested their subjects on the "switched social contract" (mathematically, the converse statement "if Q then P"), in which the correct answer by the logic of social exchange is different from that of mathematical logic (Cosmides & Tooby, 1992, pp. 187–193). As predicted, their subjects overwhelmingly chose the former over the latter: *When conflict arises, the logic of social exchange overrides mathematical logic*. In other words, in a social exchange situation, people will mostly interpret the Wason task statement in a symmetrical way, as if it were an "if and only if" statement, rather than a directional way as required by mathematical logic.

This EP theoretical and empirical framework adds a new level of support, prediction and explanation to the many findings that students are prone to confusing between mathematical propositions and their converse and, in particular, to our Lagrange's theorem data. Importantly, in the EP view, people fail not because of a weakness in their cognitive apparatus, but because of its *strength*: our impressive skill in negotiating social relationships. Unfortunately for mathematics education, this otherwise adaptive skill sometimes happens to clash with the requirements of modern mathematical thinking.

6 Conclusion

In this article, we have proposed to view mathematical thinking in the wider context of human cognition in general, specifically, focusing on situations of conflict between intuitive and analytical thinking. We have surveyed research from cognitive psychology and EP that bears on two questions of interest to mathematics education theory and practice: How is mathematical thinking enabled by our general cognitive system, and, on the other hand, why is mathematical thinking so difficult for so many people? More specifically, we have discussed four perspectives on the sources of recurring mathematical errors and on how to deal with them in the classroom. We believe that looking at these questions from several perspectives—complementary rather than competing—may enrich our understanding and offer new insights and new research directions. As for teaching and curriculum planning, these perspectives give additional weight to the importance of teaching as much as possible in ways that are consonant with intuition (more specifically, with "human nature" in its evolutionary psychological sense), and the challenge of finding ways to bridge the gap where formal mathematics clashes with this intuition. Finally, the cognitive and evolutionary perspectives shed a new light (Leron, 2003) on the question "Is mathematical thinking a natural extension of common sense?"

Acknowledgements The first author gratefully acknowledges a Bellagio Residency grant from the Rockefeller Foundation, where some of the ideas presented in this paper were developed.

References

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *The Behavioral and Brain Sciences*, *30*(3), 241–254.

Barkow, J., Cosmides, L., & Tooby, J. (Eds.) (1992). In *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press: Oxford.

Bjorklund, D. F., & Pellegrini, A. D. (2002). *The origins of human nature: Evolutionary developmental psychology*. Washington, D.C.: American Psychological Association Press.

Buss, D. M. (2005). *The handbook of evolutionary psychology*. Hoboken: Wiley.

Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, *299*, 999–1000.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.Barkow, L.Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford: Oxford University Press.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1–73. doi:10.1016/0010-0277(95)00664-8.

Cosmides, L., & Tooby, J. (1997). *Evolutionary psychology: A primer*, Center for Evolutionary Psychology, University of California, Santa Barbara. http://www.psych.ucsb.edu/research/cep/primer.html. Accessed 26 September 2008.

Davis, R. B., Maher, C. A. & Noddings, N. (Eds.) (1990). Constructivist views on the teaching and learning of mathematics. *Journal for Research in Mathematics Education*, Monograph no. 4.

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7* (10), 454–459. doi:10.1016/j.tics.2003.08.012.

Evans, J. S. B. T. (2006). The heuristic–analytic theory of reasoning: extension and evaluation. *Psychonomic Bulletin & Review*, *13*, 378–395.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 61–624.

Gallian, J. A. (1990). *Contemporary abstract algebra* (2nd ed.). Lexington: D. C. Heath.

Geary, D. (2002). Principles of evolutionary educational psychology. *Learning and Individual Differences*, *12*, 317–345. doi:10.1016/S1041-6080(02)00046-8.

Gigerenzer, G. (2005). I think, therefore I err. *Social Research*, *72*(1), 1–24. doi:10.1007/s11205-004-4512-5.

Gigerenzer, G., Todd, P. M., & ABC research group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.) (2002). In *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.

Hazzan, O., & Leron, U. (1996). Students' use and misuse of mathematical theorems: the case of Lagrange's theorem. *For the Learning of Mathematics*, *16*, 23–26.

Kahneman, D. (Nobel Prize Lecture, December 8, 2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. In T. Frangsmyr (Ed.), *Les Prix Nobel* (pp. 416–499). http://www.nobel.se/economics/laureates/2002/kahnemann-lecture.pdf. Accessed 26 September 2008.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* pp. 267–294. Cambridge: Cambridge University Press.

Leron, U. (2003). *Origins of mathematical thinking: A synthesis, Proceedings CERME3*, Bellaria, Italy, March, 2003. http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG1/TG1_leron_cerme3.pdf. Accessed 26 September 2008.

Leron, U., & Hazzan, O. (1997). The world according to Johnny: a coping perspective in mathematics education. *Educational Studies in Mathematics*, *32*, 265–292. doi:10.1023/A:1002908608251.

Leron, U., & Hazzan, O. (2006). The rationality debate: application of cognitive psychology to mathematics education. *Educational Studies in Mathematics*, *62*(2), 105–126. doi:10.1007/s10649-006-4833-1.

Looren de Jong, H. (2002). Levels of explanation in biological psychology. *Philosophical Psychology*, *15*(4), 441–462. doi:10.1080/0951508021000042003.

Nesher, P., & Teubal, E. (1975). Verbal cues as an interfering factor in verbal problem solving. *Educational Studies in Mathematics*, *6*, 41–51. doi:10.1007/BF00590023.

Nisan, N., & Schocken, S. (2005). *The elements of computing systems: Building a modern computer from first principles*. Cambridge: MIT.

Paz, T., & Leron, U. (2008) The slippery road from actions on objects to functions and variables. *Journal for Research in Mathematics Education*, in press.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Pinker, S. (2002). *The blank slate: The modern denial of human nature*. Nyborg: Viking.

Ridley, M. (2003). *Nature via nurture: Genes, experience, and what makes us human*. London: Harper Collins.

Samuels, R., Stitch, S., & Tremoulet, P. (1999). Rethinking rationality: From bleak implications to Darwinian modules. In E. LePore, & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 74–120). Oxford: Blackwell.

Stanovich, K. (2008). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans, & K. Frankish (Eds.), *In two minds: Dual processes and beyond*. Oxford: Oxford University Press.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *The Behavioral and Brain Sciences*, *23*, 645–726. doi:10.1017/S0140525X00003435.

Stanovich, K. E., & West, R. F. (2003). Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality. In D. E. Over (Ed.), *Evolution and the psychology of thinking: The debate* (pp. 171–230). Hove: Psychology Press.

Stein, R. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Oxford University Press.

Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). New York: Wiley.

Wason, P., & Johnson-Laird, P. (1972). *The psychology of reasoning: Structure and content*. Cambridge: Harvard University Press.