THE POWER OF NATURAL THINKING: APPLICATIONS OF COGNITIVE PSYCHOLOGY TO MATHEMATICS EDUCATION

Uri Leron
Technion – Israel Institute of Technology
uril@technion.ac.il

The talk begins with an inquiry into the relationship between people's natural thinking – *the suit of skills that is acquired by all people spontaneously and successfully under normal developmental conditions – and mathematical thinking. More specifically, when do these two thinking modes go together and when do they clash? The influential* dual-process theory *from cognitive psychology is applied to shed some light on this issue. A remarkable conclusion is that many of the recurring errors we make come from the* strength *of our mind rather than its weakness. The talk then proceeds to address a crucial* design issue*: In cases where natural and mathematical thinking clash, what can we as math educators do to help students create peaceful coexistence between the two? The extensively-researched* medical diagnosis problem *will be used to demonstrate how theory, design and experiment collaborate in the pursuit of this goal.*

Introduction

In the background of this talk lurks the momentous *rationality debate*: Are humans rational beings or not? Or, better, how rational are human beings? Or, still better, what kind of rationality (or irrationality) is invoked under what conditions? This question had been endlessly debated by the great philosophers through the millennia, but has become an empirical issue for cognitive psychologists in the second half of the 20th century, culminating with the 2002 Nobel Prize in economy to Daniel Kahneman for his work with Tversky on "intuitive judgment and choice" (Kahneman, 2002).

In this talk I will focus on a narrower (and more immediately relevant) facet of the rationality debate: What is the relation between people's natural thinking – the suit of skills that is acquired by all people spontaneously and successfully under normal developmental conditions – and mathematical thinking. More specifically, when do these two modes of thinking go together and when do they clash? Or, even more specifically, when can we as math educators build on the strength of students' natural thinking, and when do we need to devise ways to overcome it. (For a comprehensive discussion of the rationality debate see Gigerenzer, 2005; Samuels *et al*., 2004; Saunders and Over, 2009; Stanovich and West, 2000; Stanovich, 2004; Stein, 1996.)

To begin our one-thousand-mile journey with a small step, consider the following puzzle:

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?

This simple arithmetical puzzle would be totally devoid of interest, if it were not for the fact that it poses what I will call a *cognitive challenge*, best summarized in Kahneman's (2002) Nobel Prize lecture:

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. Frederick found that 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)

The trivial arithmetical challenge has thus turned into a non-trivial challenge for cognitive psychologists: What is it about the workings of our mind that causes so many intelligent people to err on such a simple problem, when they surely posses the necessary mathematical knowledge to solve it correctly?

Complicating this cognitive challenge even further, research in cognitive psychology has revealed that *harder* versions of the task may result in *better* performance by the subjects. For example, we can enhance the subjects' performance by making the numbers more messy (let the bat and ball cost together $1.12 and the bat cost 94 cents more than the ball), or by displaying the puzzle via hard-to-read font on a computer screen (Song and Schwarz, 2008).

This challenge and many others like it have led to one of the most influential theories in current cognitive psychology, *Dual Process Theory* (DPT), roughly positing the existence of "two minds in one brain". These two thinking modes – intuitive and analytic – mostly work together to yield useful and adaptive behaviour, but, as the long list of cognitive challenges demonstrate, they can also fail in their respective roles, yielding non-normative answers to mathematical, logical or statistical tasks. A corollary of particular interest for mathematics education is that many recurring and prevalent mathematical errors originate from general mechanisms of our mind and not from faulty mathematical knowledge. Significantly, such errors often result from the *strengths* of our mind rather than its weaknesses (hence *the power of natural thinking* in the title).

This paper is organized in two main parts. In the first part (based on Leron and Hazzan 2006, 2009) I introduce the dual process theory and demonstrate its explanatory power in math education research. The second part, which is based on work in progress with Abraham Arcavi and with Lisser Rye Ejersbo, addresses the educational challenge of bridging the gap between intuitive and analytical thinking. This is treated as a design issue: Given a problem with counter-intuitive solution (in our case, the famous and extensively-researched *medical diagnosis problem*), design a variation of the problem which brings the solution closer to intuition (or, alternatively, stretches the intuition towards the solution). I hope that by focusing on the power of students' natural thinking, this talk might contribute to the goal of this conference: *Optimizing student understanding in mathematics*!

"Doin' what comes natur'lly": Dual-process theory (DPT)

Annie Oakley's phrase "Doin' what comes natur'lly", from Irving Berlin's musical *Annie get your Gun*, touches charmingly on the ancient distinction between intuitive and analytical modes of thinking. This distinction has achieved a new level of specifity and rigor in what cognitive psychologists call *dual-process theory* (DPT). In fact, there are several such theories, but since the differences are not significant for the present discussion, we will ignore the nuances and will adopt the generic framework presented in Stanovich and West (2000), Kahneman and Frederick (2005) and Kahneman (2002). For state of the art thinking on DPT – history, empirical support, applications, criticism, adaptations, new developments – see Evans and Frankish (2009). The present concise – and much oversimplified – introduction to DPT and its applications in mathematics education is based on Leron and Hazzan (2006, 2009).

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 common sense notions of intuitive and analytical thinking. These modes operate in different ways, are activated by different parts of the brain, and have different evolutionary origins (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 applications to the rationality debate and to mathematics education research.

Like perceptions, S1 processes are characterized as being fast, automatic, effortless, non-conscious 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. 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 may generate quick automatic *non-normative* responses, while S2 may or may not intervene in its role as monitor and critic to correct or override S1's response. The relation of this framework to the concepts of intuition, cognition and meta-cognition as used in the mathematics education research literature (e.g., Fischbein, 1987; Stavy and Tirosh, 2000; Vinner, 1997) is elaborated in Leron and Hazzan (2006).

Many of the non-normative 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. Significantly, according to this framework, some of the ubiquitous mathematical misconceptions may have their origins in general mechanisms of the human mind, and not in faulty mathematical knowledge.

The bat-and-ball task is a typical example for the tendency of the insuppressible and fast-reacting S1 to "hijack" the subject's attention and lead to a non-normative answer. Specifically, the salient features of the problem cause S1 to jump 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, the effortful and slow moving S2 is not alerted, and they accept S1's output uncritically, thus in a sense "behave irrationally" (Stanovich, 2004). For others, S1 also immediately had 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). Evolutionary psychologists, who study the ancient evolutionary origins of universal human nature, stress that 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's appropriate in the given situation, would be adaptive behaviour under the natural conditions of our ancestors, such as searching for food or avoiding predators (Buss, 2005; Cosmides and Tooby, 1997; Tooby and Cosmides, 2005). Gigerenzer (2005; Gigerenzer *et al.*, 1999) claims that this is a case of *ecological rationality* being fooled by a tricky task, rather than a case of irrationality.

Evans (2009) offers a slightly different view – called *default-interventionist* approach – of the relations between the two systems. According to this approach, applied to the bat-and-ball data, only S1 has access to all the incoming data, and its role is to filter it and submit its "suggestions" for S2's scrutiny, analysis and final decision. This is a particularly efficient way to operate in view of the huge amount of incoming information the brain constantly needs to process, because it saves the scarce working memory resources that S2 depends on. On the other hand, it is error-prone, because the features that S1 selects are the most accessible but not always the most essential. In the bat-and-ball phenomenon, according to this model, the features that S1 has selected and submitted to S2 were the salient numbers 10 cents and 1 dollar, but the

condition about the difference has remained below consciousness level. Even though S2 has the authority to override S1's decision, it may not do it due to lack of access to all the pertinent data.

The seemingly paradoxical phenomena that more difficult task formulations actually enhance performance is also well explained by DPT. Making the task more difficult in the above-mentioned sense, has the effect of suppressing the automatic response of S1, thus forcing S2 to participate. Since the subjects' S2 does posses the necessary mathematical knowledge, all that is required to solve the problem correctly is suppressing S1 and activating S2, which is exactly the effect of these added complications.

It is important to note that skills can migrate between the two systems. When a person becomes an expert in some skill, perhaps after a prolonged training, this skill may become S1 for this person. For example, driving is an effortful S2 behavior for beginners, requiring deep concentration and full engagement of working memory processing. For experienced drivers, in contrast, driving becomes an S1 skill which they can perform automatically while their working memory is engaged in other tasks, such as a deep intellectual or emotional conversation. Conversely, many S1 skills (such as walking straight or talking in a familiar but non-native language), deteriorate with advancing age, or when just being tired or drunk, all of a sudden requiring conscious effort to perform successfully (behaving in effect like S2).

We have now moved well along our one-thousand-mile journey, contemplating the power of natural thinking (roughly the psychologists' S1), and its uneasy relation with analytical thinking (S2). The psychological research literature on dual-process theory is immense, and we could barely touch the surface here. Many interesting and important questions remain open, such as what are the mechanisms that cause (or fail to cause) S2 to intervene and override S1's output. Much more is known to psychologists about such questions, but equally much still remains unanswered (see Evans and Frankish, 2009). In the rest of this paper we will delve more deeply into the educational relevance of the foregoing theoretical framework.

Bridging intuitive and analytical thinking: A design approach

Our second one-thousand-mile journey begins again with a small step – this time the famous string-around-the-earth puzzle (dating back to 1702).

Imagine you have a string tightly encircling the equator of a basketball. How much extra string would you need for it to be moved one foot from the surface at all points? Hold that thought, and now think about a string tightly encircling the Earth – making it around 25,000 miles long. Same question: how much extra string would you need for it to be one foot from the surface at all points?

Everybody seems to feel strongly that the Earth would need *a lot* more extra string than the basketball. The surprising answer is that they both need the same amount: 2π, or approximately 6.28 feet. (If R is the radius of any of them, then the extra string is calculated by the formula $2\pi(R+1) - 2\pi R = 2\pi$.)

As with the bat-and-ball puzzle, *this surprise is what we are after*, for it tells us something important about how the mind works, which is why cognitive psychologists are so interested in such puzzles. This time, however, alongside with the cognitive challenge, there is also an important *educational challenge*, to which we now turn. Suppose you present this puzzle to your math class. Being a seasoned math teacher, you first let them be surprised; that is, you first let them do some guessing, bringing out the strong intuitive feeling that the required additional string is small for the basketball but huge for the earth. Then you have them carry out the easy

calculation (as above) showing that – contrary to their intuition – the additional string is actually quite small, and is in fact independent of the size of the ball.

Now, given the classroom situation just described, here is the educational challenge: *As teachers and math educators, what do we do next?* I haven't conducted a survey, but my guess is that most teachers would leave it at that, or at best discuss with the students the clash they have just experienced between the intuitive and analytical solutions. But is this the best we can do?

Taking my clue from Seymour Papert (of Logo fame), I claim that in fact we can do better. We want to avoid the "default" conclusion that students should not trust their intuition, and should abandon it in the face of conflicting analytical solution. We also want to help students deal with the uncomfortable situation, whereby their mind harbours two conflicting solutions, one intuitive but now declared illegitimate, the other correct but counter-intuitive.

Papert's (1993/1980, 146-150) answer to this educational challenge is simple but ingenious: Just imagine a cubic earth instead of a spherical one! Now follow in your mind's eye the huge square equator with two strings, one snug around its perimeter and the other running in parallel 1 foot away (Fig. 1). Then you can actually *see* that the two strings have the same length along the sides of the square (the size doesn't matter!), and that the only additional length is at the corners. In addition, you can now *see* why the extra string should be 2π: It is equal to the perimeter of the small circle (of radius 1 foot) that we get by joining together the 4 circular sectors at the corners.

The final step in Papert's ingenious construction is to bridge the gap between the square and the circle with a chain of perfect polygons, doubling the number of sides at each step. The next polygon in the chain after the square would be an octagon (Fig. 1, right). Here we have 8 circular sections at the corners, each half the size of those in the square case, so that they again can be joined to form a circle of exactly the same size as before. This demonstrates that doubling the number of sides (and getting closer to a circle) leaves us with the same length of extra string.
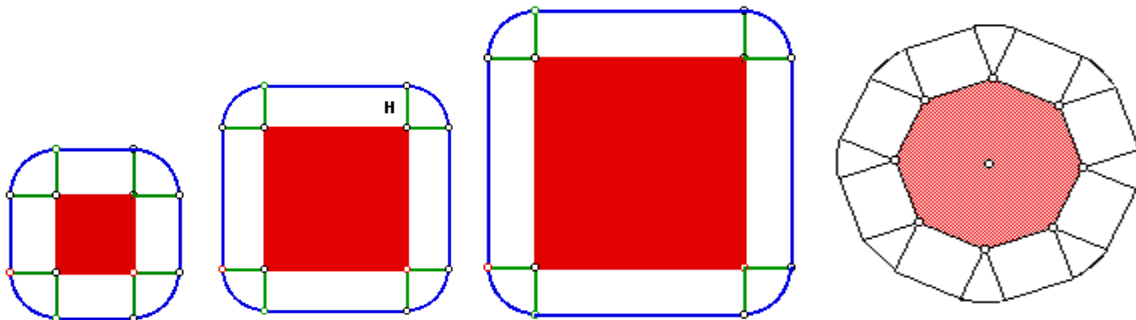

Fig. 1: A string around a square Earth – and around an octagonal one.

Can this beautiful example be generalized? When intuitive and analytical thinking clash, can we always design such "bridging tasks" that will help draw them closer? How should theory, design and experiment be put together in this search? We go more deeply into these questions in the next section.

Theory, design, experiment: The medical diagnosis problem (MDP)

Drawing inspiration from Papert's approach, and prompted by the questions closing up the last section, Lisser Rye Ejersbo and I set ourselves the challenge of designing an analogous treatment for the more advanced, relevant, and extensively-researched task from cognitive psychology: *the medical diagnosis problem* (MDP).

MDP background. Here is a standard formulation of the MDP task and data, taken from Samuels *et al.* (2004, p. 136).

> Before leaving the topic of base-rate neglect, we want to offer one further example illustrating the way in which the phenomenon might well have serious practical consequences. Here is a problem that Casscells et. al. (1978) presented to a group of faculty, staff and fourth-year students at the Harvard Medical School.
>
> > [**MDP**:] If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, 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%.

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 (the *diagnostic information*). In this case, the base rate is 1/1000, the diagnostic information is that the patient has tested positive, and the task is intended to discover how the subjects will update their estimate of the chance that the patient actually has the disease. The meaning of "5% false positive rate" is that 5% of the healthy people taking the test would test positive. *Base-rate neglect* reflects the widespread tendency among subjects to ignore the base rate, instead simply subtracting the false positive rate of 5% from 100%. Indeed, it is not at all intuitively clear why the the base rate should matter, and how it could be taken into the calculation.

A formal solution to the task is based on Bayes' theorem, but there are many complications and controversies involving mathematics, psychology and philosophy, concerning the interpretation of that theorem. Indeed, this debate – "Are humans good intuitive statisticians after all?" (Cosmides and Tooby, 1996) – is a central issue in the great rationality debate. See Barbey and Sloman (2007) for a comprehensive discussion, and a glimpse of the controversy.

Here is a simple intuitive solution for the MDP, bypassing Bayes' theorem: Assume that the population consists of 1,000 people and that all have taken the test (see Fig. 2). We know that one person will have the disease (because of the base rate) and will test positive (because no false negative rate is indicated). In addition, 5% of the remaining 999 healthy people (approximately 50) will test false-positive – a total of 51 positive results. Thus, the probability that a person who tests positive actually has the disease is 1/51, which is about 2%.

Researchers with evolutionary and ecological orientation (Cosmides and Tooby, 1996; Gigerenzer *et al.*, 1999) claim that people are "good statisticians after all" if only the input and output is 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 and Tooby, 1996, 3)

Evolutionary psychologists theorize that the brains of our hunter-gatherer ancestors developed such a module because it was vital for survival and reproduction, and because this is the statistical format that people would naturally encounter under those conditions. The statistical formats of today, in contrast, are the result of the huge amount of information that is collected, processed and shared by modern societies with modern technologies and mass media.

Indeed, Cosmides and Tooby (1996) have replicated the Casscells *et al.* (1978) experiment, 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. (Cosmides and Tooby, 1996, 58)

These results, and the evolutionary claims accompanying them, have been consequently challenged by other researchers (Evans, 2006; Barbie and Sloman, 2007). In particular, Evans (2006) claims that what makes the subjects in these experiments achieve such a high success rate is not the frequency format per se, but rather a problem structure that cues explicit mental models of nested-set relationships (see below). However, the fresh perspective offered by evolutionary psychology has been seminal in re-invigorating 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, p 101) 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.

The educational challenge as design issue. The extensive data on base rate neglect in the MDP (leading to the 95% answer) demonstrates the counter-intuitive nature of the analytical solution, as in the case of the string around the earth. As math educators, we are interested in helping students build bridges between the intuitive and analytical perspectives, hopefully establishing peaceful co-existence between these two modes of thought. As we have seen in Papert's example, achieving such reconciliation involves a design issue: Design a new *bridging task*, which is logically equivalent to, but psychologically much easier than the given task. (Compare Clements' (1993) "bridging analogies" and "anchoring intuitions" in physics education.)

From the extensive experimental and theoretical research on the MDP in psychology, we were especially influenced in our design efforts by the *nested subsets hypothesis* (Fig. 2):

> All this research suggests that what makes Bayesian inference easy are problems that provide direct cues to the *nested set* relationships involved […]
> It appears that heuristic [S1] processes cannot lead to correct integration of diagnostic and base rate information, and so Bayesian problems can only be solved analytically [i.e., by S2]. 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. (Evans, 2006, 391)

Indeed, it is not easy to form a mental representation of the subsets of sick and healthy people, and even less so for the results of the medical test. Mental images of people all look basically the same, whether they are sick or healthy or tested positive or negative. The task of finding a more intuitive version of the MDP has thus been operationalized to finding a task

which will "cue construction of a single mental model that integrates the information in the form of nested sets" (ibid).
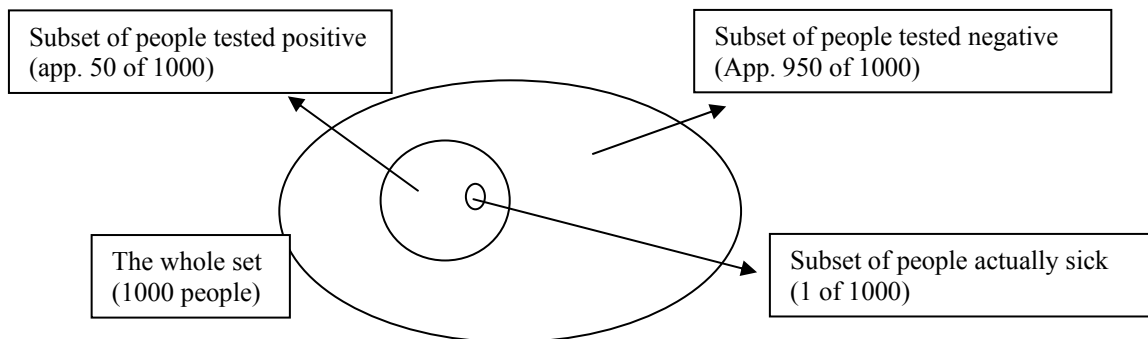

Fig. 2: Nested subsets in the medical diagnosis problem (for 1000 people)

Based on this theoretical background, we formulated three *design criteria* for the new task (which would also serve as testable predictions).

1. *Intuitively accessible*: The bridging task we will design will be easier ("more intuitive") than the original MDP (i.e., significantly more people – the term is used here in a qualitative sense –will succeed in solving it correctly).
2. *Bridging function*: Significantly more people will solve the MDP correctly, without any instruction, after having solved the new task.
3. *Nested subsets hypothesis*: Base rate neglect will be significantly reduced.

Note that the first two design criteria pull the new task in opposite directions. Criterion 1 (turning the hard task into an easy one) requires a task which is sufficiently different from the original one, while criterion 2 (the bridging function) requires a task that is sufficiently similar to the original one. The new task, then, should be an *equilibrium point* in the "design space" – sufficiently different from the original task but not *too* different.

Armed with these criteria, we set out on the search for the new bridging task. After a long process of trial and error, intermediate versions, partial successes and failures, we have finally come up with the *Robot-and-Marbles Problem* (RMP), which we felt had a good chance of satisfying the design criteria and withstanding the empirical test. The RMP is based on the idea of replacing sick and healthy people in the population by red and green marbles in a box. The medical test is then replaced by a colour-detecting robot, which can distinguish between red and green marbles via a colour sensor. The sensor is not perfect, however, and 5% of the green marbles are falsely identified as red, corresponding to the 5% healthy people in the MDP who are falsely diagnosed as sick. We also decided to make the action of the robot on the marbles more vividly imaginable by actually describing the process, not just the result. A final step in the design of the new problem was to slightly change the numbers from the original MDP, in order to make the connection less obvious. According to the bridging criterion, our subjects who solved the RMP first, should then solve more successfully the MDP. For this to happen, they would first need to recognize the *similarity* between the two problems, and we didn't want to make this too obvious by using the same numbers. (For a more detailed and nuanced description of the design process, as well as the experiments that followed, see Ejersbo and Leron, in preparation).

Here then is the final product of our design process, the version that would actually be put to the empirical test to see whether the design criteria have been satisfied.

RMP: In a box of red and green marbles, 2/1000 of the marbles are red. A robot equipped with green-marble detector with a 10% error rate (10% green marbles are identified as red), throws out all the marbles which it identifies as green, and then you are to pick a marble at random from the box. What is the probability that the marble you have picked would be red?

The experiment. The participants in the experiment were 128 students studying towards M.A degree in Educational Psychology at a Danish university, with no special background in mathematics or statistics. All the participants were assigned the two tasks – the medical diagnosis problem (MDP) and the robot-and-marbles problem (RMP) – and were given 5 minutes to complete each task. (In a pilot experiment we found that 5 minutes were enough both for those who could solve the problem and those who couldn't.) The subjects were assigned randomly into two groups of 64 students each. The order of the tasks was MDP first and RMP second for one group (called here the *MR group*), and the reverse order for the second group (the *RM group*). The results of the RM group were clearly our main interest, the MR group serving mainly as control.

	Group 1: Robot first		Group 2: Medical diagnosis first	
	RMP 1st	MDP 2nd	MDP 1st	RMP 2nd
Correct	31	17	8	20
Base-rate neglect	1	12	22	4
Incorrect other	32	35	34	40
Total	64	64	64	64

Table 1: Numbers of responses in the various categories

The results are summarized in Table 1, and it can be seen that the design criteria have been validated and the predictions confirmed. Here is a brief summary of the results for the RM group (with comparative notes in parentheses).

1. The RMP succeeded in its role as bridge between intuitive and analytical thinking: 48% (31/64) of the subjects in the RM group solved it correctly. (compared to 18% success on the MDP in the original Harvard experiment and 12% (8/64) in our MR group.)
2. The RMP succeeded in its role as stepping stone for the MDP: More than 25% (17/64) solved the MDP, without any instruction, when it followed the RMP. (Again compared to 18% in the original Harvard experiment and 12% in our MR group.)
3. The notorious base-rate neglect has all but disappeared in the RMP: it was exhibited by only 1 student out of 64 in the RM group and 4 out of 64 in the MR group. (Compared to 45% on the MDP in the original Harvard group and 34% in our MR group.)
4. Remarkably, the MDP, when given first, does not at all help in solving the RMP that follows. Worse, the MDP gets in the way: The table shows 48% success on the RMP alone, vs. 31% success on the RMP when given after the MDP.
5. Even though the performance on the RMP and the MDP has greatly improved in the RM group, still the largest number of participants appear in the "incorrect other" category. This category consists of diverse errors which do not directly relate to the MDP,

including (somewhat surprisingly for this population) many errors concerning misuse of percentages.

Conclusion

In this article we have seen how the dual-process theory from cognitive psychology highlights and help explain the power of natural thinking. We have used the medical diagnosis problem to discuss the gap between intuition (S1) and analytical thinking (S2), and to develop design principles for bridging this gap. It is my belief that bridging the gap between intuition and analytical thinking (in research, curriculum planning, learning environments, teaching methods, work with teachers and students) is a major step towards "optimizing student understanding in mathematics".

Based on the above examples and analysis, and indulging in a bit of over-optimism, here are some of the developments in the educational system I'd like to see happen in the future. (There is no claim of originality; many related ideas have appeared in the math education literature in various forms.)

- Map out the high school curriculum (and beyond) for components that could build on natural thinking and parts that would need to overcome it. For example, which aspects of functions (or fractions, or proofs) are consonant or dissonant with natural thinking?

- Design curricula, learning environments, teaching methods, that build from the power of natural thinking.

- Build a stock of puzzles and problems which challenge the intuition, and develop ways to work profitably with teachers and students on these challenges.

I wish to conclude with an even bigger educational challenge. If you ask mathematicians for examples of beautiful theorems, you will discover that many of them are *counter-intuitive*; indeed, that they are beautiful *because* they are counter-intuitive, *because* they challenge our natural thinking. Like a good joke, the beauty is in the surprise, the unexpected, the unbelievable. Like a world-class performance in classical ballet, sports, or a soprano coloratura, the beauty is (partly) in overcoming the limitations of human nature. Examples abound in the history of mathematics: The infinity of primes, the irrationality of $\sqrt{2}$, the equi-numerosity of even and natural numbers, the impossibility theorems (trisecting angles by ruler and compass, solving 5th-degree equations by radicals, enumerating the real numbers, Gödel's theorems). Recall, too, the joy of discovering that – contrary to your intuition – the extra length in the string-around-the-earth puzzle is quite small, and the beauty of Papert's cubic earth thought experiment.

Here, then, is the challenge: By all means, let us build on the power of natural thinking, but let us also look for ways to help our students feel the joy and see the beauty of going beyond it, or even against it. We thus arrive at the closing slogan – a variation on the title of the talk:

The power of natural thinking, the challenge of stretching it, the beauty of overcoming it.

Acknowledgements

References

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

Buss, D. M. (Ed., 2005). *The Handbook of Evolutionary Psychology*, Wiley.

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

Clements, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching* **30**(10), 1241-1257.

Cosmides, L. and Tooby, J. (1997). Evolutionary psychology: A primer. *Center for Evolutionary Psychology*, University of California, Santa Barbara, accessed on 4.8.10 at http://www.psych.ucsb.edu/research/cep/primer.html

Cosmides, L. and Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition* **58**, 1-73.

Evans, J.St.B.T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review* **13**(3), 378-395.

Evans, J.St.B.T. (2009). How many dual-process theories do we need? One, two, or many?, in Evans and Frankish (2009), 33-54.

Evans, J.St.B.T. and Frankish, K. (Eds., 2009). *In Two Minds: Dual Processes and Beyond*, Oxford University Press.

Fischbein, E. (1987). *Intuition in Science and Mathematics: An Educational Approach*, Reidel.

Gigerenzer, G., Todd, P.M., the ABC research group. (1999). *Simple Heuristics that Make us smart*, Oxford University Press.

Gigerenzer, G. (2005). I think, therefore I err. *Social Research* **72**(1), 1-24.

Kahneman, D. (Nobel Prize Lecture, 2002). Maps of bounded rationality: A perspective on intuitive judgment and choice, in T. Frangsmyr (Ed.), *Les Prix Nobel*, 416-499. Also accessible at http://www.nobel.se/economics/laureates/2002/kahnemann-lecture.pdf

Kahneman, D. and Frederick, S. (2005). A model of Heuristic judgment, in Holyoak, K.J. and Morisson, R.G. (Eds., 2005), *The Cambridge handbook of thinking and reasoning*, Cambridge University Press, 267-294.

Leron, U., & Hazzan, O. (2006). The rationality debate: application of cognitive psychology to mathematics education. *Educational Studies in Mathematics*, **62**(2), 105–126.

Leron, U., & Hazzan, O. (2009). Intuitive vs. analytical thinking: four perspectives. *Educational Studies in Mathematics*, **71**, 263-278.

Papert, S. (1993/1980). *Mindstorms: Children, Computers and Powerful Ideas*, Basic Books (2nd Edition).

Samuels, R., Stitch, S. and Tremoulet, P. (1999). Rethinking rationality: From bleak implications to Darwinian modules, in E. LePore and Z. Pylyshyn (Eds.), *What is Cognitive Science?,* Blackwell, 74-120.

Samuels, R., Stitch, S. and Faucher, L. (2004). Reason and rationality, in Niiniluoto, I., Sintonen, M. and Woleński, J. (Ed., 2004). *Handbook of Epistemology*, Kluwer, 131-182

Saunders, C. and Over, D.E. (2009). In two minds about rationality?, in Evans and Frankish (2009), 317-334.

Song, H. and Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social cognition* **26**(6), 791-799.

Stanovich, K. E. and West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Science* **23**, 645-726.

Stanovich, K. E. (2004). *The Robot Rebellion: Finding Meaning in the Age of Darwin*, Chicago University Press.

Stavy, R. and Tirosh, D. (2000). *How Students (Mis-)Understand Science and Mathematics: Intuitive Rules*, Teachers College Press.

Stein, R. (1996). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science,* Oxford University Press.

Tooby, J. and Cosmides, L. (2005). Conceptual foundations of evolutionary psychology, in Buss, D. M. (Ed.), *The Handbook of Evolutionary Psychology*, Wiley, 5-67.

Vinner, S. (1997). The pseudo-conceptual and the pseudo-analytical thought processes in mathematics learning, *Educational Studies in Mathematics* 34, 97–129.