Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2007/24

Better Than Conscious? The Brain, the Psyche, Behavior, and Institutions

Christoph Engel / Wolf Singer

MAX PLANCK SOCIETY

Preprints of the
Max Planck Institute
for Research on Collective Goods Bonn 2007/24

Better Than Conscious?
The Brain, the Psyche, Behavior, and Institutions

Christoph Engel / Wolf Singer

December 2007

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.coll.mpg.de

Better Than Conscious? The Brain, the Psyche, Behavior, and Institutions

Christoph Engel[1] and Wolf Singer[2]

Abstract

The title of this chapter is deliberately provocative. Intuitively, many will be inclined to see conscious control of mental process as a good thing. Yet control comes at a high price. The consciously not directly controlled, automatic, parallel processing of information is not only much faster, it also handles much more information, and it does so in a qualitatively different manner. This different mental machinery is not adequate for all tasks. The human ability to consciously deliberate has evolved for good reason. But on many more tasks than one might think at first sight, intuitive decision-making, or at least an intuitive component in a more complex mental process, does indeed improve performance. This chapter presents the issue, offers concepts to understand it, discusses the effects in terms of problem solving capacity, contrasts norms for saying when this is a good thing, and points to scientific and real world audiences for this work.

1 Max Planck Institute for Research on Collective Goods, 53113 Bonn, Germany
2 Max Planck Institute for Brain Research, 60528 Frankfurt am Main, Germany

Illustrations

Combining Quantities

2 + 2 = ? What a stupid question! As of the first grade, most everyone can easily answer this. If you do not believe it, let us prove it. Assume the answer were 4. One the axioms of arithmetics posits that addition is revertible:

$$a + b = c \Leftrightarrow c - b = a,$$

from the original question we have $a = 2$ and $b = 2$. Then according to the right-hand side of the above equation, we must have $c - 2 = 2 \Rightarrow c = 2 + 2 = 4$ ■

You are not convinced by the proof? Well, since you are still not content: here is an apple. Here is another apple. How many apples are on the table? Let's count them. One apple, two apples, period. Now let's do the same operation, with two more apples, on the desk. Now we have two apples on the table, and two on the desk, right? What if we put all apples on the table? Well, let's count again: 1, 2, 3, 4. Convincing?

Yes, but a category error. Our proof was axiomatic.We exploited the axiom that addition and subtraction are perfectly revertible. By definition, an axiom cannot be proven. One can only design proofs built upon axioms that are accepted by the community to which one is speaking. Our second argument, by contrast, utilized the counting routine. It relied on (a) a social practice that all of us have been taught in childhood, (b) a piece of language, and (c) the concept of natural numbers situated behind this language. You probably had no difficulty following the second example since you share this social practice, the meaning of the names for natural numbers, and the experience that there is a reliable mapping between the two of these and simple arithmetic. However, ultimately the second example does not constitute a proof, but instead is simply an appeal to shared cultural practice.

Of course, both this practice and the axiomatic system are very useful. What if we put two oranges on the table and two oranges on the desk? Dead easy: 2 + 2 = 4. And what if we put two apples on the table and two oranges on the desk? How many pieces of fruit do we have on the table if we move the oranges from the desk to the table? Dead easy again: 2 + 2 = 4. Try this task on a two-year-old and be prepared for a number of surprises. Which of the two tasks would you, as an adult, expect to be more difficult? You would probably guess the second, since it implies abstraction. Yet it has been shown that most two- to three-year-olds fully understand the abstract concept of numerosity (Wynn 1992). Once they know the numbers up to 20, it is even easy for them to extend their mathematical ability beyond the frontier, simply by asking them to repeat the operations of addition and subtraction (Lipton and Spelke 2005).

It is, however, much more difficult for children to map the counting routine and the names of numbers they associate with the abstract concept (Wynn 1992). Only by age four do children understand that the words of the counting routine refer to numerosity and that the routine itself functions to enumerate the items (Hauser and Spelke 2005). If children apply the counting routine to a novel set of objects, they perform well, up to the limit to which they have learned numbers. However, if you then ask them—How many of these objects are there?—only 16% of the 2½-year-olds are able to give the correct answer, and only 55% of the 3½-year-olds (Wynn 1990). If a child at age 2½ is told that a pile contains "four fish," and then two fish are removed, the child will insist that the pile still contains four fish (Condry et al. 2003). Although the child uses the counting words correctly in the counting routine, she evidently interprets each word above one as simply meaning "more than one" (Hauser and Spelke 2005).

How do other primates perform on the same tasks? Consider "Ai," a chimpanzee from the Primate Research Center in Kyoto, who for over 20 years has been involved in experiments that test her cognitive abilities (Matsuzawa et al. 2006). Initially, Ai interpreted "1" as "one" and "2" as "more than one." With further training, Ai learned to apply "2" and "3" correctly, but the amount of training needed to make this incremental move was no different from the amount of training needed for the first two integers. When the symbol "4" was introduced, Ai's performance fell to chance on "3"; she evidently interpreted "3" as "more than two." Ai has not progressed beyond the symbol "9" after 20 years of training. Chimpanzees learn the integer list through brute association, mapping each symbol to a discrete quantity. Human children, however, learn by making an induction from a limited body of evidence (Hauser and Spelke 2005).

Is solving the task 2 + 2 = ? on two apples and two oranges lying on a table a conscious activity? Before reading the above paragraphs, you may have again thought: What a stupid question! Defining consciousness is tricky business. For the moment let us agree that mental activity can be considered "conscious" if a person is aware of the activity. One is even more on the safe side if the individual is able to report this mental activity. That said, we would bet that not only is a toddler unable to report how she maps the names of natural numbers to the abstract concept of numerosity but, most likely, if you are not a cognitive scientist, you also have not been aware of the need to do that mapping before you perform the simple task of saying how many pieces of fruit are on the table. What patently appears to be a conscious activity heavily capitalises on non-conscious mental processes.

2 + 3 = ? Is the mental process for solving this problem, say again in apples, any different from solving 2 + 2 = ? God no, you will think: this is again primitive arithmetics. Yet if one measures reaction times, there is a huge difference, although the example is a bit tricky in that it is on the borderline. To make the difference net, let's compare 2 + 1 with 2 + 3. The reaction time on 2 + 1 is in the order of 40–100 ms above the baseline per item. For 2 + 3, it goes up to 250–350 ms above the baseline per item. Thus, to complete the second task, subjects need at least four times the time required to finish the first computation. This suggests that different mental mechanisms are involved. The process of counting fewer than (and sometimes up to) four items is an automatic, pre-attentive activity related to vision (Trick and Pylyshyn 1994). Humans are thus able to perform the first task (2 + 1) in a totally non-conscious way.

For a final example, consider the following: On the table as well as the desk, there is one bowl containing dried beans. You are asked to estimate whether the total amount of beans from the bowls on the table and desk is enough to fill an empty bucket on the floor. Now, if a computer would solve this problem, you would have to specify the following: how many beans of average size are necessary to fill the bucket? How many beans are, at the minimum, in either bowl? The computer would then add up these two numbers and compare them to the threshold of the bucket. Without these numbers, the computer would be helpless. Not so for human subjects, for along with the mental machinery for exact counting, humans possess a second mechanism that enables an approximate assessment of large numbers to be made. This mechanism is even able to perform simple arithmetic operations, such as addition or subtraction. The key feature, the tool for approximating quantity without counting, relies on an automatic process (Dehaene 1997).

Determination of Guilt

As long as a country is not at war, no other governmental activity interferes as severely in its citizens' lives as the criminal justice system. In some countries, a citizen accused of a crime risks being sentenced to death. In all countries, prison terms are possible. A precondition for sentencing an accused citizen (the defendant) is to prove to the requisite standard that she has indeed committed the crime to which she is accused. Defining this standard presents the legal order with a serious dilemma: unless the defendant pleads guilty, past events must be reconstructed to arrive at a decision. Reconstructing past events, however, is prone to errors. The risk of incorrect decisions is further aggravated by the adversarial nature of courtroom procedure. In their attempts to exact a decision in their favor, both the attorney general and the counsel for the defense deliberately try to influence the court; neither is likely to present a fair account of the events.

The legal order has to satisfy two goals:

1. Materially wrong decisions should be avoided. (Of course, convicting an innocent is worse than acquitting a guilty defendant. In the parlance of signal detection theory: false alarms matter more than false misses. Still, both lead to suboptimal decisions.)

2. The supreme exercise of sovereign powers should be ostensibly as correct as possible; the rules of law and principles of democracy demand this.

These two goals are often in conflict, because non consciously controlled or reportable decisions or elements that contribute to the decision are often more reliable.

Consider the process by which the legal system determines whether a witness has actually observed the defendant at the crime scene: A lineup is organized, where the accused is interspersed with a sufficient number of similar looking individuals. If the witness is able to single out the defendant, a first level of proof has been established. Of course, the legal system understands that this method is subject to error (Sporer et al. 1996), and thus it is also standard practice to ask the witness how confident she is, as well as to observe signals such as body language or the nuances of utterances. During this process, both the witness and the legal system rely on intuitive judgment, which, by its very nature, is not controllable or reportable. Only the results of this process can be reported.

Recent research by Sauerland and Sporer (2007) has analysed this more thoroughly. In a mock lineup, the reaction times of witnesses were measured and post-experiment reports matched to the results. When witnesses made an intuitive, holistic judgment, they did so at a much faster speed, and the outcome was much more reliable than if they had looked for certain criteria (e.g., hair color, shape of the nose, or thickness of the eyebrows). Thus, intuition beats deliberation.

In most criminal cases, though, conflict is not confined to an isolated aspect of the case. A whole series of events may be in dispute. To address these, the attorney general typically tells a story to explain the events and establish guilt. Tactically, the defendant and her counsel must decide whether they are better off telling a competing story or contradicting key features of the attorney general's story. Moreover, under criminal procedure, any feature of the story that is disputed must be formally proven. It is again a tactical issue for both parties to determine the order for the presentation of the evidence and counterevidence. Typically, the judge and/or the jury listen to evidence in a scrambled order, and it remains incomplete until the end. Experimental work with mock juries has shown that provisional stories are worked out as the evidence is presented. Jurors are content to convict if, after hearing all the evidence, there is a convincing story, with no conspicuous holes, and with no striking pieces of counterevidence (Hastie et al. 1983; Hastie 1993). In other words, juries and judges ultimately rely on a consciously not controlled mental process, namely intuition.

Subsequent research has deciphered the underlying mental mechanism. It has been found that a person determines whether a story is "convincing" once coherence is generated among the pieces of evidence. In this process, world knowledge defines base rates, and these rates are updated in light of the evidence. Subjects give progressively more weight to evidence that supports their preferred story while progressively underrating conflicting evidence. The process of re-weighting continues until the requisite mental standard for being convinced is reached, or until the process is stopped because the evidence is patently conflicting (Simon 1998; Simon, Pham et al. 2001).

This is troubling news for the legal order. A typical crime has far too many degrees of freedom to lend itself to any form of rigorous scientific fact finding. Typically, until the end of a court case, there remains a certain degree of uncertainty. The legal order cannot exclude the possibility that judges and juries engage in sense making. Actually, continental legal orders do so quite openly: They do not exclusively define the standard of proof in objective terms. They do not ask for a 95% or 99% probability that the defendant has committed the crime. Rather they ask a judge to rely on her *conviction intime*. A judge must be able to overcome her doubts or she must acquit the defendant (Schulz 1992). Psychological work on consistency maximization via parallel constraint satisfaction explains why this doctrinal requirement is meaningful (Glöckner, this volume). However, work on coherence-based reasoning demonstrates that the mechanism is not free of errors.

Playing against an Opponent

Workable competition is a beautiful institution; it forces actors on both sides of the market to reveal their preferences. Competition is a direct mechanism (Bolton and Dewatripont 2005). A seller has nothing to gain from withholding quantity or from demanding a price above marginal cost. A buyer has nothing to gain from refusing to deal, as other sellers or other buyers would make the profit. In game theoretic terms, workable competition forces all actors to play a game against Nature (Straffin 1993). It does not make sense for them to treat the other market side as a strategic actor.

The world looks very different if the other market side is small, or if the entry of new buyers or sellers into the market is unlikely (Baumol 1982). Here a seller is best off if she sets a high price or keeps the quantity sold low. Likewise a buyer is best off if she refuses to buy an expensive item, although the price for this one unit is below its marginal revenue. If she does, some units will not sell, but the units that are traded are sold cheaper. She cashes in a distributional gain.

Game theory offers a highly elegant technology for solving the resulting conflicts of strategic interaction. Its mainstay is the Nash equilibrium (Nash 1951). Either player plays a best response, given that the other player also plays a best response. Most experimental economics is about testing the predictions of game theory in the lab. Not surprisingly, experimental subjects systematically deviate from these predictions in a host of games, i.e., strategic situations. Typically, this is taken as evidence for cognitive biases (Kahneman and Tversky 2000), as well as for motivational dispositions that deviate from utility maximization (Kagel and Roth 1995). Less attention has been given to the fact that experimental evidence also speaks to the conscious/non-conscious divide.

A pertinent piece of evidence for this is provided by Ariel Rubinstein, one of the leading game theorists. Rubinstein has set up a website designed to help his colleagues teach game theory. A set of rather tricky games is presented, and teachers are invited to have their class play these games online in private the day before the respective topic is taught in class. The teacher then receives statistical information back from Rubinstein's site as to how the pupils performed on these games. As a byproduct, Rubinstein records reaction times. In most of these games, a characteristic picture is seen. While in theory there should be plenty of strategies taken, students utilized only a very few. More interestingly, for some of these strategies, reaction times are much shorter than for others. Apparently, only a few strategies are highly intuitive (Rubinstein 2007).

Consider the following, relatively simple example: Subjects were asked to be the row player against an anonymous opponent (the column player) in the following game:

	L	R
T	2, -2	0, 0
B	0, 0	1, -1

Figure 1
Ariel Rubinstein's game.

Read this matrix the following way: the row player has the choice of selecting either the top (T) or the bottom (B); the column player chooses between left (L) and right (R). There are four possible combinations of actions: TL, TR, BL, BR. In each of the four cells, the left number represents the row player's payoff; the right number is the payoff for the column player. If both players are fully rational, they would attempt to select the best response: What would the row player want to do if she knew that the column player would play left? Since 2 is more than 0, the row player would want to play top. Similarily, you can make the comparisons for all four cells, and for both players. In the matrix, the resulting best responses are underlined. Note that in no cell are two payoffs underlined. Whatever the other player does, the first player would want to do something different. In game theoretic parlance: the game has no Nash equilibrium in pure strategies. The only equilibrium of this game is in mixed strategies. The row player has to randomize between her two actions, anticipating that the column player will do the same. In this equilibrium, the row player would play top with probability 1/3, and bottom with the corresponding probability 2/3. Likewise, the column player would play left with probability 1/3, and right with probability 2/3.

If you have never been exposed to game theory, you will certainly not find this solution intuitive. In experiments, subjects find it pretty hard to randomize, if the possibility of an equilibrium in mixed strategies dawns on them at all. Moreover, Rubinstein has assigned you the role of the row player. This is a comfortable position. In the worst of all cases, you come out even. In the TR and BL cells, you make zero profit. However in both the TL and the BR cells, you expect a positive profit. Admittedly, it goes at the expense of your opponent. The game is a zero sum game. All you win is taken away from your opponent. Therefore, if you deeply believe in altruism, you would rather play bottom. That way you will make sure that your opponent cannot suffer a big loss. However, most of us are probably not that altruistic. We rather are attracted by the chance to make a big gain in the TL cell. That is at any rate what 67% of Ariel Rubinstein's subjects did. And those who did took less time to make up their minds. The median reaction time for those who choose top was 37 seconds, while the median for those who choose bottom was 50 seconds. Admittedly 37 seconds is still quite a bit of time. We must be careful in interpreting the difference. But it seems safe to say that subjects found playing top more intuitive, even if they have not reached that conclusion exclusively with their automatic system. In other games the differences in reaction times were even more impressive. But since most of these games require more game theoretic knowledge, they are not reported here.

For the purposes of this book, there are two implications: even at the very heart of rational choice theory, in strategic interaction, we should be prepared for actions that are influenced by nonconscious mental operations. Whether this is good news or not crucially depends on the structure of the game. In the concrete case, if the opponent anticipates you to be greedy, she will play right, which makes her best off, and you go home with nothing. She might find out since she is smart, or she might counter your intuition with her intuition, saying her that there are more egoists than altruists in the population. Now if you (intuitively) anticipate this, it is better to play bottom. A bird in the hand is worth two in the bush. Yet if your opponent is really clever, she will outwit this move, and play left, which again leaves you with nothing. Of course you could counter this by playing top and would be in the best of all worlds, ad infinitum. In this class of games, intuition does not offer an easy way out. But in many others, it does.

Concepts

How Many Agents Are Involved?

Actually, the presentation of the issue in the examples already makes an assumption that need not be uncontroversial. Conscious decision making and non-conscious decision making are seen as two distinguishable processes. Humans possess (at least) two different mental apparatuses that make decisions: one operates under conscious control; the other does not. This view does not exclude the possibility that both machineries interact, nor does it mean that either machinery has an exclusive domain, to which the other has no or only rare access. Both machineries need not be in competition at all times. Nonetheless, this perspective treats human decision makers as multiple agents.

Once one allows for this possibility, there is no need to limit the number of agents to just two. There might be an emotional agent, distinct from a more cold-blooded one. There might be a general purpose agent, distinct from a more task- or context-specific agent. There might be an agent for a decision taken by the individual alone, distinct from an agent for participating in group decision making.

If there is more than one agent, the interaction between them becomes the issue. Do the agents simultaneously work on the same task? If so, how are conflicts resolved? Is there something like a master and slave mode that gives one of the agents control over what the other does? Is it possible to split a decision task such that several agents provide well-defined inputs? If so, how are these inputs integrated, and by which agent? Is there something like a meta-decision that assigns a new task to one of the agents? If so, is this meta-decision taken by just one (superior?) agent, or by some form of competition between the potentially pertinent agents? Is the meta-decision taken centrally or decentrally? In metaphorical terms: Is the interaction of multiple mental agents better captured by a firm or by a market model?

Definitions

Defining consciousness is as unpleasant as nailing a pudding to the wall. In our experience, we do not take all decisions the same way. Not so rarely, we feel we are driven by forces we do not understand, we do not control, we do not predict. From these we feel we can distinguish other decisions. Colloquially we call them conscious. But all attempts to draw a bright line between conscious and the non-conscious meet with objections. For example:

1. A decision may be considered to have been consciously made if a person is able to report the information taken into consideration as well as the way in which it was processed. Such a definition could very well be too narrow, for the person could merely lack language or suffer from deception, as shown with the arithmetic example.

2. A decision may be considered to be conscious if a person is aware of the mental process leading to it. This definition may be too broad. If I ponder over a difficult problem, and all of a sudden I see the solution, I know it when I see it, but I have at best guesses where it came from.

3. A decision may be considered to have been consciously made when the prefrontal cortex has been engaged. Again, this explanation may be too narrow, as neuroimaging demonstrates that there is a gradual shift between almost no and very far-reaching activation. Intermediate states like "pre-consciousness" can be identified (Dehaene et al. 2006).

Clusters

Even if we can agree on one definition for consciousness, how do we describe the borderline area between related concepts: deliberate versus intuitive; explicit versus implicit; controlled versus automatic; serial versus parallel; analytic versus holistic; slow versus rapid; ad hoc versus learned; plastic versus patterned? One can certainly come up with examples that fall into the first category on pretty much all of these dimensions. For example, if I am a game theorist, and a friend approaches me with a serious problem of strategic interaction, I can sit down and work out a formal model exploiting all the many degrees of freedom game theory gives me, even though it will take me considerable time to do so. Likewise there are examples that fall into the second category on almost all of these dimensions: When I drive to work in the morning, almost all of the many decisions I make en route will be fast and automatic; I will process information in parallel, will heavily rely on prepackaged routines learned in driving school, and improved in my later driving experience.

However it is not difficult to identify examples that fall into one side of some dimensions, and to another side in others. The three illustrations given at the beginning of this chapter demonstrate this:

1. To do arithmetics, we must combine our understanding of numerosity with our knowledge of names for numbers. For most of us, neither mental process is consciously accessible. Thus, we build our conscious ability to perform mathematical operations on these non-conscious foundations.

2. When deciding whether a defendant has committed the crime, a jury certainly makes an effort to discuss the explicit relevance of the pieces of evidence heard in court. Ultimately, however, jury members cannot avoid, and should not try to avoid, making sense out of the evidence. They become aware of the result, but not of the mental machinery that generates coherence to produce the result.

3. When we decide how to act in the presence of an opponent, we may wish to exploit our knowledge of game theory. However, all of our training does not help if we fail to combine it with a sense of how this specific opponent is going to act. Only in the closed world of game theory can we take it for granted that our opponent is a rational utility maximizer. If our opponent deviates from these assumptions, and experimental economics has shown that many do, we are best off to optimize using a best guess about the type of behavior that our opponent will choose.

These illustrations raise questions; they unfortunately do not point to the answers. Behaviour can be clustered along the dimensions listed above. Fine. Behaviour need not be clustered along the dimensions listed above. Also fine. But when is it clustered? Along which dimensions are clusters more likely than along others? Are there systematic correlations between tasks or environments and certain clusters? Are there other systematic correlations between personality and certain clusters? Is clustering a matter of learning and training? Are interaction partners able to predict or influence these clusters?

Effects

Is Rational Choice Descriptively Correct?

Mental mechanism matters. Human subjects react differently to changes in their environment, depending on how they process information. Since the cognitive revolution in psychology, this view has been widely held (Chomsky 1959). Ultimately, to define the effects on performance, one needs to map the distinctions listed in the previous section to modes of decision making or problem solving. This raises the issue of clusters: Which combinations of features are at all feasible? How permanent are these combinations? Are the combinations human universals, or are they distributed in the population? How do individuals acquire problem-solving modes? How do they map onto perceived features of problems? What is the role of institutions in triggering, or in even generating, specific clusters?

Some of these questions have been addressed elsewhere (Engel and Weber 2006). For our purposes here, we will slight the clustering problem and ask a simplified question: Provided that an adequate cluster of features has formed, what is the effect of foregoing conscious control?

From the days of the grand Roman orators on, the *argumentum ad absurdum* has been a powerful rhetorical weapon. In the crusade against the rational choice model, an *argumentum ad absurdum* is frequently profferred . The power of this model rests on the ability to reduce social interaction to an exercise in arithmetic: Define the utility function for each individual who is part of the conflict. Create a common currency (e.g., money). Use this currency to quantify the initial endowments of all individuals. Translate all features of the environment into cost, again quantified by means of the common currency. Say exactly who knows what. If individuals move sequentially, define the timeline. Assume that all individuals have exclusively one goal, and that everyone realizes that all others have this goal: the maximization of their utility, given all constraints. Then, with this setup, you can calculate the equilibrium. Granted, the mathematical skills required for this were not taught in most secondary schools. However, after a few weeks of practice, a person should be able to derive the correct solution to standard optimization problems.

We all solve problems all day, don't we? Yet the percentage of humans able to solve even a simple optimization problem is far below 1%. Even to the most enthusiastic defenders of the rational choice model, it therefore seemed obvious that the model does not describe the underlying mental mechanism. Economists tend to say: it is only an “as if” model (Friedman 1953). Psychologists would use different terms for the same statement. The rational choice model is at best paramorphic, not isomorphic (Fischhoff et al. 1981). Standard reasoning has it that, in the black box of the human mind, some process is going on that one would not really want to understand, but that ultimately converges to yield the predictions of the rational choice model. Specifically, the defenders of the rational choice model claim that individuals ultimately adapt to restrictions, and to the forces of competition, in particular (Friedman 1953).

So goes the *argumentum ad absurdum*. The argument is powerful since it resonates with our experience. Most of us just cannot reach the answer via mathematics. Even those who have the prerequisite skills would agree that it takes more effort than a person can afford in daily activities. To our experience, cognitive resources are severely limited. Actually, the limitations begin before a person even learns mathematics. Most individuals have a memory span of seven plus/minus one item (Wechsler 1945). Even in simple problems of strategic interaction, this is not enough. Think back to Ariel Rubinstein’s game: Before you realize that the game does not have a solution in terms of pure strategy, you must have tried out all combinations for both players and memorized the results. To calculate each best response, you must compare two payoffs. Even if we ignore the information about structure, we need to compare the correct payoffs, or 2 × 8 items. We are only less aware of this second limitation since we are so dependent on external aids. We realize that this is going to be tricky business, and thus reach for paper and a pencil.

What if we have fallen prey to an illusion? Until Galileo, it seemed clear that the Earth was flat, and that the Sun and the Moon circled around it. Subsequently, science has convinced us that the Earth is a sphere that circles around the Sun. Will science also be able to convince us that the rational choice model pretty much captures what happens in our mind when we solve a problem of strategic interaction? Science is not there yet, and it is difficult to know whether it will arrive at exactly the mechanism assumed by the model. Still, it is much closer than most of us would find intuitive.

The main barrier to the formation of descriptively adequate intuitions about mental mechanisms is consciousness. We possess the ability to handle information consciously and to derive from this a plan for action. We are, however, misled to believe that this fully describes the underlying mental mechanism. Psychology has demonstrated that we are able to handle vast amounts of information in almost no time, non-consciously (Glöckner 2007). Neuroscience has demonstrated that we are able to record fine-grained information about positive and negative utility, and that we update this information in light of every single new experience, non-consciously (Glimcher 2003). With these ingredients, a mechanism pretty close to the assumptions of rational choice theory could be implemented.

Decisions on Incomplete Factual Basis

Even the most faithful believers in the rational choice model do not claim that it describes the underlying mental mechanisms of a chess player (de Groot 1965). Chess is a deterministic problem: the chessboard is exactly defined, as is the composition of the chessmen, their moves, and the conflict rules if two chessmen arrive at the same field. In a world without resource limits, one could thus calculate the optimal way to play a game of chess. More interestingly, one could calculate the optimal sequence of future moves, given that the opponent has moved in some known way in the past. For quite a few years, computers have succeeded in beating the world champion in chess. How do they accomplish this? Do they achieve success through brute force, making innumerable calculations? To the contrary, the computers employed heuristics, and the reason is straightforward. In a game of chess, there are so many degrees of freedom that combinatorics explode. Likewise even the stupendously powerful, non-conscious, parallel-processing mechanism in the mind is overwhelmed (Gigerenzer et al. 1999).

Many real-life problems are even worse. Often we have to decide on an issue when we do not have access to complete information. For example, in competition, we know that our competitor is working on a process innovation, but we do not know how likely she is to succeed. We also do not know how much this innovation would reduce her production costs, is she were to succeed. Should we also spend money to find a cheaper way of manufacturing the product? Or would we still make enough profit once our competitor has reduced her price (e.g., since her capacity remains limited)? In principle, there are three reasons why information may be incomplete: (a) we are attempting to know something that no one else knows, or that cannot be communicated in a credible way; (b) the price of generating or acquiring the information is prohibitive; and (c) some overriding legal or social norm prevents us from gaining access to the information.

When the factual basis is incomplete, we must decide under uncertainty. Some forms of uncertainty lend themselves to rational choice analysis, and hence to the calculation of optimal responses. For example, if uncertainty is stochastic, and we know the probability distribution, then we can calculate expected values. If we do not know the probability distribution, we can replace it by our beliefs (Savage 1954). If uncertainty is strategic in nature, and we have reason to believe that our opponents maximize utility, we can use game theory to base our decision. If the problem space is well defined, but we do not know the combination of features, we can introduce an artificial move of Nature and again use game theory (Harsanyi 1967–1968). However, many real-life problems are plagued by an element of pure ignorance (Kirzner 1994). The problem space is at least partly not defined. We are down to guesswork.

Interestingly, the non-conscious mental machinery is pretty good at that as well. It is prepared to produce decent responses to problems that are only partly defined. It is permanently engaged in making sense out of the available information and in creating meaning (Glöckner, this volume). To that end, it relies on deliberately fuzzy tools like exemplars and schemas. In an exemplar, the individual happens to know a good response to a graphically described problem. This may be an earlier good or bad experience, a good story about someone else's experiences, or a colorful counterfactual. The individual transposes the solution by way of analogy (Smith and Zárate 1992; Juslin et al. 2003). A schema is less tightly linked to a full story. In memory, the individual has stored a more or less rich description of a concept and matches the available information to this concept. In so doing, more weight is given to the presence, or the absence, of what the person sees to be the key factors (Bartlett 1932). All of this happens rapidly, and only the result is perceived consciously.

Insight

Every scientist has experienced insight: after struggling with a problem for hours, days, or even years, out of the blue, the way forward suddenly appears—fully convincing and often far simpler than expected. Lawyers are trained to achieve similar results through a method known as hermeneutics. Given an overly rich set of potentially relevant facts of the case as well as an overly rich set of legal rules, doctrinal lawyers progressively narrow down, through an iterative process, the interpretations of code as well as the interpretations of the facts, until they match. In legal jargon, this end result is called syllogism. In all but the most simple cases, there is a dose of creativity, or insight, in this hermeneutical exercise. All of a sudden, a doctrinal lawyer sees an appropriate selection of facts and an appropriate doctrinal twist, which results in a match. Afterwards, she is able, and obliged for that matter, to tell an explicit, fully consistent story. But this is only a technology for the representation of a decision that has been found otherwise. Specifically, the doctrinal lawyer is able to construct an explicit representation of a decision, but she does not have access to the mental process that has led her to find this result (Engel 2007).

Both scientists and doctrinal lawyers capitalize on insight (Bullbrook 1932). There is reason to believe that the mental process behind insight is closely related to the parallel constraint satisfaction mechanism that has already been presented in explaining how judges and juries assess the evidence. The key feature in the model, which also helps us understand insight, is attractor dynamics. The mind strives for generating meaning out of what initially looks like inconsistent evidence. It does so in deflating some while inflating other elements, until a qualitatively new, coherent picture emerges. The same holds when we search for a solution to what initially looks like an intractable problem. Not only is this process consciously not accessible; when subjects are induced to verbalize, their ability to find the solution deteriorates (Schooler 1993).

Norms

Performance

When is a decision "better than conscious"? The answer depends on the norm. The primary goal must be performance. Is the individual likely to perform better on the task at hand if she foregoes conscious deliberation and lets the non-conscious mental apparatus do all the work, or at least part of it (Hammond et al. 1987)? In deterministic problems, at least after the fact, performance can be assessed in an unequivocal way. If the problem is aggregating two apples and two oranges into four pieces of fruit, every answer that is not "four" is materially wrong. We may dispute whether this assessment follows from a violation of the axioms, or from a violation of social practice. But we have no doubt that being materially correct is the norm.

For the other two illustrations (i.e., guilt determination and competition), defining the norm is not so easy. Is the decision "guilty" below standard if, a few years later, DNA evidence proves that the defendant was not at the crime scene? The decision is certainly materially wrong. But is "truth" the appropriate standard? Ultimately, the legal order leaves this unanswered. In most legal orders, later DNA evidence on personal identity is sufficient to re-open proceedings. But as long as the new procedure is pending, the old judgment is still valid and, for example, it is possible to keep the defendant in prison until the former judgment has been formally repealed. Why is that? A radical improvement in the evidence, after the final ruling has been given, is a rare event. New, but less compelling evidence, however happens quite frequently. If this type of evidence sufficed to call earlier judgments into question, many judgments would be provisional, at best. To avoid this situation, the legal order is content with judgments that are "good enough," and procedural codes exist to define the necessary criteria. Typically, the standard has two elements. First, a final judgment must be based on evidence that, at least at face value, looks conclusive. Second, the requisite procedural safeguards for making the decision may not have been grossly violated.

Legal safety, the reliability of judgments, respect for the courts: these are all secondary goals. Yet behind each of them is the problem of defining the primary goal. Rare instances notwithstanding, for which DNA evidence is the best approximation, the legal system has no chance of ever fully establishing the truth. Instead, it must reach a decision based on an incomplete set of information. The legal system must face the very real chance of making materially wrong decisions. All it can attempt to do is to avoid patently wrong decisions. In reality, the problem goes deeper. In almost all legal cases, even after the fact, nobody will ever be able to say with certainty whether the decision has been materially right or wrong. The legal order has reacted by defining standards of proof. In criminal justice, the American legal order requires guilt to be established "beyond reasonable doubt." This obviously does not constitute certainty. It is sufficient if the judgment is based on evidence that makes guilt so highly probable that a reasonable person would exclude alternative explanations. Note the implication: if this standard was met in the original judgment, the conviction met the standard although the defendant was innocent.

Seemingly, with Ariel Rubinstein's game, one is back to determinism. Remember the game has a unique equilibrium, in mixed strategies, where the row player plays top with probability 1/3, and the column player plays left with probability 1/3. Yet this is a best response for the row player only if she believes the column player to maximize expected utility. This need not be true. The column player might be risk averse. In the interest of making sure that she is not left with her worst payoff, –2, she might always play right. From Figure 1 we already know that, if this is what the row player believes, she should play bottom, to secure her second best outcome, 1. Or the column player might be puzzled by the game, and flip a coin. If the coin is fair, this gives probability 1/2 to both left and right. If this is what the row player believes, her best response is not matching this, but playing top with certainty. In expected values, this gives her:

$$\frac{1}{2}(TL)+\frac{1}{2}(TR)=\frac{1}{2}(2)+\frac{1}{2}(0)=1.$$

Were she to match, she would only have:

$$\frac{1}{4}(TL)+\frac{1}{4}(TR)+\frac{1}{4}(BL)+\frac{1}{4}(BR)=\frac{1}{4}(2)+\frac{1}{4}(0)+\frac{1}{4}(0)+\frac{1}{4}(1)=\frac{3}{4}.$$

From this, one learns that only the calculation of the best response is fully determined, but that there is uncertainty about the behavior of the opponent. If the game is played only one time, and if the opponent is anonymous, the row player can at most rely on her world knowledge to base her selection. If she has read Ariel Rubinstein's paper, she might surmise that column players have the opposite propensity of row players. Since they are hit rather badly in the top left cell, they might, on average, be more inclined to play right. In that case, the row player is in a difficult situation. Let's indicate the subjective probability of the column player playing right by a, which we assume to be larger than 1/2. This implies the belief that she plays left with corresponding probability $(1 - a)$. If the row player plays bottom, she expects to earn $a \times 1 = a$. If she plays top, she expects $(1 - a)2$. She is indifferent between playing top and bottom if

$$a = (1 - a)2 \Rightarrow a = 2/3.$$

Actually this is exactly the probability distribution if the column player plays her best response. Making the other player indifferent is how a rational player determines probabilities in the mixed equilibrium. Consequently, if the row player maximizes utility, it is not enough for her to hypothesize that the column player would privilege right. She must believe her to privilege this action very strongly.

Rare events like frequency auctions notwithstanding, in reality people do not look up the literature on experimental games before they decide on strategic action. Instead the row player might rely on common sense to come up with the same hypothesis. Common sense basically works through the construction of a simple counterfactual: were I the column player, what would I do? Are others likely to be similar to me? To come up with these assessments, the row player would have to heavily rely on her intuition.

Secondary Goals

Until now, our analysis has ignored decision cost, which may be prohibitive. Imagine a two-person, simultaneous game where each player has to choose between four different actions. In reality, this is not a very rich setting; however, it means that a player must calculate best responses for 16 cells. Since a player must also know the best responses of the opponent, 32 results are needed. Now imagine that the game has no equilibrium in pure strategies. If a player starts calculating the mixed equilibrium over four strategies, she must solve a system of four equations in three unknowns. This will take some time and could potentially introduce a calculation error. The player may not have time or may not trust the calculations. Such a game is, of course, still manageable; however, if complexity is introduced, calculation becomes overwhelming. Recall chess. In statistical jargon, the game is NP-hard. No one could just calculate best responses.

The legal example of guilt determination epitomizes secondary goals that are, metaphorically speaking, measured in a different normative currency. Society is not content with judges merely reaching materially acceptable decisions. We want them to give reasons that the defendant can understand. We want them to demonstrate to the public at large that the legal system works properly. We want them to expose themselves to peer review by the upper courts and the wider legal community. We want them, as a side benefit of deciding the concrete case, to contribute to the evolution of the legal order (Engel 2001).

Audiences

Who should be interested in better understanding the power of non-conscious (elements in) decision making and problem solving? Psychologists, of course, since decision making and problem solving is their business. Neuroscientists as well, since searching for the neural correlates of non-conscious decision making provides them with a stimulating set of challenging research questions. But equally so, economists. Although economics has experienced a powerful behavioral revolution, this has only very tentatively evolved into a cognitive revolution. Economics, as a discipline, is not very attentive to mental mechanisms. This volume demonstrates that taking non-conscious decision making seriously would be very productive for the field.

Last, but not certainly least, non-conscious decision making and problem solving are issues of great importance for institutional analysis and design. The illustration given of guilt determination has served as one example. Relevance, however, is not confined to understanding the workings of the legal system. We must also gain a more adequate sense of those individual and social problems that institutions set out to address. To date, most institutional analysis and, in particular, most new institutional economics build on the assumption that institutional addressees decide deliberately. In many practically, and legally for that matter, relevant contexts, this is factually wrong. If the power of the non-conscious were properly taken into account, institutional designers would come up with different, more adequate interventions.

Acknowledgements

Helpful comments on an earlier version by Andreas Glöckner and Julia Lupp are gratefully acknowledged.

References

Bartlett, F. C. 1932. Remembering. A Study in Experimental and Social Psychology. Cambridge: Cambridge Univ. Press.

Baumol, W. J. 1982. Contestable markets. An uprising in the theory of industry structure. *Am. Econ. Rev.* 72:1–15.

Bolton, P., and M. Dewatripont. 2005. Contract Theory. Cambridge, MA: MIT Press.

Bullbrook, M. E. 1932. An experimental inquiry into the existence and nature of "insight." *Am. J. Psychol.* 44:409–453.

Chomsky, N. 1959. A review of B.F. Skinner's verbal behavior. *Language* 35:26–58.

Condry, K. F., E. Gramzow, and G. A. Cayton. 2003. Toddler Counting. Three-year-olds' Inferences about Addition and Subtraction.

de Groot, A. D. 1965. Thought and Choice in Chess. The Hague: Mouton.

Dehaene, S. 1997. The Number Sense. How the Mind Creates Mathematics. New York: Oxford Univ. Press.

Dehaene, S., J.-P. Changeux, L. Naccache, J. Sackur, and C. Sergent. 2006. Conscious, preconscious, and subliminal processing. A testable taxonomy. *Trends Cogn. Sci.* 10:204–211.

Engel, C. 2001. Die Grammatik des Rechts. In: Instrumente des Umweltschutzes im Wirkungsverbund, ed. H.-W. Rengeling and H. Hof, pp. 17–49. Baden-Baden: Nomos.

Engel, C. 2007. The Psychological Case for Obliging Judges to Write Reasons. In: The Impact of Court Procedure on the Psychology of Judical Decision Making, ed. C. Engel and F. Strack. Baden-Baden, Nomos: 71-109.

Engel, C., and E. U. Weber. 2006. The Impact of Institutions on the Decision How to Decide. MPI Collective Goods Preprint No. 2006/19. http://ssrn.com/abstract=935024.

Fischhoff, B., B. Goitein, and Z. Shapira. 1981. Subjected Expected Utility. A model of decision-making. *J. Am. Soc. Info. Sci.* 32:391–399.

Friedman, M. 1953. Essays in Positive Economics. Chicago: Univ. of Chicago Press.

Gigerenzer, G., P. M. Todd, and ABC Research Group. 1999. Simple Heuristics that Make Us Smart. New York: Oxford Univ. Press.

Glimcher, P. W. 2003. Decisions, Uncertainty, and the Brain. The Science of Neuroeconomics. Cambridge, MA: MIT Press.

Glöckner, A. 2007. Does intuition beat fast and frugal heuristics? A systematic empirical analysis. In: Intuition in Judgment and Decision Making, ed. H. Plessner, C. Betsch, and T. Betsch. Mahway, NJ: Lawrence Erlbaum, in press.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T., 1987. Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. IEEE Transactions on Systems, Man, & Cybernetics, 17(5), 753-770.

Harsanyi, J. 1967–1968. Games with incomplete information played by "Bayesian" players. *Management Sci.* 14:159–182, 320–224, 486–502.

Hastie, R. 1993. Inside the Juror. The Psychology of Juror Decision Making. Cambridge: Cambridge Univ. Press.

Hastie, R., S. Penrod, and N. Pennington. 1983. Inside the Jury. Cambridge, MA: Harvard Univ. Press.

Hauser, M. D., and E. S. Spelke. 2005. Evolutionary and developmental foundations of human knowledge: A case study of mathematics. In: The Cognitive Neurosciences III, ed. M. S. Gazzaniga, pp. 853–864. Cambridge, MA: MIT Press.

Juslin, P., H. Olsson, and A.-C. Olsson. 2003. Exemplar effects in categorization and multiple-cue judgment. *J. Exp. Psychol.: Gen.* 132:133–156.

Kagel, J. H., and A. E. Roth, Eds. 1995. The Handbook of Experimental Economics. Princeton, NJ: Princeton Univ. Press.

Kahneman, D., and A. Tversky. 2000. Choices, Values, and Frames. Cambridge: Russell Sage Foundation; Cambridge Univ. Press.

Kirzner, I. 1994. On the economics of time and ignorance. The market process. In: Essays in Contemporary Austrian Economics, ed. P. J. Boettke und D. L. Prychitko, pp. 38–44. Brookfield: Elgar.

Lipton, J. S., and E. S. Spelke. 2005. Preschool children master the logic of number word meanings. *Cognition* 98:57–66.

Matsuzawa, T.; Tomonaga, M.; Tanaka, M. (Eds.) 2006, Cognitive Development in Chimpanzees.

Nash, J. 1951. Non-cooperative games. *Annals Math.* 54:286–295.

Rubinstein, A. 2007. Instinctive and Cognitive Reasoning: A Study of Response Times. *Econ. J.* ***: ***.

Sauerland, M., and S. L. Sporer. 2007. Post-Decision Confidence, Decision Time, and Self-Reported Decision Processes as Postdictors of Identification Accuracy. *Psychology, Crime, & Law* ***: ***.

Savage, L. J. 1954. The Foundations of Statistics. New York: Wiley.

Schooler, J. W. 1993. Thoughts beyond words: When language overshadows insight. *J. Exp. Psychol.: Gen.* 122:166–183.

Schulz, J. 1992. Sachverhaltsfeststellung und Beweistheorie: Elemente einer Theorie strafprozessualer Sachverhaltsfeststellung. Köln: Heymanns.

Simon, D. 1998. A psychological model of judicial decision making. *Rutgers Law J.* 30:1–142.

Simon, D., L. B. Pham, Q. A. Le, and K. J. Holyoak. 2001. The emergence of coherence over the course of decision making. *J. Exp. Psychol.: Learn. Mem. Cogn.* 27:1250–1260.

Smith, E. R., and M. A. Zárate. 1992. Exemplar-based model of social judgement. *Psychol. Rev.* 99:3–21.

Sporer, S. L., R. S. Malpass, and G. Köhnken. 1996. Psychological Issues in Eyewitness Identification. Mahwah, NJ: Lawrence Erlbaum.

Straffin, P. D. 1993. Game Theory and Strategy. Washington, D.C.: Mathematical Assoc. of America.

Trick, L. M. and Z. W. Pylyshyn. 1994. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol. Rev.* 101:80–102.

Wechsler, D. 1945. A standardized memory test for clinical use. *J. Psychol.* 19:87–95.

Wynn, K. 1990. Children's understanding of counting. *Cognition* 36:155–193.

Wynn, K. 1992. Children's acquisition of the number words and the counting system. *Cogn. Psychol.* 24:220–251.