Decision behaviour – Improving expert judgement

Geir Kirkebøen

”Expert judgments have been worse than those of the simplest statistical models in virtually all domains that have been studied”
(Camerer and Johnson, 1991)

“In nearly every study of experts carried out within the judgment and decision-making approach, experience has been shown to be unrelated to the empirical accuracy of expert judgments”
(Hammond, 1996)

“Professionals are not stupid, but they are human”
(Bazerman, 2006)

In this chapter, common biases in professionals' judgement and decision-making, and how these deviations can be corrected, are explained. A principal reason for these decision biases is human beings cognitive limitations. Due to these limitations, simplifying strategies are often used, producing predictable biases. Emotional and motivational factors may also contribute to reduced decision quality. Experience does often not improve the quality of professionals' judgement and decision-making, because they do not receive accurate and timely feedback. No easy recipe for eliminating decision biases exists. However, simple cognitive strategies like "take an outsider's view", and "consider the opposite" are efficient in many circumstances. Research findings suggest a wide use of technological decision support within many professions. In particular, professionals' judgements in repeatable decision-making situations should be automated to a far greater extent than is the case today.

Introduction

People's judgements are not perfect. Even experts make mistakes. Interestingly, mistakes in judgement and decision-making are not random. Professionals deviate systematically from normative standards. Per definition, systematic biases result in reduced decision quality, and accordingly, may have serious consequences. However, the fact that judgement biases are systematic also gives hope. If the biases and their causes can be identified, it should be possible to correct them.

The typical biases in our decision behaviour are well described in the psychology of judgement and decision-making (henceforth JDM). This chapter first gives an overview of central findings and theories within JDM. Typical biases in judgement and decision-making are then presented. Finally, some acknowledged methods for improving professionals' judgements and decision-making are discussed.

Human decision-making – central findings

Decision-making is an essential part of all projects in life. Decisions often involve risk or uncertainty with respect to the outcome. A central topic within JDM is how individuals behave when faced with a risky choice. Historically, the descriptive study of decision behaviour is closely linked to the normative question of how we should behave. In their monumental work *Theory of games and economic behavior*, Von Neumann and Morgenstern (1947) put forward a very influential answer to this normative question. In 1947 a central assumption in economics had, for a long time, been that economic agents make rational decisions. But what does that mean? What is rational decision behaviour?

Von Neumann and Morgenstern put forward a mathematical theory as an answer to this question. The essence of their so-called Expected Utility (EU) theory is a small set of simple axioms. Von Neumann and Morgenstern prove that if you make choices in accordance with these basic axioms, then it is possible to attach a particular utility (or personal value) to each possible outcome of the choice, in such a way that one alternative is preferred to another, if, and only if, the expected utility (EU) of that alternative is higher than the expected utility of the other alternatives. The expected utility of each choice alternative under consideration is calculated by multiplying the (estimated) utility of each of the potential consequences of outcomes by their probabilities of occurrence, and then adding up all the component products.

The EU-theory is a model of decision under risk, i.e. the theory assumes that the probability for each outcome is objective or given. Savage's (1954) *Subjective Expected Utility* (SEU) theory extended expected utility from (objective) risk to subjective (or personal) probabilities of outcomes. Thus, SEU is the generalization of EU from risk to uncertainty in general.

Bounded rationality

The (S)EU theory was soon subject to debate. Is this or that axiom of the theory reasonable as a normative principle? Do people make their choices in accordance with the principle? Do people make decisions in accordance with what the SEU-theory as a whole prescribes?
According to the SEU theory, to make a rational choice requires one to first map out all the alternatives, find the possible outcomes of each of them, estimate the probability for each of these outcomes, make clear their consequences, estimate the utility of the consequences of each outcome relatively to each other, calculate the expected utility of each alternative, and finally select the alternative with the highest expected utility. Do people perform all this work when they make their decisions?

In the 1950s, Herbert Simon, a subsequent Nobel Prize winner in Economics, was the most pronounced critic of the assumption that the SEU theory gives a relevant description of human rationality. According to Simon, people rarely have access to all the information they need in order to make choices in the way the SEU theory prescribes. In addition, the ability to process information is usually far too limited to follow the SEU theory's prescriptions. Motivated by these insights, Simon (1955) launches the concept of bounded rationality. He argues that “the task is to replace the global rationality of economic man with a kind of rational behaviour that is compatible with the access to information and the computational capacities that are actually possessed [by humans]” (p. 99). A year later, Simon follows this up with another central idea in modern decision psychology. He claims, “However adaptive the behaviour of organisms in learning and choice situations, this adaptiveness falls far short of the ideal of maximizing in economic theory. Evidently, organisms adapt well enough to ‘satisfice’; they do not, in general, optimize” (1956, p. 129).

Simon’s bounded rationality perspective sets the agenda for later decision psychology, namely to uncover what particular “satisficing” strategies are made use of in different choice situations and how decision makers, as a consequence, deviate from normative models, such as the SEU theory.

Expert judgement versus formula

The first quotation in the chapter introduction summarizes one of the most spectacular findings in JDM research: Simple statistical models integrate information systematically better than intelligent and experienced human experts. Paul Meehl (1954) was the pioneer behind this discovery. He reviewed research comparing clinicians’ judgement or interpretation of data with simple regression equations integration of the same data. (Meehl drew a sharp distinction between collecting data and interpreting data.) The equations outperformed the clinicians in most of the 20 studies reviewed by Meehl. In none of the studies did the clinicians perform better than the equation.

One reason why Meehl’s finding was surprising is that his linear equations did not incorporate what clinicians believed characterize their own judgement, namely that the weight they put on a particular variable’s value depends on the values of the other variables involved. Still, the equations systematically outperformed or equalized with the clinicians’ “holistic” judgements.

In the comparisons undertaken by Meehl, the models and the clinicians had access to the same information. However, even if the human judge has access to more information than the equation, the equation often outperforms the judge. For example, it is common practice in the USA that a committee of experienced educational professionals is involved in the admission process to college. These committees normally have access to the student’s grades, a letter of recommendation from the student’s high school, and information on the quality of that particular high school. The committee also undertake a lengthy interview with each candidate. Dawes (1971) worked out a regression equation, based on registered data, that predicted how well students would perform in college. The only variables in that model were the student’s Grade Point Average (GPA), a letter of recommendation (scored 1 to 5) and the quality of the high school the student came from (also scored 1 to 5). Dawes applied the regression equation on 384 new applicants. He found that the model eliminated 55% of the weakest students, without eliminating any of those accepted by the committee. In addition, the model predicted more accurately than the committee how the students would perform at college. Dawes also found that only the students’ GPA was sufficient to outperform the committee’s predictions.

Since 1954, hundreds of studies have been conducted in many different professional domains, under the conditions set by Meehl for comparing expert judgement and equations. In hardly any study has the expert beaten the model. It is now a well established fact that simple models consistently outperform intelligent and experienced professionals' judgement (e.g. Dawes et al, 1989). Why?

Heuristic and biases

The prevalence of breast cancer is 1% for woman over the age of 40. A widely used test, mammography, gives a positive result in 10% of women without breast cancer, and in 80% of women with breast cancer. What is the probability that a woman in this age bracket who tests positive actually has breast cancer?

David Eddy (1982) asked experienced doctors this question. The majority of the doctors (95%) estimated the probability for breast cancer to be between 70 and 80%. The correct answer (according to Bayes' formula) is c. 7.5%. Why do highly educated and experienced doctors get the answer so wrong to a kind of question they routinely answer in their work?

Kahneman and Tversky's influential *heuristic and biases* perspective in JDM answers this question. The rules of probability prescribe the optimal way to reason with uncertainty. However, to follow these rules can easily become too complicated. It thus becomes necessary, in Simon's terminology, to "satisfice". According to Kahneman and Tversky, reasoning with uncertainty is ruled by a number of satisficing strategies, or "judgement heuristics". The doctors in Eddy's study were asked to estimate how probable it is that a woman who has a positive mammography test, also has breast cancer. Kahneman and Tversky demonstrated that, when asked to judge whether a person (event or object) belongs to a particular category, "representative heuristics" is often used. The answer is based on how typical this person is judged to be of the category in question. According to the heuristic and biases perspective, the doctors in Eddy's study confused the probability for breast cancer with how likely it is that a woman with breast cancer also has a positive mammography test (as is the case in 80% of woman with cancer).

The doctors in Eddy's study deviated strongly from Bayes' formula. This formula states that judgement of the probability of breast cancer when the mammography test is positive, has to take into account the relation between the base rates, i.e. the relation between the frequency of breast cancer and positive mammography respectively, among 40-year-old women. This relation is little less than one to ten, i.e. almost ten times more women have a positive test than the number of women who have breast cancer. When the relation among the base rates is uneven, the use of representative heuristics in judgement of category-belonging, results in systematic biases in respect to the rules of probability.

What percentage of the housework do you contribute to at home? Or, what percentage of the total work did you perform in the last project you participated in? When spouses or project participants are asked such questions separately, the total adds up to far more than 100%. The use of “availability heuristics”, another strategy that Tversky and Kahneman (1974) claim influences our probability judgement, can account for these phenomena. They suggest that judgement of how frequently something (X) takes place (relative to something else), is influenced by how easily available X is, i.e. how easy it is to notice, remember or imagine X. It is easier to register and remember one’s own work than the work of others. For several reasons, one’s own work is far more "available" in retrospect. It is therefore easy to overestimate how much work has been done personally.

The heuristic and biases perspective has been very influential, particularly after Kahneman and Tversky managed to incorporate this perspective on probability judgements into a descriptive alternative to the SEU theory.

Prospect theory

Imagine that the United States is preparing for the outbreak of an unusual Asian disease that is expected to kill 600 people. Two alternative programmes to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programmes are as follows:

If Programme A is adopted 200 people will be saved.
If Programme B is adopted, there is a one third probability that 600 people will be saved and a two-thirds probability that no people will be saved.

Which of the two programmes would you favour?

Tversky and Kahneman (1981) gave subjects this choice. Most chose Programme A (72%). They gave others the following option:
If Programme A(2) is adopted, 400 people will die.
If Programme B(2) is adopted, there is a one third probability that no one will die and a two-thirds probability that 600 people will die.

Most people (78%) then chose Programme B(2).

Close examination of the two sets of programmes shows that they are objectively the same. Programme A has the same consequence as Programme A(2), and Programme B has the same expected effect as Programme B(2). Nevertheless, most individuals chose programme A in the first set and programme B(2) in the second set. The subjects in this study were psychology students. The same study on professionals (doctors) gave similar results (McNeil et al 1982). How are these findings explained?

Before the Asian disease study, decision psychology had largely been ignored by economists. Simon’s bounded rationality perspective had not changed the economists’ core assumption that economic actors act rationally. The kinds of decision behaviour that Simon considered as evidence for bounded rationality were typically explained away as rational adaptations to the cost of searching for more information. The Asian disease findings were not so easily explained away. It is a central assumption in the SEU theory that preferences should not be influenced by irrelevant aspects, e.g. the way in which identical alternatives are described. This is an essential aspect of the concept of rationality in economic theory. Kahneman and Tversky's Asian disease study demonstrated beyond doubt that this assumption does not hold. Positive and negative presentations of objectively identical alternatives can have a dramatic impact on decision behaviour.

A couple of years before the Asian disease study, Kahneman and Tversky (1979) published a psychological theory on decision-making that predicted the findings of this study. Their *Prospect theory* specifies how decision-makers systematically deviate from the SEU theory. Many of the individual elements in the theory had been known a long time before they published the Prospect theory. Kahneman and Tversky's great contribution was that they were able to incorporate these findings in a precisely formulated theoretical alternative to the SEU theory. Kahneman described their strategy thus:

"The theory that we constructed was as conservative as possible … We did not challenge the philosophical analysis of choices in terms of beliefs and desires that underlies utility theory, nor did we question the normative [status of] models of rational choice … The goal … was to assemble the minimal set of modifications of expected utility theory that would provide a descriptive account." (Kahneman and Tversky, 2000)

The main reason why Kahneman and Tversky were "as conservative as possible" was that they considered economists to be their main target. It was important to maintain most of the established approach, so that economists could easily recognize their alternative theory. They did. The paper in which the Prospect theory was first described is the second most cited paper in economics (Laibson and Zeckhauser, 1998). In 2002 Kahneman got the Nobel Prize in Economics, mainly due to the Prospect theory.

The central components in the Prospect theory are the value function**,** the decision weight function, and the idea of a variable reference point that tells what the decision-maker (in choice situations) experiences as the division between losses and gains. The value function specifies how decision-makers typically assess objective values, or more precisely, changes in objective values. The value function has four major characteristics. Firstly, the theory incorporates an insight that goes back to David Bernoulli's (1738) *law of diminishing returns*, i.e. the idea that fixed increments in cash (objective value) lead to ever smaller increments of perceived wealth (or utility). Secondly, the value function describes a similar declining curve related to experienced losses. Thirdly, the function shows that losses are twice as painful as gains are pleasurable. Fourthly, the Prospect theory assumes that when decisions are made, gains and losses are experienced relative to a subjective reference point, which is usually the status quo. The theory also assumes that this individual reference point is easily changed.

The Prospect theory's assumption on a subjective and changeable reference point explains the findings in the Asian disease study. If the subjects accept Programme A in the first choice, 200 people are definitely saved. This outcome usually becomes the subjects' reference point. When they then consider Programme B, they will experience a risk of losing these 200 lives. In the second option, subjects typically frame the choice so that no death becomes the reference point. Decision-makers then experience that they may avoid the loss of 400 lives, by choosing alternative B(2). The value function suggests that losses feel twice as painful as gains feel pleasurable. This, together with the change in reference point, explains why subjects typically choose Programme A and Programme B(2) in the Asian disease study.

Thaler's (1980) discovery of the *endowment effect* was another early example on how people systematically deviate from the SEU theory in the way predicted by the Prospect theory. Thaler showed that the maximum amount people were willing to pay for particular goods was far less than the minimum amount they demand for selling the same goods. Buying price typically surpasses the selling price by a factor somewhat larger than two, exactly as Prospect theory's value function predicts.

Prospect theory is, like the SEU theory, a theory on choice under uncertainty. The SEU theory does not distinguish between probabilities, and how probabilities or uncertainties are experienced. Prospect theory's decision weight function specifies the systematic biases in people's (subjective) experience of objective uncertainty. The function suggests that people have a strong tendency to attribute far too much weight to small probabilities. In particular, they are very sensitive to the difference between impossibility and a tiny possibility. This sensitivity is named the *possibility effect*. There is usually a good fit between objective probabilities around 0.2 and experienced uncertainty. However, people are insensitive to differences in intermediate probabilities, with a tendency to underestimate probabilities in the interval 0.2 to 0.8. Moreover, they are very sensitive to changes from something almost certain, to completely certain, and vice versa. For example, when the objective probability of an outcome changes from certain (1.0) to 0.99, this has a far greater psychological impact than a similar change in probability from 0.7 to 0.69, a bias named the *certainty effect*.

Prospect theory's decision weight function predicts several typical deviations from what the SEU theory prescribes. For example, the possibility effect can account for people's willingness to participate in lotteries, a phenomenon that for a long time had been a paradox, when considered from a normative perspective on decision behaviour. This is one example of many that shows the relevance of the Prospect theory in the real world (e.g Camerer, 2001).

Emotions

In the 1980s it was quite obvious that the SEU theory was not able to account for descriptive decision-making. Prospect theory was one response to this observation, and Kahneman and Tversky were "as conservative as possible" when they constructed the theory. One consequence of their conservatism was that emotions, which had no explicit role in the SEU theory, were not incorporated into the Prospect theory either. An alternative strategy to explain deviations from the SEU theory was to supply the SEU framework with emotional variables. This was done by Loomes and Sugden (1982) in their *Regret theory*. They included expected or anticipated emotions (e.g., regret and rejoice) in the possible decision outcomes when explaining decision-makers' estimation of future expected utilities. They hoped to thereby account for people's systematic deviations from the SEU theory. Even if the Regret theory was not very successful, it is now beyond doubt that affective phenomena strongly influence our decision behaviour (e.g. Loewenstein and Lerner, 2003).

Much research evidence indicates that every stimulus evokes an affective evaluation (e.g. Zajonc, 1980). On meeting a person, viewing a house, or reading the first chapter of a book, the mind is immediately made up, as to whether one likes it or dislikes it. This first, affective (good/bad) reaction is swift, and is not always conscious. It may be difficult to explain, but the feeling is there, and it will often influence one's judgement of the person, the house, or the book. Another example: when the boss has decided to employ a new person, he will probably claim that, after serious examination, he has chosen the best qualified applicant. Psychological research shows that the person who finally gets the job is not necessarily the best qualified, but is very often one the boss immediately liked. Decisions to appoint new employees, as well as other decisions, are, to a large extent, influenced by affective feelings.

In a paper based on his Nobel lecture, Kahneman (2003) considers the idea of an *affect heuristic* (Slovic et al, 2002), i.e. that basic affective reactions replace more complex evaluations, as the third main heuristic (together with representative and availability heuristics) in judgement and decision-making. He also offers a new understanding of the process of heuristic reasoning, namely that, without noticing it, people simply replace a difficult question with an easier one. For example, when asked what the probability is that X belongs to the category Y (X could be a woman with a positive mammography test, and Y the category breast cancer) they instead answer the question: How representative is X (positive test) to Y (breast cancer)? This utilizes representative heuristic. In a similar way, the use of affect heuristic involves the unconscious replacement of a complicated question such as, "Who is best qualified?" with the far easier question, "Who do I like best?"

Intuition versus analysis

Quickly answer this question:

A bat and a ball cost $ 1.10 in total.
The bat costs $1 more than the ball.
How much does the ball cost?

It is now widely accepted that the use of heuristics is distinctive for *System 1* reasoning, one of two main kinds of reasoning characterized as associative, affective, largely automized**,** implicit (inaccessible to introspection) fast (information-processing in parallel) and requiring little cognitive effort. *System 2* reasoning, the other main kind of reasoning, is affectively neutral, analytic, deliberate, slow (serial information-processing), controlled, rule-governed, flexible, and requires large cognitive efforts. One main function of System 2 is to monitor the quality of fast System 1 responses.

By the way, what was the price of the ball above? If you came up with the wrong answer “10 cents”, you are not alone. Almost everyone reports an initial impulse to give this answer, and 50% of Princeton students actually replied ,“10 cents” (Kahneman, 2003). The high rate of errors in this easy problem illustrates how lightly the output of System 1 is monitored by System 2. Deliberate thought is often too demanding, and it is easier to trust the seemingly plausible answers spontaneously supplied by System 1. From this System 1/System 2 perspective, intuitive judgements may be said to correspond with judgements not modified by System 2.

In the examples above, intuition is linked to bad performance. However, intuitive thinking may also be forceful and precise. In “kind” learning environments (cf. the conditions necessary for learning from experience below), a high level of accurate, fast and effortless judgemental performance can be acquired through praxis (e.g. Hogarth, 2005). Grand masters in chess are one obvious example. The quality of their play in speed chess is very high, even if their thinking is largely automatized. Some studies even indicate that experienced decision-makers in specific professions perform better when they trust their intuition than when they enter into deliberate analyses (Klein, 1998).

Decision biases – overview

A decision bias can be defined as a way of thinking that contributes to a systematic deviation from rational or normative decision-making. Decision biases thus contribute per definition to reduced decision quality. The number of demonstrated decision biases is large, and rapidly growing. It is not possible to give a full overview here, but some common biases are described and classified below.

Biases in information processing

Examples of a few kinds of decision biases have already been given. The doctors in Eddy’s study (1982) systematically overlooked the relevant base rates when estimating the probability for breast cancer based on mammography test results. *Insensitivity to base rates* is one of many systematic biases in probability judgement (of category membership) due to the use of the representative heuristic. Another bias typically caused by representative thinking is *conjunction fallacy*, where the events or properties A & B are judged to be more probable than A alone. In a classic example (Tversky and Kahneman,1983) subjects were asked to look at a description of a person named Linda. They were then asked to range the most probable statement of, among others, “Linda is a bank teller”, and “Linda is a bank teller and active in the feminist movement.” The majority judged the second statement to be more probable than the first, because the description of Linda was representative of a typical feminist.

The use of availability heuristic also gives rise to predictable biases in probability judgement. Bazerman (2006) gives an example where subjects, very few of whom, when asked what they judged as the most probable cause of death in the USA between drug use, guns, traffic accidents, bad diet/physical inactivity and tobacco, came up with the correct answer. The correct answer lists the causes in the opposite order of that listed above. It is common to overestimate the frequency of the causes that are most easily noticed. How easy or difficult a phenomenon is to imagine or remember, often leads to biases in frequency judgement due to the use of availability heuristics

Availability and representative heuristics are strategies that simplify the processing of information when probabilities are estimated. Tversky and Kahneman (1974) named the third satisficing strategy “anchoring and adjustment”. Estimates are often made in two steps. The starting point is a more or less justified value (*the anchor*), based on easily accessible information. This anchor value is then adjusted to a more plausible estimate. Even anchors that have no relevance to the judgement being made can have a huge impact on the final estimate. In a classic study, Tversky and Kahneman asked subjects to estimate what percentage of the UN member states are African. Before the participants put forward their estimates, they were asked to throw a roulette ball. The roulette was manipulated so that it stopped on 65 for one group of subjects, and on 10 for the second group. The average answer given by the group who saw the roulette ball stopping on 65 was 45%; but 25% by the group where the ball stopped on 10. Even though the anchor in this case was completely irrelevant, it still strongly influenced the subjects’ final estimates. Anchoring is a very robust phenomenon that systematically influences professionals’ estimates too (Chapman and Johnson, 2002).

A decision bias that is often a contributory cause when projects are not finished on time, is the general tendency to overestimate the probability for conjunctive events, i.e., the probability that, for example, all the events A *and* B *and* C will take place, when the probability is known that each of them will take place separately. Similarly, there is a tendency to underestimate the probability for disjunctive events, i.e. the probability that A or B or C will take place. Therefore the probability that complex systems, like human bodies or nuclear plants, will break down is often underestimated. Even if the probability is tiny that each essential component in the system will break down, the probability that the whole system will break down can still be large, if many essential components are involved.

Preference reversals and biases associated with the presentation of data

Some examples have already been given of biases associated with how information is presented. In the Asian disease study, the subjects’ decision behaviour was strongly influenced by the way choice alternatives were described in terms of gains or losses. This is one example of *preference reversal*, namely, that alternative A is preferred to alternative B, when the alternatives are presented in one way, but that preferences change when A and B are presented in a different (although from a normative perspective, equivalent) way. Tversky et al (1988) demonstrated another common kind of preference reversal, between choice and matching. Participants were presented with the following scenario:

About 600 people are killed each year in Israel in traffic accidents. The ministry of transportation investigates various programmes to reduce the number of casualties. Consider the following two programmes, described in terms of yearly costs (in millions of dollars), and the number of casualties per year that is expected, following the implementation of each programme:

Programme	*Expected number of casualties*	*Cost*
X	*500*	*$55 M*
Y	*570*	*$12 M*

When subjects were asked to *choose* which programme they favoured, 67% favoured Programme X, which saved more lives, but at a higher cost per life saved, than Programme Y. Subjects in the "matching condition" were given the same description of the two programmes but with one of the numbers missing. They were asked to *match* the two programmes, i.e. to fill in the missing number so that the two programmes would be equally desirable. Only 4% of the subjects then favoured program X, i.e. filled in a number equal to or larger than $55 million when the cost of Programme X was missing.

Tversky et al proposed the *prominence hypothesis* to explain this finding. The hypothesis simply says that the more prominent attribute will weigh more heavily in choice than in matching. Subjects regard casualties as more important, or *prominent*, than cost. In making a choice, the most important goal is often the main consideration, while in matching, the two goals are given more equal consideration. The reason is probably that a choice will often have to be justified to others (or ourselves) later on, and when justifying choice the most prominent attribute is typically in focus. An overview of other preference reversals is given by Hsee et al (2004).

Motivational causes to decision biases

So far, mainly cognitive reasons have been given for the deviation from rational thinking. The tendency to overestimate personal contributions to projects participated in has been explained in purely cognitive terms. It is easier to notice and remember one's own work than that of other participants. However, not only is there a tendency to take a disproportionately large share of the credit for collective successes, there is also an inclination to accept too little responsibility for collective failures. It is harder to explain this last fact in purely cognitive terms.

One non-cognitive reason for systematic deviation from rational thinking is the consistent motivation to come to a desired conclusion (e.g. Kunda, 1990). Reasons are sought to support one's own views, and counter-arguments that support opposing conclusions are neglected. There is a tendency to evaluate ambiguous information in a way that is beneficial to one's own interests. Such *motivated reasoning* can explain both the tendency to claim responsibility for successes rather than failures, and the reasoning behind other *self serving biases*.

It may not be surprising that judgements are often influenced by self-interest. More surprising, and more problematic, is the tendency to ignore certain important information in which motivational or affective factors are apparently not involved.

Confirmation trap

"You will be given three numbers which conform to a simple rule that I have in mind. Your aim is to discover this rule by writing down sets of three numbers. After you have written down each set, I shall tell you whether your numbers conform to the rule or not. You should try to discover this rule by citing the minimum sets of numbers. When you feel highly confident that you have discovered [the rule], and not before, you are to write it down".

Wason (1960) gave these instructions to 29 college students, along with the sample set of numbers: [2, 4, 6]. What is the rule? Which sequence of three numbers would *you* like to test out? Think about it before proceeding.

The rule Wason had in mind was: "Three numbers in increasing order of magnitude." Only six subjects discovered the correct rule without first naming an incorrect one. The others suggested more complex rules than the correct one. Commonly proposed rules included "Numbers that go up by two", and "The difference between the first two numbers equals the difference between the last two numbers." Wason found that subjects tried to confirm the rule they assumed the experimenter had in mind, far more often than they tried to disconfirm the rule. However, to find the rule requires the accumulation of disconfirming, rather than confirming, evidence. Wason concluded that few had the attitude necessary to succeed: "a willingness to attempt to falsify hypotheses, and thus to test those intuitive ideas that so often carry the feeling of certitude".

So, do *you* have this attitude? Were the three numbers you would have tested out, an attempt to invalidate your own hypothesis on the rule in the experimenter's mind?

Wason's study illuminates a strong tendency everyone has to look for information that supports their beliefs. The belief that one should own a particular kind of car, employ a particular person, etc. leads to a strong tendency to seek out information that strengthens that belief. This tendency is known as the *confirmation bias*.

To look for information that can eventually invalidate what one believes in, is often the most effective way to test out both the weaknesses and the strengths of personal opinions. It is far more beneficial to hire a consultant company that acts as the devil's advocate, and comes up with the best reasons to give up or change project plans, than one that, for the most part, advocates existing ideas.

Overconfidence

In 1957, the cost of the Sydney Opera House was estimated at seven million dollars with a completion date in 1963. An adapted version of the planned building was finished in 1973, at a cost of 102 million dollars. This is a prime example of *planning fallacy*, i.e. the tendency to be over-optimistic with regard to when planned projects will be finished (Buehler et al, 2002). In a more mundane example, Buehler et al asked students in the final stage of their psychology study when they “realistically” expected to submit their theses. They were also asked when they would submit them if “everything went as poorly as it possibly could.” The students’ “realistic” predictions were overly optimistic. Only 30% of them finished the project by the predicted time. On average, the students took 55 days to complete their theses, 22 days longer than they had anticipated, and seven days longer than the average worst-case prediction.

The planning fallacy is an example of the of the *overconfidence bias*, i.e. the strong tendency to be more certain about one’s judgements and conclusions than one has reason to be: the *overconfidence bias.* Overconfidence depends on the difficulty of the judgemental task. Tasks resulting in correct answers of about 75% or less tend to produce overconfidence, whereas easier tasks tend to produce lack of confidence. For example, in a series of experiments conducted in the 1970s, Lichtenstein and Fischoff found that people were 65-70% confident on being right, when they were actually correct about 50% of the time (Hoffrage, 2004).

No bias in judgement and decision-making is more prevalent, and more potentially catastrophic, than overconfidence. Historians have for a long time emphasized overconfidence as a substantial cause of war. In wars since 1500AD, 50-75% of the attacking sides have lost. There is also a likelihood that the winning side in a war will find the victory to be more costly than expected before the war started. Moreover, if the belligerent parties’ judgement as to the chances of victory before beginning a war is summarized in percentage terms, the numbers usually add up to far more than hundred (Johnson 2004).

Hindsight bias

The fall of the Berlin Wall in 1989; the planes crashing into the World Trade Center on 11/9/2001: these are events that no-one could have failed to notice. But try to imagine the time before these events took place. How probable was it that the Wall would fall, and terrorists fly into the WTC? Can knowledge of past events be overlooked when giving an answer?

After an event has taken place, there is a strong tendency to overestimate the extent to which it could have been foreseen. This bias is called the *hindsight bias* (Fischhoff, 1975), and is often explained as an anchoring effect. Knowledge on an event easily becomes an anchor on how we believe that we judged the probability for that event in advance. Adjustments related to such anchors have a tendency to be insufficient. So, knowledge we receive after an event has taking place influence how we remember that we judged the probability of that event in advance.

An unfortunate consequence of the hindsight bias is that it reduces the possibility to learn from misjudgements. The bias detracts from mistakes made in the past, and thus contributes to overconfidence in future predictions.

Closely related to the hindsight bias is a bias known as the *curse of knowledge* (Camerer et al, 1989). In the business world it is important to be able to foresee the market. Imagine knowing exactly what information other participants in the market have. Camerer et al demonstrated in several experiments that the more information about the market that is acquired beyond that, the harder it is to predict accurately how these other actors will behave, even if with the knowledge that they do not have the additional information. Better informed agents are unable to ignore private information, even when it is in their interest to do so. More information is not always better. Once something is known, it is hard to imagine what it was like not to know it: hence the "curse" of knowledge.

Why do people reason in ways that lead to systematic biases?

So far, the focus has been on the decision biases resulting from satisficing strategies. However, from a bounded rationality perspective, such non-optimal strategies also have a positive side. They contribute to decision efficiency, and decision-making in a complex world often requires the use of such simplifying strategies. Many of the satisficing strategies used contribute to adequate rather than inadequate decisions, and the loss of decision quality will usually be counterbalanced by increased efficiency. Tversky and Kahneman (1974) warned, "In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors". For example, the use of representative heuristic is a good strategy when there is no access to base rate information, as is often the case.

Correspondingly, availability heuristic works well when the correlation is strong between availability and frequency. This heuristic was probably useful in a pre-modern world, when what was important in order to survive, was also easy to notice and to remember. In our modern world however, statistical information is often far more reliable than easily accessible personal experiences. Availability heuristic contributes to the strong tendency to put exaggerated weight on the latest piece of information. Instead of thinking in statistical terms, the inclination is to think "dramatically", in personal terms.

Seen from a bounded rationality perspective, ways of thinking that systematically lead to decision biases simultaneously contribute to decision efficiency. However, so far, the main focus within decision psychology has been on the biases, and not on the efficiency. One exception is Gigerenzer et al (1999), who, through simulations, have demonstrated that, under certain circumstances (large uncertainty, little knowledge etc.), some very simple heuristics (satisficing strategies) often give results as good as more optimal methods.

The other ways of thinking, mentioned above, that lead to systematic decision biases, also have positive functions. A marked tendency to seek confirmation, contributes to the strengthening and stabilizing of beliefs and opinions. Doubt is certainly not always a good thing. In many circumstances, it will be almost impossible, or at least not very practical, to go after information that contradicts one's beliefs. Moreover, many beliefs are hard to disconfirm.

It is well known that positive illusions and unrealistic optimism strengthen our health, creativity and performances in both physical and mental exercises (Taylor, 1989). Overconfidence as a positive illusion can undoubtedly have positive functions. A doctor who is overconfident in the treatment she provides will (because of placebo effects etc.) get better treatment results than a more realistic doctor. It is also "overconfident" editors and contributors who manage to publish books like this one. If we had had a realistic picture of our time schedules when we were asked to contribute to this book, then many of us would probably have answered, "No" …

How decision-making can be improved – some efficient methods

It is a common belief that experience improves the quality of judgements and decision-making. But this is far from the truth. In fact, "experience has been shown to be unrelated to the empirical accuracy of expert judgments" (Hammond, 1996).

Why does experience (often) not improve the quality of professionals' judgements? The short answer is that professional work is usually performed in 'wicked' learning environments, where the necessary conditions for learning from experience are not present (Hogarth, 2005). The first condition is that immediate, unambiguous and consistent feedback is offered where mistakes have been made. The second condition is that the content of the feedback gives a clear understanding of precisely what was done wrong. One reason why these two conditions are seldom present is that professionals usually perform their judgement and decision-making in probabilistic contexts, i.e. in circumstances where the same kind of judgement or choice can give different feedback. (In deterministic contexts, on the other hand, the same kind of judgement/choice will always elicit the same kind of feedback). This means that professionals can make the right (over time) decision, but still occasionally achieve negative outcome feedback (result). Conversely, wrong choices sometimes may lead to positive results.

Another reason why decisions do not improve with experience is that often there is no feedback on the decisions per se. Frequently the only feedback given is on the decision in conjunction with the actions that follow the decision. For example, imagine a psychiatrist who decides to commit a person to a mental hospital. If the patient has calmed down after a week, how will the psychiatrist evaluate his decision to commit her? What if, after a week, the patient has turned completely crazy? The psychiatrist will, in both cases, probably interpret positively the outcome feedback on his decision to commit the patient. So, it is well documented that experience is not sufficient to improve judgements and decisions. How then to improve decision behaviour?

Debiasing – some main strategies

In the JDM literature several methods are suggested to improve or *debias* judgement and decision-making. Larrick (2004) distinguishes between three main kinds of debiasing strategies: motivational, cognitive and technological. Motivational strategies are based on the critical assumption that people possess the necessary normative strategies ('cognitive capital'), and will use them when the benefits exceed the costs. There are two main ways to motivate decision-makers: incentives to reward decision quality, and accountability, i.e. demanding that decision-makers will have to explain and defend their decisions to others.

Research shows that incentives motivate people to put more effort into their decision-making. It is thus natural to believe that incentives also improve decision quality. However, there is hardly any evidence to show that incentives consistently improve mean decision performance. A recent review study on the effect of incentives, concludes, "there is no replicated study in which a theory of rational choice was rejected at low stakes in favour of a well-specified behavioural alternative, and accepted at high stakes" (Camerer and Hogarth, 1999).

The principal reason why incentives do not usually have any effect, is simply that people do not possess the cognitive capital necessary to make better decisions. For example, if good decisions require knowledge of Bayesian reasoning, incentives will be of no help if the subject has not heard of Bayes' formula. Incentives thus contribute to what Larrick (2004) names the 'lost pilot' effect: "I don't know where I'm going but I'm making good time".

Like monetary incentives, accountability motivates subjects to put more effort into their decision-making, but this does not necessarily lead to better decisions. In addition, accountability often evokes a strong need to look consistent to others. This can both improve and deteriorate the quality of decision-making. Accountability often leads to biases, because decisions are so easily adapted to suit the audience. Another problem is that accountability is likely to strengthen reliance on easily justified aspects of the decision.

Cognitive strategies to improve decision-making include: drawing people's attention to decision biases; training in normative decision rules, e.g. Bayes' formula; learning to re-formulate decision problems; the use of other special strategies. To make people aware of decision biases, e.g. by reading chapters like this one, seems to have little debiasing effect (Fischhoff,1982). Training on normative rules in logic and probability theory has, in certain circumstances, some effect (Larrick, 2004). To re-formulate decision problems can, in some cases, have a dramatic effect. For example, the doctors faced with Eddy's (1982) question on breast cancer formulated in frequency terms instead of probability terms, i.e. 10 out of 1000 women aged 40 have breast cancer, 8 out of 10 women with cancer test positively, 1 out of 10 women without cancer also get a positive test result. More than half of the doctors then answered correctly the question on the probability that a woman who has a positive test also has breast cancer (Gigerenzer,1996). However, such re-formulations require quite a lot on behalf of the decision-maker.

Contrary to motivational and cognitive strategies, 'technological' strategies (in Larrick's use of the term) involve improving decision-making by going beyond the individual decision-maker, either by incorporating a group process, or by using different kinds of technological means. A large number of methods have been developed, both to support decision-makers in different aspects of the decision-making process, and to replace the decision-maker for part of the decision process. From a JDM perspective, there is unanimity with regard to one characteristic of decision support systems: they are grossly under-used**.**

Before focusing on two different technological strategies, a couple of simple but efficient cognitive strategies are suggested below.

Take an outsider's view!

When planning a new project, a major concern is how long it will take and how much it will cost. Estimates are usually far too optimistic. Kahneman and Lovallo (1993) claim that one reason for this is the neglecting of past statistics, and the tendency to consider new projects as unique. An 'inside' view, is taken, rather than an 'outside' view.

In 1976, Kahneman was involved in a project designed to develop a curriculum for the study of JDM for high schools. He asked each group member to indicate their best estimate of the number of months that would be needed to finish the project. The estimates ranged from 18 to 30 months. Kahneman then asked one member of the group, a distinguished expert in curriculum development, "We are surely not the only team to have tried to develop a curriculum when none existed before. Please try to recall as many such cases as you can. Think of them as they were in a stage comparable to ours at present. How long did it take them … to complete their projects?" (Kahneman and Lovallo, 1993) The experienced curriculum developer estimated that 40% of such projects were eventually abandoned, and that no such curriculum, as far as he knew, had been completed in less than seven years. The team took eight years to complete the JDM curriculum project.

In this example, all the participants in the meeting, Kahneman and the curriculum expert included, spontaneously adopted an inside view of the problem. This is typical. The inside view is overwhelmingly preferred in intuitive forecasting. Kahneman encouraged the curriculum expert to adopt an outside view. This ignores the details of the case in hand, and instead, focuses on the statistics of a class of cases similar to the present one.

Even if there is massive evidence that the outsider makes better estimates and decisions than the insider, decision-makers tend to believe and act on the insider view. However, it is not uncommon to switch between an inside and an outside view. Estimates based on detailed knowledge of a particular project are usually thought to be quite accurate (inside view), but at the same time there may be an awareness that most similar projects usually take a much longer time (outside view). For example, someone who builds their own house may have a particular opinion on what the price of the finished house will be, but at the same time acknowledge that their own estimate is probably far too low. They then use the knowledge of other house builders who had underestimated the costs of their finished houses. In general, a way to proceed is to make an 'inside' estimation of a project's costs first, and then correct this estimate by taking an outside view, discovering the ratio of actual spending to planned costs across other similar projects. Finally, the inside view estimate should be multiplied by this ratio.

Consider the opposite!

'Consider the opposite' is another efficient strategy for improving decision-making, and is closely related to 'take an outsider's view.' The strategy consists of simply asking the question, "What are some of the reasons that my initial judgement might be wrong?" Just to ask oneself this question has shown to effectively reduce overconfidence, hindsight biases and anchoring effects (e.g. Larrick, 2004).

The strategy is efficient because it counters the strong tendency to seek for information that confirms one's own opinions. The consideration of alternative options directs the attention to contrary evidence that would not otherwise be considered. The information base for decision-making is thus extended and made more representative. In this way, the fundamental problem with associative System 1 thinking is counteracted, namely that too narrow a selection of information is taken into account.

Group decision-making

In 1906, the multi-scientist Francis Galton, a half cousin of Charles Darwin, visited a cattle show, where a huge bull was being exhibited. The visitors were invited to guess the weight of the bull. A prize was put up for the best bet. After the exhibition, Galton collected all the bets on the weight of the bull. He calculated the average guess to be 1197 pounds. The weight of the bull was, in fact, 1198 pounds. The bets of the cattle experts were all far off the precision of the average bet of the unskilled mob. This is an early example of an acknowledged strategy to improve judgement, by combining the independent estimates of each individual member of the group.

Another example of the efficiency of this strategy was shown in an analysis of the quality of advice received by participants in the TV show, "Who wants to be a millionaire?". In this show the participants are able to ask for advice. They can place a call to a person who, they believe, knows a lot about the topic, or they can poll the studio audience. The analysis of the advice showed that the "experts" offered the right answer approximately 65% of the time. The advice from the audience, which was the multiple choice option voted by the majority in the TV studio, chose the right answer 91% of the time (Surowiecki 2004).

In both of these examples, the best judgements are a result of combining the judgement of each individual member of a group. When groups reach a conclusion through deliberation, the quality of the conclusion is often far from impressive. Deliberation often produces a series of unfortunate results: group polarization, amplification of errors, ‘cascade’ effects etc. (e.g Sunstein, 2005).

Surowiecki emphasizes four requirements to create a judicious group. The first requirement is *diversity of opinion*: each person should have private information on the case discussed. The second is *independence*: people's opinions must not be determined by the opinions of those around them. The third requirement is *decentralization*: people must be able to specialize and draw on local knowledge. The fourth is *aggregation*: some ‘mechanism’ has to be used for turning private judgements into a collective decision.

The *Delphi technique,* an interactive method for obtaining forecasts from a panel of independent experts, satisfies these requirements. This is a structured group process, in which individuals are required to give numerical judgements over a number of rounds. After each round, a facilitator provides an anonymous summary of the forecasts from the previous round, together with the reasons the experts provided for their judgements. Participants are thus encouraged to revise their earlier answers in light of the replies of other members of the group. During this process, the range of the answers will usually decrease, and the group will converge towards the "correct" answer. The process comes to an end after a pre-defined stop criterion (e.g. number of rounds, stability of results etc.), and the final aggregate is taken as the process output.

The Delphi method was developed by the RAND Corporation in the 1950s, and has been applied extensively. In a recent evaluation, Rowe and Wright (2001) found that the method gave more accurate estimates than traditional groups in five of their studies, less in one of them, and equalised in two. Overall, they found that the Delphi method improved accuracy in 71% of the cases, and reduced it in 12%.

The idea that the average of several independent estimates is better than the individual best estimate, is the fundamental concept behind an increasingly important information-aggregation tool, known as *prediction markets*. These are “opinion markets”, and work in a similar way to commodity markets. Prediction markets channel inputs from all traders into a single dynamic stock price. Instead of determining the value of a particular good, a prediction market is used to determine the probability of a particular event occurring. The participants trade in contracts whose pay-off depends on unknown future events. The price of these contracts can be directly interpreted as a market-generated forecast of some unknown quantity. Prediction markets are extremely useful for estimating the market’s expectation of the probability of a particular event occurring. For example, they have yielded very accurate predictions for elections, and outperformed the major pollsters. At Hewlett-Packard, a prediction market produced more accurate forecasts of printer sales than the firm's internal processes. Google utilised prediction markets successfully to obtain better forecasting of what the company and its competitors were planning to do (e.g Wolfers and Zitzewitz, 2004).

Automate judgements

Simple linear models have been shown to systematically outperform experts' judgements, under certain circumstances. This would seem to indicate that professionals' judgements in repeatable decision-making situations should be automated. One reason for automation is to improve judgement and decision quality. Another good reason is to reduce costs. Dawes (1971) estimated that if his model had been used in the admission of students to college, not only would the normal admission quality have been better, but the USA public sector would also have saved 18 million dollars annually. Bazerman (2006) suggests that, adjusted for the current number of graduate-school applications and today's dollar, this number would exceed 500 million dollars.

However, even where a particular decision is automated, it is still necessary to depend on experienced professionals. A classic study by Einhorn (1972) highlights both the strength and weakness in professionals' judgemental abilities. Einhorn studied physicians who coded biopsies of patients with Hodgkin's disease, and then made an overall rating of severity. The individual ratings had no predictive power of the survival time of the patients, all of whom died of the disease. In contrast, a multiple regression analysis based on the nine biopsy characteristics scaled by the doctors succeeded, to some extent, in predicting how long the patients would live. The general point from this is that experts knew what information to consider, but that a linear model combines this information in a way that is superior to the global judgements of these very same experts.

Meehl (1954) compared clinicians' judgements with actuarial or '*proper*' equations, i.e. the regression weights (coefficients) in the equations were based on statistical techniques that optimize prediction. However, even when the models are '*improper*', this means that the variable weights are determined by non-optimal means (e.g. unity or random weights, by 'bootstrapping' etc.), the models' integration of data normally outperform professionals' judgements of the same data. For example, a thousand applicants have to be judged and ranked, based on the same kind of information. If a hundred of the cases are judged, and then a regression analysis is performed on these judgements, the resulting linear model (which is a bootstrapping model of those judgements) will judge the last nine hundred applicants with greater efficiency.

The overall conclusion is that if the variables relevant for making a particular judgement are known, the weights given to the variables in the linear equations are of less importance. How then should professionals in repeatable judgemental situations be replaced? Dawes and Corrigan's (1974) much-quoted recipe is, "the whole trick is to decide what variables to look at and then to know how to add".

Conclusion

People's decision-making behaviour deviates systematically from normative models. Professionals are no exception. Their decision-making shows the same kind of biases as lay people's decision- making, and simple models consistently outperform intelligent and experienced professionals' judgements. A principal reason for biases in judgement and decision-making is limited cognitive capacity, which often makes it necessary to use simplifying reasoning strategies. Motivational and emotional factors can also create biases.

The use of simplifying strategies that sometimes lead to decision biases, also contributes to decision efficiency. The loss of decision quality will often be counterbalanced by increased efficiency. One main problem is a lack of awareness of what kind of reasoning strategies have been applied, and how these may negatively influence decision-making. No distinction is made between situations where particular reasoning strategies are advantageous, and situations where the same strategies are potentially damaging. A better understanding of this division is a key to improved decision-making.

The quality of professionals' judgement and decision-making is rarely improved through experience, mainly because professionals do not receive accurate and timely feedback. No simple recipe for eliminating decision biases exists. However, simple cognitive strategies like 'take an outsider's view', and 'consider the opposite' are efficient in many circumstances. Findings in decision psychology suggest a far wider use of technological decision support within many professions. In particular, professionals' judgements in repeatable decision-making situations should be automated to a far greater extent than is the case today.

References

Bazerman M (2006). *Judgment in managerial decision making* (6th.ed.). Wiley: Hoboken, NJ.

Bernoulli D (1738/1954). Exposition of a new theory on the measurement of risk. *Econometrica* 22: 23-36.

Buehler R, Griffin D and Ross M (2002). Inside the planning fallacy: The causes and consequences of optimistic time predictions. In: T Gilovich, D Griffin, and D Kahneman (eds). *Heuristics and biases. The psychology of intuitive judgment* (pp.250-70). Cambridge University Press: Cambridge.

Camerer C (2001). Prospect theory in the wild: Evidence from the field. In: D Kahneman and A Tversky (eds). *Choices, Values, and Frames* (pp.288-300). Cambridge University Press: Cambridge.

Camerer C, Loewenstein G and Weber M (1989). The curse of knowledge in economic settings - an experimental-analysis. *Journal of political economy* 97 (5): 1232-1254.

Camerer C and Hogarth R (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty* 19 (1-3): 7-42.

Camerer C and Johnson E (1991). The process-performance paradox in expert judgment: How can experts know so much and predict so badly? In: K Anders Ericsson and J Smith (eds). *Toward a general theory of expertise*. (pp. 195-217). Cambridge University Press: Cambridge.

Chapman G and Johnson E (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. In: T Gilovich, D Griffin and D Kahneman (eds). *Heuristics and biases. The psychology of intuitive judgment.* (pp. 120-38). Cambridge University Press: Cambridge.

Dawes R (1971). A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making. *American Psychologist* 26: 180-88.

Dawes R and Corrigan B (1974). Linear Models in Decision Making. *Psychological Bulletin* 81: 95-106.

Dawes R. Faust D and Meehl P (1989). Clinical versus Actuarial judgment. *Science* 243: 1668-1673.

Eddy D (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In: D Kahneman, P Slovic and A Tversky (eds). *Judgment under uncertainty: Heuristics and biases*. (pp. 249–67). Cambridge University Press: Cambridge, England.

Einhorn H (1972). Expert measurement and mechanical combination. *Organizational behavior and human performance* 7:86-106.

Fischhoff B (1975). Hindsight foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance* 1: 288-99.

Fischhoff B (1982). Debiasing. In: D Kahneman, P Slovic and A Tversky (eds). *Judgment under uncertainty: Heuristics and biases*. (pp. 422-444). Cambridge University Press: Cambridge, England.

Gigerenzer G (1996). The psychology of good judgment: Frequency formats and simple algorithms. *Medical decision making* 16: 273-280.

Gigerenzer G, Todd P and the ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford University Press: New York.

Hammond K (1996). *Human judgment and social policy. Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford University Press: New York.

Hoffrage U (2004). Overconfidence. In: Pohl, Rüdiger F (ed). *Cognitive illusions*. (pp. 235-54). Psychology Press: New York.

Hogarth R (2005). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. In: T Betsch and S Haberstroh (eds). *The routines of decision making (*pp. 67-82). Lawrence Erlbaum Associates, Inc: Mahwah, NJ.

Hsee C, Zhang J and Chen J (2004). Internal and substanstive inconsistencies in decision making. In: D Koehler and N Harvey (eds). *Blackwell handbook of judgment and decision making*. (pp. 360-78). Blackwell: Malden, MA.

Johnson D (2004). *Overconfidence and War: the Havoc and Glory of Positive Illusions*. Harvard University Press: Cambridge, MA.

Kahneman D (2003). A perspective on judgment and choice. Mapping bounded rationality. *American psychologist* 58: 697-720.

Kahneman D and Tversky A (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47: 263-91.

Kahneman D and Tversky A (2000). *Choices, Values, and Frames*. Cambridge University Press: New York.

Kahneman D and Lovallo D (1993). Timid choices and bold forecasts - a cognitive perspective on risk-taking. *Management science* 39 (1): 17-31.

Klein G (1998). *Sources of power. How people make decisions*. MIT-press: Cambridge, MASS.

Kunda Z (1990). The case for motivated reasoning. *Psychological Bulletin* 108: 480-98.

Laibson D and Zeckhauser R (1998). Amos Tversky and the ascent of behavioural economics. *Journal of Risk and Uncertainty* 16: 7-47.

Larrick R (2004). Debiasing. In: D Koehler and N Harvey (eds). *Blackwell handbook of judgment and decision making*. (pp. 316-337). Blackwell: Malden, MA.

Loomes G and Sugden R (1982). Regret theory - an alternative theory of rational choice under uncertainty. *Economic journal* 92 (368): 805-824.

Loewenstein G and Lerner J (2003). The role of affect in decision making. In: R Davidson et al (eds). *Handbook of Affective Sciences* (pp. 619-42). Oxford University Press: Oxford New York.

McNeil B, Pauker S, Sox H jr and Tversky A (1982). On the elicitation of preferences for alternative. *New England Journal of Medicine* 306: 1259-1562.

Meehl P (1954). *Clinical vs. statistical prediction.* University of Minnesota Press: Minneapolis.

Rowe G and Wright G (2001). Expert opinions in forecasting: the role of the Delphi technique. In: J Armstrong (ed). *Principles of forecasting* (pp. 125-144). Kluwer: Boston .

Savage L (1954). *The foundations of statistics.* Wiley: New York.

Simon H (1955). A behavioral model of rational choice. *Quarterly Journal of Economics* 69: 99-118.

Simon H (1956). Rational choice and the structure of the environment. *Psychological Review* 63: 129-138.

Slovic P, Finucane M, Peters E and MacGregor D (2002). The affect heuristic. In: T Gilovich, D Griffin and D Kahneman (eds). *Heuristics and biases. The psychology of intuitive judgment* (pp. 397-420). Cambridge University Press: Cambridge.

Soll J and Klayman J (2004). Overconfidence in interval estimates. *Journal of experimental psychology-learning memory and cognition* 30 (2): 299-314.

Sunstein C (2005). Group judgments: Statistical means, deliberation, and information markets. *New York University Law Review* 80 (3): 962-1049.

Surowiecki J (2004). *The wisdom of crowds*. Doubleday: New York

Taylor S (1989). *Positive illusions.* Basic Books: New York

Thaler R (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and organization* 1: 39-60.

Tversky A and Kahneman D (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185: 1124-1131.

Tversky A and Kahneman D (1981). The framing of decisions and the psychology of choice. *Science* 211: 453-63.

Tversky A and Kahneman D (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgments. *Psychological Review* 90: 293-315.

Tversky A, Sattath S and Slovic P (1988). Contingent weighting in judgment and choice. *Psychological review* 95 (3): 371-384.

Von Neumann J and Morgenstern O (1947). *Theory of games and economic behavior*. Princeton University Press: Princeton, NJ.

Wason P (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12: 129-40.

Wolfers J and Zitzewitz E (2004). Prediction markets. *Journal of economic perspectives* 18 (2): 107-126.

Zajonc R (1980. Feeling and thinking: Preferences need no inferences. *American Psychologist* 35: h151-175.