About bias - A short summary of "Thinking, Fast and Slow"

Introduction

There are books that you forgets, others, of which you keep some memories throughout your life. And then, there are books about which you wish you remembered everything, at all time. To me, "Thinking, Fast and Slow"[1], by Daniel Kahneman, belongs to that latter category.

Have you ever been in a situation where two people present the very same idea, in the same manner, to the very same group, and yet meet very different level of success in convincing the group? Where some irrational decision was made based on insignificant last-minute detail while ignoring the big picture? Where people systematically overestimated the probability of an event based on their fears? I guess you have, and that you had a gut feeling of what was going on. But I can also guess that, if you tried to discuss the event with your friend or colleagues, you felt let down by your words and had a hard time conceptualising what had happened. You may have noticed already how language and vocabulary format the thought... and what you observed at the time can be boiled down to a limitation of our vocabulary.

This is precisely the issue that the book is trying to address: break down the intellectual biases that we all have, define them more precisely by referring to minimal experiments that were made to identify them, and give them names, so that anyone can discuss them with more ease.

"Thinking, Fast and Slow" to my mind, is thus a book of vocabulary, which aims at giving names to precise common factors at play in our biases, and as such is a reading I would recommend to anyone. It will help you understand the mechanics at play in decision making, and why they can commit systematic errors. More important, maybe, it will help you put names on those mechanics, and hopefully discuss them with more ease when being confronted to them when part of a group.

Reading through this book can be quite thorough, which is a pity when you think that so people -just like me- probably want to re-read it from time to time. This is the reason why I started writing the document at hand. The first motivation is quite personal: being able to recall myself all the key ideas from the book without having to read it through again. The second -and ultimately most important- motivation is to be able to share more easily the key concepts put forward by D.Kahneman with anyone, as I think they are worth knowing.

In a nutshell, this document thus aims at presenting in a condensed and readable manner the key ideas of "Thinking, Fast and Slow". As a rule of thumb, it will present them in the same order as they can be found in the book (this is primarily due to the writing process of this document). Though I trust you will enjoy it, here are several disclaimers:

- Though I will try my best not to transform the ideas from the book, you have to remain aware that what I present here is my own understanding of them, which may be different from the one originally intended by D.Kahneman. In case of doubt, do not refrain from using your esprit

critique to get tune your own representation of the concepts presented.

- As this document is only intended as a summary, it will leave out many ideas, and references, from the original book. Thus, if you liked my summary and want to know more or develop a deeper view on the matter, please go read "Thinking, Fast and Slow". It is definitely worth it.

[NDLR 15/12/20: Now that I've advanced the writting, I see that following the order of the book is not quite the best idea. What I should have done is create a map (or data base) of the ideas presented in the book (and the various experiments?) and use it to reorganise the book semantically]

That being said, let's get to it. Enjoy the read!

Part 1: Two systems of thought

A story with two characters

The idea of "Slow thinking" and "Fast thinking" is key in formalising thought processes. Though it is obviously a simplification of mechanisms that we do not yet comprehend, they are very handy in modelling what is going on in our brains when we think.

Take two different tasks. In the first one, you are shown a picture of someone's face, and have to determine whether that person is angry. In the other one, you are tasked with solving in your head a two-digits multiplication, say 17×12. When shown a picture in task 1, it is very likely that an answer popped in your head immediately. When asked to solve task 2, you probably had to stop for a second before even starting the computation, if you only tried it. We call "Fast Thinking" processes like the one at play in task 1: Fast Thinking is the continuous process that gives you quick first impressions, and processes immediate information. We call "Slow thinking" processes like the one at play it task two: the slow and deliberate reflection that takes place when the Fast Thinking is not sufficient and more resources are required to form an opinion.

You can see those two "systems" of thinking as two characters that live in your brain and determine how you will process information. "Fast Thinking" is... fast. It is easy and cheap for Fast Thinking to form an opinion on anything, based on previous experience, and it'll always be the first to come up with an answer to a question. Slow thinking, in the other hand, is much less reactive. It is actually "lazy", and only comes in action when it is explicitly called by system one to solve complex tasks or focus on something specific. This happens when no quick answer can be brought to a question. Moreover, Slow Thinking is very limited and can only process a limited amount of information at the same time. It easily gets distracted. Think of doing the multiplication above while juggling. You'll probably have a hard time.

Those two characters are of course fictional, and encompass for a lot of different processes, but you'll see through the next chapters that they allow to model many empirical phenomenons: when it comes to understanding intellectual biases, the main difficulty comes from the fact that we tend to identify to our "Slow thinking" (very self-aware) while it is actually mainly our "Fast Thinking" that rules our perceptions.

Attention and Effort

Let's explore some key characteristics of Slow Thinking.

- Slow Thinking is fairly easy to recognise. This is in parts what allowed to study it: performing a difficult mental task requires a lot of efforts from the organism, and doesn't go without material symptoms: when computing the multiplication from the previous chapters, your pupil dilated, and your heartbeat accelerated slightly (or at least it would have if the effort had been long enough).

- Slow thinking requires a lot of efforts, but learns quite quickly to minimise it when a task is repeated: you can fairly imagine that performing a two-digits multiplication requires less efforts from an engineering student than from someone who doesn't use them often. A task that is made very regularly will eventually even resort to Fast Thinking rather that to Slow Thinking: a chess master will be able to see good moves "from instinct" without any effort.

- Slow Thinking can only retain a certain amount of information at the time, and be attentive to a limited number of event. Try doing a to digits multiplication while reciting a poem you just learned, you'll see that you quickly reach its limits.

The lazy supervisor

In addition to limited spot capacity (eg: you can't do a two digits multiplication while driving in Paris), Slow thinking (or more precisely any explicit mental effort) is very tiring, and can "exhaust" itself: you'll be more irritable of subject to limited attention after you made a strong mental effort (a phenomenon referred to as "exhaustion of ego"). Research suggests this is due to high consumption of glucose in the brain during mental effort (empirical evidence: consuming glucose strongly limits the exhaustion). This has strong empirical repercussion (eg: judges tend to be harsher after a long session of work than right after a break).

Slow Thinking also has the role of verifying the intuitive answers from Fast Thinking: quite often the first answer that comes to mind is wrong, and an easy check can show it. But this check is not always performed: depending on how "lazy" Slow Thinking is, the answer might never be checked. Lack of verification from Slow Thinking is empirically linked to lack of self-control (facing a temptation etc) and in the long run to generic intelligence and general patience tests. Working on verification tasks also improve metrics on the others. Some authors further suggest that Slow thinking (remember: Slow Thinking is only a conceptual tool) can be split in two parts: "computing" (performing difficult reasoning) and "rationality" (tendency to always check intuitions from Fast Thinking). We tend to associate intelligence only to computing capacities, but rationality is actually what assesses the tendency to resist cognitive biases.

NDLR: following this last remark, I tend to think that "Slow Thinking" is better understood as an encompassing concept for all "lazy" and "demanding" mental processes, rather than as the collection presented by the book ("computing power" + "checking first intuition" + "focus")... I hope that a more comprehensive/satisfying understanding has been or will be derived. (This reflection is probably part of what is meant in the last part of Chapter 1 from the book, though).

The Association Machine

About association of ideas: the simple fact of reading some words (limited simulation) can provoke a large cognitive - and physical - reaction (by reading a "disgusting" word, your body will show symptoms

of disgust; you brain will start proactively to look for other ideas that explain the presence of that word). The name proposed for that phenomenon is "associative activation". People usually model the structure of ideas as a graph, where edges as possible associations between ideas (eg: resemblance, causality, contiguity, properties etc). A classical model was to think see attention as a walker on that graph, going from on idea to the other etc. This is now seen as too simplistic: from one idea, a lot of underlying (unconscious) processes walk on the graph, and the ideas they cross are "randomly" brought to our attention. [NDLR: I would be curious to see if there is some in depth (eg: at least experimental) studies of that model]

A quite interesting aspect of association of ideas is the priming effect ("effet d'amorçage" in french): your interpretation of an information depends greatly on the ideas you had instants before, and on parallel processes in your brain. As an example, when presented the incomplete word SO_P, you would more likely to complete it into "SOAP" if you thought about taking a shower before, and into "SOUP" if you are hungry. This priming effect also have physical, and unconscious consequences: when presented words linked with old age ("grey", "Florida" etc), participants to an experiment tended to walk more slowly than those confronted to words linked to youth. More interestingly, participants acknowledged that they were totally unaware of that tendency.

What's more, the priming effect can have strong - and unconscious - social consequences, that one might want to be aware of: studies showed that people voting in schools were more likely to support education-friendly policies than those voting in other buildings; people "primed" with money were more likely to take selfish decisions etc... These results are very disturbing as they contradict our intuition that we are "aware" of our decision making.

Cognitive Ease

Call "Cognitive Ease" the state of mind when no ambiguous or hard to process information is encountered, and all tasks are processed by Fast Thinking. Factors creating cognitive ease include the familiarity, clarity, priming etc. Re-reading a text you know, listening to a clear speaker can provoke it, or smiling while doing a task favour cognitive ease. Cognitive ease will in turn affect your processing of information: since Slow Thinking checks are "disabled", you're more likely to like what you do, believe what you are told, be more creative etc. In the other hand, when not "cognitively at ease", you'll be more vigilant, commit less mistakes and feel less at ease. [NDLR: I feel like this last idea "se mord la queue": the consequences of cognitive ease are very close to its causes. Does this mean that there is a positive retro-action loop, or that we failed to conceptualise the dynamics and simply rephrase the description of that state of mind?]

Cognitive Ease can provoke what is called "the illusion of souvenir": at word or idea with which you fell "cognitively at ease" (eg: you've been primed to it) will sound more familiar to you. Think of the following experiment: first, you see a list of random names. Then, you are presented a list of names -

some famous, others not -. When asked to spot the famous names within that list, you'll probably also select the names you saw earlier, even if you are not able to say where you have seen them: as you've already encountered them, you'll feel more at ease with them, and associate this easiness to them being good answers to the question. [NDLR: this example is very specific and its hard to really conceptualise what happens - it could simply be a bias to think that familiar names/ideas are famous ones]

Not only do ideas encountered in the past resonate with more clarity when met again, but in turn we also tend to associate the clarity with which we can conceptualise an idea as an indicator for its veracity. This results in what is called "the illusion of truth": if an idea (let's say "chicken's corporal temperature") is familiar to you (for example if it has been presented to you repeatedly), you'll tend to believe more easily facts related to that idea ("chicken's corporal temperature is lower than human's"), only because it will ring a bell when it appears to you. This is an idea often used in politics and marketing: repeatedly bringing up a subject unconsciously makes people more inclined to believe messages related to it.

"Illusion of truth" phenomenons can have other sources (NDLR: which might sound obvious, but its good to know that they were actually studied): for example, you'll be more likely to believe a message if its written in a readable font, or if the name of its author can be pronounced easily. Finally, your mood also has a very strong implication on your cognitive ease (good mood => ease). This doesn't mean that your appreciation of veracity is not deeply rooted in logic or previous knowledge, but that the feeling of cognitive ease that those arise can also be provoked by other parasite factors.

Simple exposition, even unconscious, to something is enough to create cognitive ease with the exposed object, and the phenomenon was also shown in animals. Biologically, this can be explained by the fact that a familiar stimulus (even unconscious) is intuitively linked to a feeling of security.

Experimenting with cognitive ease can also go the other way around: when asked to do a questionnaire, people are less incline to fall for cheap tricks if the questionnaire is hardly readable than if it is, because their lack of ease to read questions will trigger Slow Thinking.

Norms, Surprises and Causes

[NDLR: there are not many concepts in this section, much more about experiments, I'll do my best] "How many animals of each specie did Moses take with him on the Arch?" You likely remembered very quickly the answer to that question, but did you spot the issue within it? (the Arch is Noa's, not Moses'). This illusion is likely provoked by the fact that both names are biblical and similar in length, so Moses' name did not sound awkward in that context (had you been asked about Freddy Mercury, you'd probably have noticed the trick much more easily). The human brain is quite good at picking anomalies, but any kind of similarity between a could-be anomaly and a known situation will drive the brain to find an explanation for the anomaly rather than spot it.

Such an automatic and unconscious work of the brain is also at play when we assess intentions: in

a famous experiment, Albert Michotte made a film where drawings of triangles and circles move around a schematic house. Despite these geometric figures being obviously deprived of intention, most people find intent in the way they behave. This tendency to find motives also generalises to sentiments of causality: our brains unconsciously find motives or causality (ie to create a story) in most situation, even when there is none.

A Machine for Hasty Conclusions

When interpreting a proposition (or a stimulus more generally), Fast Thinking will tend to look in recent memories (and then older one) to look for related content that will help the interpretation. This is useful as it helps evacuate ambiguity in certain situation - think of double meaning sentences: "take the leap" means different things depending if you are about to quit your job or if you are base-jumping -. What is worrying is that in most cases we remain unaware of the ambiguity, because Fast Thinking handled things unconsciously.

Fast Thinking will always try to find related ideas that will make sense of a new stimulus, and this can also be quite dangerous: it has an empirical "Confirmation Bias", eg a bias towards evaluating as "true" the propositions that it is presented. This is all the more the case when Slow Thinking is occupied by another task or tired (take this example: when asked to retain a list of numbers while presented some propositions, the confirmation bias will be reinforced and people will tend to "accept" any proposition that is proposed to them).

"The Halo Effect" - this is a concept that was named in psychology decades ago, that you likely observe every day, but whose name hasn't percolated in the common vocabulary. I'll start by describing it with an example from the book to be more clear. Take two men named Fred and Georges. Fred is intelligent, hard working, impulsive, jealous. Georges, in the other hand, is jealous, impulsive, hard working and intelligent. Even though you must have noticed the similarities in Fred and Georges description (they are the very same description), the intuitive images you have of them are probably very different - and Fred made a better impression. The first element you learned about each ("intelligent" or "jealous") served as a context used for Fast Thinking to interpret the description, and subsequent words had a better connotation when they regarded Fred. This is known as the "Halo Effect": the first impression you have on someone will strongly bias all subsequent information you have about them, and more importantly "fill the gap" for any information that you DON'T have about them. It talked here about individuals, but this is of course also true about ideas etc.

Close to the Halo Effect is the "Framing Effect": your opinion on a subject depends on the frame in which it was presented to you (this might sound obvious nowadays), even if the facts are the same. You are more likely to say yes to an operation that "has a 90% survival rate", than to one that has "a10%

mortality rate. [NDLR: I'm not quite sure how this relates to the rest of the chapter, but I respected the author's placement. Maybe move that to another chapter]

Another key concept introduced in the book is the one of "What You See Is All There Is" (WYSIATI). At this point you may have begun to get an intuition of it: as Fast Thinking has a tendency to jump to conclusions and fill lacking or ambiguous information using the context at hand, it often prevents you from becoming aware of missing information and instead bases its judgement only on the little information at hand. What matters to it is not the quantity of information you have, but the coherence of that information, and it will interpret things so as to maximise that coherence. A striking consequence of that is that people with little - but coherent - information on a subject will have stronger opinions on the subject than those with superior amount of information: if the information is lacking, those who don't have it will assume it is in line with what they previously know. Khaneman points out three specific consequences of WYSIATI:

- Sufficiency: when people have little but coherent information on something, they won't check whether they lack important parts of the information - they will unconsciously assume it from what they know.

- Neglecting empirical rates: "What is the probability for a very smart kid from your preschool to go to the Ivy League?" when thinking of the answer, but you probably never asked yourself what was the real empirical rate of kids going to the Ivy League: when told a coherent story, people will often forget compare their answer to the general case.

[NDLR: I wonder if the results presented could not have a relevant interpretation with information theory and Bayes rule: when lacking information, people will assume the prior that maximises the likelihood of the observed information, instead of choosing the most likely posterior - Now that I write it, I remember that it is mentioned somewhere else in the book -. In the same principle, the neglection of the empirical rates might be linked to lack of conditioned data: if I'm not wondering about the empirical Ivy League rate, it is because I know that I would need the "rate conditioned to being a kid from my preschool" (which I probably don't have due to lack of samples), and not the generic one. In a nutshell, maybe WYSIATI is optimal with respect to some information criterion.]

How judgements work

In the general situation, a new information will be processed by Fast Thinking alone, which will call Slow Thinking for help only if needed. In the case where a judgement must be emitted, as Fast Thinking is looking to provide an answer whatever the case, it will have a tendency to simplify the question at hand if it can help it provide an answer. An example of that can be found at the frontier between psychology and politics. You might have noticed that Fast Thinking is very good at intuiting intent and temperament

from the picture of a face alone (it is sometimes mistaken, of course, but still). This "judgement" (eg "is he friendly?", etc) is often substituted to other questions. In politics, studied showed a high correlation between the impression made by a candidate's face and his scores in elections, particularly for poorly-informed voters. For them, the question "does he look like a guy I can trust?" likely substituted to "Would he make a good elected-person?"

Psychologists studied a little more deeply how Fast Thinking worked: it is very good at comparing features, deriving means and clustering objects (even when Slow Thinking is occupied by another task). This is why it is easy for you to categorise a person, an idea etc. It is also very easy for it to perceive intensities, and establish equivalences between these intensities (between a smell and a sound, etc).
But it is very poor at deriving sums (experience to reproduce: draw several random segments on a blank sheet of paper; you'll fair rather well at drawing a segment of "average" length; but rather badly (and it'll cost you more efforts) at drawing a segment as long as the sum of the lengths). It is also ignorant of numbers: in certain situations it'll give the same answer regardless the quantities presented to it (people are willing to pay the same price to save 2000 birds as they'll pay to save 20000).

Fast Thinking also has a "buckshot" tendency: when asked presented a task, it cannot avoid to preform neighbouring tasks if they are easy for him. The results from those auxiliary - an often unconscious - tasks will be taken into account when processing the original one. As an example, if presented a task relying only on the pronunciation of some words, people can't avoid being influenced by the orthography of words: Fast Thinking cannot prevent itself from running orthography comparison and include it in its judgement.

Answering a Simple Question

Fast Thinking is incredibly good at providing simple answers to questions that are objectively tricky. How can form very quickly an opinion on whether you can trust a person, or on the economic perspectives of a company, while you have only a limited information at hand. This is due to what is called the "Substitution Effect". It is a combination of the buckshot effect and the easiness of equivalence of intensities: when asked a difficult question (eg "What are the chances that Mr.A will become president next year?"), Fast Thinking will unconsciously evaluate other questions (eg "Do I find that Mr.A is a skilled politician?") - this is the buckshot effect -. And if an answer to one of those questions is found easily, it will convert the "intensity" of your answer (eg "very skilled" to "very likely") to make it fit the original question. Of course Slow Thinking should be checking that answer, but as it is "lazy" it often does not. The Substitution effect is a simple heuristic that explains how you can easily form opinions on difficult questions. The same idea is at play in some optical illusions, where the context raised by the drawing tricks you into answering the wrong question.

Another funny example of the buckshot effect is the following: If you ask people how happy they are, and then how many dates they had in the last month, the two answer won't be correlated. But if you flip the questions (asking first about dates, then about happiness), then a strong correlation appears. You can replace the question about dates with one about their economic situation, friendships etc, the result will be the same. This is likely because having thought recently about dates makes that answer easy to access for Fat Thinking, triggering a substitution: the question about happiness is replaced with the easier one about dates.

In a general manner, your feelings about a subject strongly affects your judgement and opinions about that subject. Slow Thinking, which is usually tasked with checking first impressions from Fast Thinking, seems to be all the more negligent if the proposed answer if strongly emotional (it is more likely to forget to check it. [NDLR: what is that paragraph doing here? This seems to be directly related to the Halo Effect]

Part 2: Key Cognitive Biases

The Law of Small Numbers

Most people know, even imperfectly, a theorem known as the Law of Large Numbers (LLN). In a nutshell, it states that when a random experiment is repeated several times, its empirical average result will come closer to the true expectation for that experiment, and that the more time the experiment is repeated, the closer you can expect the empirical average to be from the expectation. But very few people think of the LLN the other way around: if you repeat an experiment only a small number of times, you are more likely to get an average result that is far from the real expectation. In practice, this simply means that you are more likely to observe extreme results on experiments made on small samples. This is what makes pools on small populations unreliable, for example.
The issue is that we actually overlook that simple fact most of the time. If I tell you that the places with the smallest cancer rates are all small villages, your first reflex will be to look for an logical explanation, not to interpret those as statistical artefacts. This is the result of Fast Thinking trying whatsoever to find a story behind the facts, and thus to "justify" random results with logical arguments rather than seeing them as they are.

Most likely, this fact seems obvious to you. Now try to summarise in three words "a pool conducted over phone over 300 elderly people resulted in 60% of them saying they supported the president". The words you chose are likely close to "elderly support president". After reading the previous paragraph you probably thought about the significance of the pool, but the summary wouldn't have changed if the sample was 200 people, or 1000, or conducted over the internet, which might change greatly the reli-

ability of the pool. Your brain is much more focused on the content of the message than on its reliability.

Unfortunately this negligence is found everywhere: the author state that researchers in psychology often don't compute the necessary sample sizes for their experiments, and that as a result of this numerous psychology study have more than 50% chances of failing to detect the expected result.

Anchor Effect

The example illustrating the Anchor Effect in the book is really efficient so I'm gonna restate it (simplified) here: in an experiment, two groups of students were presented a wheel ranging from 1 to 100. After a turn of the wheel (yielding a "random" number X between 1 and 100), they where asked "Is the proportion of African countries at United Nations (UN) greater than X? What is this proportion?". The proposed number X obviously yielded no information on the proportion of African countries in the UN; the students knew it was random. Yet a very strong correlation was observed between that first number and students estimations. This is known as the Anchor Effect: when evaluating something, the first proposition (the "anchor") that comes to your mind strongly affects all subsequent evaluation. It can also be observed in housing prices etc. There are two different mechanisms that can provoke it:

The first phenomenon at play is that Slow Thinking will first check the answer, and if it doesn't seem right, modify this value according to the possible biases it has detected. In doing so, it think of reasonable reasons to move the value and apply them one by one, but will be conservative in its approach, and thus probably stop in between the anchor and the value it would have thought of had it been independently.
The second explanation is that when presented the anchor, Fast Thinking, which always tries to make sense of the information presented to it, will think of bring up (unconsciously) reasons why the anchor might be true. Even if the anchor is absurd, these reasons brought up by Fast Thinking will prime it towards an answer that is closer to the anchor.

Examples of the Anchor Effect can be found everywhere in daily life, from evaluating the height of a tree to proposing a price for a house. It also affects experts in their fields, but is rarely acknowledged: when people are asked to detail their answers, they rarely mention the anchor. [NDLR: the author proposes here a way to measure the Anchor Effect, but I chose not to mention -I am not quite sure how relevant it is-]

The Science of Availability

How do people evaluate a probability? One heuristic is to assess the easiness with which examples of the phenomenon at stake can be found - their availability -. This method is referred to as the "Heuristic of availability". You don't have to recall explicitly the examples - the feeling that you could find some easily is enough.

Other factors that are used in the heuristic (and bias it) are: striking occurrence of events (politics, media), dramatic occurrence of events, personal events. It is hard to get rid of those biases: ideally you should take them one by one and validate they are not at play in your evaluation. Since personal examples bias the evaluation, we can explain why people tend to take credit for collective actions, and why the sum of people's evaluation of their contributions to a project is often above 100%: each one evaluates his/her contribution based on positive examples of what they did. Acknowledging those biases can thus help projects and communities be at peace.

It's the easiness that matter: people asked to retrieve 12 examples (quite a lot) before answering lower the probability compared to those that had to recall 6 (they had less examples, but it has easier for them to get them). Smiling (reinforcing the feeling of easiness) leads to higher evaluated probability. Practical examples include: - people are less sure of a choice after finding arguments for it; - people are less sure that an event could be avoided after asked how it could have; - people are less impressed by a car after having been asked to list its advantages. If the easiness/hardship to recall examples has an explanation for the subject, its influence is reduced - it's unexpected easiness or hardship that matters the most.

Studies suggest that this whole "heuristic of availability" is primarily linked to Fast Thinking - Slow Thinking on the other hands has people concentrate more on the content of the examples they recall: people tend to use more the heuristic of availability if they are cognitively at ease.

Availability, Emotion and Risk

The heuristic of availability also explain how people view risk: the probability that they associate to each risk is related to the availability of examples of such risks (ex: accidents, more covered by the media, are highly overblown, as compared to AVCs). Their degree of preparation is correlated with the available memory of a risk [NDLR: can this simply denote their awareness of the risk?]

The heuristic of availability can create, when it comes to social interactions, what is called "a waterfall of availability": the occurrence or mention of an event primes everyone with it, driving people (eg media) to talk more about it, which reinforces its availability for everyone, etc... finally leading to a strong over-evaluation around this event (a phenomenon often observed in politics).

The perception of risk also implies the "Heuristic of Affect": it is an example of substitution where the question "what I think about that" (a rather difficult question for Fast Thinking) with "How do I feel about that" (an easier question). That heuristic plays an important role in the perception of risks, and explain why striking images (spectacular incidents, deaths) often lead to an over-evaluation of the risks.

The study of risks perception also puts forward the effects of what is called "Associative Coherence": as Fast Thinking tries to quickly make sense of any new information, it hates ambiguity and will interpret information so that it fits the current perception. It tries to make associated ideas "coherent", and thus to emphasise strong opinions. That may sound obvious -even though it should not-, but it is empirical that people who find many advantages to a technology will turn a blind eye on its defaults (and vice versa). This interestingly, it was shown that by changing the view of someone on the advantages of a technology, you will also change his/her views on its risks - without needing to mention them -.

[There is a whole section here about the evaluation of risk in politics: the metrics used to assess risks can trigger different emotions, so their choice is very political. There is a debate about whether the "people" should choose the metrics used (so that they reflect what it is subjectively concerned about) or whether it should be "experts" choosing (to limit the bias from the people, but could introduce a shift from their expectations). But I'm really not sure this section is relevant in this chapter -and in this summary-]

About Bayesian Statistics ("Tom W's Major")

If 2% of all cars in New York are taxis, and that 90% of all taxis, as well as 5% of all other cars, are yellow, the probability of a yellow car been a taxi can is sometimes called "probability of the car been a taxi, conditioned by the fact that it is yellow". Bayes rule is a simple formula that helps to compute those conditional probabilities (if you did statistics at what point in you life, you must have heard of it). It resorts on what is called the "basic rate" of the event whose probability we want to compute: here the "basic rate" is the percentage of all cars that are taxis. This basic rate is key in the computation: if you divide the overall number of taxis by 10, you'll agree that the conditional probability mentioned above is strongly impacted.

When trying to assess a probability, you would thus expect people -especially those with a formation in statistics- to always take into account that basic rate. Yet empirical results show that most people, regardless of the statistical knowledge, are blind to it in most situation. Admit it: if I told you about a yellow car in New York, you think highly probable that it is a taxi, without ever wondering about the overall proportion of taxis in NY. With the (fake) figures mentioned above, the probability is actually 27%. How can intuition be so wrong?

This is because the heuristic that we use is incorrect: when asked to evaluate the conditional probability we replace the real evaluation, which implies applying the Bayes rule ("complicated", in the sense that it implies slow thinking), by a much simpler one of "plausibility": "which option is the most likely to produce what I observe?". The very fact that you have to think to understand the difference between the two shows how likely it is to interchange them - while they are two completely different things -.

[NDLR: formally, the heuristic is replacing the evaluation of $P(B|A)$ by the evaluation of $P(A|B)$ (and selecting the maximum of that second evaluation)]

This is one important cognitive bias, which is often observe in practice: ignorance of the basic rate (and more formally, of Bayes rule). One way to fight it is too always keep in mind that it exists (and must be corrected) when thinking of statistics: studies also showed that people were less subject to it when asked to "think like a statistician".

More is Less

The replacement of a logical reasoning by the evaluation of plausibility (see previous section) is also at play in two classical mistakes:

- Studies showed that when people are asked to evaluate the probability of different elements, they tend to give higher probabilities to elements with more details than to simpler ones (for example, they will give a higher probability to the conjunction of two event that to each event individually). A possible explanation is that more details means that the element is deemed more plausible, and that plausibility is used as a proxy for probability, without Slow Thinking noticing that the answer provided is in blatant violation of logic. The name proposed for this mistake is "a conjunction mistake".

- When asked to evaluate a group of objects, subjects will focus on the average value of the group, and not its sum. This can lead to absurd situations, where a strictly larger group is deemed less valuable that one of its components. [NDLR: is this really that absurd? In intricate or social situation a component can be more "valuable" that the lot (think of a wolf pack with/without younglings), and maybe we've evolved to take that into account]. This is know as the "More is Less" effect, when adding less valuable items leads to a decreased evaluation. [NDLR: I'm not actually sure this relates to plausibility]

One thing to note to avoid those mistakes is that they are far less common when people replace an abstract task (evaluating a percentage, blind evaluation) by one where they can visualise what is at stakes (eg taking an explicit sample and asking for average values in it; explicitly comparing different group of objects instead of evaluating them separately).

When causes overwhelm statistics

Further studies looked into why people tend to ignore the basic rates. One of their first findings is that basic rates are processed differently if they can be linked to the question by a causal relation (we refer to those as "causal" basic rates) compared to when there is no such relation and the rate is just considered as a statistical fact (we refer to them as "statistical" basic rates). Causal basic rates appeal to

the tendency of Fast Thinking to find causes to everything, and are thus much less frequently "ignored" than statistical basic rates. In the example of the taxis (see **??**, saying that only 2% of all vehicles in New York are taxis because people there love to own their car makes that basic rate causal and would reduce our propensity to forget it.

In general, what is key to taking into account the basic rates is our capacity to visualise them [NDLR: how appealing they are to our Fast Thinking, I guess]: specific studies showed that even causal basic rates were ignored when they contradicted anchored opinions, but that using specific examples that illustrate the basic rate (eg giving explicit examples) was much more efficient to have people remember to take them into account.

In a nutshell, striking the imagination of someone by examples and stories is much more efficient to have him/her remember basic rates that actually giving the figure. [NDLR: Which is not very reassuring as it shows the importance of "story telling" over facts...]

Regression towards the Mean

Do you know of the phenomenon of "regression towards them mean"? Most people don't, and even when they do they are blind when confronted to it. You see the effects of that ignorance every day.

Take sportsmen (say footballers). You'll agree with me that footballers that have exceptional performances at one season often prove disappointing during the next one. It is easy to find explanations and causes for that. You are probably thinking of laziness or of debauchery lifestyles. You need not search so far, as even in the absence of any cause this phenomenon will happen. You'll agree that performance on one season is not only in the hands of the player and rely on external factors (we will simplify by saying its "luck"). Thus the best players in one season are probably not only good, but also at least slightly lucky. As luck is a random phenomenon, in average they will be less lucky next season, and see their performance degrade. This is an illustration of regression towards the mean: the most extreme samples in a phenomenon at a given time are likely to get closer to the mean on the next iteration. [NDLR: formally, I think there are hypotheses for this to be true? I guess here we consider them true]

While this in general, and can be observed in many situations, it is very hard to conceptualise and visualise: it is far more easy for Fast Thinking to think of reasonable causes than to interpret it as an artefact of statistical behaviours. In consequence, most people will find causes in the lack of performance of their stars, or in the fact that "the most intelligent women marry men that are less intelligent than them" (the phenomenon happens as soon as the two variables of interest are not perfectly correlated).

Even when aware of the regression towards the mean, most people can't help to turn to causal explanation (not that of course they may exist, but one should first check for that effect). As a proof of our difficulty to form an intuition on regression towards the mean, be aware that regression towards the mean, even though observed in our daily lives, was discovered about 200 years (ish) after gravitation theory.

Taming intuitive predictions

When asked to make a prediction, people often substitute the task with an evaluation of intensity: they'll find any related information (eg high school grades to evaluate chances of earning a Nobel) and assume that there is an equivalence of intensity between that information and the one to predict (eg translate in another unit the one information they could link, no matter the complexity or relevance of the link). They'll thus ignore most statistical reasoning, like basic rates and regression toward the mean.

You can try to correct this tendency by explicitly wondering about the correlation between your proxy information and the variable to predict, and weight the "equivalence of intensity" estimation by that correlation (using a weighted sum between that estimation and the basic rate). Making the effort to correct those biases will lead to less extreme predictions - you tend to regress to the mean each time -. This is not always what people are looking for, as we feel more comfortable making strong "guts" decisions. So the question of whether to correct or not the biases is actually worth asking.

In general correcting biases caused by overlooking the basic rate or regression toward the mean is very difficult, both because it means fighting our tendency to form strong opinions and because the intellectual gymnastic necessary to make the correction is very counter-intuitive to our brains.

Part 3: Excessive Self-Confidence

The Illusion of Understanding

When thinking back of past events, we have a strong tendency to think that were predictable and inevitable, while we are quite unable to predict future events in practice. This illustrates our bias towards constructing stories and intentions, and then reinforcing them. Most people would say that the 2008's crisis, or Google's success could easily have been predicted, but that's ignoring the proportion of luck and un-predicted side-events or non-events (events that could have occurred), that are evacuated from our simplifying story afterwards (note that we do not only speak about the "social" story that we retain, but also about our own internal perceptions.

We can call this process the "retrospective bias", or "result bias" (as it also applies to results of decisions: if the output is bad, the decision is deemed bad even if it made sense when it was taken). This is especially an issue for decision makers: if the output of their decisions is positive, people tend to think that taking the decision was obvious. If it is negative, they are deemed as idiots. It is important to notice that this bias also exerts its influence when people change opinions: studies showed that they tend not to remember their previous opinions. When proven wrong, they tend to think (unconsciously) that they knew the right answer since the beginning.

This tendency to find causes and stories around phenomenons leads us to oversimplify some questions, and thus over-empower (in our representations) some decisions. A great deal of books have been published that analyse and draw conclusions from famous CEO's styles or decisions, whereas empirically the influence of a CEO over his/her company's performance is actually very reduced (Kahneman quotes some studies that showed a correlation of 0.3, but I don't know exactly what was compared).

The Illusion of Validity

Once our belief in a fact/cause is well established, it is very hard to debunk, even with very strong empricial evidence: the strengh of the stories we built for ourselves (that tend to reinforce each other) is too great.

An example is trading: figures show that the performance of the traders is completely random and do not show any role in skills market knowledge (in a nutshell, professionals make money out of the loss of amateurs but from one professional to another the performance is completely random (not at all consistent in time, which would denote a real importance of skill). Yet the belief that professional traders can really predict the ups and downs of the market stays anchored even in the traders themselves, who know about statistics and know the figure: they unconsciously ignore "basic rate" that this information brings them and concentrate on their own perception of their talent.

The same illusion is also true bout experts in several fields (eg politics): statistically, their predictions aren't more accurate that random ones. They keep there social "relevance" as they help to build stories and find (possibly fallacious) explanations to events that have already happened. This can for example explain why the most assertive experts are the most mediatized ones, even though their predictions aren't relevant. The one result to look at when asserting the relevance of a prediction is the empirical correlation between the source of the prediction and the results. If it low (as in most cases), the prediction should be handled carefully. Thinking that a phenomenon or causality is true despite evidence is referred to as "the illusion of validity".

Intuition versus formulas

In many fields - from prediction of wine prices to evaluation treatments length -, the intuition of experts was proven weaker than simple formulas based on statistics, and this even when experts had access to the simple statistical prediction. Part of the issue is that experts are pretty inconsistent in their predictions: a study showed that radiologists contradict themselves about 20% of the time when shown a same thorax radio twice in different occasions. [NDLR: this should be taken with a grain of salt, it may

only show that 20% of all thorax radios are heavily ambiguous (eg a classifier would give similar scores to each class)] This can be explained by the fact that experts are heavily influenced by their context when making predictions, as Fast Thinking primes their opinions with that context: judges tend to soften their verdicts after their breaks.

Kahneman quotes a personal example where the intuitive prediction of the expert actually added to the predictive power, if it was asked after having the expert use the formula.

The formula doesn't have to be complicated to be relevant, and in fact intricate algorithms are often less robust than simple "intuitive" formulas, at least in social sciences.

Despite its empirical relevance, the "formulas" approach is still frowned upon most of the time: you are well more likely to trust an expert than the result of a formula, despite the basic rate giving a strong advantage to the formula's prediction. [NDLR: I see 2 possible explanations. The first one is a priming effect; if the expert looks nice and speaks convincingly, you're primed to trust his/her opinion. This can't happen with a formula. The second is about responsibility: when an expert makes a mistake, the loss of trust is exerted on this expert specifically: he/she "screwed up" (and not experts in general), since he is distinct from his colleagues. You'll not put the discredit on the "experts" as a whole. If it is an algorithm that leads to a bad decision, the whole algorithmic approach is "discredited".]

Applying the implicit recommendations from this chapter is quite hard: before making a decision, you should first design a an objective formula, and then stick to it whatever your intuition tells you. (of course, this is only true if you can come up with a relevant formula).

When to trust experts

Despite biases, it is indisputable that in some fields, experts intuition should be respected. But how is it formed, and when should you trust it?

Recognition-Primed Decision Model (RPDM): a model for experts' decision making (based on studies of firefighters leaders). When facing a situation, system 1 proposes, based on previous experiences and association of ideas, a possible answer ("intuitively"). System 2 then "plays it" mentally and check how well it fits the situation and possible modifications to make. If the solution is deemed irrelevant, System 1 brings up another answer. In this model, only 1 answer is considered each time, and different possible answers are never explicitly compared.

Those intuitions of System 1 can be very simple to learn, especially when linked to emotions such as fear (a single experience, or description of an experience is enough to anchor intuition). Expertise in complex fields takes longer to develop as it requires the coordination of several skills [and because patterns are harder to detect from experience - compare for example chess with car driving - NDLR - ie

useful features are much higher level and farther from the ones we are used to]

How to explain the difference in the accuracy of experts' predictions in different fields? Most of it is related to the field itself: in order for an expert - or even an algorithm - to develop a true expertise, its environment (the field of prediction) must be regular enough as time passes to allow the expert to learn from its experiences, and give a readable feedback to the expert (as close as possible to the expert's decision). This explains why firefighters have a strong predictive skill -their field doesn't change much through time - whereas traders or politicians don't. But both "categories" suffer from the illusion of validity and have the same overconfidence in their own predictions (the overconfidence can also come from the fact that their are good a certain tasks but not in all of them (eg good at short-term predictions, not at long-term ones)). Those two facts should be taken into account when dealing with expertise.

Vocabulary

This section lists the important concepts developed in the book and the chapters related to them (maybe I'll add a comprehensive definition latter). I've after reading this document these ideas are still unfamiliar to you, then I guess that I did a poor job... or you may want to CTRL+F it!

- Fast Thinking

- Slow Thinking

- Halo Effect

- WYSIATI

- Priming Effect

- Buckshot Effect

- Substitution Effect (Answering a Simple Question)

- Law of Small Numbers (The Law of Small Numbers)

- Anchor Effect (Anchor Effect)

- Heuristic of Availability (The Science of Availability)

- Heuristic of Affect (Availability, Emotion and Risk)

- Conjunction Mistake

- More is Less Effect

- Regression towards the Mean (Regression towards the Mean)

- Retrospective Bias, Result Bias (The Illusion of Understanding)

- Illusion of Validity (The Illusion of Validity)

- RPDM (When to trust experts)

References

[1] Daniel Kahneman. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, 2011.