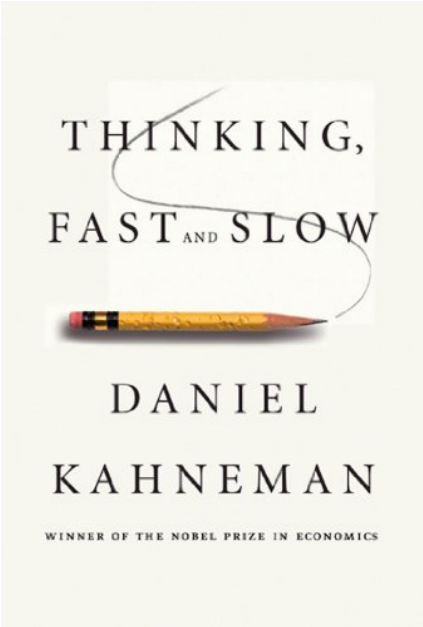
Best Summary+PDF: Thinking Fast and Slow, by Daniel Kahneman

by Allen Cheng

https://www.allencheng.com/thinking-fast-and-slow-summary-pdf-daniel-kahneman/



We're so self-confident in our rationality that we think all our decisions are well-considered. When we choose a job, decide how to spend our time, or buy something, we think we've considered all the relevant factors and are making the optimal choice.

In reality, our minds are besieged by deep-rooted evolutionary biases. Where they once enhanced survival, they now lead to poor decision making.

Thinking, Fast and Slow is a masterful book on psychology and behavioral economics by Nobel Prize laureate Daniel Kahneman. You might already be aware of biases, from books like Poor Charlie's Almanack or Cialdini's *Influence*. *Thinking, Fast and Slow* exceeds these books by presenting an overall framework of cognition that explains ***why*** **the biases exist and how they work**.

This is a book to take your time and chew on. You should apply each learning to mistakes you've made and decisions you're mulling over. The rewards are important: you'll make better decisions and maximize your happiness.

In this *Thinking, Fast and Slow* summary, learn:

- How something as simple as reordering a list can change your decision
- Why most pundits and stock analysts are worthless
- Why losses are more painful than gains, and when you inappropriately take on risk

- Why we fantasize about hail-mary shots at big wins (lottery tickets and investments), while also obsessing about tiny chances of bad outcomes (disasters)
- Why 90% of drivers state they're above average drivers

Full Title: Thinking, Fast and Slow by Daniel Kahneman

Caveats:

***Thinking, Fast and Slow* deserves to be read in full** - of all the books I've read, I'd say less than 10% qualify for this recommendation. One of the best reasons to read the original *Thinking, Fast and Slow* is to engage in the decision puzzles Kahneman presents, without having the answers spoiled for you. Committing cognitive errors yourself drives home the point that your mind is fallible.

Thinking, Fast and Slow is a relatively dense book compared to other nonfiction pop psychology books - it covers three major domains (System 1 and 2 and cognitive biases; prospect theory; and happiness). It's full of experimental results, and Kahneman does not add fluff.

Because of the book's density, this *Thinking, Fast and Slow* summary is one of my longest. I preserve the key experiments and puzzles that demonstrate our minds' fallibility, and I also expand on the book's applications and examples.

Finally, the chapter on priming has been disavowed by Kahneman himself, since the original papers showed signs of p-hacking and irreproducibility. More on this later.

1-Page Thinking, Fast and Slow Summary

Thinking, Fast and Slow covers three major areas of Daniel Kahneman's research: cognitive biases, prospect theory, and happiness.

Cognitive Biases

- System 1 and 2 definition
    - System 1: operates automatically and quickly, with little or no effort, and no sense of voluntary control
        - Examples: Detect that one object is farther than another; detect sadness in a voice; read words on billboards; understand simple sentences; drive a car on an empty road.
    - System 2: allocates attention to the effortful mental activities that demand it, including complex computations. Often associated with the subjective experience of agency, choice and concentration
        - Examples: Focus attention on a particular person in a crowd; exercise faster than is normal for you; monitor your behavior in a social situation; park in a narrow space; multiple 17 x 24.
- System 1 automatically generates suggestions, feelings, intuitions for System 2. If endorsed by

System 2, intuitions turn into beliefs, and impulses turn into voluntary actions. A lazy System 2 accepts what the faulty System 1 gives it, without questioning.

- Important biases/heuristics
    - Narrative fallacy - we seek to explain stories that may be due to randomness. This gives us unjustified confidence about predicting the future.
    - Narrow framing/"What you see is all there is" (WYSIATI) - we don't consider the global set of alternatives or data. We don't realize the data that are missing.
        - Planning fallacy - ignoring the many ways things could go wrong
        - Sunk cost fallacy - we separate life into separate accounts, instead of considering the global account
    - Intensity matching - to answer a more difficult question, we gauge how we feel about an easier question, then substitute the answer
    - Ignoring Reversion to the mean - if randomness is a major factor in outcomes, high performers will suffer and low performers will improve, for no meaningful reason
    - Representativeness - just because something looks like a prototype does not give it the same characteristics (eg stereotype of a tech company founder)
    - Availability bias. Vivid images and stronger emotions make items easier to recall and are overweighted.
- The best we can do is a compromise: learn to recognize situations in which mistakes are likely and try harder to avoid significant mistakes when the stakes are high." It's easier to recognize other people's mistakes than our own.
- Antidotes
    - Invert the question.
        - How can this project fail?
        - Turn a loss into a gain and re-ask the question, and vice versa.
    - When estimating probability, begin with the basal probability for the category. Then adjust from this rate based on new data. Do NOT start with your independent guess of probability since you ignore the data you don't have.
    - Use formulas and checklists to avoid biases.

Prospect Theory

- Rather than utility on an absolute scale, our utility depends on **changes from one's reference point**. A who goes from 4 to 2 million is much worse feeling than someone who goes from 1 to 2 million.
- Losses are more painful than equivalent gains
- People are risk seeking to avoid certain losses, and risk avoiding to lock in certain gains.
- **Possibility effect**: we overweight small probabilities (eg 1% is much more significant than its low likelihood suggests - both in terms of lottery ticket and chance of injury). **Certainty effect**: we underweight near-certain probabilities (eg 98% is much worse than its high likelihood suggests).
- Status quo bias, endowment effect - you like what you have and don't want to lose it, even if your prior indifference curve suggested you would.
- Vividness and triggering emotions makes you less sensitive to probability (you overestimate low chances and underestimate high chances).
    - 80 out of 10000 seems more likely than 1 out of 100.
- Antidotes:

    - Strip away irrelevant details from a rosy story to regain sensitivity to probabiltiies. (eg don’t picture a lottery ticket win)
    - To avoid denominator neglect, always reduce two statistics to the same denominator for comparison (eg mortality rate per 100k people). Apples to apples.
- Broad framing - don’t hold narrow accounts as loss aversion and sunk cost fallacy will cause subpar decisions in the long run.
    - Novice stock traders see each stock in isolation, selling to capture gains and keeping losers to recover losses, when the opposite is a more profitable strategy.
    - In a company, individual project leaders can be risk averse (since their incentives are heavily tied to project success). Yet the CEO may wish that all project leaders take on maximum appropriate risk to maximize the EV of the total portfolio.
- Antidote:
    - Reduce all decisions down to a single fungible metric, in the broadest account possible, to allow global calibration.
        - Benefits to human life can be assessed in terms of QALYs/$, so that repairing glaucoma can be compared to childhood cancers.
        - Projects and spending should be considered in terms of ROI or IRR or happiness/$, so that all possible spending can be assessed on the same scale from a global account
    - Deliberately consider the broad range of alternatives for scarce resources - time, labor, money. Be aggressive with defining the global set of opportunity costs.

Happiness and the Two Selves

- **The experiencing self:** the person who enjoys pleasures and feels pain moment to moment. Experienced utility would best be measured by measuring happiness over time, and integrating the area under the curve.
- **The remembering self:** the person who reflects on past experiences and evaluates it overall.
- The remembering self evaluates differently from the experiencing self in two critical ways:
    - Peak-end rule: The overall rating is determined by the peak intensity of the experience and its end. It does not care much about the averages throughout the experience.
    - Duration neglect: The duration of the experience has little effect on the memory of the event.
- We tend to prioritize the remembering self (where we book vacations, willingness to endure pain we will forget later) and don’t give enough to the experiencing self.

Part 1: Two Systems

This *Thinking, Fast and Slow* summary begins with explaining the two systems of our mind, what they each do, and how they interact.

We believe we’re being rational most of the time, but really much of our thinking is automatic, done subconsciously by instinct. Most impressions arise without your knowing how they got there. Can you pinpoint how you knew a man was angry from his facial expression, or how you could tell that one object was farther than another, or why you laughed at a funny joke?

This becomes more practically important for the decisions we make. Often, we've decided what we're going to do before we even realize it. Only *after* this subconscious decision does our rational mind try to justify it.

The brain does this to save on effort, substituting easier questions for harder questions. Instead of thinking, "should I invest in Tesla stock? Is it priced correctly?" you might instead think, "do I like Tesla cars?" The insidious part is, you often don't notice the substitution. This type of substitution produces systematic errors, also called biases. We are blind to our blindness.

System 1 and System 2 Thinking

In *Thinking, Fast and Slow*, Kahneman defines two systems of the mind:

- System 1: operates automatically and quickly, with little or no effort, and no sense of voluntary control
    - Examples: Detect that one object is farther than another; detect sadness in a voice; read words on billboards; understand simple sentences; drive a car on an empty road.
- System 2: allocates attention to the effortful mental activities that demand it, including complex computations. Often associated with the subjective experience of agency, choice and concentration
    - Examples: Focus attention on a particular person in a crowd; exercise faster than is normal for you; monitor your behavior in a social situation; park in a narrow space; multiple 17 x 24.

System 1 can be completely involuntary. You can't stop your brain from completing 2 + 2 = ?, or from considering a cheesecake as delicious. You can't unsee optical illusions, even if you rationally know what's going on.

System 1 can arise from expert intuition, trained over many hours of learning. In this way a chess master can recognize a strong move within a second, where it would take a novice minutes of System 2 thinking.

System 2 requires attention and is disrupted when attention is drawn away. More on this next.

System 1 automatically generates suggestions, feelings, intuitions for System 2. If endorsed by System 2, intuitions turn into beliefs, and impulses turn into voluntary actions.

System 1 can detect errors, provoking surprise, and recruits System 2 for additional firepower.

- Kahneman tells a story of a veteran firefighter who entered a burning house with his crew, felt something was wrong, and called for them to get out. The house collapsed shortly after. He only later realized that his ears were unusually hot but the fire was unusually quiet, indicating the fire was in the basement.

Because System 1 operates automatically and can't be turned off, biases are difficult to prevent. Yet it's also not wise (or energetically possible) to constantly question System 1, and System 2 is too slow to

substitute in routine decisions. **The best we can do is a compromise: learn to recognize situations in which mistakes are likely and try harder to avoid significant mistakes when the stakes are high."**

In summary, "most of what you (your System 2) thinks and does originates in System 1, but System 2 takes over when things get difficult, and it normally has the last word."

[The automaticity of System 1 has profound implications for stereotype processing (eg race, gender, body type). If existing stereotypes in your mind are activated automatically by System 1, we can only consciously battle them with System 2 - but to some extent we cannot turn them off.]

A Lazy System 2 Accepts the Errors of System 1

Consider these questions, and go through them quickly, trusting your intuition.

A bat and ball cost $1.10. The bat costs one dollar more than the ball. How much does the ball cost?

How many murders occur in the state of Michigan in one year?

Does the conclusion follow from the premises?

- All roses are flowers.
- Some flowers fade quickly.
- Therefore some roses fade quickly.

The answer to the first question is $0.05. The common intuitive (and wrong) answer is $0.10.

The trick to the second question is whether you remember that Detroit is in Michigan. People who remember this estimate a number that is much higher than those who forget.

The last answer is no - all roses may not fit into the subcategory of flowers that fade quickly.

===

All of your answers, if you really spent time on it, could be verified by deliberate System 2 thinking. In the first question, it's easy to see that if the ball cost $0.10, the total would be $1.20, which is clearly incompatible with the question. For the second question, if you had to enumerate the major cities of Michigan, you would likely list Detroit.

For some, spending enough time would be sufficient to get the answers right. But for many, even when given infinite time, might not think to apply their System 2 to question their answers and find different approaches to the question. Over 50% of students at Harvard and MIT gave the wrong answer to the bat-and-ball question; over 80% at less selective universities.

This is the insidious problem of a “lazy System 2.” System 1 surfaces the intuitive answer for System 2 to evaluate. **But a lazy System 2 doesn’t properly do its job** - it accepts what System 1 offers without expending the small investment of effort that could have rejected the wrong answer.

Even worse, this aggravates confirmation bias - a piece of information that fits your prior beliefs might evoke a positive System 1 feeling, while your System 2 might never pause to evaluate the validity of the piece of information. **If you believe a conclusion is true, you might believe arguments that support it, even when the arguments are unsound.** [No wonder today’s media can be so polarizing! I also suspect that the misleading labeling of news as “we report, you decide” makes viewers believe they’re thinking hard, when really they’re digesting prepackaged opinions.]

It’s useful then to distinguish between intelligence and rationality.

- Intelligence might be considered the full computational horsepower of a person’s brain.
- Rationality is resistance to mental laziness; **not accepting a superficially plausible answer**; being more skeptical of intuitions; tending to put in the hard work of checking the logic; and thus immunity to biases.
- In other words, a powerful system 2 is useless if the person doesn’t recognize the need to override their system 1 response.

The theme here, that will recur through the book, is that **people are overconfident and place too much faith in their intuitions**. Further, they find cognitive effort unpleasant and avoid it as much as possible.

A Maximum Capacity for System 2

System 2 thinking has a limited budget of attention - you can only do so many cognitively difficult things at once.

- This is true in parallel - if you’re navigating traffic on a busy highway, it becomes far harder to solve a multiplication problem.
- This is also true serially - ego depletion earlier in the day can lower inhibitions later (eg buying from late night infomercials).
- All forms of voluntary effort - cognitive, emotional, physical - seem to draw at least partly on a shared pool of mental energy.

- Stifling emotions during a sad film worsens physical stamina later.
- Memorizing a list of seven digits makes subjects more likely to yield to more decadent desserts.

The law of least effort states that **“if there are several ways of achieving the same goal, people will eventually gravitate to the least demanding course of action.”**

What makes some cognitive operations more demanding than others?

- Holding in memory several ideas that require separate actions (like memorizing a supermarket shopping list). System 1 is not capable of dealing with multiple distinct topics at once.

- Obeying an instruction that overrides habitual responses, aka “executive control.”
    - Again, this covers cognitive, emotional, and physical impulses of all kinds.
- Switching between tasks.
- Time pressure.

The strain of a cognitive task can be measured by pupil size - the harder the task, the more the pupil dilates, in real time. Heart rate also increases.

- Kahneman cites one task as the limit of what most people can do in the lab, dilating the pupil by 50% and increasing heart rate by 7bpm.
    - The task is “Add-3”: write several 4 digit numbers on separate index cards. To the beat of a metronome, read the four digits aloud. Wait for two beats, then report a string in which each of the original digits is incremented by 3 (eg 4829 goes to 7152). Any harder than this and most people give up.
    - Relatively speaking, this is sprinting as hard as you can, whereas casual conversation is a leisurely stroll.
- [I wonder if two people engaged in demanding conversation, with dilated pupils, might thus become more attracted to each other.]

Because System 2 has limited resources, stressful situations such as the below make it harder to think clearly:

- In intense exercise, where you need to exercise resistance to the urge to slow down. (Thus the most complex arguments might be done while sitting.)
- In the presence of distractions.
- Exercising self-control or willpower in general.

Because of the fixed capacity, you cannot will yourself to think harder in the moment than Add-3, even with a gun to your head. But there are some alleviations:

- The more skilled you get at a task, the less energy it consumes. Thus experts can solve situations that require far more intense work for novices (eg diagnosing a disease, making a chess move).
- In a state of flow, you can concentrate energy without having to exert willpower, thus freeing your entire mental faculty to the task. Flow has been described as one of the most enjoyable states of life.
    - [Flow theory suggests three conditions have to be met: 1) Clear set of goals and progress. 2) Clear and immediate feedback, to adjust performance to maintain flow. 3) Good balance between perceived challenges and perceived skills - confidence in ability to complete the task.]
- When given proper incentives, people resisted ego depletion.
- Consuming more glucose seems to undo ego depletion, since cognitive effort depletes blood glucose.
    - Remember the Israeli judge study showing more parole leniency after meals.
- Playing computer games requiring attention and control may improve executive control and intelligence scores.
- [I suspect that mental capacity can be trained over time, like stressing a mental muscle. Whether

this has a systematic effect on increasing System 2 bandwidth I'm not sure about.]

As we'll find later, when System 2 is taxed, it has less firepower to question the conclusions of System 1.

The Link Between System 2 Capacity and Intelligence

Given that self-control and cognitive tasks draw from the same pool of energy, is there a relationship between self-control and intelligence?

Walter Mischel's famous marshmallow experiment showed that children who better endured delayed gratification showed significantly better life outcomes, measured by SAT scores, education attainment, and BMI.

Inversely, those who score lower on the Cognitive Reflection Test show more impulsive behavior - more willing to pay for overnight delivery, less willing to wait for more time to receive more money - as well as tendency to fall prey to fallacies like gambler's and sunk cost.

As with most things in life, it appears that executive control has been attributed to both genetics and environment (parenting techniques).

Associative System 1

Think of your brain as a vast network of ideas connected to each other. These ideas can be concrete or abstract; can involve memories, emotions, and physical sensations. When one node in the network is activated, say by seeing a word or image, **it automatically activates its surrounding nodes**, rippling outward like a pebble thrown in water.

As an example, consider the following two words:

"Bananas Vomit"

Suddenly, within a second, you might have pictured yellow fruits; felt a physiological aversion in the pit of your stomach; remembered the last time you vomited; thought about other diseases - all done automatically without your conscious control.

The evocations can be self-reinforcing - a word evokes memories, which evoke emotions, which evoke facial expressions, which evoke other reactions; which reinforce other ideas.

The purpose of this association is to prepare you for events that had become more likely, and to evaluate how surprising the event is. The more external inputs associate with each other, and the more they associate with your internal mind, the less surprising it is, and the harder it is to detect errors.

Consider the sentence: "how many animals of each kind did Moses take into the ark?"

- The answer is none - it was Noah who took animals into the ark.

- But the idea of animals, Moses, and the ark all set up a biblical context that associated together. Moses was not a surprising name in this context.
- However, say “how many animals of each kind did Kanye West take into the ark?” and the illusion falls apart.

Links between ideas consist of several forms:

- Cause -> Effect
- Belonging to the Same Category (lemon -> fruit)
- Things to their properties (lemon -> yellow, sour)

Normative

System 1 maintains a model of your world by what is normal and not.

Violations of normality can be detected extremely quickly, within fractions of a second. If an upper class English accent says, “I have a large tattoo on my back,” brain activity occurs within 2/10 of a second, even though a large amount of world knowledge needs to be invoked to recognize the discrepancy (that rich people don’t typically get tattoos).

(Tangent: We also communicate by norms and shared knowledge. When I mention a table, you know it’s a solid object with a level surface and fewer than 25 legs.

[This also explains why so many moral arguments are based around semantics. In different insular communities, people will have different conceptions of what the same word means, like “life” in the abortion debate. The norms are entirely different, but often people don’t realize this.])

Association is Fast and Subconscious

Consider these three words, and think about which word fits with them as a phrase:

cottage Swiss cake

You might have thought of “cheese.”

Now look at these two sets of words, and within seconds decide which one feels better:

sleep mail switch

salt deep foam

You might have found that the second one felt better. There is a very faint signal from the associative machine of System 1, long before you consciously found the word (which is salt).

For another example, consider the sentence “Ana approached the bank.”

You automatically pictured a lot of things. The bank as a financial institution, Ana walking toward it.

Now let’s add a sentence to the front: “The group canoed down the river. Ana approached the bank.”

The context changes your interpretation automatically. Now you see how automatic your first reading of the sentence was, and how little you questioned the meaning of the word “bank.”

Building a Narrative / Seeing Causes and Intentions

The brain wants to make sense of the world. It wants large events to cause effects, and it wants effects to have causes. It tries to bring coherence to a set of data points and sees interpretations that may not be explicitly mentioned.

For example: “After spending a day exploring sites in the crowded streets of New York, Jane discovered that her wallet was missing.”

Immediately, you likely pictured a pickpocket, and you would likely recall this word even if it wasn’t stated in the text.

Once you receive a surprising data point, you also interpret new data to fit the narrative. Imagine you’re observing a restaurant, and a man tastes a soup and suddenly yelps. This is surprising. Now two things can happen that will change your interpretation of the event:

- The server touches the man’s shoulder and he yelps.
- A woman at a different table drinks her soup, and she also yelps.

In the first case, you’ll think the man is hyper-reactive. In the second, there’s something wrong with the soup. In both cases, System 1 assigns cause and effect without any conscious thought.

This works much of the time, but **it fails when you apply causal thinking to situations that require statistical thinking.**

Conversely, **maintaining multiple incompatible explanations requires mental effort and System 2**. It’s far easier to see the world in black and white than in shades of gray.

(Tangent: note that you interpret your own actions differently from other people. When you act, you experience the action as a decision that a disembodied *you* made. The world of your mind is separate from the world of objects, which makes it possible to envision soulless bodies and bodiless souls, which then can give rise to religions.)

Priming and Associations (Questionable)

[Warning: This chapter of *Thinking, Fast and Slow* cites the highly controversial literature on priming,

which has failed to replicate in follow-up studies and has been accused of p-hacking/publishing only positive results. Kahneman himself admitted: "I placed too much faith in underpowered studies"; "the experimental evidence for the ideas I presented in that chapter was significantly weaker than I believed when I wrote it," the "size of behavioral priming effects...cannot be as large and as robust as my chapter suggested." Even Nobel Laureates are subject to biases!]

The concept of **priming** took association to the functional level of **ideomotor activation**. When an idea is triggered, its associations can cause you to *behave* in a meaningfully different way without your consciously realizing it. Examples include:

- People who were primed with money (by images or words) acted more selfishly and independently.
- Voters at a school location were more likely to vote in favor of educational propositions.
- People who were primed with phrases containing of the elderly (Florida, forgetful, wrinkle) walked more slowly.
- Near a "by your honor" cafeteria collection jar for shared milk, an image showing human eyes looking at the person gathered more donations than pictures of flowers.

In the reverse direction, behaving in a certain way can trigger ideas, emotions:

- People were instructed to listen to messages through headphones and were instructed to move their heads to check for distortions of sound. THose instructed to nod up and down were more likely to agree with the messages than those shaking heads side to side.
- People who were instructed to think of a shameful thought were more likely to complete the word SO_P as SOAP rather than SOUP. Feeling morally dirty triggers the desire to cleanse: the "Lady Macbeth Effect"

The implications of priming are profound - if we are surrounded everyday by deliberately constructed images, how can that not affect our behavior?

- Being surrounded by images of money and consumption make us more selfish.
- Seeing images of filial respect make us nicer to our parents.
- Seeing propaganda of an autocratic leader makes you feel watched and reduces spontaneous thought and independent action.

And likewise, if we are required to behave in a certain way (eg in a workplace, social community, as citizens), does that not affect our cognition and beliefs?

The effects may not be huge - but a few percentage points by swinging marginal voters can make a difference.

[Many of the studies cited were later found to have insufficient power, such that either the studies were being p-hacked, or only cherry-picked positive results were being published. It appears the field was too eager to jump on evidence that fit their view of the world.

Note the irony about being biased *about biases*. When priming came out, the field of

psychology/behavioral economics had just undergone a paradigm change of humans being subject to systematic biases. The field hungered for confirming evidence itself, becoming too ready to accept a neat story (priming) without employing its System 2 thinking to question whether the evidence was valid!

Kahneman notes that he's still a believer, "There is adequate evidence for all the building blocks: semantic priming, significant processing of stimuli that are not consciously perceived, and ideo-motor activation....I am still attached to every study that I cited, and have not unbelieved them, to use Daniel Gilbert's phrase."]

Cognitive Ease

Cognitive ease is an internal measure of how easy or strained your cognitive load is.

In a state of cognitive ease, you're probably in a good mood, believe what you hear, trust your intuitions, feel the situation is familiar, are more creative, and are superficial in your thinking. System 1 is humming along, and System 2 accepts what System 1 says.

In a state of cognitive strain, you're more vigilant and suspicious, feel less comfortable, invest more effort, trigger System 2, make fewer errors, but also less intuitive and less creative.

Cognitive ease is affected by a range of inputs, including:

- Repeated experience
- Clear display
- Primed idea
- Good mood

The ways of inducing ease are interchangeable - you don't know exactly what it is that makes it easy or strained. Rather, it gets compressed into a single "is this easy" factor.

Drilling down into each input:

Repeated Experience / Mere Exposure / Familiarity

- Merely exposing someone to an input repeatedly makes them like it more. Having a memory of a word, phrase, or idea makes it easier to see again.
- Example experiments:
    - Exposing Turkish-looking words to people at randomly different frequency makes people perceive the more repeated words often.
    - Exposing people to random names beforehand makes them more likely to identify them as celebrities in a list of names.
- Even a single occurrence can make an idea familiar.
    - Say a car lights on fire a block away from your house. The second time you see it, it is far less surprising than the first time it happened.
- Familiarity is not easily distinguished from the truth.

- If you trust a source, you are more likely to believe what is said.
- If the new idea fits your existing mental framework, you will digest it more easily.
- [This is partially why nonspecific brand advertising, like what Coca-Cola does, works.]

Evolutionarily, this benefits the organism by discarding stimuli that don't cause danger over repeated experiences, to save energy for new things that might.

[Thus, be aware of "common sense" intuitions that seem true merely because it's repeated often, like "searing meat seals in the juices."]

Clear display

In general, to be more persuasive, make the message as easy to digest as possible.

> Consider the two statements:

- **Adolf Hitler was born in 1892.**

> Adolf Hitler was born in 1887.
> You likely found the first one a bit more believable, because it stood out.

- To make your message more persuasive, increase the contrast and make it easier to read. Use high quality paper. If colored, use vivid colors of blue or red rather than pale neutral colors.
- Use simpler language. Pretentious language is interpreted as a sign of poor intelligence and low credibility.
- Make your ideas rhyme and in verse if you can.
    - "Woes unite foes." is interpreted as more meaningful as "Woes unite enemies."
- When citing a source, choose one with an easier name to pronounce.
- [Scott Adams takes this further in Win Bigly, citing visual imagery as especially effective.]

[Mere exposure and clear display are used to great effect in catchphrases like "If the glove doesn't fit, you must acquit."]

In contrast, making cognition difficult actually activates System 2.

- For the bat and ball problem, making print smaller and difficult to read increased the accuracy significantly.

Emotions

Cognitive ease is associated with good feelings, and the inverse for cognitive strain.

- Stocks with grammatically reasonable symbols (like KAR) outperform non-words like PXG or RDO.

The causality also reverses: your emotional state affects your thinking.

- In the three-word creative association test above (salt deep foam), people instructed to think about sad memories were half as accurate in their intuition as those who thought about happy memories.
- **When we are uncomfortable and unhappy, we lose touch with our intuition**. (but are more engaged with our System 2).

You can't always tell why something is easy.

Again, cognitive ease is a summary feeling that takes in multiple inputs and squishes them together to form an impression. When you feel cognitively at ease, you're not always aware about why - it might be that the idea is actually sound and fits your correct view of the world, or that it's simply printed with high contrast and has a nice rhyme.

Cognitive ease is associated with a "**pleasant feeling of truth**." Things that seem intuitively true may actually be false on inspection.

How Judgments Happen

System 1 continuously monitors what's going on outside and inside the mind and generates assessments with little effort and without intention. The basic assessments include language, facial recognition, social hierarchy, similarity, causality, associations, and exemplars.

- In this way, you can look at a male face and consider him competent (with a strong chin and a slight confident smile).
- The survival purpose is to monitor surroundings for threats.

However, not every attribute of the situation is measured. System 1 is much better at determining comparisons between things and the *average* of things, not the *sum* of things.

In the below picture, try to quickly picture what the average length of the lines is.

Now try to determine the sum of the length of the lines.

This is less intuitive and requires System 2.

These basic assessments are not impaired when the observer is cognitively busy (eg with a memory task).

In addition to basic assessments: System 1 also has two other characteristics:

Translating values across dimensions, or intensity matching

- Consider the athletic skill of a minor league baseball player relative to the rest of the population. Now consider that as a level in the range of year-round weather temperature. Just as a minor league player is above average but not the top tier, the correlative temperature might be something like 80 Fahrenheit.
- Consider crimes and punishments, each expressed as music volume. Large mismatches between volumes will show a sense of injustice.

Mental shotgun

System 1 often carries out more computations than is needed. Example:

- Consider whether each of the following three statements is literally true:
    - Some roads are snakes.
    - Some jobs are snakes.
    - Some jobs are jails.
- The second statement likely registered more quickly as false to you. The other two sentences are metaphorically true, and you couldn't help noticing this connection, even though it wasn't relevant to the task and slowed you down.

Heuristics: Answering an Easier Question

Despite all the complexities of life, notice that you're rarely stumped. You rarely face situations as mentally taxing as having to solve 9382 x 7491 in your head. Isn't it profound how we can make decisions at all without realizing it? You like or dislike people before you know much about them; you feel a company will succeed or fail without analyzing it.

When faced with a difficult question, System 1 substitutes an easier question, or the **heuristic question**.

The answer is often adequate, though imperfect.

Consider the following examples of heuristics:

- Target question: How much is saving an endangered species worth to me?
    - Heuristic question: How much emotion do I feel when I think of dying dolphins?
- Target question: How happy are you with your life?
    - Heuristic question: What's my current mood?

- Target question: How should financial advisers who commit fraud be punished?
    - Heuristic question: How much anger do I feel when I think of financial predators?
- Target question: How far will this political candidate get in her party?
    - Heuristic question: Does this person *look* like a political winner?

Based on what we learned in the last chapter about characteristics of System 1, here's how heuristics are generated:

- System 1 applies a mental shotgun with a slew of basic assessments about the problem. These impressions are readily available without thought.
    - Many impressions are more emotion-based than we would like.
- **System 1 uses intensity matching to scale the relevant assessments to the target question**.
    - How painful the image of a dying dolphin feels is scaled to how much money you're willing to contribute.

System 2 has the opportunity to reject this answer, but **a lazy System 2 often endorses the heuristic without much scrutiny**. Even more insidiously, a lazy System 2 will feel as though it's applied tremendous brainpower to the question.

In one experiment, students were asked:

- How happy are you these days?
- How many dates did you have last month?

When presented in that order, there was no correlation between the answers.

When reversed, the correlation was very high. The first question prompted an emotional response, which was then used to answer the happiness question.

[How long lasting is this effect? Can you improve employee ratings by asking them first if they like the snacks in the office, then asking how much they like their job?]

Biases of System 1

Putting it all together, we are most vulnerable to biases when:

- System 1 forms a narrative that conveniently connects the dots and doesn't express surprise.
- Because of the cognitive ease by System 1, System 2 is not invoked to question the data. It merely accepts the conclusions of System 1.

This is acceptable if the conclusions are likely to be correct, the costs of a mistake are acceptable, and if the jump saves time and effort. This is risky when the stakes are high and there's no time to collect more information.

What You See is All There Is: WYSIATI

When presented with evidence, especially those that confirm your mental model, you do not question what evidence might be missing. System 1 seeks to build the most coherent story it can **- it does not stop to examine the quality and the quantity of information**.

In an experiment, groups were given background to a legal case. Then one group were given just the plaintiff's argument, another the defendant's argument, and the last both arguments. Those given only one side gave a more skewed judgment, and **were more confident of their judgments** than those given both sides, *even though they were fully aware of the setup*

We often fail to account for critical evidence that is missing.

Halo Effect

If you think positively about something, it extends to everything else you can think about that thing.

Say you find someone visually attractive and you like this person. You more likely find her intelligent or capable, even if you have no evidence of this. You tend to like intelligent people, and now that you think she's intelligent, you like her better than you did before. In other words, your emotional response fills in the blanks for what's cognitively missing from your understanding.

This forms a simpler, more coherent story by generalizing one attribute to the entire person. Inconsistencies are easier to understand: "Hitler loved dogs and little children" is troubling for many to comprehend.

Ordering Effect

First impressions matter. They form the "trunk of the tree" on which later impressions are attached to like branches. It takes a lot of work to reorder the impressions to form a new trunk.

Consider two people who are described as follows:

Amos: intelligent, hard-working, strategic, suspicious, selfish

Barry: selfish, suspicious, strategic, hard-working, intelligent

Most likely you viewed Amos as the more likable person, even though the five words used are identical, just differently ordered. The initial traits change your interpretation of the traits that appear later.

[This might occur because we expect lists to be rank-ordered, but our judgment fails when they are not.]

This explains a number of effects:

- Pygmalion effect: A person's expectation of a target person affects the target person's performance. Students who randomly carry a positive expectation of being brilliant end up performing better, possibly due to teacher's expectations.
- Kahneman once graded exams by going through an entire student's test before going to the next.

He found that the first essay dramatically influenced his interpretation of later essays - an excellent first essay would earn the student benefit of the doubt on a poor second essay. He subverted this by batching by essay and iterating through all students.
- Work meetings often polarize around the first and most vocal people to speak. They would better yield the best ideas if people could write down opinions beforehand [likely party of Jeff Bezos's technique to have people prepare 6-pagers to read before discussion.]
- Witnesses are not allowed to discuss events in a trial before testimony.

Antidote:

- Elicit opinions from the group on writing confidentially, before having a public discussion.

Confirmation Bias

This materializes in a few ways:

- We selectively pay attention to data that fit our prior beliefs and discard data that don't.
- We seek out sources that tend to

Mere Exposure Effect

Already described above.

[Narrative Fallacy

Explained more in Part 2. People want to believe a story and will seek cause-and-effect explanations in times of uncertainty. This helps explain the following:

- Stock market movements are explained like horoscopes, where the same explanation can be used to justify both rises and drops (eg capturing Saddam Hussein).
- Most religions explain the creation of earth, of humans, and of the afterlife.
- Famous people are given origin stories - Steve Jobs reached his success because of his abandonment by his birth parents. Sports stars who lose a championship have the loss attributed to a host of reasons.

Once a story is established, it becomes difficult to overwrite. This helps explain why frauds like Theranos and Enron were allowed to perpetuate - observers believed the story they wanted to hear.

A good treatment of this here.**]**

Affect Heuristic

How you like or dislike something determines your beliefs about the world.

- If you like a particular option, you'll believe the benefits are better and the costs/risks more manageable than those of alternatives.

    - Interestingly, if you get a new piece of information about the benefits, you will also decrease your assessment of the risks, even though you haven't gotten any information about the risks.

We're more vulnerable to biases when System 2 is taxed.

Daniel Gilbert has a model of believing ideas:

- System 1 constructs the best possible interpretation of the belief - if the idea were true, what does it mean?
- System 2 evaluates *whether* to believe the idea - "unbelieving" false ideas.

Experiments show that when System 2 is taxed (eg when holding digits in memory), you become more susceptible to false sentences. You'll believe almost anything.

- This might explain why infomercials are effective late at night, and why societies in turmoil might apply less logical thinking to persuasive arguments (eg Germany during Hitler's rise).

Part 2: Heuristics and Biases

Part 2 of this *Thinking, Fast and Slow* summary dives further into cognitive biases.

The general theme of these biases: we prefer certainty over doubt. We prefer coherent stories of the world, clear causes and effects. Sustaining incompatible viewpoints at once is harder work than sliding into certainty. A message, if it is not immediately rejected as a lie, will affect our thinking, regardless of how reliable the message is.

Statistical Mistakes

The general theme is that **we pay more attention to the content of the story than to the reliability of the data**. We prefer simpler and coherent views of the world and overlook why those views are not deserved. We overestimate causal explanations and ignore base statistical rates. Often, these intuitive predictions are too extreme, and you will put too much faith in them.

The Law of Small Numbers

The smaller your sample size, the more likely you are to have extreme results. Do NOT fall for outliers.

- Facetiously - in a series of 2 coin tosses, you are likely to get 100% heads. This doesn't mean the coin is rigged.
- **A study found that certain rural counties in the South had the lowest rates of kidney cancer.** What was special about these counties - something about the rigorous hard work of farming, or the free open air?
    - Invert the question: where were the areas with the lowest rates of kidney cancer? They

were also rural areas! The outliers appeared merely because the populations were so small.

- The Gates Foundation studied educational outcomes in schools and found small schools were habitually at the top of the list. It tried to apply small-school practices at large (lower student:teacher ratio, smaller student body sizes).
    - They did not invert the question - what are the characteristics of the worst schools? They'd find these to be smaller than average as well.

Here System 1 is finding spurious causal connections between events, too ready to jump to conclusions that make logical sense. With a surprising result, we immediately skip to understanding causality rather than questioning the result itself.

Even professional academics are bad at understanding this - they often trust the results of underpowered studies, especially when the conclusions fit their view of the world. [Hence see Kahneman's own problem with priming studies.]

The only way to get statistical robustness is to compute sample size.

The name of this law comes from the facetious idea that "the law of large numbers applies to small numbers as well."

Focusing on the Story Rather than the Reliability

"In a telephone poll of 300 seniors, 60% support the president." If you were asked to summarize this in a few words, you'd likely end with "old people like the president."

You don't react much differently if the sample were with 150 people or 3000 people. You are not adequately sensitive to sample size.

Obviously, if the figures are way off (6 seniors were asked, or 600 million were asked) System 1 detects a surprise and kicks it to System 2 to reject. [But this can also be easily disguised, as in "6 out of 10 seniors".]

Extending this further, you don't always discriminate between "I heard from a smart friend" and "I read in the New York Times." As long as you don't immediately reject the story, you tend to accept it as 100% true.

A Misunderstanding of Randomness

People tend to expect randomness to occur regularly. For coin flips following sequences all have equal probability:

HHHTTT

TTTTTT

HTHTTH

However, sequence 3 looks far more random. Sequence 1 and 2 are more likely to trigger a desire for alternative explanations. [The illusion also occurs because there is only one such sequence of TTTTTT, but hundreds of the type like the third that we don't strongly distinguish between.]

Corollary: we look for patterns where none exist.

Other examples:

- In World War II, London was bombed evenly throughout except for a few conspicuous gaps. Some suspected German spies were in those areas and thus those areas were deliberately saved. In reality, the distribution of hits were typical of a random process.
- There is no such thing as a hot hand in basketball - these are entirely random.

Evolutionarily, this might have arisen out of a margin of safety for hazardous situations. That is, if a pack of lions seems to double, you don't think about whether this is just a random fluctuation. You assume there's a cause and you leave.

Causal Situations

Individual cases are often overweighted relative to mere statistics. The taxicab problem shows how causal base rates are overweighted relative to statistical base rates.

This was shown to great effect when psychology students were taught about troubling experiments like the Milgram shocking experiment, where 26 of 40 participants delivered the highest voltage shock.

Students were then shown videos of two normal-seeming people, not the type to voluntarily shock a stranger. How likely were these individuals to have delivered the highest voltage shock?

The students far underestimated the % chance below 26/40.

This is odd. The students hadn't learned anything at all! They had exempted themselves from the conclusions of experiments. "Surely the people who administered the shocks were deprave in some way - I would have behaved better, and normal people would as well."

The antidote to this was to reverse the order - students were told about the experimental setup, shown the videos, and only then told the outcome of how the two ordinary people they saw had failed. Their estimate of the failure rate was much more accurate.

Reversion to the Mean

Over repeated sampling periods, outliers tend to revert to the mean. High performers show disappointing results, and strugglers show sudden improvement.

Here are examples, all with their own cute causal explanations:

- A military commander has two units return, one with 20% casualties and another with 50% casualties. He praise the first and berates the second. The next time, the two units return with the opposite results. He “learns” that praise weakens performance and berating increases performance.
- Likely in the Small Numbers example above, if the Gates Foundation had measured the schools again in the next years, they’d find the top performing small schools would have lost their luster.
- Patients see a doctor/do some superstitious act/eat medicine and get better, ignoring that they would likely have returned to full health without any action. (Hence why treatments must be compared to placebo.)
- Athletes on the cover of Sports Illustrated seem to do much worse, explained by buckling under the spotlight, but really because the athlete had had an outlier year and had now reverted to the mean.
- Intelligent women tend to marry men who are less intelligent than they are, because smart men are threatened by intelligence, or smart women want to be dominant, etc.
- The majority of mutual funds do not return better than market after fees, even though individually they may do well one year or another.
- Companies at the top of Fortune’s “Most Admired Companies” finds that the firms with the worst ratings earn much higher stock returns than the most admired firms (because the top firms rest on laurels and worst companies light a fire).
- [Few startup founders, who make it big once, found another company as successful as their first.]
- How many years should you wait before concluding an investment fund or a CEO are earning above market returns?

This occurs when the correlation between two measures is imperfect. One example above can be restated, “the correlation between the intelligence scores of spouses is imperfect,” or “the correlation between year 1 and year 2 of an athlete’s career is imperfect.”

When we have a cute causal explanation for regression to the mean, we can come up with superstitions and misleading rules.

Antidote:

- When looking at high and low performers, question what fundamental aspects are correlated with their performance, and how much will revert to the mean.

Anchors

Anchoring: When you are exposed to a number, the number affects your estimate of an unknown quantity. This happens even when the number has no meaningful relevance to the quantity to be estimated.

- Students who are asked if Gandhi died before or after age 144 estimate higher ages than those asked about age 32.

- Students were shown 10 or 65 on a rigged wheel of fortune and were then asked to estimate the % of African nations in the UN. The average estimates came to 25% and 45%, respectively.
- When donations for a nonprofit requested $400, the average was $143; when requesting $5, the average was $20.
- Buy Now prices anchor online auctions.
- Arbitrary rationing, like supermarkets with “limit of 12 per person” makes people buy more cans. (This is also confounded as a signal of demand.)

[The setup seems to be regular:

- Is the [quantity to estimate] more or less than [anchor]?
- What’s your best guess about the [quantity to estimate]?

I suppose the first question is required for the anchor to be lodged in the student’s mind.]

Sometimes, the anchor works because you infer the number is given for a reason, and it’s a reasonable place to adjust from. But again even meaningless numbers, even dice rolls, can anchor you.

The **anchoring index** measures how effective the anchor is, defined as (the difference between the average guesses when exposed to two different anchors) / (the difference between the two anchors). Studies show this index can be over 50%!

- Asking about the height of the redwood anchored to 180 feet and got an answer of 282 feet; or anchored to 1200 feet and got an answer of 1020 feet.

Insidiously, **people take pride in their supposed immunity to anchored numbers**. You really don’t have full command of your cognition.

- Real estate agents claim to be immune to listing prices, when the opposite is true.

How does anchoring have its effect in the brain? There are two mechanisms:

- **System 2**: You start with the exposed number as an initial guess, then adjust in one direction until you’re not confident you should adjust further. At this point you’ve reached the edge of your confidence interval, not the middle of it.
    - Say you were given a piece of paper and asked to draw from the bottom up until you reached 2.5 inches. Then you were asked to draw from the top down until 2.5 inches were left. The first line would likely be shorter than the space on the second line - **you stop when you reach the edge of the confidence range, not the middle of it**.
    - You drive much faster coming off the highway than you would otherwise.
    - If asked about the boiling temperature of water at top of Mount Everest, you start with 100C and work downwards.
    - Evidence of System 2: People adjust less from the anchor when their mental resources are depleted.
- **System 1**: The anchor invokes associations that influence your thinking. System 1 tries to construct a world in which the anchor is the true number.

    - Thinking of Gandhi as age 144 primes associations of old age. This causes a higher estimate.
    - Evidence: Asking participants whether the average temperature was higher or lower than 68F made it easier to recognize summer words (like beach). The same was true of asking about 41F.

In negotiations, when someone offers an outrageous anchor, don't engage with an equally outrageous counteroffer. Instead, threaten to end the negotiation if that number is still on the table.

[Scott Adams takes this further into the idea space - when you hear an extreme idea (like "build the wall", you anchor to that idea so that qualifications seem like concessions.]

[Antidotes:

- **If you're given one extreme number to adjust from, repeat your reasoning with an extreme value from the other direction.** Adjust from there, then average your final two results.
- When estimating, first adjust from the anchor to where you feel like you should stop. Then deliberately go much further to the point that you want to dial it back. Average the two points.
- Move your estimate from the anchor to the minimal or maximal amount it could be.]

Availability Bias

When trying to answer the question "what do I think about X," you think about the easier heuristic questions, "what do I remember about X, and how easily do I remember it?"

More quantitatively, when trying to estimate the size of a category or the frequency of an event, you instead **report the heuristic of the ease with which instances come to mind**.

In other words:

- Items that are easier to recall may take on greater weight than they should.
- If retrieval for a category (eg "dangerous animals") is easy and fluent, the category will be judged to be large.
- You judge frequency by the ease with which instances come to mind.

This manifests in a number of ways:

- Items that trigger stronger emotions (like terrorist attacks) are more readily available than those that don't (like diabetes), causing overestimation of the former's importance.
- More recent events are more available than past events.
- **More vivid examples and personal experiences are more available than mere words or statistics**.
    - Famously, spouses were asked for their contributions to different tasks, and the sum of both spouses' answers totaled more than 100%.
    - [This might counterintuitively call to discount an opinion on a subject from someone

who's had personal experience.]
- Items that are covered more in media take on greater importance than those that aren't (even if the latter have more practical importance).

A line of experiments asked people to list a specified number of instances to a situation (such as 6 examples of when they were assertive) than to answer a question ("evaluate how assertive you are").

- Question: what has a greater effect on their perception of assertiveness, the number of examples or the ease of recall?
- Counterintuitively, it's **the ease of recall**. When people are asked to name 6 examples of their assertiveness, they feel more assertive than those asked to name 12 examples. The difficulty of scraping up the last few examples dampens one's confidence.
- [People seem to suffer from 1) overestimation of their recall ability, and 2) not knowing how many examples is needed to prove assertiveness. They could very well say, "this single example shows enough, and I know my memory is poor."]
- Similarly, **people are less confident in a position when they're asked to produce more arguments to support it**.

- [This bias can be co-opted in both directions:
    - To dampen the strength of someone's feeling, ask them to name a lot of items (eg "list 15 reasons you think climate change isn't happening")
    - To strengthen their feeling, ask them to name only a few items (eg "list 3 ways your alma mater has contributed to your success")]

There are some exceptions to the above effect:

- This effect dissipates when people are given a cover story for their lack of recall (eg listening to this music will impair your recall ability -> rated themselves as equally assertive)
    - Profoundly, System 2 can change the surprise of System 1 (much like being told beforehand "the man you're about to meet has 7 fingers, don't mind it.")
- This effect reverses when someone has personal experience with the situation.
    - When asked about preventative behaviors for disease, people with a family of heart disease felt safer when they retrieved many instances, rather than being subject by ease of recall.

The conclusion is that System 1 uses ease of recall as a heuristic, while System 2 focuses on content.

- Thus, you're more susceptible to availability bias when you're in a good mood, when System 2 is being taxed, when you're a novice rather than an expert, and **if you're made to feel powerful**.
- [Thus when making a heavy decision, don't focus on the times that you succeeded as the good feelings might make you feel overconfident.]

Availability bias interacts with the media, and it can cause a vicious cycle where something minor gets blown out of proportion:

- A minor curious event is reported. A group of people overreact to it.
- News about the overreaction triggers more attention and coverage of the event.

    - Media is looking for eyeballs.
    - The reporters themselves are subject to availability bias.
- This continues snowballing as increasingly more people see this as a crisis.
- The naysayers are rejected as participating in a coverup.
- This can in turn affect real policy, where scarce resources are used to solve an overreaction rather than a quantitatively more important problem.
    - (Kahneman notes that even irrational fear is debilitating, and policymakers need to protect the public from fear, not just real dangers.)

Thus we tend to weigh small risks either as zero or far too large.

- When you're waiting for your teenage child to come home at night, **you're thinking about the numerator, not the denominator**.

[Antidotes:

- Instead of judging the number of instances or frequency, assign weights to their significance.
    - Eg when thinking about reasons to and against quitting your job, score them by significance rather than counting the reasons or judging ease of recall.
    - Eg when thinking about deaths by lightning streaks or diabetes, estimate the # of people from first principle.
- Think about the denominator, not just the numerator.]

Representativeness

"Tom W. is meek and keeps to himself. He likes soft music and wears glasses. Which profession is Tom W. more likely to be? 1) Librarian. 2) Construction worker."

If you picked librarian without thinking too hard, you used the **representativeness** heuristic - you matched the description to the stereotype, without thinking about the base rates. More generally, representativeness is when we estimate the likelihood of an event by comparing it to an existing prototype in our minds - matching like to like.

Just because something is plausible does not make it more probable.

Ideally, you should have examined the **base rate** of both professions in the male population, then adjusted based on his description. Construction workers outnumber librarians by 10:1 in the US - there are likely to be more shy construction workers than all librarians.

And **even when base rate data are given**, the representativeness is still weighted more than the statistics. And even when people are told the information isn't trustworthy, because of WYSIATI, people don't discount the information.

[Even after reading this, you might think - but what about self-selection of meek people into libraries and away from construction? That goes to show how entrenched the representativeness heuristic is.]

Another one: “Someone on the subway is reading the New York Times. Is the stranger more likely to have a PhD, or to not have a college degree?”

Again, by raw count, there are many more people in the latter group than the former.

Representativeness is used because System 1 desires coherence, and matching like to like forms a coherent story that is simply irresistible. **Take away this convenient story, and System 2 is engaged** - when students are asked “estimate the % of employees in these lines of work” without any narrative context, the guesses are far more accurate.

Insidiously, the representativeness heuristic works much of the time. An athlete who’s thin and tall is more likely to be in the NBA than the NFL; people who act friendly are indeed friendly by nature.

The taxicab problem is another fun example where subjects confuse the probability of events.

[Other examples that come to mind:

- For a role, you hire based on stereotype rather than real performance. Like punishing an engineer candidate for being fashionable and gregarious.
- Any test situation that is only partially representative of the situation you care about, like how a job interview is only partially representative of on-job performance.
    - Google once used brain teasers in hiring engineers, assuming that creativity under stress in an interview would also be relevant on the job. They later found brain teasers poorly predictive.
    - While in the Israeli army, Kahneman was tasked with evaluating officer candidates in a teamwork exercise (getting a log over a wall without touching the wall). He felt so sure about his predictions (“this person is a certain star” “this person is mediocre”) were only slightly better than chance.
- Invocation of any stereotype. Eg only women can be rape victims, not men. Candidates are judged by their facial structure rather than their policy.
- The startup by analogy: “Uber for X” companies or “Y on the blockchain.”
- Pattern matching in general - patterns of successful and failed companies.

]

Antidote:

- **Examine the base probability of the situation.** Try to reason from first principle on the factors that matter and whether the instance has those factors (eg “what skills do we need in this job, and does the candidate display them, and how do we know?”).
- Apply Bayesian statistics. A simple form:
    - Start by predicting the base rates.
    - Consider how the new data should adjust the base rates.
        - Eg Tom W is 4 times more likely to be a librarian than construction worker.

Conjunction fallacy

“Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?

1. Linda is a bank teller.
2. Linda is a bank teller and is active in the feminist movement.”

If you guessed 2, you fell for the conjunction fallacy. 1 is clearly a superset of 2, so 1 should always be more likely. However, 2 explicitly mentioned a coherent story and thus seemed more representative of Linda, even though being more statistically unlikely.

(Don’t feel bad - over 85% of undergrads chose the second option. Even more fail when the options are surrounded by other vocations, like “teacher” and “insurance salesperson.”)

Compare these two events:

“A massive flood in North America in which more than 1,000 people drown.

An earthquake in California that causes a flood in which more than 1,000 people drown.”

The latter is more plausible because of the vividness of its detail, but it is certainly less probable. This is a problem for forecasters - **adding details to scenarios makes them more persuasive, but less likely to come true**.

Yet this fallacy is not invoked in a more detailed story without coherence:

“Which is more probable?

Mark has hair.

Mark has blond hair.”

More examples here.

Interestingly, there is a way to debias people by discrete quantities:

“There are 100 people who fit the description above. How many of them are:

- Bank tellers? __ of 100
- Bank tellers and active in the feminist movement? __ of 100”

Antidotes:

- When hearing a complicated explanation that has too many convenient assumptions or vivid details (like an investment pitch), beware of how each additional assumption lowers the likelihood.
- When asked to estimate probability or percentages, think about a finite number, like 100 people.

Sets and Averages

Which is more valuable?

- 24 dinner plates
- 30 dinner plates, 6 of them broken

When viewed together, the question is easy. #2 should be strictly more valuable.

But when viewed separately, people are willing to pay less for #2 than #1. The presence of broken dinner plates “pollutes” the set, and people average the whole set less. This was also shown when mixing common cards with star cards lowered their price.

Antidote:

- Build value from the ground up, from first principle. Think about the sum of each individual item, rather than the average.

Overcoming the Heuristics

“Julie is currently a senior in a state university. She read fluently when she was four years old. What is her grade point average (GPA)?”

People often compute this as follows:

- Reading fluently at 4 puts her say at the 80th percentile.
- The 80th percentile GPA is somewhere around a 3.7.
- Thus Julie likely has a 3.7 GPA.

Notice how misled this is! People feel confident to predict an outcome 2 decades later because of appearances of causality and intensity matching.

Antidote: The proper way to answer questions like these is:

- Start with an estimate of average GPA.
- Determine the GPA that matches your impression of the GPA.
- Estimate the correlation between your evidence and GPA.
- Adjust from the average to your estimate by the correlation strength (eg if correlation is 1.0, move 100% toward it).

This approach is generalizable. It avoids overly extreme results from intuition, uses base rates, and assesses the quality of information. It allows for regression toward the mean (eg replace “average GPA and student’s GPA” with “day 1 golf score and day 2 golf score”).

However, Kahneman notes that absence of bias is not always what matters most. A reliance in statistics would avoid the prediction of rare or extreme events on shaky information. This would paralyze venture capitalists, who lose only 1x their money on a bad investment but make 1000x their money on a Google.

Similarly, when the evidence is against you but it feels right (eg divorce rate or startup success), you might be happier deluding yourself with extreme predictions. Following our intuitions is more natural and sometimes more pleasant than acting against them. **But at the least, be aware of how much you know**.

Part 3: Overconfidence

Part 3 of this *Thinking, Fast and Slow* summary explores biases that lead to overconfidence. When we have satisfying stories about the world, we overestimate our ability to predict the future.

The Illusion of Understanding

WIth all the heuristics and biases described above working against us, we vastly overestimate how much we understand about the past, present, and future.

The general principle has been this: we desire a coherent story of the world. This comforts us in a world that may be **largely random**. If it’s a good story, you believe it.

Insidiously, the fewer data points you receive, the more coherent the story you can form (if you don’t notice how little information you actually have and wonder about what is missing). We focus on the data we have, and don’t imagine all the events that failed to happen (nonevents). You ignore your ignorance.

And **even if you’re aware of the biases**, you are nowhere near immune to them, and often exempt yourself for being smart enough to avoid them.

The ultimate test of an explanation is whether it can predict future events accurately.

Narrative Fallacy

We desire packaging up a messy world into a clean-cut story. It is unsatisfying to believe that outcomes are based largely on chance, partially because this makes the future unpredictable. But in the presence of randomness, “regular patterns can only be mirages.”

Examples:

- History is presented as an inevitable march, rather than a chaotic mishmash of influences and people. (If the past were so easily predictable in hindsight, then why can’t you predict the future?)
- Management literature profiles the rise and fall of companies, attributing the growth to key

decisions and leadership styles, even founder childhood traits. The stories **ignore all the other things that *didn't* happen** that could have caused the company to fail (and that did happen to the many failed companies that aren't profiled - survivorship bias). The stories are presented as inevitabilities.
    - The book *Built to Last* profiled companies at the top of their field; these companies did not outperform the market after the book was published.
- Management literature also tries to find patterns to management systems that predict success.
    - The success of the firm and the quality of its CEO might be as high as .30, which is lower than what most people might guess. This .30 suggests that the stronger CEO would lead the stronger firm in 60% of pairs, just 10% better than chance.
- We readily trust our judgments in situations that are poor proxies for real performance (like in job interviews).

Funnily, in some situations the same explanation can be applied to either outcome.

- An event might happen (the fed lowers interest rates). If the stock market goes up, people say investors are emboldened. If the stock market goes down, people say they were expecting greater movement, or the market had already priced it in.
- A CEO is known for his methodical, rigid style. If his company does well, they compliment him for sticking to his guns with clearly the right strategy. If the company falters, they blame his inflexibility to adapt to new situations.

Even knowing the narrative fallacy, you might *still* be tempted to write a narrative - eg successful companies become complacent, while underdogs try harder, so that's why reversion to the mean happens. Kahneman says this is the wrong way to think about it - the gap must shrink because part of the outcome was due to luck. It is pure statistics.

[When an explanation reads like a horoscope in this way, then the explanation should be discounted heavily. You can detect it by inverting the situation and seeing if the same explanation applies.]

There are obviously factors that correlate somewhat with outcomes. The founders of Google and Facebook are likely more skilled than the lower quartile of startups. Warren Buffett's massive reading is likely a good contributor to success, and you'd be more successful if you did it too. **The key question is - how strong is the correlation?**

Clearly a professional golfer can beat a novice perhaps 100% of the time. However, skill is the dominant factor here - the correlation is very high, so predictability is very high. If you took the management principles espoused in the business literature and tried to predict company outcomes, you might find they predict little.

Antidotes:

- **Be wary of highly consistent patterns** from comparison of successful and less successful examples. You don't know lots of things - eg whether the samples were cherrypicked, whether the failed companies are not part of the dataset.
- Be wary of people who declare very high confidence. This suggests they've constructed a

coherent story, not that the story is true.

An Aside about Startup “Best Practices”

[I enjoy reading startup narratives, and confronting narrative fallacy is straining. Much of startup advice is “standard practice” - “you should have at least 2 cofounders;” “culture is critical;” “raise money per the standard seed - series A - B;” “don’t charge money, this will stifle your growth.”

Among successful startups, certain attributes are nearly universal (the founders have personal experience with the problem) while others show a wide diversity (number of founders, funding method).

If luck is the primary factor, then the distribution of attributes should reflect the natural population of founders, and no more. If the attribute is nearly universal, then it likely means people with the attribute self-select as founders, and the attribute isn’t actually causative.

- Founders often search widely for a problem to solve, so it makes sense for them to notice problems they have experience with. This doesn’t necessarily mean that solving your own problem is necessary for success.
- Founder count ranges widely (from solo to two to more), suggesting each can work in its own way.

Conversely, there are obviously thousands of companies that follow best practices (hire well, promote a strong culture, be located in SF) that fail. So these attributes are clearly not sufficient, and possibly not at all predictive.

There remain just a few attributes that I believe are predictive, but so much so to be tautological - you need to execute, you need to make something people want to buy, and you can’t run out of money.

What does this mean as a founder? Focus on the factors that make you happy; realize that a lot of things are unpredictable and out of your hands.]

Hindsight Bias/Outcome Bias

Once we know the outcome, we connect the dots in the past that make the outcome seem inevitable and predictable.

Insidiously, **you don’t remember how uncertain you were in the past** - once the outcome is revealed, you believe your past self was much more certain than it actually was! It might be even difficult to believe you ever felt differently. In other words, “I knew it all along.” You rewrite the history of your mind.

- Imagine all the people who believe they saw the 2000 dotcom bubble burst, or the 2008 financial crisis.
- Professional forecasters (eg experts who show up on talk shows) perform no better than chance in predicting events. This might also be because they’re hired for their vocalness, not accuracy.

This is a problem because **it inflates our confidence about predicting the future**. If we are certain that our past selves were amazing predictors of the future, we believe our present selves to be no worse.

Another problem is that **it rewards and punishes people based on outcome**, not on their prior beliefs and their appropriate actions. **People who made the right decision that led to failure are punished more than those who took irresponsible risks that worked out**.

- An experiment used a legal case to ask subjects whether the city should have done a certain preventative action. When exposed to the evidence at the time, 24% of people felt they should have taken the action. When exposed to the *outcome*, this rose to 56% of people!
- The government was castigated for ignoring information about al-Qaeda in July 2001 [but the popular story ignores the volume of other threats they had at the time.]
- The natural consequence to this is bureaucracy - **if you are to be judged by your decisions, then it's better to follow procedure, and be reluctant to take risks**.
    - [In large companies, this suggests that good risk taking should be promoted, not the outcome. Perhaps a rating of people as they launch projects, before the outcomes are known.]

[Antidotes:

- Keep a journal of your current beliefs and your estimation of the outcomes.
- Reward people based on the decisions they make at the time, before the outcomes come out.

]

We Mostly Don't Accept These Biases

Even when presented with data of your poor predictions, you do not tend to adjust your confidence in your predictions. You forge on ahead, confident as always, discarding the news.

- Kahneman argues the entire industry of the stock market is built on an illusion of skill. People know that on average, investors do not beat market returns (by definition). And studies show that retail investors trade poorly - they sell stocks to lock in the gains and hang on to their losers out of hope. Large professional investors are happy to take advantage of these mistakes. But retail traders continue marching on.

Here are several reasons it's so difficult to believe randomness is the primary factor in your outcomes, and that your skill is worse than you think:

- Pride and ego are at stake.
    - The more famous the forecaster, the more overconfident and flamboyant the predictions.
    - Experts resist admitting they're wrong, instead giving excuses: wrong only in timing, an unforeseeable event had intervened, wrong but for the right reasons.
- You take deliberate, skillful steps to guide the outcome. A lot of motion can't be wrong!
    - Stock analysts pore over financial statements and build models. This requires lots of training. But this doesn't answer the real question - is the information about the form

already priced into the stock?
    - **Managers focus on the strength of their strategy and how good their company seems**, discounting what their competitors are doing and market changes ("competition neglect").
    - Your experience shows many instances where your predictions came true - partially because those are most available to you and you explain away your mistakes.
- You focus on the causal role of skill and neglect the role of luck - the illusion of control.
- You don't know what you don't know.
- Large monetary incentives at stake. You are being paid for your skill - if your skill turns out to be irrelevant, you'll lose your job.
    - CFOs tried to predict the 80% confidence interval of stock market returns. This expects 20% surprises, but the actual count was 67%. The real answer was "between -10% and +30%" - but any CFO who says this would be penalized for admitting it.
    - It's a sign of weakness to be unsure. It can have material marketing consequences (eg to customers or investors) or to staff (who want more certainty).
- Strong social proof in your community - it can't be a fraud if all these other smart people believe it too.
- Emergencies call for decisive action. (People's fear of uncertainty is exacerbated in stressful situations.) Ambivalent decision making is anathema during these times.
- [Much of your identity is based around mastery and achievement. If this is largely a farce, then what purpose did all that work serve? A well-meaning mutual fund manager has to wonder why she's even doing what she's doing, and whether she's hurting customers.]
- [If outcomes are really random, then we may owe more to the unlucky to help them out. It's not strongly their fault they have their circumstances.]

Intuitions vs Formulas

Humans have to make decisions from complicated datasets frequently. Doctors make diagnoses, social workers decide if foster parents are good, bank lenders measure business risk, employers have to hire employees.

According to *Thinking, Fast and Slow*, humans are also surprisingly bad at making the right prediction. **Universally in all studies, algorithms have beaten or matched humans** (and even when matching, it's a win because it's so much cheaper).

Why are humans so bad? Simply, **they overcomplicate things.**

- They inappropriately weigh factors that are not predictive of performance (like whether they like the person in an interview).
- They try too hard to be clever, considering complex combinations of features when simply weighted features are sufficient.
- Their judgment varies moment to moment without realizing. System 1 is very susceptible to influences without the conscious mind realizing.
    - The person's environment, current mood, state of hunger, recent exposure to information can all influence decisions. Algorithms don't feel hunger.
    - Radiologists read the same X-ray twice and give different answers 20% of the time.

- Even when given data from the formula, humans are bad! They feel falsely they can “override” the formula because they see something that’s not accounted for.

Simple algorithms are surprisingly good predictors. Even equally weighted formulas can be as accurate as multiple-regression formulas, since they avoid accidents of sampling.

- How do you predict marital stability? Frequency of sex - frequency of arguments.
- Susan Apgar pioneered a simple quantitative score to detect troubled newborns. Before this, doctors used their (poor judgment).
- To predict wine prices, an economist used the summer temperature and rainfall. This was accurate, but oenophiles were aghast - “how can you price the wine without tasting it?!” The formula was in fact better because he did *not* taste it.

There is still some stigma about life being pervaded by robotic algorithms, but this is eroding as algorithms recommend us useful things to buy and forming winning baseball teams.

- Professionals feel outrage because many of their predictions do turn out right (also because of confirmation bias), and they are skilled, but they don’t know the boundaries of their skill.
- It seems more heart-wrenching to lose a child because of an algorithm’s mistake than because of a human error.

Antidote: When hiring, Kahneman recommends **standardizing the interview**:

- Make a list of up to 6 traits important for success in the role. Ideally these are orthogonal.
- Create direct questions to assess each trait.
    - For conscientiousness, you might ask, “how many times were you late to work in the past month?”
- Create a rubric from 1 to 5, with notes on what you’re looking for in each grade.
- When interviewing, collect information on one trait at a time, completing it before moving on.
- Strictly hire the person with the highest score. Do not override this - you’re letting the halo effect and liking bias override you.

[This is similar in concept to the Job Scorecard, espoused in the book Who]

When Can You Trust Human Intuition?

Clearly people develop skilled intuitions. Chess players do spot meaningful moves; doctors do make correct diagnoses. The school of Naturalistic Decision Making puts faith in human intuition.

When can you trust human intuition? In *Thinking, Fast and Slow*, Kahneman argues accurate human intuition is developed in situations with **two requirements**:

- **An environment that is sufficiently regular to be predictable, with fast feedback**

- Prolonged practice to learn these regularities

Examples:

- Most people can learn to drive. The input -> output relationship is clear and immediate.
- Training can even occur theoretically, through words or thoughts. You can simulate situations in your brain to learn, much like a young military commander can feel tension when going through a ravine for a first time, because he learned this favored an ambush.

Not all supposed experts have real predictive skill. **The problem with pundits and stock pickers is that they don't train in predictable environments.** When noise dominates the outcomes and feedback cycles are long, any confidence in the intuition's validity is largely illusory.

- Even worse, there can be "wicked" environments, where you learn the wrong lessons from experience if you influence the outcome. For example, Lewis Thomas felt he could predict typhoid by touching the patient's tongue, without realizing he carried typhoid on his hands.

In the brain, how do accurate intuitive decisions? They first arise from pattern matching in memory - System 1 retrieves a solution that fits the situation. System 2 then analyzes it, modifying it to overcome shortcomings until it seems appropriate. If the solution fails, another solution is retrieved and the process restarts.

To the NDM camp, Kahneman concedes that in the presence of clear signals, the formula does not identify critical factors that humans do not - humans are efficient learners and generally don't miss obvious predictors. However, algorithms do win at detecting useful signals within noisy environments.

Supreme confidence without good explanation may also be a sign of untrustworthy intuition.

- True experts know the limits of their knowledge. People who are unshakingly confident (like pundits on TV) may be selected more for their boisterousness than their accuracy.
- How confident you feel about your intuition is not a reliable guide to its validity. You need to learn to identify situations in which intuition will betray you.
- You might conflate short-term results with long-term results. A psychiatrist may feel skilled in building rapport with patients, but this doesn't correlate strongly with long-term outcomes, which take in many more factors.

The Outside View

We are often better at analyzing external situations (the "outside view") than our own. When you look inward at yourself (the "inside view"), it's too tempting to consider ourselves exceptional - "the average rules and statistics don't apply to me!" Information about relevant population statistics are readily discarded.

"People who have information about an individual case rarely feel the need to know the statistics of the class to which the case belongs." And when the statistics are given, they are routinely discarded when they conflict with one's personal impressions.

Examples:

- 90% of drivers state they're above average drivers. Here they don't necessarily think about what "average" means - they think about whether the skill is easy for them, and intensity match to where they fit the population.
    - [This might also be a semantics problem. When strangers ask me how my day's going, I might say "it's an average day." Alarmed, they ask, "what's wrong?" clearly thinking "average" means "bad." It could be that people are conditioned to think of "average" as a C in school, and a C as bad. Thus most people think they're "not bad" drivers.]
- Lawyers may refuse to comment on the outcome of a case, saying "every case is unique."
- Business owners know that only 35% of new businesses survive after 5 years. Despite this, 82% of entrepreneurs put their personal odds of success at 70% or higher, and 33% said their chance of failing was zero!
    - Kahneman tells a story of meeting motel owners who said they bought the motel cheaply because "the previous owners failed to make a go of it."
- Most people believe they are superior to most others on most desirable traits.
- CEOs make large M&A despite a poor historical track record.

The **planning fallacy**: you habitually underestimate the amount of time and resources required to finish a project. You don't know what you don't know - the emergencies, loss of motivation, obstacles that will delay. You tend to give "best case scenario" estimates, rather than confidence ranges.

- In *Thinking, Fast and Slow*, Kahneman gives an example of a curriculum committee meeting to plan a book. They happily estimate 2 years for completion of the book. He asks the editor how long other teams have taken (7-10 years, with 40% of teams failing to finish). He asks how their team skill compares to the other teams (below average).
    - This was an astounding example of how a person may have relevant statistics in her head, but **completely fails to recall this data as relevant for the situation**. (The book later did indeed take 8 years).
    - Furthermore, **the team didn't feel they needed information about other teams to make their guess**! They looked only at their own data situation.
- Government projects have a funny pattern of being universally under budget and delayed. (This may be incentivized, as projects that are lower cost and shorter time are easier to get through.)

Antidote:

- Similar to the correction for heuristics in the last section:
    - Identify the appropriate category the situation belongs in.
    - Consider the population distribution for the category. Start here with this baseline prediction.
    - Use specific information about the case that allows you to adjust from the baseline prediction.
- **Premortem**: "Imagine that it's a year from now. We implemented the plan. It was a disaster. Write a brief history of the disaster."
    - This overcomes the public groupthink coming from public discussion; it encourages creative dissent, which is normally suppressed when a project begins.

- When evaluating project execution, reward people who finish according to their original deadlines, not those who finish much earlier or later than planned. [Though this may incentivize sandbagging for people who were overly conservative.]

On Optimism and Entrepreneurs

Optimism is largely genetic (though Seligman believes it can be learned). Optimistic people are happier, recover from setbacks more easily, have greater self-confidence, feel healthier and live longer.

Optimistic people play a disproportionate role in shaping the world - the inventors, entrepreneurs, political leaders. They take risks, seek challenges, are talented but are also lucky (luckier than they acknowledge). Their success confirms their faith in their judgment and their ability to control events. [In *Principles*, Ray Dalio describes these people as shapers, with traits like "are resilient and gritty, since their need to achieve is stronger than the pain they experience."]

As described above, most founders know the statistics - most startups fail, and the higher expected value path is to sell their services to an employer. To forsake the latter path and start a company, you need to be overoptimistic or deluded [or not be maximizing for income].

One drawback to optimism is that it encourages people to take outsized risks because they overestimate their chances of success.

- Of inventors who were told their inventions were destined to fail, half ignored the advice and continued, doubling their original losses.
- More optimistic CEOs (own more equity in the company) take excessive risks and lower company performance. Furthermore, prestigious press awards cause companies to underperform.
- Movie studios release big budget movies on the same day (eg July 4th) because they ignore the competition - "we've got a great story and marketing department, so it's going to do great." The real question is - "considering the market, how many people will see our film?"
- Because of outcome bias, people who take extraordinary risks and have it work are lauded for their prescience (and serve as a model for new entrepreneurs) - all the failures are forgotten.

Kahneman notes that yet there still might be value here in the legions of entrepreneurs who try and fail. They perform a market discovery service, figuring out pockets of opportunity that larger companies can later service - dying as "optimistic martyrs."

And overall, mixing high optimism with good implementation is positive. It allows endurance through setbacks and belief in what one is doing.

Part 4: Choices

Part 4 of this *Thinking, Fast and Slow* summary departs from cognitive biases and toward Prospect Theory, which covers risk aversion and risk seeking, our inaccurate weighting of probabilities, and sunk cost fallacy.

Prospect Theory

Prior Work on Utility

How do people make decisions in the face of uncertainty?

Traditional “expected utility theory” asserts that people are rational agents that calculate the utility of each situation and makes the optimum choice each time.

- If you preferred apples to bananas, would you rather have a 10% chance of winning an apple, or 10% chance of winning a banana? Clearly you’d prefer the former.
- Similarly, when taking bets, this model assumes that people calculate the expected value (EV) and choose the best option.

This is a simple, elegant theory that by and large works and is still taught in intro economics. But it failed to explain the phenomenon of **risk aversion**, where in some situations a lower-EV choice was preferred

- Consider: Would you rather have a 80% chance of gaining $100 and a 20% chance to win $10, or a certain gain of $80? The expected value of the former is greater (at $82) but most people choose the latter. This makes no sense in classic utility theory - you should be willing to take a positive EV gamble every time.

To address this, in the 1700s, Bernoulli argued that 1) people dislike risk, and that 2) people evaluate gambles not based on dollar outcomes, but on their psychological values of outcomes - their utilities.

Bernoulli then stated argued **utility and wealth had a logarithmic relationship**. The difference in happiness between someone with $1,000 and someone with $100 was the same as $100 vs $10. On a linear scale, money has diminishing marginal utility.

- Further, this meant that $10 was worth more to someone with $100 than to someone with $10.
- This explained the value of certainty in gamble problems.
    - On a logarithmic scale for utility, having 100% of $80 was better than having 80% of $100.
- This explained insurance - people with less wealth were willing to sell risk to the wealthier, who would suffer less relative utility loss in the insured loss.

But this still presented intuitive problems in other cases.

- Say Anthony has $1 million and Beth has $4 million. Anthony gains $1 million and Beth loses $2 million, so they each now have $2 million. Are Anthony and Beth equally happy?
    - Obviously not - Beth lost, while Anthony gained. Yet Bernoulli would argue they end up at the same logarithmic utility.
- Now give Anthony and Beth the following choice:
    - 1) 50% chance of ending with 1 million or 50% chance of ending with 4 million
    - 2) 100% chance of ending with 2 million

- Option 1 has an expected value of $2.5 million. Bernoulli would argue that's the clear choice for both parties.
- But Anthony is more inclined to choose option 2, while Beth is more likely to choose option 1. Anthony sees the certain doubling of his wealth as attractive, while Beth sees the certain loss of half her wealth as very unattractive. She would rather take the gamble to preserve her wealth.

Prospect Theory

The key insight from the above example is that evaluations of utility are not purely dependent on the current state. **Utility depends on changes from one's reference point.** Utility is attached to *changes* of wealth, not *states* of wealth. **And losses hurt more than gains.**

- Consider that you probably don't know your wealth to the nearest hundred, or even thousand. But the loss of $100 - from an overcharge, or a parking ticket - is very acute. Isn't this odd?
- (I personally find the experimental problems more meaningful when I multiply the numbers to get from $100 to $10,000. In low numbers, it's easier for me to do EV, but I really become risk averse when the numbers get higher.)

Consider these problems:

1. You have been given $1,000. Which do you choose:
    1. 50% chance to win $1,000, or get $500 for sure
2. You have been given $2,000. Which do you choose:
    1. 50% chance to lose $1,000, or get $500 for sure

Note these are completely identical problems. In both cases you have the certainty of ending with $1,500, or equal chances of having $1,000 or $2,000. Yet you probably chose different answers.

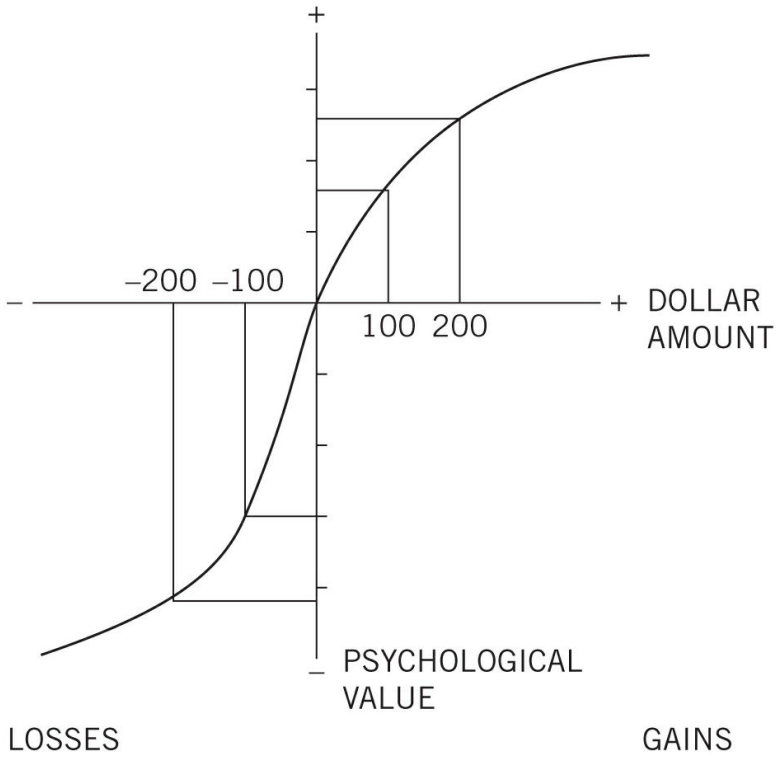
You were probably risk averse in problem 1 (choosing the sure bet), and risk seeking in problem 2 (choosing the chance). This is because your reference points were different - from one point you were gaining, and in the other you were losing. And losses hurt more than gains. Prospect theory explains why.

Prospect theory in 3 points:

- **Evaluation is relative to a neutral reference point.**

    Usually this is the status quo, but it can also be a outcome you expect or feel entitled to (like an annual raise).
- **Diminishing marginal utility applies to changes in wealth** (and sensory inputs).
    - Going from $100 to $200 feels much better than going from $900 to $1,000.
- **Losses loom larger than gains**.
    - Evolutionarily, the organisms that treated threats more urgently than opportunities tended to survive and reproduce better. We have evolved to react extremely quickly to bad news.
    - A friendship that takes years to develop can be ruined by a single action.

The master image of prospect theory is this:



Note that loss aversion is shown as a steeper slope to the left of the y-axis. The “loss aversion ratio” for most people ranges between 1.5 to 2.5 - people would have to gain $200 to offset a loss of $100.

- Professional risk takers (eg traders) are more tolerant of losses, possibly because they are inured to fluctuations.

The diminishing marginal utility of losses explains different risk preferences for gains and losses:

- 100% of 100 is larger than 50% of 200, so people are risk averse to get gains.
- 50% of -200 is less negative than 100% of -100, so people are risk seeking to avoid losses.

In other words, while Bernoulli presented utility as an absolute scale starting from 0, prospect theory calibrates the curve to the reference point.

[Note this is true of experience in general - I find that my mood is heavy influenced by the [reality - expectations]. In other words, surprises (positive and negative) are what change my mood, not the outcome itself.]

Probabilities are Overweighted at the Edges

Consider which is more meaningful to you:

- Going from 0% chance of winning $1 million to 5% chance
- Going from 5% chance of winning $1 million to 10%

Most likely you felt better about the first than the second. **The mere *possibility* of winning something (that may still be highly unlikely) is overweighted**. As Jim Carrey said in Dumb and Dumber, in response to a woman who gave him a 1 in million shot at being with her: “so you’re telling me there’s a chance!”

Examples:

- We fantasize about small chances of big gains
    - Lottery tickets and gambling in general
    - A small sliver of chance to rescue a company is given outsized weight
- We obsess about tiny chances of very bad outcomes
    - The risk of nuclear disasters and natural disasters is overweighted
    - We worry about our child coming home late at night, though rationally we know there’s little chance anything bad happened
- People are willing to pay disproportionately more to reduce risk entirely
    - Parents were willing to pay 24% more per bottle of insect spray to reduce risk of poisoning by ?, but were willing to pay 80% as much to reduce it to 0.
    - People are not willing to pay less than half price for half the insurance coverage (eg earthquake insurance only on odd days)
    - People would rather take a 25% chance to lose $200 than a sure loss of $50. However they are willing to buy $50 insurance to guard against the loss.

Now consider the opposite end of probability:

- In a surgical procedure, going from 90% success rate to 95% success rate.
- In a surgical procedure, going from 95% success rate to 100% success rate

Most likely you felt better about the second than the first. Outcomes that are almost certain are given less weight than their probability justifies.

Examples:

- Taking settlements for strong cases, because the small chance of a loss is overweighted.

[Each scenario can be inverted, converting the possibility of a loss into a gain and vice versa. For example:

- A company has a 100% chance of failure, but a new project reduces that to 99%. The sensation is that the chance of failure is reduced much more than 1%. Inversely, a project that increases the rate of success from 0% to 1% seems much more likely to work than 1% suggests.

]

Below is the empirical decision weight for each probability.

Probab	0	1	2	5	10	20	50	80	90	95	98	99	100

ility (%)												
Decision weight	0	5.5	8.1	13.2	18.6	26.1	42.1	60.1	71.2	79.3	87.1	91.2	100

- See that the weight for 1% is 5.5 (when it should rationally be 1.0) and the weight for 99% is 91.2 (far from 99). People overestimate low probabilities and discount high probabilities that aren't certain.
    - Imagine that you're told your operation is 99% successful with a 1% rate of permanent disability. You will certainly be more anxious than the miniscule 1% suggests!
- Also note that 50% is weighted at 42.1 - so 50-50 coin tosses really aren't valued at 50-50! The loss aversion is higher.
- The range between 5% and 95% ranges in weight from 13 to 79 (range of 66 vs the rational 90) - we are not adequately sensitive to changes in probability!
    - We can't easily tell the difference between a 0.001% cancer risk and 0.00001% risk, even though this translates to 3000 cancers for the US population vs 30.

The Fourfold Pattern

	GAINS	LOSSES
HIGH PROBABILITY Certainty Effect	95% chance to win $10,000 Fear of disappointment RISK AVERSE Accept unfavorable settlement Example: Lawsuit settlement	95% chance to lose $10,000 Hope to avoid loss RISK SEEKING Reject favorable settlement Example: Hail mary to save failing company
LOW PROBABILITY Possibility Effect	5% chance to win $10,000 Hope of large gain RISK SEEKING Reject favorable settlement Example: Lottery	5% chance to lose $10,000 Fear of large loss RISK AVERSE Accept unfavorable settlement Example: Insurance

Putting it all together - two factors are at work in evaluating gambles:

- Diminishing sensitivity (the slope decreases as you move away from the y-axis)
- Inaccurate weighting of probabilities

In the first row of this table, the two factors work in the same direction:

- Loss aversion already occurs because 100% of -900 is more negative than 90% of 1,000 on the utility graph (diminishing sensitivity makes the sure loss more aversive).
- Even worse, while 100% is weighted at 100, 90% is weighted only at 71! (The certainty effect reduces the aversiveness of the gamble; or makes the certain loss more grave than it should.)
- Similarly, in the positive, diminishing sensitivity makes a sure lower gain more attractive, and the certainty effect reduces the attractiveness of the gamble.

In the bottom row, **the two factors work in opposite directions**:

- In the lower left corner, diminishing sensitivity still makes the sure gain more attractive than the *chance* of a gain. But the overweighting of low probabilities overcomes this effect.

Miscellaneous

Opposing Incentives in Litigation

Litigation is a nice example of where all of the above can cause tumult:

- Plaintiffs file frivolous lawsuits. They have a low chance of winning, and they overweigh the probability (lower left corner in fourfold pattern).
- Defendants prefer settling to lower the risk of a more expensive loss (lower right corner)
- Yet if the defendant does this habitually for each lawsuit, it can be costly in the long run.

Reverse the situations:

- A plaintiff has a strong case may and an almost certain chance to win, but wants to avoid the small chance of a loss, so is prone to risk-aversively take a settlement.
- The defendant knows it’s likely to lose, but has a chance of winning. They’re willing to drag the case on because the offer of a settlement from the plaintiff is worse than the worst outcome possible.
- In this case, the defendant holds the stronger hand, and plaintiff will settle for less than the case strength suggests.

Theory-Accepted Blindness

Why did it take so long for someone to notice the problems with Bernoulli’s conception of utility? Kahneman notes that you have accepted a theory and use it as a tool in your thinking, it is very difficult to notice its flaws. Inconsistencies are reasoned away with the impression that the prevailing model somehow takes care of it, and that so many smart people over time can’t all be wrong.

Even when people noticed inconsistencies, as with Allais’s paradox, they tried to bend utility theory to fit the problem. Kahneman and Tversky instead abandoned the idea that people are rational choosers, and took a psychological bent.

Blind Spots in Prospect Theory

Prospect theory has holes in its reasoning as well. Kahneman argues that it can't handle disappointment - that not all zeroes are the same. Consider:

- 1% chance to win $1 million and 99% chance to win nothing
- 99% chance to win $1 million and 1% chance to win nothing.

In both these cases, prospect theory would assign the same value to "winning nothing." But losing in case 2 clearly feels worse. The high probability has set up a new reference point - possibly at say $800k.

It also can't handle regret, in which failing to win gambles causes losses to become more painful.

More complicated models do factor in regret and disappointment, but they haven't yielded enough novel findings to justify the extra complexity.

Variations on a Theme of Prospect Theory

Indifference Curves and the Endowment Effect

Basic theory suggests that people have indifference curves when relating two dimensions - like salary and number of vacation days. Theoretically, you should be willing to trade for any other portion of the indifference curve at any time.

As Kahneman says in *Thinking, Fast and Slow*, this doesn't ring true in intuition. People have inertia to change.

Thinking, Fast and Slow calls this the **endowment effect**. Before you have something, you might have a certain indifference curve. But once you get it, a new reference point is set, and from this point the utility of a gain is less than a corresponding loss on the previous indifference curve. [In other words, you plot a new indifference curve.]

- This can lead to odd situations where you're only willing to sell something for higher than you can buy it for (like a collector's item).
- Famous experiment: half of subjects were randomly given mugs and were asked to quote prices to sell; the other half were asked to bid for the cups. The cup owners asked for $7.12, over double the buyers' bid of $2.87!
    - Notably, a third group (Choosers) could either receive a mug or a sum of money, and they bid $3.12. The mug owners had the identical choice (getting money or the mug) but they vastly embellished their ask!
- Owners who buy a house at higher prices spend longer trying to sell their home and set a higher listing price - even though the current market value is rationally all that matters.
- [This also suggests that companies with retail stockholders with an emotional bent - like TSLA or AAPL - may tend to be overvalued (corrected for shorting).]

The endowment effect doesn't occur in all cases - people are willing to exchange $5 for five $1 bills, and vendors are happy to exchange a table for money. When the asset under question is **held for exchange**, the endowment effect doesn't apply. You only feel endowed with items that are **planned for consumption or use** (like a bottle of wine or vacation days).

- Furthermore, experienced traders show less attachment to the endowment effect.

Goals are Reference Points

We are driven more to avoid failing a goal than to exceed it. Failing a goal is perceived as a loss; exceeding the goal is a gain.

- Cabdrivers have a daily target income. When the going is good, they clock out early, rather than making the most of the day to offset less profitable days.
- Golfers are more accurate when they putt for par (to avoid a bogey) than when they aim for a birdie, with a difference in success of 3.6%.

Status Quo Bias

In another reframing of loss aversion, we are biased toward keeping the status quo. The endowment effect exaggerates the value of what you have, warping your prior indifference curve; loss aversion makes you hesitant to take on risky bets, since losses are more painful than gains.

Examples:

- In negotiations, concessions are painful because they represent losses from the status quo. Both parties are trying to make concessions, and getting gains from the other side that don't make up for concessions (because losses outweigh gains) - thus everyone walks away unhappy! This is aggravated in situations where parties are allocating losses, not gains in a growing pie.
- Political reform is difficult because the people who have something to lose mobilize more aggressively than those who win. Reforms thus often include grandfather clauses that protect existing stakeholders.

Fairness

- The normal price of a good in a store is the reference point. It's considered unfair to exploit market power to increase prices to match increased demand, unless the store must do so to retain profits.
    - [The rise of Uber and surge/realtime pricing may blunt this effect over time. Soon it might seem normal for shovels to be more expensive during snowstorms.]
- An employee is hired at $9 per hour, but the market rate drops to $7 per hour. It's considered unfair to change the employee's rate. But if employee leaves, it's acceptable to pay the new employee $7 per hour.
    - The old employee had a personal entitlement to the wage. The new employee has no entitlement to the previous worker's wage. The firm has entitlement to *retain* its current

profit, but not to *increase* it by encroaching on others' entitlements.
- Employers who violate rules of fairness are punished by reduced productivity by employees and sales by customers.

Regret and Responsibility

In another twist, the feeling of regret depends on your default action and whether you deviate from it. If you do something uncharacteristic and fail, you're more likely to feel regret and others are less likely to blame you.

Consider: Alan never picks up hitchhikers. Yesterday he gave a man a ride and was robbed.

Barry frequently picks up hitchhikers. Yesterday he gave a man a ride and was robbed.

Who will experience greater regret?

Who will be criticized most severely by others?

Common answer: Alan will experience greater regret, but Barry will be criticized most severely. Alan will have wished he stayed the normal path and would want to undo the event. Barry will be seen as habitually taking unreasonable risks, and thus "deserving of his fate."

- [Sadly this might drive blaming rape victims who seem to be "asking for it" through their typical dress or behavior.]

In some cases, the default option can be to do nothing (not sell your stock), while the alternative is to do something.

Corollaries:

- If you do the normal thing and get a bad outcome, this feels better than doing the unusual thing and getting a bad outcome.
    - Selling stock X and missing out on stock X's big gain feels worse than keeping stock X and missing out on different stock Y's big gain.
    - Volunteering for a trial and getting a disease feels much worse than just living and getting the same disease.
- If you do an unusual thing and get a good outcome, this feels better than doing a normal thing and getting a good outcome.
    - People are happier if they gamble and win, than if they refrain and get the same amount.
- If you anticipate a bad outcome and regret, you will tend to do the normal thing.
    - Consumers who are reminded they may feel regret favor brand names over generics.
    - A surgeon believes an experimental treatment is better for a patient, but is more open to liability for a bad outcome than if following protocol.

    - Picturing a startup failing will push people into staying in their salaried jobs.
- If you anticipate a good outcome, you will tend to do the unusual thing.

Taboo tradeoff: there are certain things that people hold as sacrosanct (notably, health and safety), against which they would be unwilling to trade anything. This is driven by the prediction of regret if the trade were made, and harm were caused.

- Europe uses the “precautionary principle,” which prohibits any action that might cause harm. This could have precluded innovations like airplanes, antibiotics, X-rays.
- Also see: organic food and pesticides.

The risk of regret is that it causes inappropriate loss aversion.

Antidotes to regret:

- Note the possibility of regret before deciding. Then in the case of a bad outcome, remember that you considered the possibility of regret before deciding. This avoids hindsight bias. Avoid the feeling of “I almost made a better choice.”
- Daniel Gilbert: Downweight future regret - you will deploy psychological defenses to this pain.

Specificity and Emotion

When an event is made specific or vivid, people become less sensitive to probability (lower chances are overestimated and higher chances are underestimated).

When an event is specifically defined, your mind constructs a plausible scenario in which it can happen. And because the judgment of probability depends on the fluency to which it comes to mind, the probability is overestimated. Furthermore, the possibility of its *not* occurring is also vivid, and overweighted. Thus, the probability weighting is further compressed.

- People were asked the chances that each one of 8 NBA playoffs teams would win the championship. The sum of all 8 teams was 240%! Each team had a plausible path to winning, while the alternative of (7 other teams) was a diffuse possibility. Thus each individual team was overestimated in turn.
    - This disappeared when the subjects were asked to estimate the chance of the winning team coming from the Eastern vs Western conference.
- Planning fallacy explained: the successful case is very available to the mind, so its likelihood is embellished. The many failures cases are diffuse and not concentrated on.
- People become less sensitive to probability when the outcomes are emotional (like “getting a painful shock” or “kissing your celebrity crush”) rather than cash-based.
- Part of the value of a lottery ticket is indulging in the fantasy of winning.
- [In cognitive behavior therapy, patients are instructed to picture instances when they were competent. Making this vivid allows a depressed patient to increase their estimation of competence.]

A related effect: **Denominator neglect.** People focus on the size of the numerator, rather than examining

the basal rate.

- Subjects are given a choice of drawing a marble from an urn. They have two choices: Urn A with 10 marbles, with 1 red; Urn B with 100 marbles, with 8 red. 30% of subjects choose the urn with the larger number of marbles.
    - [Carnival games employ this]
- 0.001% of permanent disability seems low. “One out of 100,000 children will be disabled” seems high. The other 99,999 have faded into the background.
- People who saw a disease that kills “24 people out of 100” rated it less dangerous than “a disease that kills 1,286 people out of every 10,000.” Insane! (Note this is a between-subjects study; direct comparison would invoke System 2 and likely cause less error.)
- [I think people who are trained to think in terms of percentages (like financial analysts) can grok that 0.0002% is really low. Many people can’t though.]

[Antidotes:

- When estimating the chances of a project working, also picture the failure cases vividly and estimate their probabilities.
- Strip away irrelevant details from a rosy story to regain sensitivity to probabilities.
- To avoid denominator neglect, always reduce two statistics to the same denominator for comparison (eg mortality rate per 100k people). Apples to apples.
- Overweighting is not observed in choice from experience. If you feel you’re overweighting a low chance (eg embarrassing yourself at a party), then force yourself to engage - lack of event may reduce your weighting and reduce your probability estimation!

Applications:

- To overweight small chances (startup succeeding), paint a vivid picture of the event.
- To discount a likely bad event (chemo not working), make the inverse event more vivid (celebrating your grandson’s birth).
- To correct overoptimism (likelihood of getting promotion) or overpessimism (fear of terrorism), then avoid pictures and use statistics, or build your own vivid picture against it (picture a stadium of 10k people. Only 2 will die of terrorism).
- To embellish the risk of a number, make the denominator large and focus on the numerator. To make a risk look less dangerous, make the denominator small and focus on the fraction.

]

[Let’s say, because of media exposure, people overestimate the risk of dying from a disease (say they predict 1% when it’s really 0.01%). Which option would be better to convince people that the risk is much lower?

- Picture a stadium of 10,000 people. Only 1 of these people will die from the disease.
- The risk is 0.01%.

The first option seems more compelling to me in demonstrating how low the number is, by making the

probability real and countable. Yet because of the vividness of pointing out one person, it would also cause overestimation above 0.01%. How to reconcile this?

Maybe there are two axes here when people consider information. One is estimation of probability, the other is believability or persuasive power. The former has a higher estimated probability but also higher persuasive power that compensates for the higher apparent probability.]

Broad Framing and Global Thinking

When you evaluate a decision, you're prone to focus on the individual instance, rather than the big picture of all instances of that type. A decision that might make sense in isolation can become very costly when repeated many times.

For a theoretical illustration, consider both decision pairs, then decide what you would choose in each:

1. A certain gain of $240.
2. 25% chance of gaining $1000 and 75% chance of nothing.
3. A certain loss of $750.
4. 75% chance of losing $1000 and 25% chance of losing nothing.

As we know already, you likely gravitated to Option 1 and Option 4.

But let's actually combine those two options, and weigh against the other
1+4: 75% chance of losing $760 and 25% chance of gaining $240
2+3: 75% chance of losing $750 and 25% chance of gaining $250

2+3 is clearly superior to 1+4. Yet you didn't think to combine all unique pairings and combine them with each other!

This is the difference between **narrow framing** and **broad framing**. The ideal broad framing is to consider every combination of options to find the optimum. This is obviously more cognitively taxing, so you use the narrow heuristic - what is best for each decision at each point. Analogy: single bet vs portfolio of bets.

Yet because each single decision in isolation can be subject to probability misestimations and inappropriate risk aversion/seeking, this can have large costs. Practical examples:

- If given a highly profitable gamble (eg 50-50 to lose 1x or gain 1.5x), you may be tempted to reject the gamble once. But you should gladly play 100 times in a row, for you are almost certain to come out ahead.
- In a company, individual project leaders can be risk averse (since their incentives are heavily tied to project success). Yet the CEO may wish that all project leaders take on maximum appropriate

risk to maximize the EV of the total portfolio.
- The opposite: in a company, individual project leaders running a loss may be tempted to run an expensive hail-mary, to seek the small chance of a rescue. Globally, the CEO may prefer to shut down projects and redirect resources to the winning projects.
- A risk averse defendant of frivolous lawsuits may be tempted to settle each one individually, but this can be globally disproportionately costly compared to the rate of winning lawsuits (let alone inviting more lawsuits).
- Appliance buyers may buy individual appliance insurance, when the global view of all historical appliances may show this is clearly unprofitable.

To overcome narrow framing, adopt **risk policies - simple rules to follow that give a better global outcome**. Examples:

- Check your stocks only once a quarter.
- Always take the highest possible deductible for insurance.
- Never buy extended warranties.
- [Treat your time as costing $x per hour. Avoid things that are unprofitable at this rate, even if subject to sunk cost.
- Always prioritize projects by ROI or IRR, not by other confounding reasons.
- Calibrate health outcomes to the same scale, like QALYs.]
- You win a few, you lose a few.

[Antidotes:

- To avoid the agency problem, give people equity in the company to reduce individual risk aversion.
- Be wary of people telling you to risk it all - they may have a portfolio of investments that smooth out outcomes, while you only have one.

]

Mental Accounts and Sunk Costs

Related to narrow framing - we have a tendency to bin outcomes into separate accounts. This provides organization and simplifies calculations, but can become costly. People don’t like to close accounts with a negative balance.

- Financial accounts are separated into spending, savings, investment, retirement, emergency. Theoretically you could simply use one giant account, but the person might run into self-control problems.
- Golfers who putt more accurately to avoid bogey than to get a birdie keep a separate mental account of each hole, rather than considering the overall score.
- Novice stock traders see each stock in isolation, selling to capture gains and keeping losers to recover losses, when the opposite is a more profitable strategy.
- Sunk cost calculation is done with the item at hand, rather than a global view of all assets.
    - Eg someone who buys a movie ticket is more willing to drive through a snowstorm to see

the movie, than someone who was gifted the ticket.
- A CEO who has made sunk costs unsuccessfully is often replaced, as the new CEO is unencumbered by the mental accounting and can objectively cut losses.

Reversals

Because of a narrow frame, System 1 doesn't know what information it's missing. **It doesn't consider the global set of possibilities and make decisions accordingly.** This can lead to odd situations where you make one set of decisions when considered two cases individually, than a contradictory decision when considered jointly.

Consider a nonprofit that reduces plastic in the ocean to avoid poisoning and killing baby dolphins. How much are you willing to donate to support the cause?

Now consider the discovery that farmers are at significantly higher risk of skin cancer because of the time they spend outdoors. How much are you willing to donate to a nonprofit that reduces cancer for at-risk groups?

You are likely more willing to pay a higher amount for the dolphins than the farmers. In isolation, each scenario engages a different "account" and intensity of feeling. Saving baby dolphins employs the "save the animals" account, and baby dolphins is high up (relative to saving slugs). You intensity match the high emotion you feel to an amount you're willing to pay.

In contrast, the farmer skin cancer example evokes a "human public health" account, in which farmer skin cancer is likely not as high a priority as baby dolphins is on the saving animals account. Emotion is lower, so the donation is lower.

Now consider the two examples together. Which one are you more willing to pay for? Likely you now realize that humans are more important than animals, and you recalibrate your donations so the second amount exceeds the first.

Judgment and preferences are coherent within categories, but may be incoherent when comparing across categories. And because of WYSIATI, alternative categories may not be available. And much of life is a between-subjects trial - we only get exposed to one major event at a time, and we don't think to compare across instances.

Dangers arise when you don't have a clear calibration of a case to the global set of cases. Examples:

- When considered in isolation, a court case involving a poisoned child is assigned a lower judgment than embezzlement of $10 million. When considered jointly, the poisoned child amount exceeds $10 million, since there is now an anchor. Yet in the US legal system, jurors are not allowed to consider other cases when assessing damages!
- Government agency fines make sense within the context of the agency, but are inconsistent

between agencies. A worker safety fine may be $7k while a wild bird conservation violation may cost $25k.
- [Salespeople compare what they're selling to poorer alternatives favorably, but omit superior alternatives or entirely different ways to spend your money that might be better.]
- [Within a business, separate accounts lead to disproportionate attention. In the "office supplies" account, reducing spend by $1,000/year seems like a big win; relatively less attention may be paid to the "projects" account that might save $200,000/year.]
- [This explains why a percentage is often meaningless, if you do not regularly engage with the meanings of percentages. 0.01% does not feel much different from 0.0001%, unless you grok how much lower the latter is.]

[Antidotes:

- Reduce all decisions down to a single fungible metric, in the broadest account possible, to allow global calibration.
    - Benefits to human life can be assessed in terms of QALYs/$, so that repairing glaucoma can be compared to childhood cancers.
    - Projects and spending should be considered in terms of ROI or IRR or happiness/$, so that all possible spending can be assessed on the same scale from a global account
- Deliberately consider the broad range of alternatives for scarce resources - time, labor, money. Be aggressive with defining the global set of opportunity costs.

]

Framing Effects

The context in which a decision is made makes a big difference in the emotions invoked and the decision made. In particular, even though a gain can be logically equivalently defined as a loss, because losses are so much more painful, the decisions may be contradictory.

Consider how enthusiastic you are about each opportunity:

- A 10% chance to win $100, and a 90% chance to lose $5.
- Buying a $5 raffle ticket that gives you a 10% chance to win $105 and a 90% chance of winning nothing.

These are logically identical situations - yet the latter opportunity is much more attractive! Foregone gains are less painful than losses.

Most of us accept problems a sthey are framed, without considering alternative framings. [Hence Charlie Munger's advice, **INVERT - ALWAYS INVERT.**]

Examples of framing that leads to distorted interpretations:

- Describing a large % chance of survival is more appealing than a small inverse % chance of mortality. A medical procedure is much favored when presented in the former than the latter.
- Consider two framings of two vaccine programs that can save 600 people affected by a virus:
    - Program A will save 200 people. Program B has ? chance of saving 600 and ? chance of saving none.
    - Program A will leave 400 people dead. Program B has ? chance that nobody will die, and ? chance that 600 will die.
    - Per prospect theory, you already know what most people prefer. But again, these framings are logically equivalent.
- Describing a discount off an inflated price is more attractive than an extra fee on the normal price.
- Saying “the Patriots lost” evokes different thoughts than “the Seahawks won,” even though these two are logically equivalent.
- Consider which is better for the environment:
    - Adam switches from a car with 12mpg to a car with 14mpg.
    - Barry switches from a car with 30mpg to a car with 40mpg.
    - Barry’s move looks like a bigger jump (10mpg and 33% jump) but this is deceptive. The real metric that matters is gallons per mile, which inverts the mpg, and differences in fractions are less intuitive.
    - For an extreme, consider going from 1mpg to 2mpg vs 100mpg to 200mpg. Over 1000 miles, the former will save 500 gallons; the latter, only 5!
- For organ donors, changing from an opt-in to an opt-out can increase donation rates from 10% to 90%. When action is required to overcome the default, a lazy system 2 finds it easier to follow the default.
    - [This is also confounded by social inferences where the default is assumed to be more popular.]

Even experts like doctors and public health officials suffer from the same biases. Even more troubling, when presented with the inconsistency, people cannot explain the moral basis for their decisions. System 1 used a simple rule - saving lives is good, deaths are bad - but System 2 has no moral rule to easily solve the question.

Conclusions

It’s been shown that humans are not rational. Unfortunately, belief in human rationality also pushes a libertarian ideology in which it is immoral to protect people against their choices. “Rational people make the best decisions for themselves - who are we to think we’re better?”

- The view of addiction is that a rational agent strongly prefers immediate gratification and accepts future addiction as a consequence.
- People know what they’re doing when they choose not to save for retirement or eat to obesity.

This also leads to a harsher conclusion: people deserve little sympathy for putting themselves in worse situations. Elderly people who don’t save get little more sympathy than people who complain about a bill after ordering at a restaurant. Rational agents don’t make mistakes.

Behavioral economists believe people do make mistakes, and they need help to make more accurate judgments. Freedom is a virtue, but it has a cost borne by individuals who make bad choices and by a society that feels obligated to help them.

The middle ground might be **libertarian paternalism**, in which the state nudges people to making better decisions and give the freedom for people to opt out. This includes retirement savings, health insurance, organ donation, and easy-to-understand legal contracts.

Part 5: Two Selves

Part 5 of this *Thinking, Fast and Slow* summary covers the nature of happiness. [I felt this was the weakest part of the book with less experimental evidence and finality on the conclusions, but this is still more of a work in progress.]

Happiness is tricky. There is in the moment happiness, and there is overall well being. There is happiness we experience, and happiness we remember.

Consider having to get a number of painful shots a day. There is no habituation, so each shot is as painful as the last. Which one represents a more meaningful change?

- Decreasing from 20 shots to 18 shots
- Decreasing from 6 shots to 4 shots

You likely thought the latter was far more meaningful (especially as it drives more closely toward zero pain). But Kahneman found this incomprehensible. Two shots is two shots! There is a quantum of pain that is being removed, and the two choices should be evaluated as much closer. Someone who pays different amounts for the same gain of experienced utility is making a mistake.

- [This relates to another bias: being more sensitive to relative changes rather than absolute changes. Thus saving $0.50 on a $1.00 candy bar feels like a bigger win than saving $90 off a $2000 TV purchase.]

This thought experiment kicked off Kahneman’s investigation into happiness.

Experiencing Self vs Remembering Self

Kahneman presents two selves:

- **The experiencing self:** the person who enjoys pleasures and feels pain moment to moment. Experienced utility would best be measured by measuring happiness over time, and integrating the area under the curve.
- **The remembering self:** the person who reflects on past experiences and evaluates it overall.

The remembering self factors heavily in our thinking. After a moment has passed, only the remembering self exists when thinking about our past lives. The remembering self is often the one making future

decisions.

But the remembering self evaluates differently from the experiencing self in two critical ways:

- Peak-end rule: The overall rating is determined by the peak intensity of the experience and its end. It does not care much about the averages throughout the experience.
- Duration neglect: The duration of the experience has little effect on the memory of the event.

This operates in classic System 1: by averages and norms, not by sums.

This leads to preferences that the experiencing self would find odd, and show that we cannot trust our preferences to reflect our interests:

- The ice-water experiment
    - Participants were exposed to a short episode (60 seconds of hand in 14C water) and a long episode (60 seconds in 14C, but in the last 30 seconds, the temperature increased to 15C). They were then asked which they would repeat for a third trial.
    - The experiencing self would clearly consider the long episode worse. But 80% of participants preferred the long episode, thus suffering "30 seconds of needless pain." They picked the option they liked more.
    - Oddly, people would prescribe the shorter episode for others, since they care about the experiencing self of others.
- Would you take a vacation that was very enjoyable, but at the end of which you took a pill that gave you total amnesia of the event? Most would decline, suggesting that memories are a key - perhaps dominant - portion of vacations. **The remembering self chooses vacations**!
- Inversely - say you have two options for an operation - 1) one in which you undergo general anesthesia, but will endure a month of mild pain after; 2) one in which you remain conscious and feel intense pain, but are given amnesia to remove any memory of the episode.
    - Many prefer #2, indifferent to the pains of their experiencing self. "**I am my remembering self, and the experiencing self, who does my living, is like a stranger to me.**"
- We tend to want to capture a moment (through a camera) rather than fully experiencing it - even if we never look at the video again.
- The end of someone's life matters a lot more than duration. For someone who dies at age 57, you don't care as much if they lived a whole extra year to age 58; but you do care about whether she was surrounded by her family in her last 10 minutes, or whether she was alone.
    - [Related: Make a Wish foundation rightly wants to make the last days of a child happier - would the funds be better used to extend lives of other children?]
    - In a psych experiment, people were told of Jen, an extremely happy person throughout life who died painlessly at 60 in an accident. Another group was told of a similar Jen who lived to 65, with the first 60 years identical but the last 5 years being less pleasant than before. The 5 only "slightly happy" years caused a drop in perceived total happiness in that life.
- We can enjoy a movie all the way throughout, but if the ending is poor or interrupted, we have a worse impression.
- We favor peak highs with short duration, rather than moderate highs over long duration. [A model

for drug addiction?]

- After a person dies, we feel moved by changes to their stories that they clearly did not experience
    - A man may have died happily believing his wife was faithful, only for us to find out postmortem that she had a lover.
    - We feel redemption in an artist who died in poverty but achieved posthumous renown.
    - [This might be because people live on after death in our minds.]
- An experiencing self declares “I will never forget this moment”; the remembering self frequently forgets.

These examples challenge the idea that humans have consistent preferences and know how to maximize them (the rational agent model). **We will consciously articulate that we prefer pain to be brief and pleasure to last, but our remembering self has different ideas.**

[There is evolutionary history for this: experiments show duration neglect and peak intensity in rats. Why would this behavior be evolutionarily advantageous?

Maybe it makes us more resilient to painful episodes. If we strictly assessed utility by integrating area under the curve, a traumatic experience could leave us in “happiness debt” that would take considerable time to overcome.

Finally, these are general influences, not absolutes. Odysseus’s story is poignant for his 10 years of trials - had it been only 5 minutes, it would have lost much of its drama]

Seeming Exceptions

Why, then, does a painful labor that lasts 24 hours seem worse than one lasting 6 hours? Why does a 6 day vacation seem better than 3?

Kahneman argues the mechanism of the longer duration is in changing the end state - a mother is more helpless after 24 hours than after 6; the vacationer is more relaxed after 6 days.

[Another one: if the remembering self makes the decisions tyrannically, why are good behaviors (like flossing, losing weight, and saving money) so difficult? The experiencing self endures short-term pain, but the remembering self should discount the pain and remember the benefits - and the current experiencing self enjoys the fruits of labor of the past experiencing self!

It could be that for the most difficult behaviors, the remembering self doesn’t actually experience the pleasure of long-term gains. Better health in your 30’s rewards you when you’re 70, which you’ve never experienced. Meanwhile the memory of a delicious hamburger looms larger. Same with saving money for retirement.

This suggests that positive feedback loops are important for “good behaviors”; the remembering self must recall enough pleasure from a behavior to guide future behavior. Hence working hard for a raise is easier than saving money for your 70s. People without the ability to gain positive feedback reinforce short-term behaviors (eg the poverty trap).]

Experienced Well-Being vs Judgment of Life

Measuring Experienced Well-Being

How do you measure well-being? The traditional survey question was: “All things considered, how satisfied are you with your life as a whole these days?”

Kahneman was suspicious that the remembering self would dominate the question, and that people were terrible at “considering all things.” The question tends to trigger the one thing that gives immense pleasurable (a new partner) or pain (workplace conflicts).

To measure experienced well-being, he led a team to develop Day Reconstruction Method, which prompts people to relive the day in detailed episodes, then to rate the feelings. Under the “area under the curve” philosophy, they conceived of the metric **U-index: the percentage of time an individual spends in an unpleasant state**.

Findings:

- There was large inequality in the distribution of pain. 50% of people reported going through a day without an unpleasant episode. But a minority experience considerable emotional distress for much of the day - because of illness, disposition, or misfortune.
- Different activities have different U-indices. Morning commute: 29%; childcare: 24%; TV watching: 12%; sex: 5%. Weekend is 6% lower than weekdays (because of control over pleasurable personal activities).
- Difference across cultures for the same activities. Frenchwomen spend less time with children but enjoy it more, perhaps because of access to child care. Frenchwomen spend the same amount of time eating, but enjoy it more, possibly because they focus on it.
- The mood depends largely on the current situation, not general factors influencing general satisfaction. Affects mood: coworker relations, loud noise, time pressure, boss presence. Does not affect mood: benefits, status, pay.
- Some activities generally seen as positive (like having a romantic partner) don’t improve experienced well-being. Partially because of tradeoffs - women in relationships spend less time alone, but less time with friends; they spend more time having sex, but also more time doing housework and caring for children.

Suggestions for improving experienced well-being:

- Focus your time on what you enjoy. Commute less.
    - [I suspect Kahneman would be a supporter of guaranteed basic income (all else equal) as it’d allow swaths of the population to decrease their U-index.]
- To get pleasure from an activity, you must notice that you’re doing it. Spend leisure time to active forms, like socializing and exercise, and awy from passive forms like TV.
- A reduction of a U-index of society by 1% would be a huge achievement, with millions of hours of avoided suffering.

Experienced Well-Being vs Life Evaluations

Where well-being is measured by methods like the Day Reconstruction Method, life evaluation (or life satisfaction) is measured by the Cantril Self-Anchoring Striving Scale:

“Please imagine a ladder with steps numbered from zero at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?”

Some things affect evaluation and experience differently:

- Affecting evaluation more than experience
    - Education
    - Money: after about $75k of income, it increases evaluation without increasing experience
        - Theory: Higher income might reduce ability to enjoy small pleasures of life.
- Affecting experience more than evaluation
    - Ill health
    - Living with children
    - Religion

Severe poverty amplifies experienced effects of misfortunes.

- For the top ? of individuals, a headache increases negative experience from 19% to 38%. For the poorest tenth, it starts at 38% and moves to 70%. Same with divorce and loneliness.
- The beneficial effects of the weekend are smaller for the poor.
- [This might be because of lacking mitigating factors - rich people with a headache can afford to take off work. Also, a person’s experience is relative to the base, so a larger denominator of satisfaction can reduce the impact of a single negative event (“I have a headache now, but at least I’m enjoying my comfy bed in a big house.”]

Temperament, which is largely determined by genetics, affects both experienced well-being and life satisfaction.

- This can explain why certain changes like marriage show low correlations with well-being - it works for some and not for others.

Goals make a big difference in satisfaction.

- People who care about money and get it are more satisfied than those who wanted money and didn’t get it, or those who didn’t care about money and didn’t get it.
- One recipe for dissatisfaction: setting goals that are especially difficult to attain. (The worst: “becoming accomplished in a performing art.”)

[For overall happiness, do people load differently on experienced vs evaluated happiness? Goal-driven people seem to weigh the latter more, despite possibly not improving their experienced well-being over time.

Are there correlations? Are people with unsatisfied dispositions high on evaluation and low on experience? Are hedonic people high on experience and low on satisfaction?

Is this even a valid question? Experience and life satisfaction may be just two related but separate sensations.]

Focusing Illusion

Considering overall life satisfaction is a difficult System 2 question. System 1 substitutes the answer to an easier question, including "what is my mood right now?", focusing on significant events (achievements or failures), or recurrent concerns (illness). It is more difficult to consider all the factors in your life, weigh those factors accurately, then score your factors.

Nothing in life is as important as you think it is when you are thinking about it. Our mood is largely determined by what we attend to. You get pleasure/displeasure from something when you think about it.

- Even though Northerners despise their weather and Californians enjoy theirs, **climate makes no difference in life satisfaction**.
    - When you're asked the question, "are Californians more satisfied with life than Northerners because of the weather," you overweight the climate factor in the life satisfaction question; you conjure the available image of hiking, rather than the reality that lives are similar throughout; you overestimate how often Californians think about the weather when asked about a global evaluation.
- Consider the question: "how much pleasure do you get from your car?" Now a different question: "when do you get pleasure from your car?" Answer: when you think about it - which is not often, including when you're driving it.
    - You substituted the narrower question: "how much pleasure do you get from your car when you think about it?"
- Certain activities that engage you - like social activities - retain your focus and give you more.
- Chronic stressors cause more dissatisfaction than you would predict - chronic pain, chronic loud noise, depression.
- [Smartphone use may affect your mood by focusing your attention on what's on the phone, even when eating, commuting, getting ready to sleep. And if you focus on negatives (other people's highlight reels on Facebook, outrage from news) - you can increase your U-index.]

The focusing illusion leads to **mispredictions of happiness**, for ourselves and others:

- When you forecast your own future happiness, you overestimate the effect a change will have on you (like getting a promotion), because you overestimate how salient the thought will be in future you's mind. Future you has gotten used to the new environment and now has other problems to worry about.
- You may pay significant amounts for improvements in life satisfaction, even though it has no effect on experienced happiness.
    - Colostomy patients show no difference in experienced happiness compared to healthy people. Yet they would trade away years of life for no longer having a colostomy.

    - The remembering self has a focusing illusion about life that the experiencing self endures comfortably.
- When you predict the happiness of others, you focus on the aspects of their experience that are most salient to you. You ignore that the person may have habituated to her circumstances, or that the aspect has counterbalancing benefits or drawbacks.
    - If you're preoccupied with lacking money, you predict wealthier people are happier than they are, even though they've adjusted to wealth.
    - People predict paraplegics have a higher u-index than they really do. After adjustment, paraplegics stop thinking about their condition. They enjoy friends and get mad about politics, just like you.

Adaptation to a new situation consists in large part of thinking less and less about it.

There are exceptions, like when a recurrent thought is your focus (eg being in love while sitting in traffic).

[Antidote:

- Create a life rubric, where you list the factors and their weightings, then score your life. This might make you happier by relieving the focus on single sore points.
- Choose to spend time on pleasurable activities that will engage your focus. You will pay little attention to the car as you drive, but you will focus in social interaction or playing tennis.
- Reflect on past experiences to inform future ones - how much did buying X make you happier today? What activity Y are you really thankful to your past self for doing? Do more of Y.]

==

Conclusions

Putting it all together - which self should we cater to, the remembering self or the experiencing self?

In *Thinking, Fast and Slow*, Kahneman doesn't have a clear answer, but he rules out that either should be focused on exclusively.

Catering only to the remembering self invites unnecessary suffering. Our memories are fallible, being subject to duration neglect and peak-end rule.

- The remembering self favors short periods of intense joy over long periods of moderate happiness.
- The remembering self avoids long happy periods if it knows a poor ending will come.
- In the facetious extreme, the remembering self might endure decades of pain ending in one jubilee, remembering that moment for the last second of life.

Catering only to the experiencing self treats all moments of like alike, regardless of the future benefit.

- Some instances have memories that give more value than others.

- A moment can also alter the experience of subsequent moments.
    - Learning music can enhance future playing or listening for years.
    - A traumatic event can cause PTSD and misery over a lifetime.
    - Making money today can unlock new experiences for your experiencing self.
- In the extreme, the experiencing self would greedily pursue whatever maximized happiness this moment, with little regard for the future self.

Both must be considered - their interests do not always coincide.

For a population, it’s not clear which to maximize, say for treating health conditions. Should we minimize experienced pain, or maximize the amount people would be willing to pay to sacrifice to be relieved from their condition?

[We now know that after achieving goals like more money, the experiencing self does not become meaningfully happier. In response, one option is to believe this and shift your preferences away from more money. But most likely you won’t - you’ll keep working to be able to earn more money. How do we reconcile this?

- You may simply reject the data - other people don’t get happier, but you believe you’re an exception; or the experiment is somehow flawed. Reflecting on your own experiences, you find that you were less happy with less money.
- You may simply be making a mistake, and you may realize this later.
- Instead of experienced well-being, you are prioritizing life satisfaction, which may go up over time (like a scorecard for achievement).

]