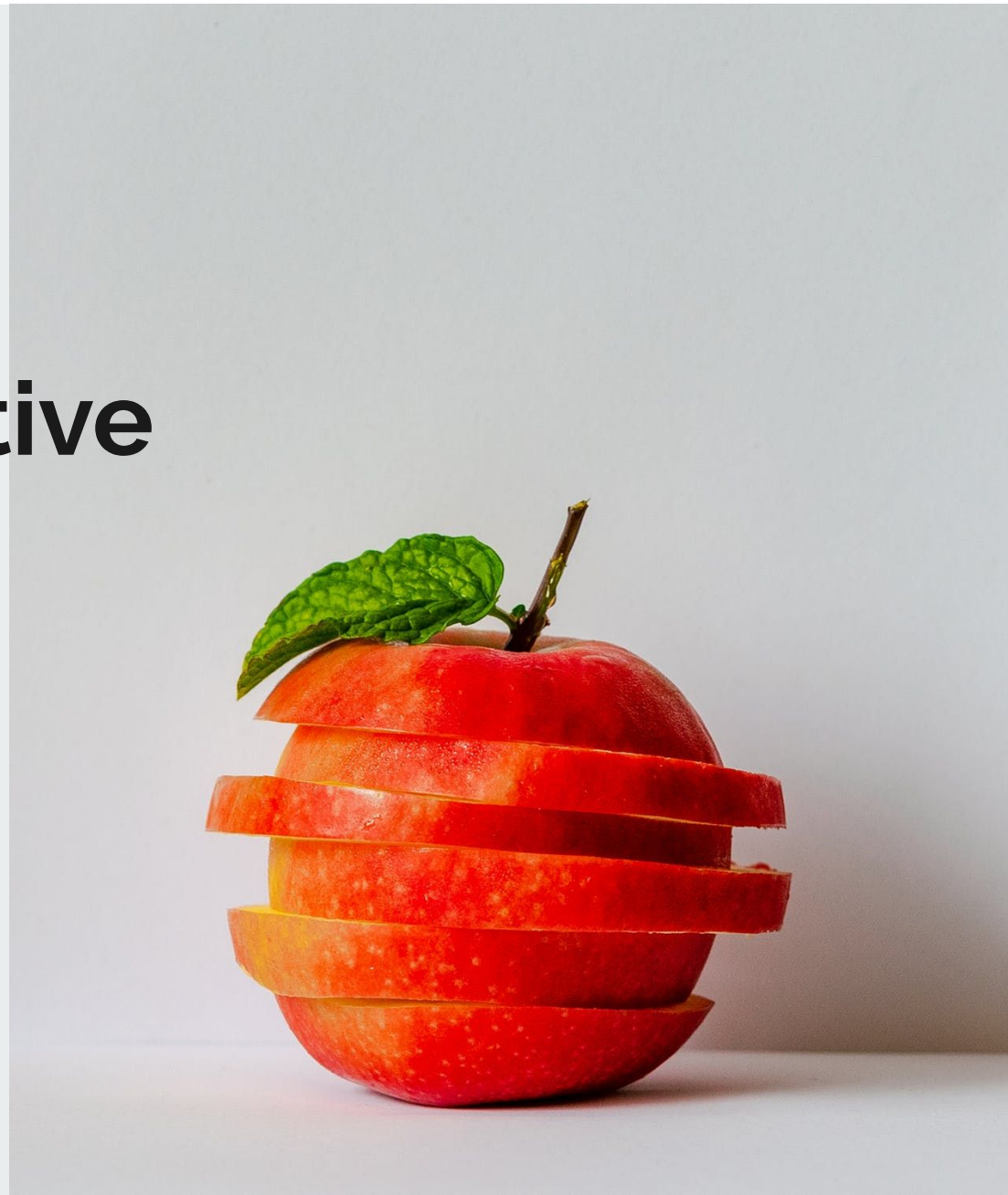
Learning to avoid cognitive biases

Mihai Verzan

How do I lie with statistics?

Supervisor: Prof. Dr. Ullrich Köthe



Mars Climate Orbiter

Lost in 1991 after it “encountered Mars at an improperly low altitude”.

Described by NASA as an “organizational failure”

Multiple mission teams; human error led to metric and Imperial units being confused

Sullivan Mine accident, Canada

2006: a worker died in the anoxic environment in a shed; the employee who found him as well as two paramedics who responded to the emergency call also lost their lives in the low oxygen environment.

All specifically trained in safety measures.

Skipping steps in medical procedures

A test run at a major US hospital showed that doctors often skipped at least one step in a 5-step procedure.

Nurses were given the authority and responsibility to catch doctors missing any steps.

After 15-months: infection rates down from 11% to 0%, 8 deaths avoided, $2 million saved.

cognitive biases

Cognitive biases

A bias is an unconscious, automatic, reliable error in human judgement

Prof. Daniel Kahneman: psychologist and economist at Princeton University

Book: “Thinking, Fast and Slow” (2011)

Name	Description
Ambiguity effect	The tendency to avoid options for which the probability of a favorable outcome is unknown.[10]
Anchoring or focalism	The tendency to rely too heavily, or "anchor", on one trait or piece of information when making decisions (usually the first piece of information acquired on that subject).[11][12]
Anthropocentric thinking	The tendency to use human analogies as a basis for reasoning about other, less familiar, biological phenomena.[13]
Anthropomorphism or personification	The tendency to characterize animals, objects, and abstract concepts as possessing human-like traits, emotions, and intentions.[14]
Attentional bias	The tendency of perception to be affected by recurring thoughts.[15]
Attribute substitution	Occurs when a judgment has to be made (of a target attribute) that is computationally complex, and instead a more easily calculated heuristic attribute is substituted. This substitution is thought of as taking place in the automatic intuitive judgment system, rather than the more self-aware reflective system.
Automation bias	The tendency to depend excessively on automated systems which can lead to erroneous automated information overriding correct decisions.[16]
Availability heuristic	The tendency to overestimate the likelihood of events with greater "availability" in memory, which can be influenced by how recent the memories are or how unusual or emotionally charged they may be.[17]
Availability cascade	A self-reinforcing process in which a collective belief gains more and more plausibility through its increasing repetition in public discourse (or "repeat something long enough and it will become true").[18]
Backfire effect	The reaction to disconfirming evidence by strengthening one's previous beliefs.[19] cf. Continued influence effect.
Bandwagon effect	The tendency to do (or believe) things because many other people do (or believe) the same. Related to groupthink and herd behavior.[20]

Wikipedia's "List of cognitive biases"

System 1

Intuitive, automatic

Fast, less effort

But also: vulnerable to bias

Relies on heuristics, pattern matching

System 2

Deliberate, analytical

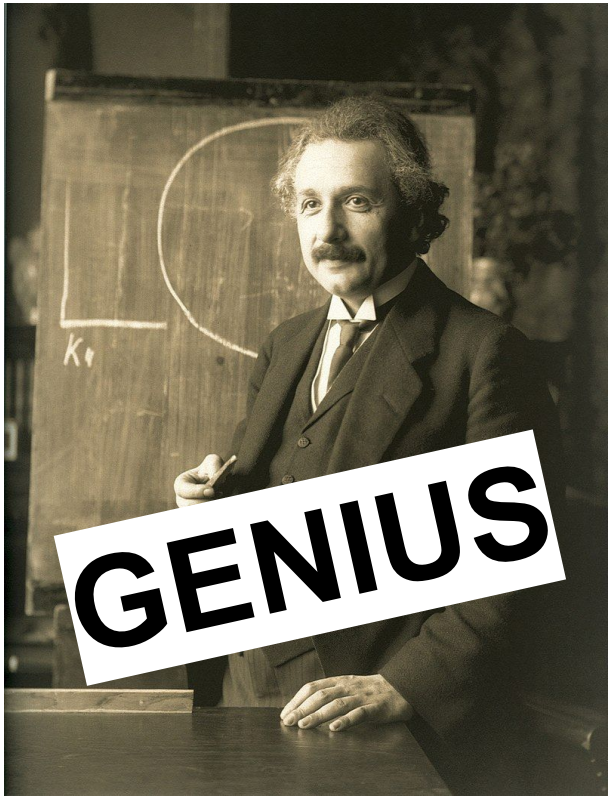
Less vulnerable to bias

But also: very slow, a lot of effort

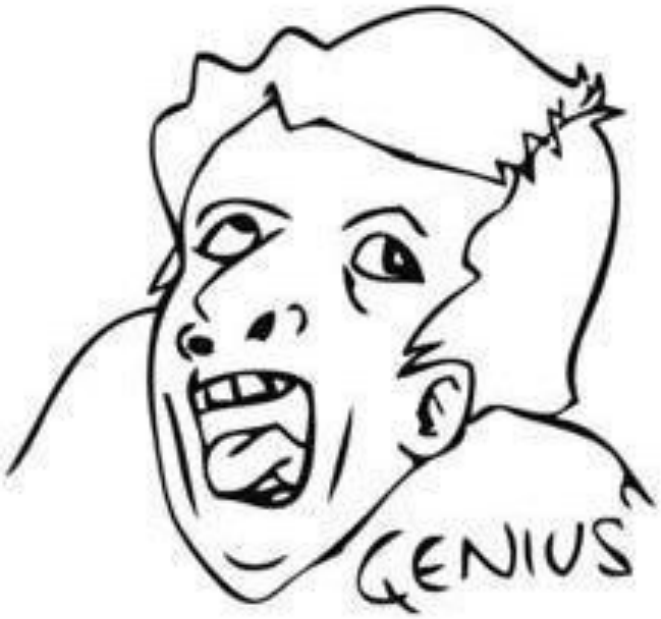
Is a conscious, intentional process

System 1

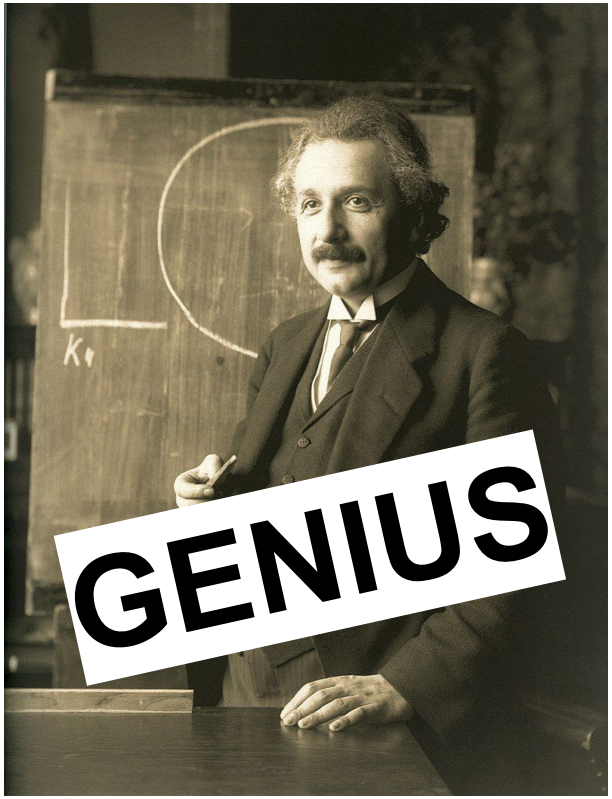
System 2



System 1



System 2



System 1

System 2

System 1 vs System 2

System 1 plays a crucial role

Rule of thumb: humans spend 95% of their time in system 1

Kahneman: system 1 is “the real hero of this book”

Not systems in the “real” sense, but rather fictitious characters, a **useful metaphor**, especially to our “biased brains”.

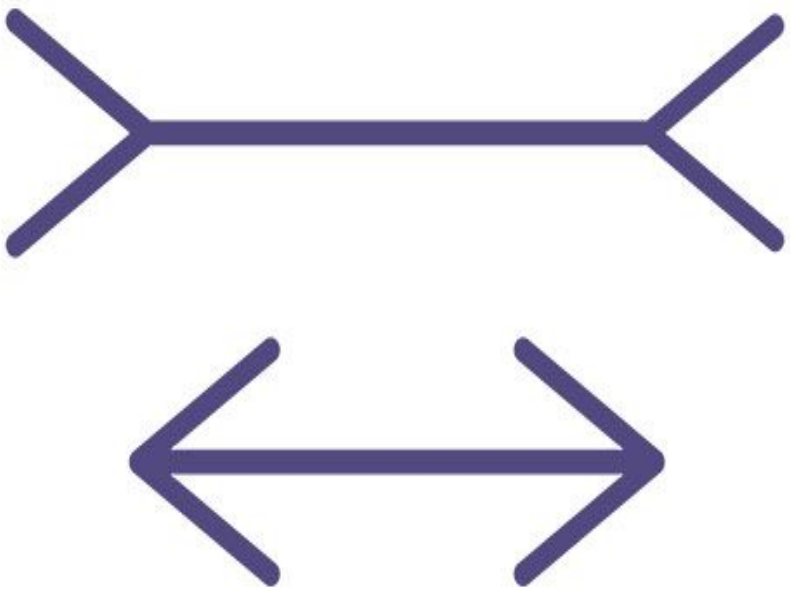
Kahneman says: you can read this book as “a psychodrama with two characters”.

Kahneman's section about priming

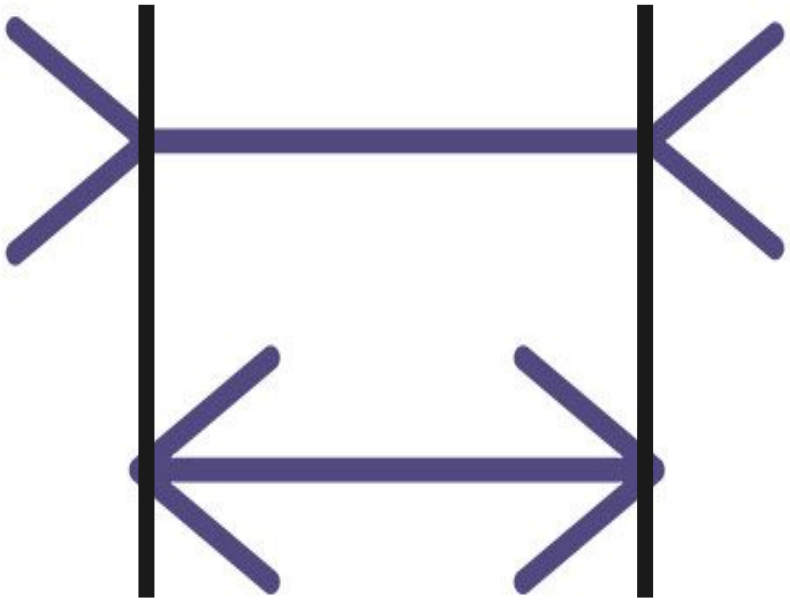
After *Thinking: Fast and Slow* was published, a section was found to be biased. Kahneman admitted as such in a blog post:

> *What the blog gets absolutely right is that I placed too much faith in underpowered studies.*
>
> *[...] There is a special irony in my mistake because the first paper that Amos Tversky and I published was about (the belief in the "law of small numbers," which allows researchers to trust the results of underpowered studies with unreasonably small samples.)*
>
> *[...] Our article was written in 1969 and published in 1971, but I failed to internalize its message.*
>
> *This was simply an error: I knew all I needed to know to moderate my enthusiasm for the surprising and elegant findings that I cited, but I did not think it through.*

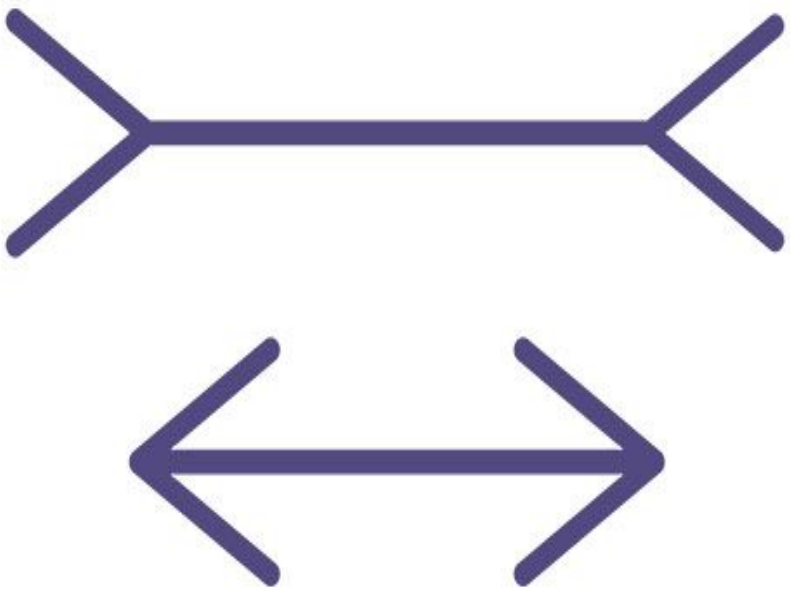
Müller-Lyer illusion



Müller-Lyer illusion



Müller-Lyer illusion



The illusion persists even after you are aware of it. **So do biases**.

teaching bias mitigation & critical thinking

Teaching critical thinking

Teaching critical thinking

Teaching critical thinking is one of the primary functions of the school system.

Teaching critical thinking

Teaching critical thinking is one of the primary functions of the school system.

The effectiveness of these initiatives is limited.

Teaching critical thinking

Teaching critical thinking is one of the primary functions of the school system.

The effectiveness of these initiatives is limited.

This is not for lack of trying.

> *nearly 40 percent [of 17 year-old students] could not draw inferences from written material and only one fifth could write a persuasive essay* [report "A Nation At Risk", 1983, USA]

After this report, programs were implemented across the country; their effectiveness was still limited.

What is critical thinking?

"critical thinking consists of seeing both sides of an issue, being open to new evidence that disconfirms your ideas, reasoning dispassionately, demanding that claims be backed by evidence, deducing and inferring conclusions from available facts, solving problems, and so forth"

3 characteristics:

- Effective (avoids pitfalls)
- Novel (not just applying an algorithm)
- Self-directed (the thinker is doing the thinking)

Why is critical thinking so hard to teach?

Critical thinking is **domain specific**.

E.g. thinking like a historian and thinking like a physicist require different skillsets

Why is critical thinking so hard to teach?

Critical thinking is **domain specific**.

E.g. thinking like a historian and thinking like a physicist require different skillsets

"Knowing that a letter was written by a Confederate private to his wife in New Orleans just after the Battle of Vicksburg won't help the student interpret the letter—unless he knows something of Civil War history."

Thought processes are intertwined with what is being thought about.

Why is critical thinking so hard to teach?

Critical thinking **requires practice**.

It cannot be taught in a vacuum. Students need to apply their skills on real problems

Why is critical thinking so hard to teach?

The brain is lazy and avoids effort whenever possible.

> *"After years of pressure from the film and television industry, the President has filed a formal complaint with China over what U.S. firms say is copyright infringement. These firms assert that the Chinese government sets stringent trade restrictions for U.S. entertainment products, even as it turns a blind eye to Chinese companies that copy American movies and sell them on the black market."*

Surface structure easy to see (news article about geopolitics)

Deep structure harder to recognize

In mathematics, word problems are similar. Students focus on the surface structure at the detriment of the deep structure (the mathematical meaning behind the problem)

How to teach critical thinking?

- Special programs aren't worth it
- Teach in context
- Teach to everyone, not just "advanced" students
- Draw on student experiences
- Make explicit and teach in stages

and New Atlantis. Oxford: Clarendon.

Bruer, J. T. 1993. *Schools for thought: A science of learning in the classroom*. Cambridge, MA: MIT Press.

Chapman, G., and M. Python. 1989. *The complete Monty Python's flying circus: All the words, volume 2*. New York: Pantheon.

Detterman, D. K. 1993. The case for the prosecution: Transfer as an epiphenomenon. In *Transfer on trial: Intelligence, cognition and instruction*, ed. D. K. Detterman and R.

diagnostic errors

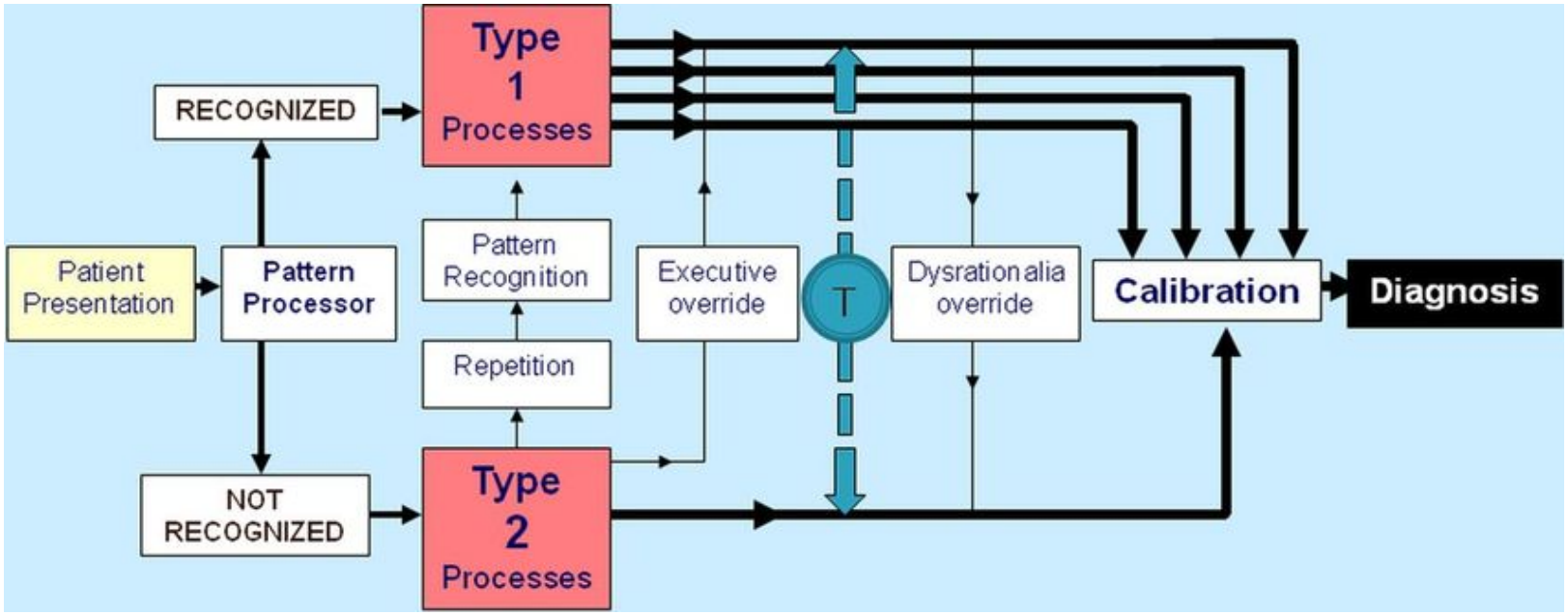
Biases in medicine: diagnostic errors

Doctors are human and therefore biased.

Analysis of diagnostic errors suggests that they often occur not because of lack of knowledge, but because of flawed (biased) judgement.

Details are hard to determine (like whether system 1 or system 2 was used).

System 1; System 2



Origins of System 1 processes

- hard-wired (darwinian)
- emotions (like fear of snakes)
- overlearning (cultural, social, frequent flyer)
- implicit learning (acquired during work or through "contagion" from others)

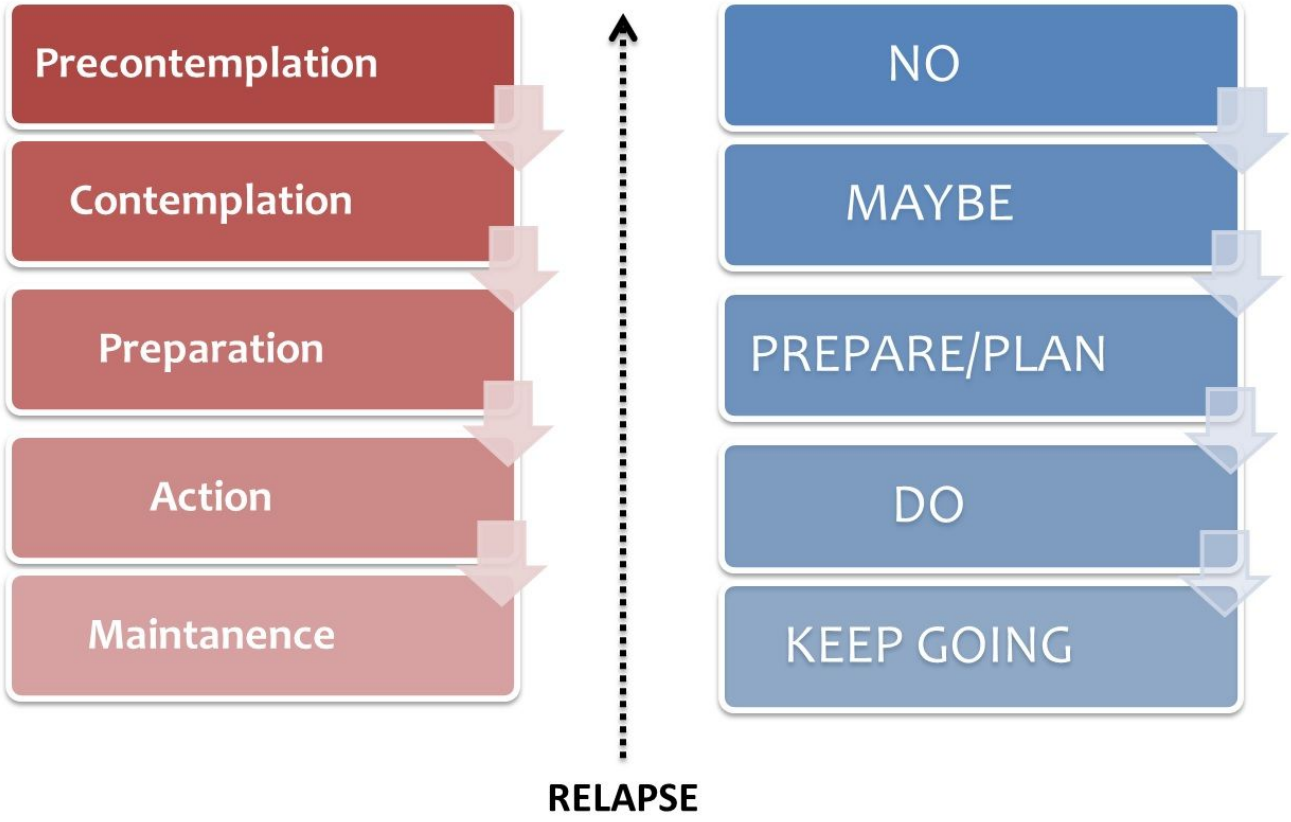
How to protect against biases?

Increased intelligence **does not protect you**

What works: ability to **identify** bias, then apply specific debiasing strategies

Emotional situations help reinforce learning.

Transtheoretical Model
Stages of change



Forcing strategies

Old examples: proverbs, aphorisms, mnemonics

- measure twice, cut once
- a bird in the hand is worth two in the bush

Scientific evidence on their effectiveness is lacking. Debiasing is an “inexact science”

Putting your keys in a specific place to find them more easily

Forcing strategies

Teach about cognitive and affective biases, tools to test for them, strategies for avoidance

get more information

slow down

be more skeptical

seek other opinions

personal accountability

supportive environments

decision support systems, airline checklists

simulation

serious games and debiasing

Missing: The Pursuit of Terry Hughes

“Serious game”: a video game designed to teach, rather than for entertainment.

Adventure game, mystery story, noir detective aesthetic. The player can examine objects, talk to characters, use a smartphone to take pictures, navigate and interact with the world

The researchers have attempted to create an engaging storyline. The player is tasked with finding Terry Hughes, their missing neighbor

Teaches three biases via “bias vignettes”, situations meant to evoke bias.

After each game, an “after action review” takes place: a video explicitly explains the bias, provides feedback on the player’s own performance, and reinforces the lesson with additional examples

Confirmation bias

People tend to search for evidence that **confirms, rather than denies**, pre-existing hypotheses.

Bias vignette: the player is primed with the hypothesis “Terry was kidnapped”, then is asked to search her apartment for clues.

Mitigation: explicitly ask participants to consider alternative hypotheses.

Fundamental attribution error

People tend to **focus on personal, rather than environmental, characteristics**.

Bias vignette: Terry is shown fidgeting in her building's elevator; the player is asked whether Terry seems to be an anxious person.

Mitigation: explicitly ask participants to consider the environment, rather than personal characteristics.

Bias blind spot

Meta-bias: people tend to **see themselves as less biased than others.**

Bias vignette: the player is asked how biased they thought their own performance in the game was compared to other participants.

Mitigation: provide feedback about real performance.

Missing: The Final Secret

Sequel to the first game, teaching players about the **anchoring bias**, the **representativeness heuristic**, and the **projection bias**.

Andragogy

How do adults learn?

Inverse of pedagogy.

- **need to know** — why is this important?
- **self-concept** — self directed learning, not dependent on a teacher
- skills that are **relevant** to their lives
- for children: content oriented, for adults: **problem or task oriented**.

"Learning is more likely to occur when learners are presented with an authentic situation that enables, simulates, or reconstructs real-life complexities and events than from experiences that are irrelevant or meaningless to learning." → video games may have an advantage

Missing: how well does it do?

"an external test campaign was conducted by an independent verification and validation team not affiliated with the authors" → standard methods to test for bias

Control group: videos, current standard for teaching bias mitigation.

Pre-test and **post-test** about biases to assess improvement. A subset: **follow-up test** 8 weeks later

2 demographics — **college students** and **intelligence analysts**, Washington DC/Baltimore region

Missing: how well does it do?

Improvement in bias **recognition**: 37% (students, N=54), 44% (analysts, N=29).

Knowledge retention faded during the 8 weeks leading up to the follow-up test, however there was still a statistically significant improvement.

Caveat: sample size small, especially for analyst sample (only 11 of them completed follow-up testing).

For bias **mitigation**: smaller decline in retention after the 8 week period.

Both *Missing* games **outperformed educational videos**.

some specific biases

Rhyming = truth?

“Put your ideas in verse if you can; they will be more likely to be taken as truth. Participants in a much cited experiment read dozens of unfamiliar aphorisms, such as:

- Woes unite foes.
- Little strokes will tumble great oaks.
- A fault confessed is half redressed.

Other students read some of the same proverbs transformed into nonrhyming versions:

- Woes unite enemies.
- Little strokes will tumble great trees.
- A fault admitted is half redressed.

The rhyming versions were judged as more insightful.

Pronounceability = truth?

Two conflicting reports, from two companies: **Artan** (easy to pronounce) and **Taahhut** (harder to pronounce): the report from Artan was given much more weight.

The brain is lazy; having to put in more effort is judged as negative.

Pronounceability = money?

“Stocks with pronounceable trading symbols (like **KAR** or **LUNMOO**) outperform those with tongue-twisting tickers like **PXG** or **RDO**—and they appear to retain a small advantage over some time.”

Cognitive strain mobilizes system 2

Participants in an experiment were shown problems such as:

> *If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?*

- 100 minutes
- 5 minutes

Some participants read the problem in a normal font, others in a small, grey, hard to read font.

90% of students with the normal font made a mistake.

35% of students with the hard to read font made a mistake.

Mere-exposure effect

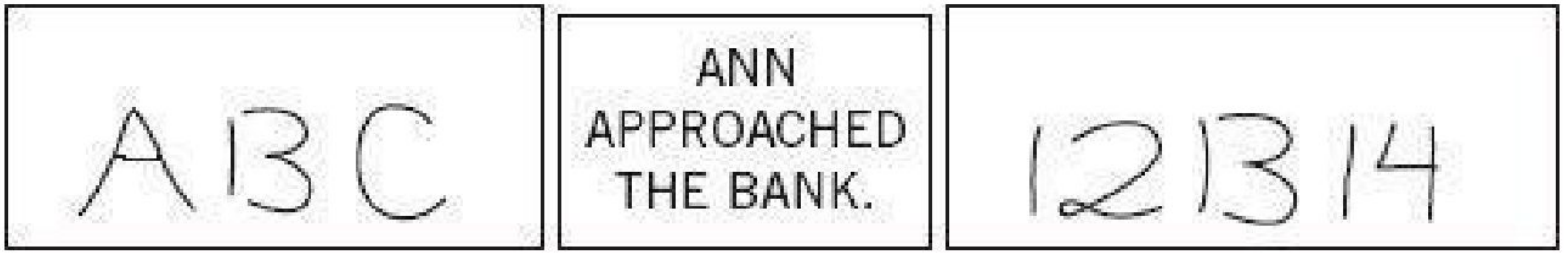
People develop positive feelings simply after encountering something repeatedly

> Experiment: ads containing one mysterious word: kadirga, saricik, biwonjni, nansoma, and iktitaf.
>
> The ads were shown at different frequencies. Those more often seen were judged much more favorably.

Applies to Chinese characters, faces, randomly shaped polygons, ...

Applies even if the subject is consciously unaware of the exposure

Neglect of ambiguity



Mental shotgun

Just like aiming precisely with a shotgun is impossible, System 1 does more work than necessary

1. Some roads are snakes.
2. Some jobs are snakes.
3. Some jobs are jails.

All are false, but 1 and 3 can be metaphorically true, so slower response times.

Availability heuristic

Listing many specific instances can make them seem harder to access.

> *First, list six instances in which you behaved assertively.*
>
> *Next, evaluate how assertive you are.*

6 vs 12 instances

“People who had just listed twelve instances rated themselves as less assertive than people who had listed only six”.

Base rate fallacy

Steve was selected at random from a representative sample

> *“Steve is very shy and withdrawn, invariably helpful but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.” Is Steve more likely to be a librarian or a farmer?*

Conjunction fallacy

Which is more likely to happen?

- A massive flood somewhere in North America next year, in which more than 1,000 people drown
- An earthquake in California sometime next year, causing a flood in which more than 1,000 people drown

Base rate fallacy

Steve was selected at random from a representative sample

> *“Steve is very shy and withdrawn, invariably helpful but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.” Is Steve more likely to be a librarian or a farmer?*

There are more than 20 male farmers for each male librarian in the United States —> Bayes’ rule

references

Daniel Kahnemann (2011): "Thinking, fast and slow".

“I placed too much faith in underpowered studies:” Nobel Prize winner admits mistakes: https://retractionwatch.com/2017/02/20/placed-much-faith-underpowered-studies-nobel-prize-winner-admits-mistakes/.

Van Gelder, T. (2005): Teaching critical thinking: some lessons from cognitive science. *College Teaching (53), 41-48.*

Willingham, D. T. Critical Thinking: Why Is It So Hard to Teach? *Arts Education Policy Review, Volume 109, 2008 - Issue 4, 21-32.*

Croskerry P, Singhal G, Mamede S. Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Qual Saf 2013;22:ii58–ii64.*

Croskerry P, Singhal G, Mamede S. Cognitive debiasing 2: impediments to and strategies for change. *BMJ Qual Saf 2013;22:ii65–ii72.*

Symborski, C., Barton, M., Quinn, M.M., Morewedge, C.K., Kassam, K.S., & Korris, J. (2014). Missing: A Serious Game for the Mitigation of Cognitive Biases.

Barton, Meg & Symborski, Carl & Quinn, Mary & Morewedge, Carey & Kassam, Karim & Korris, James. (2016). The Use of Theory in Designing a Serious Game for the Reduction of Cognitive Biases. *Transactions of the Digital Games Research Association. 2. 10.26503/todigra.v2i3.53.*

Donaldson, D. and Macpherson F. (July 2017) "Müller-Lyer Illusion" in F. Macpherson (ed.), The Illusions Index. Retrieved from https://www.illusionsindex.org/ir/mueller-lyer.