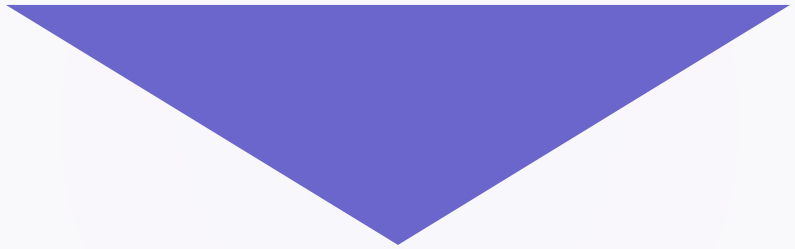
BEHAVIORAL MODELS OF INFOSEC

Industry irrationality & what to do about it

Kelly Shortridge

August 17, 2016

“Markets can stay irrational longer than you can stay solvent”



“You can stay irrational longer than you can stay uncompromised”

What is behavioral economics?

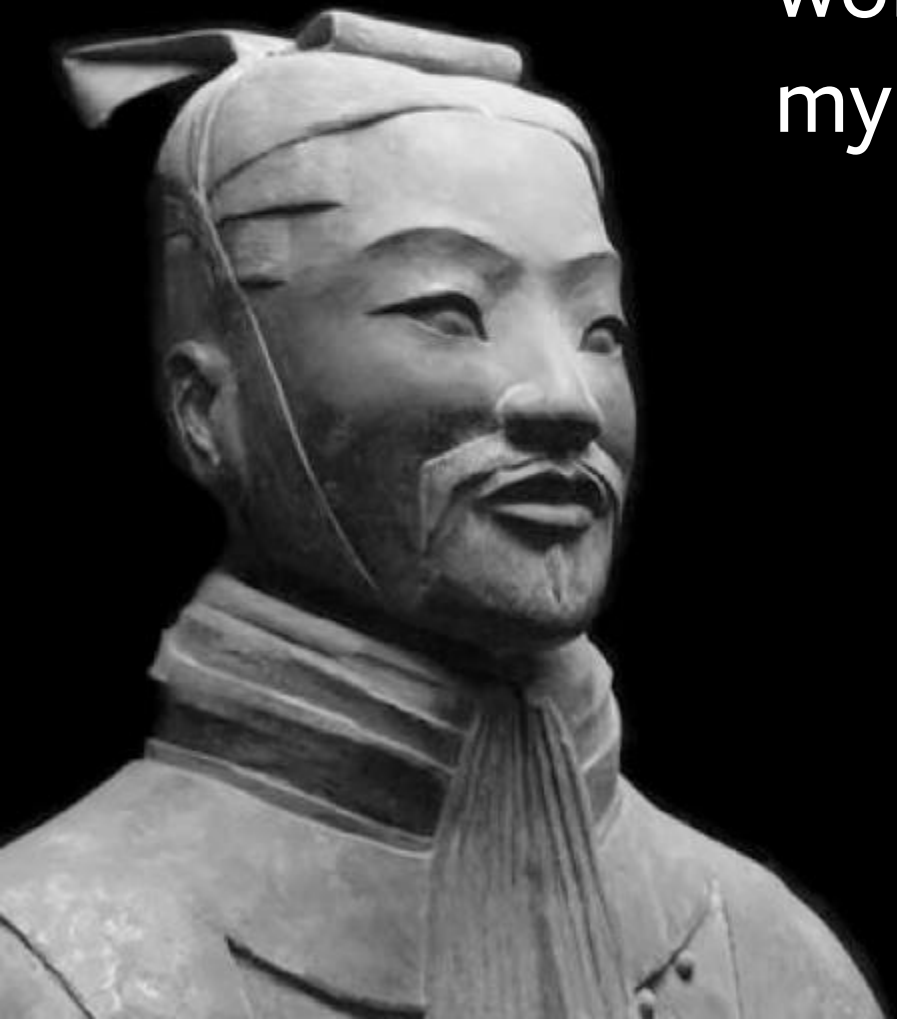
- Old school model = homo economicus (perfectly rational humans)
- Behavioral econ = measure how we *actually* behave, not how we should
- Evolutionarily viable thinking $\neq$ rational thinking
- Neckbeards wouldn't survive long in the wild

Cognitive biases

- People are “bad” at evaluating decision inputs
- They’re also “bad” at evaluating potential outcomes
- In general, lots of quirks & short-cuts (heuristics) in decision-making
- You’re probably familiar with things like confirmation bias, short-termism, Dunning-Kruger, illusion of control

Common complaints about infosec

- “Snake oil served over word salads”
- Hype over APT vs. actual attacks
- Not learning from mistakes
- Not using data to inform strategy
- Playing cat-and-mouse



"If you can't handle me at my worst, you don't deserve me at my best"

– Sun Tzu

My goal

- Start a different type of discussion on how to fix the industry, based on empirical behavior vs. how people "should" behave
- Focus on the framework; my conclusions are just a starting point
- Stop shaming defenders for common human biases; you probably suck at dieting, bro
- (also I'll show off some ~~bad~~ amazing cyber art)

What will I cover?

- Prospect Theory & Loss Aversion
- Time Inconsistency / Hyperbolic Discounting
- Less-is-better Effect
- Sunk Cost Fallacy
- Dual-system Theory
- ...and what to do about all this

Prospect Theory

Prospect theory

- People choose by evaluating potential gains and losses via probability, NOT the objective outcome
- Consistently inconsistent based on being in the domain of losses or domain of gains
- Care about relative outcomes instead of objective ones
- Prefer a smaller, more certain gain and less-certain chance of a smaller loss

Core tenets of Prospect Theory

- Reference point is set against which to measure outcomes
- Losses hurt 2.25x more than gains feel good
- Overweight small probabilities and underweight big ones
- Diminishing sensitivity to losses or gains the farther away from the reference point

Offense vs. Defense

Offense

- Risk averse
- Quickly updates reference point
- Focus on probabilistic vs. absolute outcome

Defense

- Risk-seeking
- Slow to update reference point
- Focus on absolute vs. probabilistic outcome

InfoSec reference points

- Defenders: we can withstand Z set of attacks and not experience material breaches, spending $X
  - Domain of losses
- Attackers: we can compromise a target for $X without being caught, achieving goal of value $Y
  - Domain of gains

Implications of reference points

- Defenders: loss when breached with Z set of attacks; gain from stopping harder-than-Z attacks
- Attackers: gain when spend less than $X or have outcome > $Y; loss when caught or when $X > $Y

Prospect theory in InfoSec

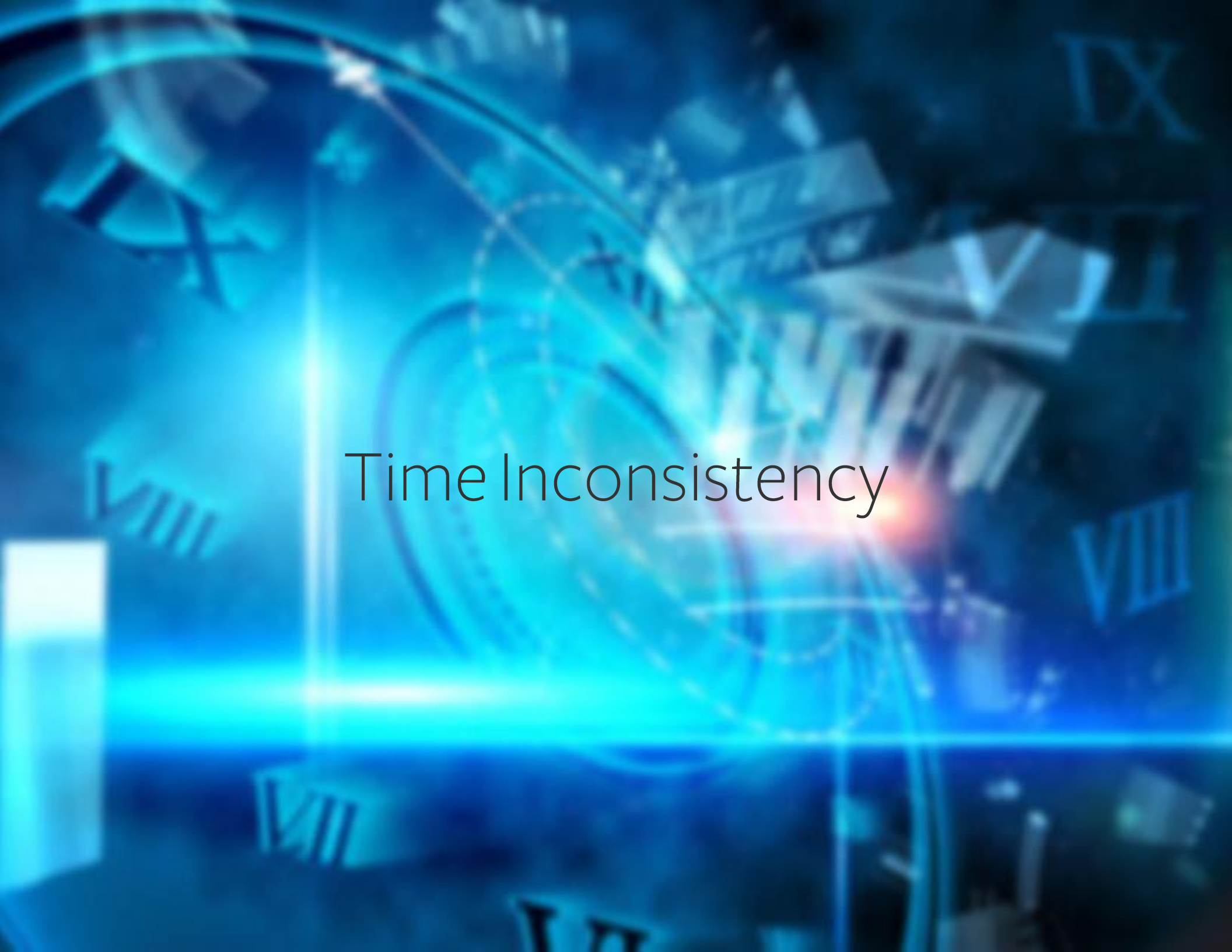
- Defenders overweight small probability attacks (APT) and underweight common ones (phishing)
- Defenders also prefer a slim chance of a smaller loss or getting a "gain" (stopping a hard attack)
- Attackers avoid hard targets and prefer repeatable / repackagable attacks (e.g. malicious macros vs. bypassing EMET)

What are the outcomes?

- Criminally under-adopted tools: EMET, 2FA, canaries, white-listing
- Criminally over-adopted tools: anti-APT, threat intelligence, IPS/IDS, dark-web anything

Incentive problems

- Defenders can't easily evaluate their current security posture, risk level, probabilities and impacts of attack
- Defenders only feel pain in the massive breach instance, otherwise "meh"
- Attackers mostly can calculate their position; their weakness is they feel losses 3x as much as defenders


Time Inconsistency

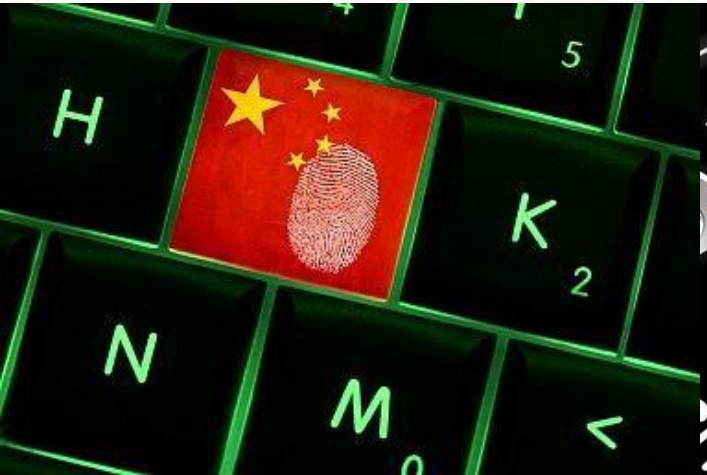
Time inconsistency

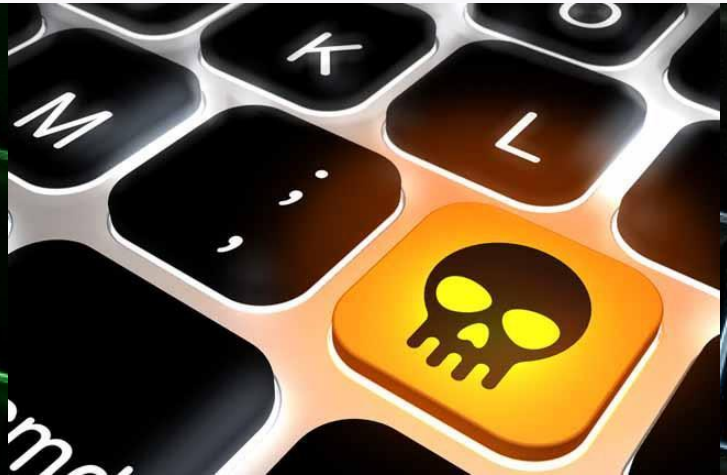
- People should choose the best outcomes, regardless of time period
- In reality: rewards in the future are less valuable (follows a hyperbolic discount)
- Classic example: kids with marshmallows; have one now or wait and get two later (they choose the marshmallow now)
- Sometimes it can be good, like with financial risk

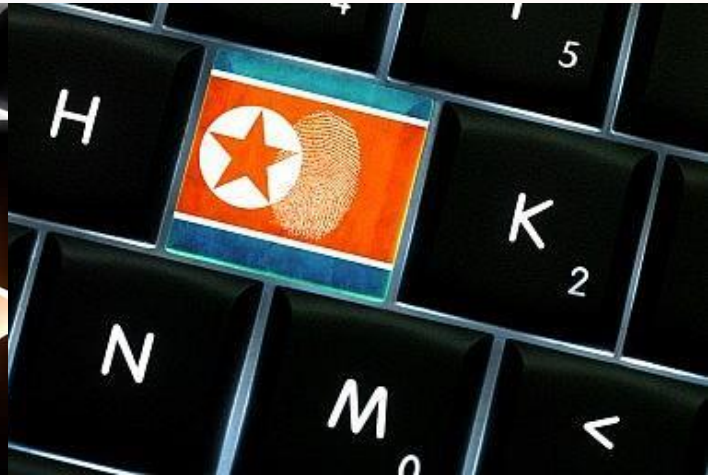
Time inconsistency in InfoSec

- Technical debt: “We’ll make this thing secure…later”
- Preferring out-of-the-box solutions vs. ones that take upfront investment (e.g. white listing)
- Looking only at current attacks vs. building in resilience for the future (even worse with stale reference points from Prospect Theory)

Less-is-better Effect







Less-is-better effect

- Evaluating things separately = lesser option
- Evaluating things together = greater option
- e.g. choose 7 oz of ice cream in an overflowing cup vs. 8 oz in a larger cup when considered apart
- Why? People focus on things that are easier to evaluate when judging separately (attribute substitution)

Attribute substitution

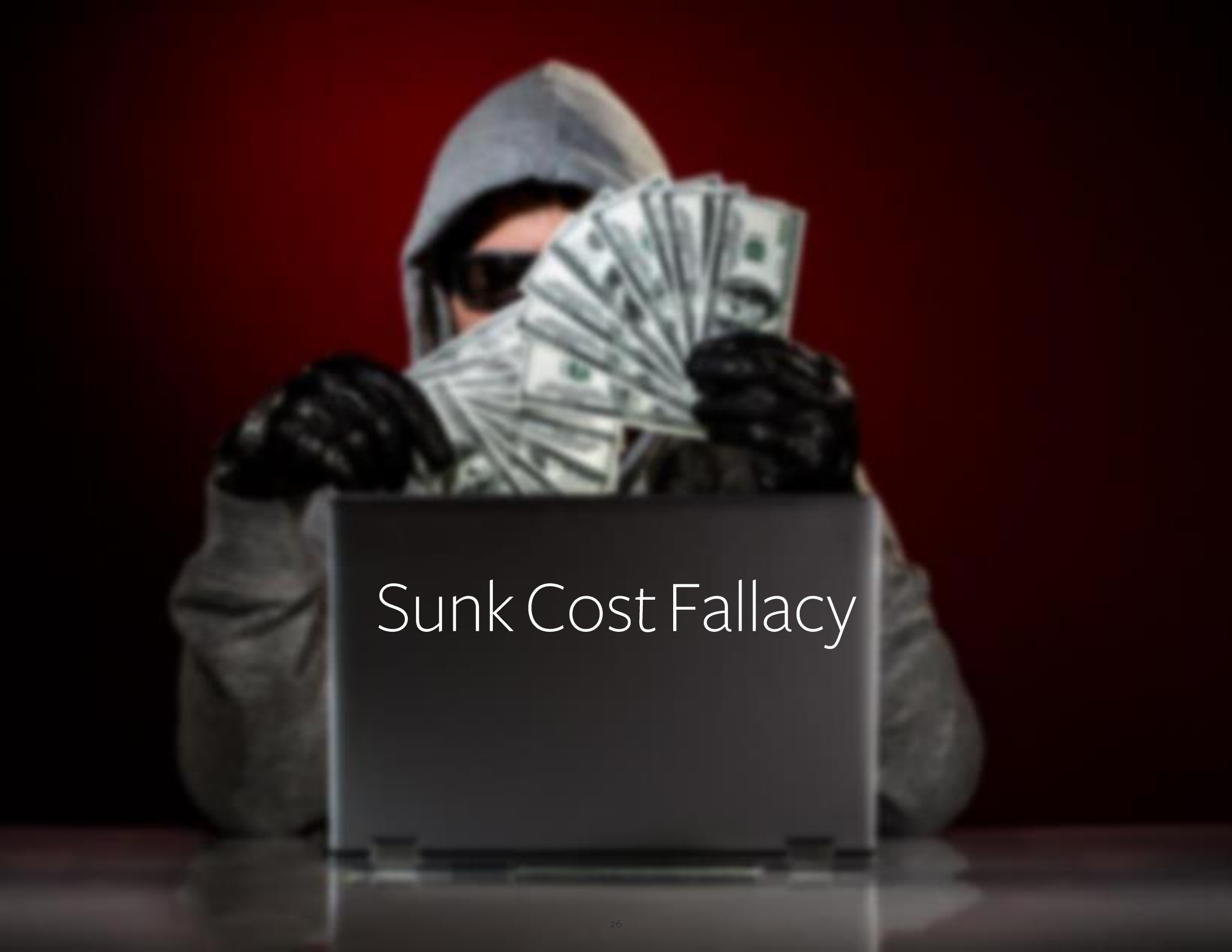
- Substitute an attribute requiring thinky-thinky for a heuristic attribute
- People do this all the time, and generally don't realize they're doing it (unconscious bias)
- Ice cream example: cup is overflowing = better
- Social example: it's hard to evaluate intelligence, so judge people based on stereotypes of relative intelligence of their race

Attribute substitution in InfoSec

- Evaluating the efficacy of a security product is really, really hard (same with security expertise)
- Easier to look for:
  - Social proof (logos on a page)
  - Representativeness (does it look like products we already use / attacks we've seen)
  - Availability (ability to recall an example, e.g. recently hyped attacks)

Less-is-better in InfoSec

- Anti-APT looks like a good deal because it probably appears low cost relative to the "high cost," unclear-riskiness attacks it's stopping
- 2FA, canaries, et al look less impressive since they're stopping most lower cost attacks, and risk you can more easily measure
- This gets even worse when you take Prospect Theory into account –defenders are really bad at estimating probabilities & impact of attacks


Sunk Cost Fallacy

Mental accounting

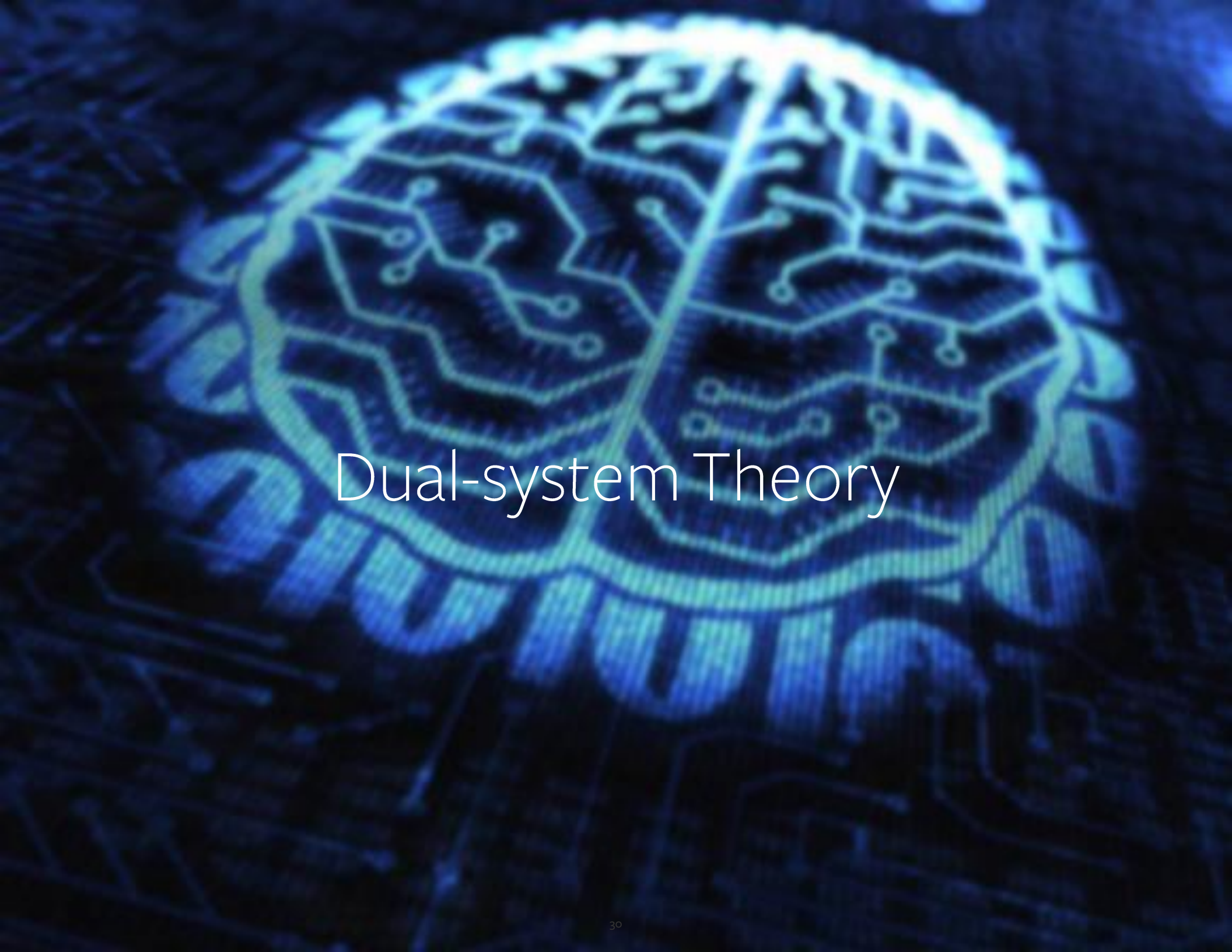
- People think about value as relative vs. absolute
- Not just about the value of an outcome or good, but also its “quality”
- People also think about money in different ways, depending on the amount, its origin and its purpose

Sunk cost fallacy

- You've bought a $20 movie ticket. It starts storming and now you don't want to go...
- ...but you do, because you "already paid for it" and "need to get your money's worth"
- This is irrational! Costs now outweigh benefits, but you're treating the costs of your time & inconvenience in a different mental account

Sunk cost fallacy in InfoSec

- Just because you spent $250k on a fancy blinky box, shouldn't keep using it if it doesn't work
- Throwing good money after bad strategies rather than pivoting to something else
- Or, "we spent all this money and still got breached, it isn't worth it to spend more now"



Dual-system Theory

Dual-system theory

- Mind System 1: automatic, fast, non-conscious
- Mind System 2: controlled, slow, conscious
- System 1 is often dominant in decision-making, esp. with time pressure, busyness, positivity
- System 2 is more dominant when it's personal and / or the person is held accountable

Dual-system theory in InfoSec

- System 1 buys products based on flashy demos at conferences and sexy word salads
- System 1 prefers established vendors vs. taking the time to evaluate all options based on efficacy
- System 1 prefers sticking with known strategies and product categories
- System 1 also cares about ego


So, what do we do about it?


Contemplate
the cybers

Improving heuristics: industry-level

- Only hype "legit" bugs / attacks (availability): very unlikely
- Proportionally reflect frequency of different types of attacks (familiarity): unlikely, but easier
- Publish accurate threat data and share security metrics (anchoring): more likely, but difficult
- Talk more about 1) the "boring" part of defense / unsexy tech that really works 2) cool internally-developed tools (social proof): easy enough

Changing incentives: defender-level

- Raise the stakes of attack + decrease value of outcome
- Find commonalities between types of attacks & defend against lowest common denominator 1st
- Erode attacker's information advantage
- Data-driven approach to stay "honest"

Leveraging attacker weaknesses

- Attackers are risk averse and won't attack if:
  - Too much uncertainty
  - Costs too much
  - Payoff is too low
- Block low-cost attacks first, minimize ability for recon, stop lateral movement and ability to "one-stop-shop" for data

How to promote System 2

- Hold defenders extra accountable for strategic and product decisions they make
- Make it personal: don't just check boxes, don't settle for the status quo, don't be a sheeple
- Leverage the "IKEA effect" – people value things more when they've put labor into them (e.g. build internal tooling)

Inequity aversion

- People really don't like being treated unfairly
- e.g. A is given $10 and can share some portion $X with B, who will get $X * 2. B then has the same option back
  - Nash Equilibrium says A gives $0 (self-interest)
  - Actual people send ~50% to player B, and B generally sends more back to A than received

Inequity aversion in infosec

- May mean defenders will be willing to share data, metrics, strategies
- Not necessarily the "as long as I'm faster than you" mentality that is commonly assumed
- Key is to set expectations of an ongoing "game"; repeated interactions promotes fairness
- So, foster a closer-knit defensive community like there exists for vuln researchers

Conclusion

Final thoughts

- Stop with the game theory 101 analyses – there are ultimately flawed, irrational people on both sides
- Understand your biases to be vigilant in recognizing & countering them
- Let's not call defenders stupid, let's walk them through how their decision-making can be improved

Questions?

- Email: kelly@greywire.net
- Twitter: @swagitda_
- Prospect Theory post: https://medium.com/@kshortridge/behavioral-models-of-infosec-prospect-theory-c6bb49902768