The Art of Explanation

Behavioral models of infosec

Kelly Shortridge

Hacktivity 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 ≠ 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 (or attributing to “sophisticated attackers” when it was really just basic phishing)
- Not learning from mistakes (see prior point)
- Not using data to inform strategy
- Playing cat-and-mouse

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 assumptions / conclusions are just a starting point
- Stop shaming defenders for common human biases
- Maybe someone will want to collaborate on an empirical study with me :)

What will I cover?

- Prospect Theory & Loss Aversion
- Time Inconsistency
- Dual-system Theory
- Groups vs. Individuals
- ...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	Defense
Risk averse	Risk-seeking
Quickly updates reference point	Slow to update reference point
Focus on probabilistic vs. absolute outcome	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 ahead of desired outcome or when \$X > \$Y

- Note: this can apply to different types of attackers – spam all the malware types want to keep ROI high via low costs; nation-state actors want ROI high via targeted, high-value assets or persistence

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 (corporate) 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

- In theory: 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)

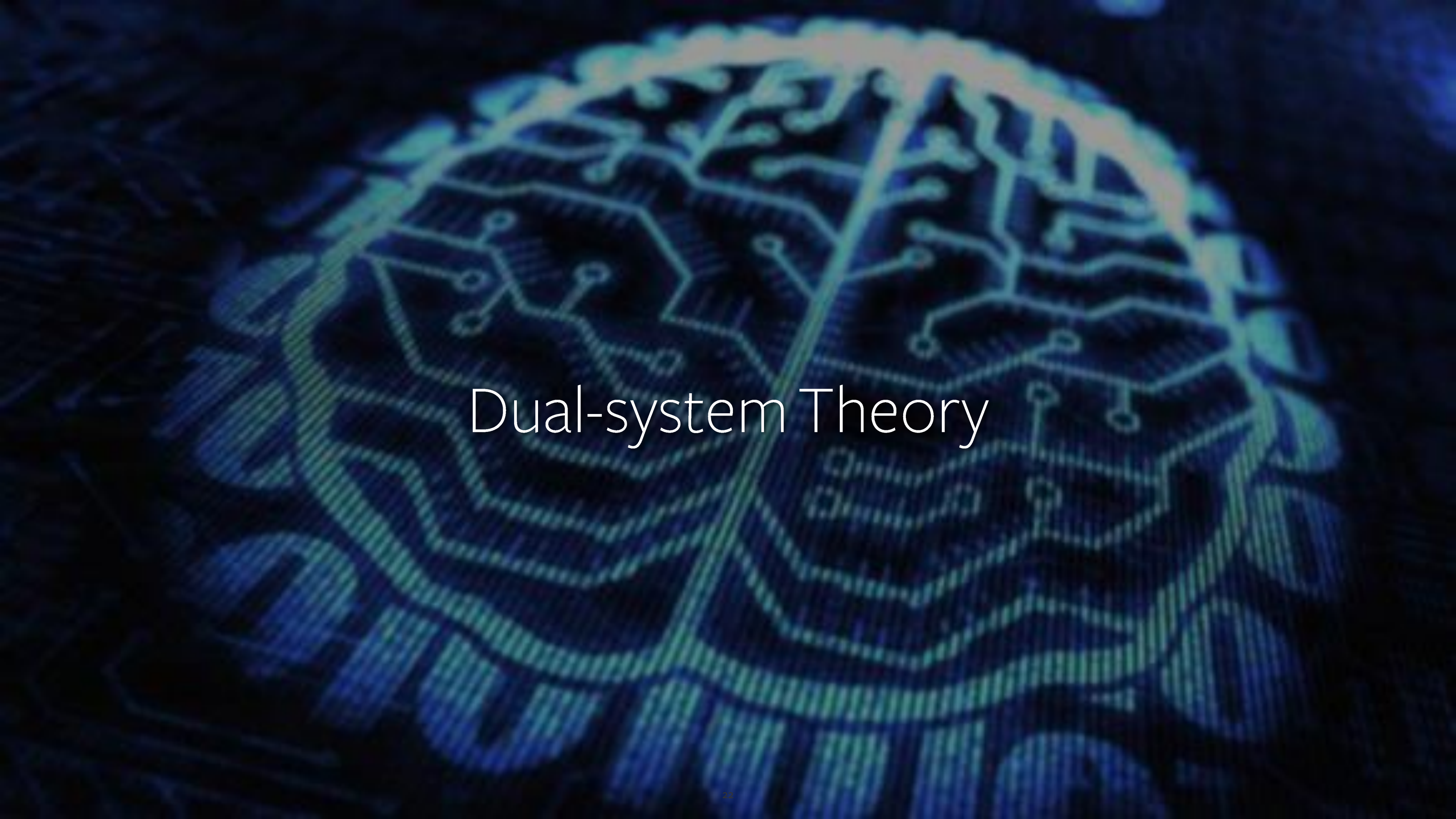
InfoSec as a public good?

- InfoSec is arguably somewhat of a public good, in that the decision makers don’t bear the full cost of the problem

- Quite a bit of research performed on time inconsistency as it relates to environmentalism (hint: delayed benefits have few fans)
  - People don’t penalize a 6 year vs. a 2 year delay much more
  - Those who like nature are less tolerant of delayed outcomes
  - Those involved in environmental orgs are more supportive of incurring costs for improvement & possess more patience

What could this mean?

- If infosec is somewhat of a public good, could imply:
  - Might as well pursue longer term, high payoff projects on a 2+ year time scale rather than “shorter” long-term time horizons
  - Employee turnover will only exacerbate the problem
  - Those who use security tools more are less tolerant of delayed outcomes to its improvement?
  - Infosec orgs could be worthwhile after all, if it increases patience with the time & money necessary for improvement



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 (attributing "advanced attackers")

What about groups?

Group vs. Individual Biases

- Infosec attackers / defenders operate on teams, so this matters
- But, the short answer is there’s less research on group behavior, so hard to say definitively what the differences are
  - Can either exacerbate biases or help reduce them ¯_(ツ)_/¯
- Depends on decision making process, type of biases, strength of biases and preference distribution among the group’s members
- Who sets the reference point for the group?

Potential risks of groups

- A leader creates new social issues – if the leader's biases are stated before a discussion, that tends to set the decision

- Some evidence that groups have a stronger "escalation of commitment" effect (doubling down)

- The term "groupthink" exists for a reason

- Groups are potentially even better at self-justification, as each individual feels the outcome is beyond their control

So, what do we do about it?

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

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 individual 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)

Other ideas

- Research has shown thinking about each side's decision trees can improve decision making (longer topic for another time)
- The more people identify with a certain cause, the less impatient they'll be with solutions to improve it (e.g. environmental groups)
- Try to shift more of the burden of the outcome onto the decision-maker – e.g. from end users to the company itself (another longer topic for another time)

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

Further research

- More research is needing on group vs. individual behavior in behavioral economics in general
- Mapping out how different types of motivations might amplify or reduce these biases
- I’d love to work with someone on empirical testing of infosec defender behaviors – get in touch if you’re game (get it?)

Questions?

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