CS 523: Social, Economic, and Legal Aspects of Security

Human Errors
Errors made inadvertently

• When intent was to not make the error
  – People who want to make decisions that are good for their own and/or their employer’s security
  – But they all too often make bad decisions

• How and why are humans so error-prone? What types of errors do they tend to make?
  – Everyone in this class (including the instructor) is inherently prone to such errors, but by becoming aware of them there is hope of mitigating them
Incomplete info, bounded rationality

• Errors are caused by both lack of info, and the limited ability of humans to process info
  – Computer analogy: Outputs computed with incomplete inputs and a limited/faulty CPU

• Humans use simple and efficient heuristics to both learn and make decisions
  – Decisions can be much worse than those resulting from a process of informed rational choice
Valence effect

• Tendency to over-estimate the likelihood of favorable outcomes
• E.g., recall the issue of insiders who commit crimes: They over-estimate both
  – The likelihood of getting away with the crimes
  – The rewards to them of committing the crimes
• This makes the fear of punishment less effective against the “insider threat”
Representativeness heuristic

• Assessing similarity of events/objects
  – Group by similarity
  – Category prototype (and judgement by prototype)

• Easy to use, but gives rise to many fallacies
  – Illusionary correlations (when none in fact exist)
  – Associating trustworthy behavior with neat appearance and design of a website
  – The above, for people’s appearance, demeanor
Endowment effect

• Value something more when you already have it
  – Alice buys mug for $5, then wants $10 to sell it
  – Alice values equally X and Y, won’t trade her X for a Y

• E.g., people assign a higher “sell” value to their private information than its “buy” value
  – Even if info has no financial consequence if released
  – Sell value = amount they request to give their info
  – Buy value = amount they’re willing to spend to make sure that their info is not released to others
Self-serving bias

• Belief that the odds for oneself are better than the odds for others
  – Like valence effect, but “self relative to others”

• Example: Users of online social networks who think personal information on social networks causes privacy problems to other users, but are not concerned about it for their own info
  – Well documented in experiments
Overconfidence

• Tendency to be more confident in one’s knowledge/abilities than what would be warranted by facts, for example:
  • Subjective confidence in knowing probabilities that are notoriously very difficult to estimate
  • E.g., in security it leads to misestimating
    – exposure to various security risks
    – ability to respond to incidents
    – consequences of a break-in
Rational ignorance

• When is it rational to choose ignorance?
• If the cost of becoming informed exceeds the benefit derived from being more informed
  – Reading privacy policy of entity holding your data
  – Reading the risk disclosures, made by a service or product provider, for using their product/service
• Can harm security and privacy
  – Failure to read fine-print that turns out to be bad
Hyperbolic discounting

• Inconsistent discounting of future events
  – Human preference for what arrives sooner rather than later (favors short term gratification even if leading to long term damages)
  – Make choices that future version of self will regret
  – Addictive behavior, gambling, reckless risk-taking

• In security and privacy, tendency of cost-benefit trade-offs to be time-inconsistent
  – Favor immediate gratification over longer term future
Inequity aversion

• Aversion to undeserved rewards
  – Whether by others or by oneself
• E.g., high sensitivity to some privacy invasions by companies, but high tolerance for others
  – Based on perceived fairness (i.e., not in relation to their seriousness and potential for future damage)
  – Tolerance for companies (like social networks) offering a free service, but not for ones that profit from users’ data without user compensation
Survivorship bias

• Concentrating on surviving entities
  – Ignoring those that perished along the way
• Leads to overly optimistic assessments
  – Mistaking luck for talent, correlation for causation
  – E.g., some career choices (college major), aspiring athletes/actors/musicians/artists
• Wrong security actions
  – Aircraft armor added in the wrong places
  – Animal injury studies based on vet data
Gambler’s fallacy

• The mistaken belief that if something happens more frequently than normal during some period, it will happen less frequently in future
  – And vice-versa (first less frequently, then more)
• Monte Carlo Casino on August 18, 1913
  – Ball fell in black 26 times in a row
  – Gamblers lost fortunes betting against black
• Also: Asylum judges, loan officers, baseball umpires, lotto players, grading quals, ... etc
Anchoring

• Tendency to attach (“anchor”) thoughts to a reference point (even apparently irrelevant)
  – The random wheel experiment of 1974

• Examples
  – List price (of software, hardware, car, house, …)
  – Diamond engagement ring (“2 month rule”)
  – Recent high of a stock that has fallen much since
Availability heuristic

• Rely on immediate examples that come to mind (believed important because recalled)
  – People’s judgments are weighed toward more recent info (opinions biased toward latest news)

• In original experiment, people were asked:
  – “If a random English word is picked, is it more likely to begin with K of to have K as third letter?”
  – Answered wrong because it’s easier to think of words beginning with K (but 3 times more words have K as third letter)
Status quo bias

• Preference for things to stay relatively same
• Example: Vast majority of users do not change their default privacy/security settings
  – Even when the settings are too permissive
  – Even when there is very little effort in changing the settings
  – Even when there is very little learning required to understand the meaning of the settings
Reciprocity and fairness

• Innate human desire to
  – Act fairly in transactions with others
  – Reward others’ behavior if deemed appropriate
  – Punish others who behave inappropriately

• People are more likely to respond to surveys when paid ahead of time rather than after

• Web sites that force registration before they provide service often receive bad (or no) info
Correlation related fallacies

• Inferring causality from correlation of A and B
• Getting the direction wrong
  – A = fast rotation of windmills, B = strong wind
  – A = being a professional swimmer, B = having an athletic and streamlined body
  – A = being in advertisements for consumer products, B = being good-looking / attractive
  – A = a country’s debt rises above 90% of GDP, B = economic growth slows down
Correlation related fallacies (cont’d)

• Missing a lurking variable C
  – A = sleeping with shoes on, B = morning headache (lurking variable C: going to sleep while drunk)
  – A = ice cream sales, B = drowning deaths (lurking variable C = hot weather)

• Correlation is spurious
  – A = amount of government spending on science and technology, B = number of suicides
Correlation related fallacies: Exercises

• What can be said if A and B correlate when
  – A = deployment of cyber-protection software (intrusion detection, anti-virus, etc), B = rate of attempted cyber-attacks (break-ins, etc)
  – A = extensive security screening prior to hiring employees, B = incidence of insider attacks
  – A = bad cholesterol numbers, B = likelihood of having a heart attack
  – A = smoking, B = likelihood of having lung cancer
Correlation related fallacies: Exercises

• What can be said if A and B correlate when:
  – A = predator numbers in a geographic area, B = prey numbers in the same geographic area
  – A = technical sophistication of malware, B = technical sophistication of anti-malware software
  – A = salary of personal assistant to the boss, B = salary of the boss
Hindsight bias

• Tendency to see a past event as having been predictable, even when it was not
  – Information after the event influences recollection
  – “I knew it all along”, “I could see it coming”, ...
  – Documented in many experiments

• Often occurs in subtle, hard-to-detect ways
  – Historians writing about battles, crises, elections
  – In trials: witnesses, juries
Hindsight bias (cont’d)

- Occurs to people who know of its existence
- Has many negative security consequences:
  - Causes memory distortions that Interfere with a victim’s ability to do forensic analysis of attack
  - Causes people to not acknowledge their responsibility for mistakes they made
  - Lowers the ability to learn from past experiences
  - Causes over-confidence, which can result in increased risk-taking
Illusion of control

• Unfounded belief in ability to control events
  – The light switch experiment (+ countless others)

• People often need this illusion, e.g.,
  – To sustain hope in bad situations
  – To get the dice to roll favorably (they roll strong if in need of a 6 or a 5, gentle if in need of a 1 or a 2)

• Placebo controls in products
  – “Elevator close” button, street crossing button, thermostat dial in hotel rooms (all are no-ops)
Confirmation bias

• Tendency to seek and interpret information in a way that confirms one's preexisting beliefs
  – Far less consideration to alternative possibilities
  – Effect is stronger for emotionally charged issues and for deeply entrenched beliefs

• Beliefs are very stubborn (esp. early ones)
  – *Continued influence effect* = tendency to believe previously learned misinformation even after it has been corrected (it still influence inferences)
Illusion of attention

• Tendency to only see what is expected, what is the focus of attention
  – In “The monkey business illusion” (you can see it on YouTube) 50% did not notice the gorilla

• Can be deadly
  – Using phone when driving impairs as much as alcohol/drugs, even with hands-free kit; usually OK if car in front slows (expected) but not if kid darts across the street (unexpected)
Base rate fallacy

• Representativeness relies on similarity
• Leads people to equate $P(A|B)$ with $P(B|A)$
  – Recall that $P(A|B) P(B) = P(B|A) P(A)$
• Example
  – Disease X’s incidence is 1 per 10,000 people. You tested positive for X in a test whose accuracy is 99%. What’s the probability that you have X?
  – Correct answer is 1% (most people get it wrong)
Inductive generalization

• Drawing general conclusions from a small number of observations
  – About a group of people based on a few members
  – About a farmer based on the fact that it fed and nurtured the turkey for the last N days
  – About your resilience to cyber-attack based on the fact that all such past attacks against you failed
  – About a species’ future survival based on the fact that it has so far avoided extinction
Law of small numbers

• Bias arising from the tendency to generalize from small amounts of data
  – People tend to underestimate the high variability of small samples
  – Even experienced scientists have fallen for it

• Example
  – Computing the average IQ of a very small class at Purdue, and using it as an estimate of the average IQ of the Purdue student population
Law of small numbers: Example

• A company has stores in many locations
  – Half the stores are very large, half are very small
• CEO examines the statistics for rates of theft at each store (as a % of sales at that store), notices that stores with highest theft rates are mainly small ones (very few are large stores)
• What can the CEO conclude from this?
  – Should extra security be deployed at small stores?
Social bias

• Abandoning one’s own information or belief under the influence of other people's actions
• Positive feedback: The more people follow an idea, the more it is deemed true by others
  – Also evident in behaviors, fashions, management techniques, hobbies, diets, religions, faculty-hiring areas, favored thesis topics, preferred employers, employment status (whether unemployed, or in high demand), choice of restaurants, suicides, ...
Social bias (cont’d)

• Advertisers cleverly exploit social bias
  – They brag about their product or service being “the most popular” or “the fastest-selling”

• Social bias also influences technology product choices (including security products)
  – Sometimes through the “halo effect” – when one positive trait of an entity X causes belief in other (unrelated) traits of X also being positive
  – Being a maverick (deviating from herd) is risky
Social bias (cont’d)

• Example of how following the herd decreases the risk of being fired:
  – Suppose decision-maker Bob follows the herd rather than his convictions, buys popular product X
  – Even if X malfunctions spectacularly, Bob gets to keep his job because “every one of our competitors also uses X”
  – There is better job security in being wrong with the crowd than right by yourself, especially when being right is a non-event (“there was no break-in”)
Groupthink

• Special case of social bias that afflicts decision-making groups that form a cohesive, close-knit, and consensus-oriented team
  – Tend to unanimously make disastrous decisions
  – Not enough dissenting voices (no one wants to disrupt the harmony, all enjoy the team’s warmth)
  – E.g., corporate or government leaders that surround themselves with “yes people” who do not question their assumptions and decisions
Groupthink (cont’d)

• Beware of unanimity of opinion and decision
  – Especially when reached without much discussion
• E.g., a security team that quickly/unanimously selects the anti-virus or intrusion detection system championed by its leader Bob
  – Without much due diligence, or discussion
  – Team member Alice had doubts, but refrained from speaking up (wanted to be a “team player”)

Groupthink (cont’d)

• Historical groupthink-caused disasters include:
  – Wrecking a country’s economy (or a sector of it)
  – Starting unnecessary and counterproductive wars
  – Causing massive corporate losses, bankruptcy

• How can a cohesive team best avoid it?
  – Appoint a designated “devil’s advocate” tasked with presenting/defending opposite point of view
  – Foster a culture that encourages questioning of assumptions, expressing disagreement
Sunk cost (a.k.a. Concorde) fallacy

• Sunk cost = cost already incurred
  – Cannot be recovered

• Bob bought a $50 non-refundable ticket
  – He no longer feels like attending it, but attends anyway because of his wish to “not waste $50”

• Tech projects that are carried through even though the commercial case for them is gone
  – Concorde airliner
Loss aversion

• Fear of loss of $X exceeds value of gain of $X
  – Partial cause of endowment effect

• Causes risk aversion, over-cautious decisions
  – E.g., preference for investments with low but certain payoffs (e.g., bonds) to risky ones with higher expected payoffs (e.g., stocks)

• Society is risk-averse
  – Laws and regulations emphasize risk reduction at the expense of losing utility of the risky activity
Loss aversion (cont’d)

• Marketers are well aware of loss aversion, and make use of it
  – E.g., the “free trial period” approach: At the expiration of the period, the user experiences having to discontinue using the software as a loss (pays the $25 to avoid that “loss”, an amount that he was unwilling to pay upfront)
Loss aversion exercises

• Which of A or B is likely to be more effective for
  – Convincing homeowner Bob to buy a new furnace (or a new A/C unit, or new insulation)
    A = “New would save you $100/month in utility bills”
    B = “Old is costing you $100/month in utility bills”
  – Convincing Alice to do self-exams for cancer
    A = “Self-exams increase the chances of finding a tumor in the early and treatable state of the disease”
    B = “Not doing self-exams decreases chances of finding a tumor in the early and treatable state of the disease”
Principal-agent problem

• Agent B makes decisions on behalf of principal A
  – Misaligned incentives: The actions of B are in his own self-interest, might not be in the best interest of A (a decision that’s optimal for B may not be optimal for A)

• Can occur when B has more information than A
  – A cannot tell the sub-optimality of B’s decisions (i.e., there is information asymmetry)

• Examples
  – B = corporate manager, A = shareholders
  – B = lawyer, A = client
Example 1 of principal-agent problem

• B gets no reward for successful risk-taking
• B is contemplating a risky investment whose outcome can result in a maximum loss of 40%, or a maximum gain of 100% (with a uniform distribution between these extremes)
  – Note that the expected gain is 30%
• B might overcome his risk aversion and make the investment for his private investment account, but not if it is part of his job as an employee of A
Example 2 of principal-agent problem

• B gets huge reward for successful risk-taking
• B is contemplating a risky investment whose outcome can result in a maximum loss of 100%, or a maximum gain of 100% (with a uniform distribution between these extremes)
  – Note that the expected gain is 0%
• Risk-averse B is unlikely to make the investment for his private investment account, but might make it if it is part of his job as an employee of A
Example 3 of principal agent problem

• B is an executive employed by A
• B is looking at deploying at A a cyber-security technology of cost X and *expected* benefit Y
  – So the net expected benefit is Y–X
• Poorly designed incentives for B could cause B to deploy the technology even when Y=X/2, or to not deploy the technology even when Y=2X
  – Incentivize B to do neither (as both are bad for A)
Mitigation of principal-agent problems

• Reward system that aligns interests of both
• Performance-related pay (narrowly economic)
  – E.g., waiter pay that heavily relies on tips (selects for friendly waiters, drives grumpy waiters away, but can also drive waiters to serve greater portions, poor service to known bad tippers, etc)
• Psycho-social compensation (pride in work)
  – Can be hurt by performance pay
Mitigation of principal-agent (cont’d)

– Measure/monitor the agent to decrease or eliminate the information asymmetry
– Threat of termination
– Profit-sharing (but: all agents benefit regardless of their respective efforts, so possibility of loafing)
– Deferred compensation (selects for workers who intend to stay longer at the firm, mitigates “not in my term of office” behavior)
Cobra effect (a.k.a. rat effect)

• Enacting solutions that worsen the problem
  – Failure to evaluate effects of “solution” (typically by ignoring perverse incentives, feedback effects)

• Rewards offered for dead cobras in Delhi, for dead rats in Hanoi (rat tail was enough)
  – Caused people to breed them for the reward
  – Scrapping the rewards worsened the problem (as now-worthless animals were set free)
Cobra effect (cont’d)

• Related “Hydra effect” is when actions, aimed at reducing a problem, instead multiply it
  – Shut down Torrent sites that come back in multiple incarnations
  – Increasing food supply of prey can lead to decrease in predator populations (first an unsustainable boom occurs, then a crash)
  – Adding a road to a road network can increase total travel time (vice-versa for a road closure)
Cobra effect (cont’d)

• Use of legal means to censor a book, suppress information (article, bad Yelp review, ...)
  – “Streisand effect”, “Banned in Boston”

• Finder’s fee offered by archeologists for each ancient parchment (1947, Dead Sea scrolls)
  – Finders tore whole parchments into pieces

• Finder’s fee offered by paleontologists for dinosaur bones (19th century in China)
  – Finders broke a large find into smaller pieces
Cobra effect (cont’d)

• Schemes’ designers erroneously assumed that their own grand intentions (advancing science, controlling pests) would be adopted by all
  – People respond more to incentives

• To avoid, consider full ramifications of the proposed “solutions”
  – Pay particular attention to incentives
  – Brainstorm with others and invite their criticism
Cobra effect (cont’d)

• Caution: Good incentive designs can be illegal and/or unethical
  – Ancient Rome forced engineers to stand below the bridge or structure that they built, at its opening ceremony
  – Reckless driving can be lowered by requiring that a sharp dagger be attached to every car’s steering wheels, pointed at the heart of the driver
Cobra effect (cont’d)

• Exercises: State the likely impact of ...
  – Paying a lawyer by the hour
  – Rewarding bank employees by the number of loans they make
  – Rewarding managers according to whether they meet their own performance targets
  – Rewarding cyber-security staff based on a low number of successful attacks
  – Making all CEO pay public (transparency)
Social loafing

• Tendency of individuals to exert less effort towards a goal when they are part of a group
  – Even when there are no coordination problems
  – Think of your group programming projects

• Effect is more (less) pronounced when individual effort is hard (easy) to distinguish from group’s

• Effort decreases as group size N increases
  – In rope-pulling experiments, with a baseline of 100% for N=1, it was 93% for N=2, 85% for N=3, 49% for N=8
Social loafing (cont’d)

• Why not zero effort?
  – Because it would be distinguishable (easier to tell that you’re not pulling at all at the rope, than to tell that you’re pulling with less than full strength)

• Also occurs in online communities
  – About 90% of participants are social-loafers (also called lurkers) who visit regularly and read what others have posted, but contribute very little
Social loafing (cont’d)

• Holds for accountability as well as for effort
• Diffusion of responsibility (not just effort)
  – Hiding behind group decisions
  – Can cause groups to take on more extreme positions than their individual members would take on individually (extreme risk if members’ individual inclination is towards risk, caution if their individual inclination is towards caution)
Regression to the mean

• When a variable that is extreme in a first measurement, tends to be less so on second

• Let $m$ be the average measured rate of security incidents per month at organization X
  – Fluctuates from month to month, but averages $m$
  – If the measured rate for one month is $2m$, the next month it is more likely to be closer to $m$
  – Similarly for a month that measures $m/2$
Regression fallacy

• Regression to the mean causes people to misattribute success or failure to measures taken in response to an extreme event
  – The “improvement” observed after the measures are taken, would have occurred anyway without them

• Examples
  – Manager Bob yells at employees who most underperformed during the past month, and praises those who most outperformed; Bob wrongly concludes that rebuke works, and that praise is counterproductive
Regression to the mean (cont’d)

• More examples
  – Teachers conclude that stick is better than carrot
  – Underperforming companies overstate the benefits of consultants hired to help improve
  – Athletes believe that appearing on the front pages of magazines “jinxes” them (it does set them up to disappoint, but not because of some “jinx”)  
  – Overstatement of the impact on a road’s safety of speed cameras installed after a surge of accidents
Outcome bias

• Judging the quality of a past decision based on its already known outcome
  – Decision should be evaluated based only on the information available at the time it was made (excluding knowledge of all other info/outcome)

• Examples
  – Judging decision to hire Bob after his cyber-fraud
  – Policy violations that increase profit are judged less harshly than those resulting in a loss
Scarcity heuristic

• Valuation based on scarcity
• The cookies experiment
  – Group A members given a box, B given 2 cookies
  – Quality rating of the cookies much higher in B
• Gmail in 2005
  – Accounts were “by invitation only”
  – More desirable (even by users who already had many email accounts of equivalent quality)
• Called “Romeo and Juliet” effect in relationships
Scarcity heuristic (cont’d)

• Other examples:
  – Pricing is “today only” or “while supplies last”
  – Red “sold” dots placed under most paintings in art galleries (makes the few available look “rare”)
  – Realtors’ “a doctor from Chicago looked at the house yesterday, are you still interested in it?”
  – The poster attractiveness experiment (students to rank 10, keep 1 as reward, re-rank after being told that “third-ranked is now unavailable”)
Omission bias

• Tendency to judge harmful actions as worse than equally harmful inactions
  – Difficult to detect the bias (hard to notice inaction)

• Exercise: Would you make decision A or B?
  – A = Approve a drug for the terminally ill, whose side effects instantly kill 20% of patients but saves the life of the other 80% (causes full remission)
  – B = Refuse to approve the drug, and 100% of the patients die
Omission bias (cont’d)

• The omission bias is ingrained in society and in the legal system, for example:
  • Euthanasia (action) vs DNR orders (inaction)
    – Euthanasia is illegal even at the explicit request of the terminally ill
    – Deliberate refusal of life-saving measures is legal by following DNR (“do not resuscitate”) orders
  • Stating untrue facts about the software, vs deliberately omitting important facts about it
Strategic misrepresentation

• Exaggeration when something big is at stake
  – Candidate for executive position who over-promises what (s)he will achieve
  – Champion for a technology or project who understates the costs and overstates the benefits

• Has many negative consequences, e.g.,
  – Project delays
  – Cost overruns
Strategic misrepresentation (cont’d)

• Particularly operative when
  – Attempts are unique, such as when a candidate for an important hire is interviewed (e.g., CEO, security firm, consulting firm)
  – Accountability is diffuse, so that many entities can plausibly be blamed for an eventual failure (i.e., a possibility of “finger pointing” exists)
  – Success or failure is determined only in the longer term (=> “circumstances have changed” excuse)
Neglect of probability

• Subjects completely ignore probabilities, instead they focus entirely on the outcomes’ magnitudes

• Subjects were partitioned into groups A and B
  – A told “you’ll get a mild electric shock”
  – B told “50% chance you’ll get a mild electric shock”
  – The same anxiety level was measured for both (even when probability was reduced to 20%, 10%, 5%)
  – But anxiety levels for both shot up if “mild shock” was replaced by “moderate shock” in the wording
Neglect of probability (cont’d)

• Researchers have shown that people are equally afraid of a 99% chance as of a 1% chance of contamination by toxic chemicals
  – They fixate on consequence, not probabilities
• The “better to wear seatbelt or not?” question
  – Subjects were told that seatbelts prevent being thrown out in case of an accident, but may prevent getting out on time if car fell into a lake or caught fire; they decided without using relative likelihoods
Neglect of probability (cont’d)

• Users who cite “fear of losing the encryption key” as the main reason for not encrypting personal confidential information that could be accessed by others (roommates, co-workers, family members, physical or cyber thieves)
  – Even when (i) the probability of data compromise was far greater than the probability of losing their encryption key; and (ii) data compromise had far more serious consequences than data loss
Conjunction fallacy

• Taking $P(X,Y)$ to be greater than both $P(X), P(Y)$
  – In fact $P(X,Y)$ must be less than both $P(X)$ and $P(Y)$

• Subjects were given a description of Alice, then asked whether it is more likely that Alice is a feminist, a bank teller, or both
  – Alice’s given description had similarity to both feminists and bank tellers
  – Often caused by representativeness heuristic
Disjunction fallacy

• Taking $P( X \cup Y )$ to be smaller than $P(X)$
• Subjects were given a description of Alice, with more similarity to that of a physicist than a biologist, then asked which of is more likely:
  – That Alice is a physicist (event $X$)
  – That Alice is a physicist (event $X$) or a biologist (event $Y$)
  – Fallacy does not occur when no similarity