# **Build Software Safely!**

17-313 Fall 2023

Foundations of Software Engineering

https://cmu-313.github.io

Michael Hilton and Edwardo Feo Flushing



# Learning Goals

- Learn to discuss risk in a project
- Strategize about ways to mitigate risk
- Learn to get early feedback to reduce risk
- Find ways to catch our technical errors







#### Administrivia

- Midterm Tues Feb 27<sup>th</sup> in class
  - We will approve ODR requests, reach out if you have concerns
- Review Session: Saturday Feb 24th from 1-3:30pm

# Risk

#### Risk



#### **Definition: Risk**

Risk is a measure of the potential inability to achieve overall program objectives within defined cost, schedule, and technical constraints.



6

# Risk is defined by two key components



The probability (or likelihood) of failing to achieve a particular outcome



The consequences (or impact) of failing to achieve that outcomes

### Internal vs. External Risk



Risks that we can control



Risks that we cannot control

## **Levels of Risk Management**

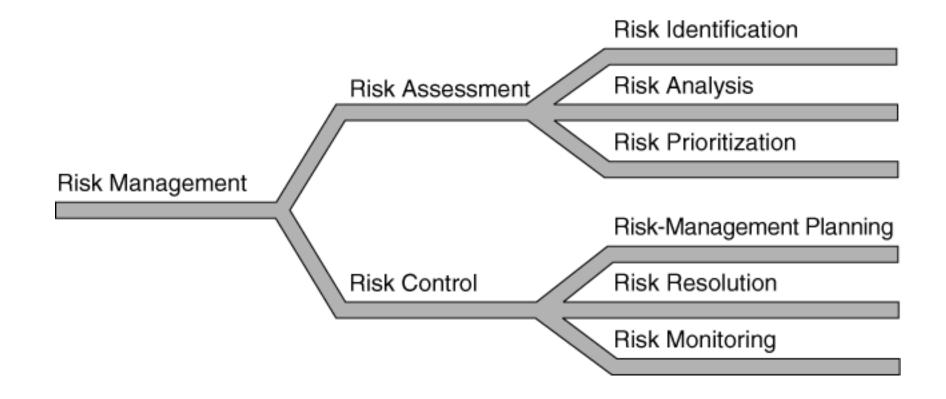
- 1. **Crisis management:** Fire fighting; address risks only after they have become problems.
- 2. **Fix on failure:** Detect and react to risks quickly, but only after they have occurred.
- 3. **Risk mitigation:** Plan ahead of time to provide resources to cover risks if they occur, but do nothing to eliminate them in the first place.
- 4. **Prevention:** Implement and execute a plan as part of the software project to identify risks and prevent them from becoming problems.
- 5. **Elimination of root causes:** Identify and eliminate factors that make it possible for risks to exist at all.



# **Levels of Risk Management**

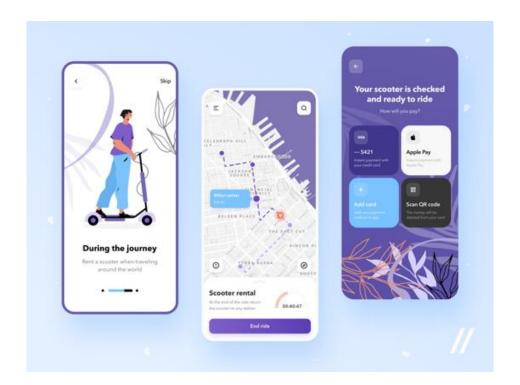
- 1. **Crisis management:** Fire fighting; address risks only after they have become problems.
- 2. **Fix on failure:** Detect and react to risks quickly, but only after they have occurred.
- 3. **Risk mitigation:** Plan ahead of time to provide resources to cover risks if they occur, but do nothing to eliminate them in the first place.
- 4. **Prevention:** Implement and execute a plan as part of the software project to identify risks and prevent them from begoming problems.
- 5. **Elimination of root causes:** Identify and eliminate factors that make it possible for risks to exist at all.

# Risk Management



#### **Team Exercise: Risk Identification**

What risks exist for the scooter app?



12



#### Risk assessment matrix



TABLE III. Risk assessment matrix

RISK ASSESSMENT MATRIX						
SEVERITY PROBABILITY	Catastrophic (1)	Critical (2)	Marginal (3)	Negligible (4)		
Frequent (A)	High	High	Serious	Medium		
Probable (B)	High	High	Serious	Medium		
Occasional (C)	High	Serious	Medium	Low		
Remote (D)	Serious	Medium	Medium	Low		
Improbable (E)	Medium	Medium	Medium	Low		
Eliminated (F)	Eliminated					



## Aviation failure impact categories

- No effect failure has no impact on safety, aircraft operation, or crew workload
- Minor failure is noticeable, causing passenger inconvenience or flight plan change
- **Major** failure is significant, causing passenger discomfort and slight workload increase
- Hazardous high workload, serious or fatal injuries
- Catastrophic loss of critical function to safely fly and land



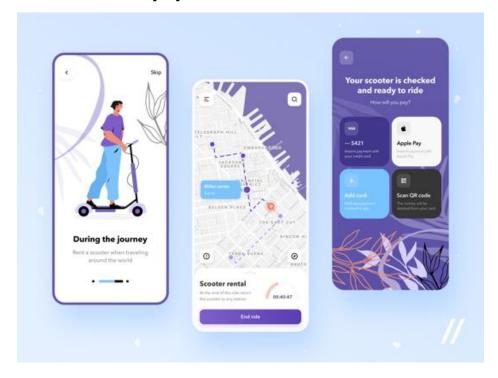


# **Risk Analysis**

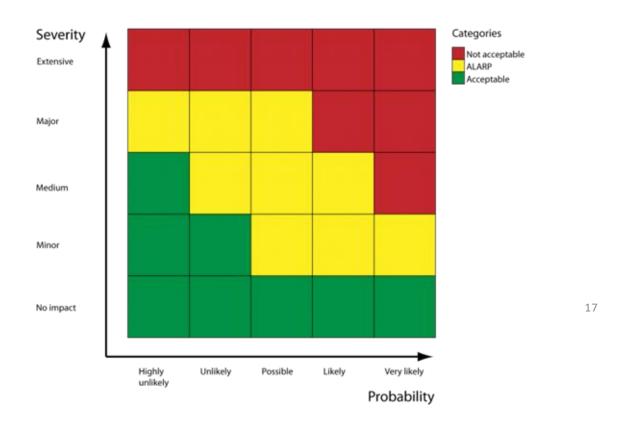
Risk	Probability (%)	Size of Loss (weeks)	Risk Exposure (weeks)
Overly optimistic schedule	50%	5	2.5
Additional features added by marketing (specific features unknown)	35%	8	2.8
Project approval takes longer than expected	25%	4	1.0
Management-level progress reporting takes more developer time than expected	10%	1	0.1
New programming tools do not produce the promised savings	30%	5	1.5
•••			
Total			12

# **Exercise: Risk Analysis**

What is the risk severity for your scooter app?



# Risk Prioritization Focus on risks with the highest exposure



#### **Risk Control**

- What steps can be taken to avoid or mitigate the risk?
- Can you better understand and forecast the risk?
- Who will be responsible for monitoring and addressing the risk?
- Have risks evolved over time?
- Bake risks into your schedule
  - Don't assume that nothing will go wrong between now and the end of the semester!

#### **DECIDE Model**

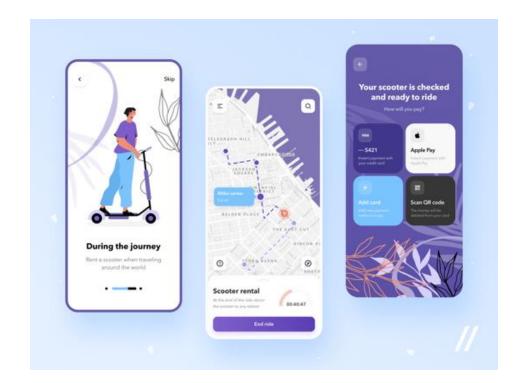
**Detect** that the action necessary Estimate the significance of the action **Choose** a desirable outcome **Identify** actions needed in order to achieve the chosen option **Do** the necessary action to achieve change **Evaluate** the effects of the action

https://www.faa.gov/regulations\_policies/handbooks\_manuals/aviation/media/FAA-H-8083-2.pdf



# Discussion: Risk Elimination and Mitigation

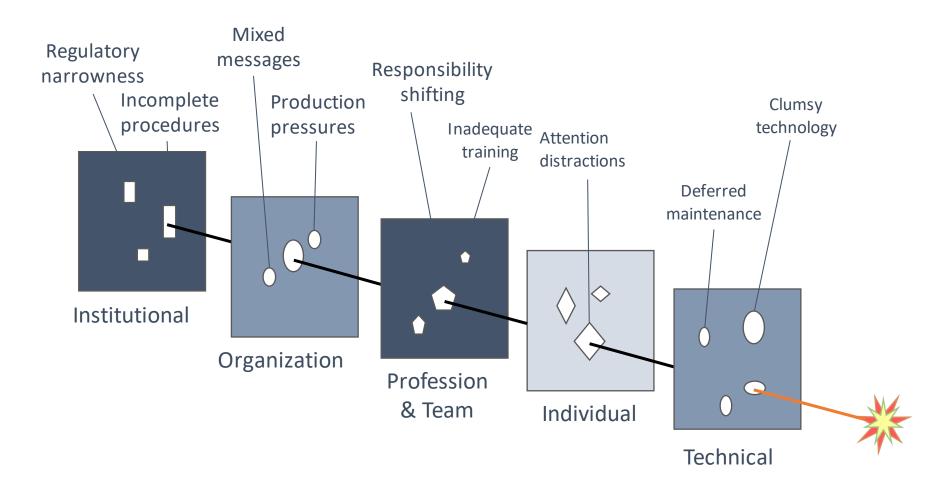
How can you eliminate/mitigate risk for your scooter app?



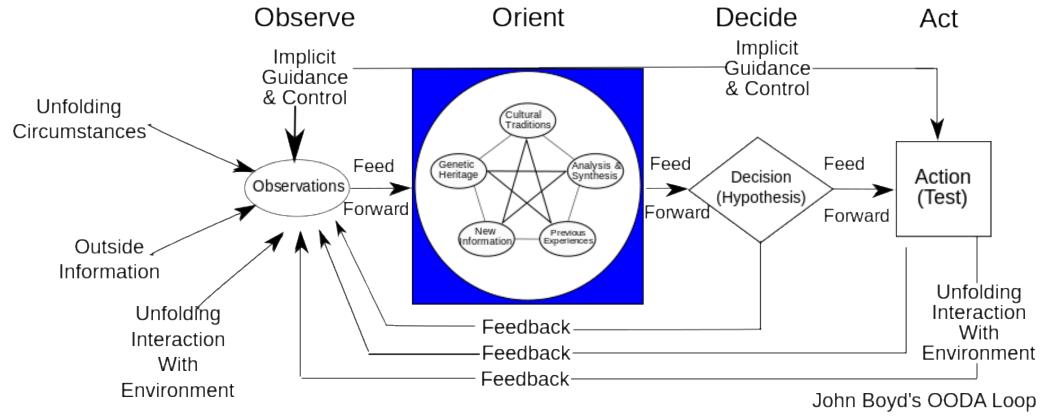
20



#### The Swiss cheese model



### OODA Loop



By Patrick Edwin Moran - Own work, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=3904554



# No matter what you do

• Some idiots won't follow your rules ©



#### Pre-mortems

 "unlike a typical critiquing session, in which project team members are asked what might go wrong, the premortem operates on the assumption that the 'patient' has died, and so asks what did go wrong."

# Why do we make misakes?

"Chock-full of fascinating examples." Why We Make Miztakes ARE ALL PRETTY SURE WE ARE WAY ABOVE AVERAGE JOSEPH T. HALLINAN

25

#### Generalization

• ...in the words of psychologist Tom Stafford, we can't find our typos because we're engaging in a high-level task in writing.

# Our brains generalize simple, component parts to focus on complex

**tasks**, so essentially we can't catch the small details because we're focused on a large task.

https://medium.com/swlh/why-we-miss-our-own-typos-96ab2f06afb7





Boredom can give rise to errors, adverse patient events, and decreased productivity—costly and unnecessary outcomes for consumers, employees, and organizations alike. As a function of boredom, individuals may feel over-worked or under-employed, and become distracted, stressed, or disillusioned. Staff who are bored also are less likely to engage with or focus on their work.

Original Articles

# Boredom in the Workplace: Reasons, Impact, and Solutions



#### Abstract

Boredom in the workplace is not uncommon, and has been discussed widely in the academic literature in relation to the associated costs to individuals and organizations. Boredom can give rise to errors, adverse patient events, and decreased productivity—costly and unnecessary outcomes for consumers, employees, and organizations alike. As a function of boredom, individuals may

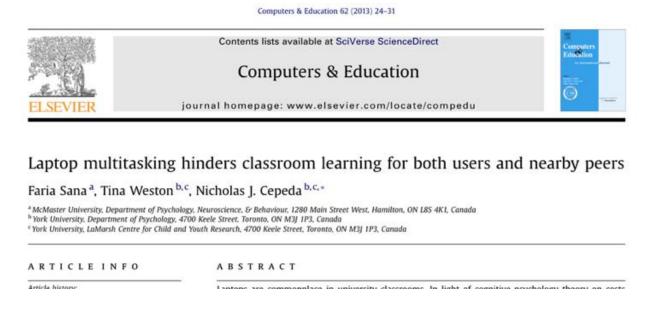




# Cognitive Load

 ..." students who switch back and forth between attending to a classroom lecture and checking e-mail, Facebook, and IMing with friends"





# Can we remove human error?

29





# Can we remove human error?

Can we catch human error before we ship our code? On we automate tasks to prevent problems?







#### Biffinctio ferta Tractatus fccundus.

Da notare circa a la figura q Dacaro posta e B stu pudi li numeri d la linea prima quale comença da n e na fin a no dicendo cost. 2.3.4.5.6.7.8.9.10.fel nuero del quate comença dal l'en a fin a fonte en contra la parima force de la pro-feccondo spacio al numero del primo referira i bancra i la primo rede de la pro-pozzió multiplici cioe dupla e se al primo referira i el numero del terço spacio ba-rai la seconda specie de la multiplici cioe tripla. E cossi sequendo in tutte le altre rigide de se to el simile trouera. Al a se al numero del secondo spacio el numero del terço spacio coparai cioc. 3.a.2. bauerai la prima ípecie de la propozitione inperparticulare cioe ferqui altera £ fe al terço el quarro cioc. 4.a.3. ferquitertia. £ fe al quarto lo quito cioc. 5.a.4. Serqui quarta e cofi in laltre fequi. £ fe al numero di terço spacio compererai el numero del qu to spacio cioe.5.a.3. Danerai la prima specie de la proportione superpartiente cioe superb partiente tertiate. Le se al numero del quarto spatio el numero del septimo refererai cio -7.a.4. bauerai la seconda specie de la proportione superpartiète cide supertripartiès qu tas. É se al numero del quinto el muo del nono cide none a.s. saras l'aterça speçie de la sup partiète cioe super quadriparties quintas. E se al numero del secondo el numero del q to cloe.5.a.2. farai la prima specie de la proportione multiplici superparticulare cloe du sesquialtera. E se a quel medessimo el septimo cloe. 7.a.2. tripla sequialtera. De la se al nu ro del terço si compara el numero de loctano cloe.8.a.3. sira la prima specie de la propo ne multiplici superpartière cioe oupla superbiparties tertias. É cost porrai p te piu ou cedere se la tauola sur a magiore. É se al quarto lundecimo quando in sosse faresti laltra fpecie vetta oupla supertri parties quartas cioca i quando più oltra volesse proceder E cosi commo babiamo vetto ve li numeri posti nela prima riga secondo li medesin glvi comparado ancora le linee inferiori quelle medesime specie te varanno che sin

la prima bai baunte, epero tu per te fequirai. 7c. Tel termino elqual se usa in venominare molte specie ve proportioni vi qui non importa altro (a te pratico) senó a piu cómodamete proferire vi cie trouato. Et est (vt supra ve multipleando integros numeros vicimi dam syllabica adiectio. Ei cómo vicemo vel via e vel sia elve susano al s

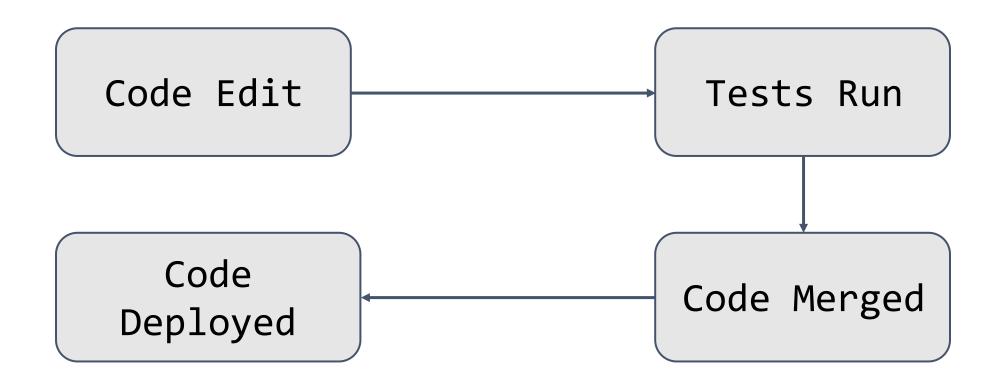
care zé. Aba el sub cha cansare lespecte o la menore inequalita si prepone a quelle giore inequalita. Est mera ppositio e così li super in più specie interposto io no mir De proportionalitatibus tra. 2º sexte vistin arti.p.

Anendo a bafrança de le proportioni parlato e quelle dinife fine a le l

# Approach: Automate what we can Review what we cannot



# CI/CD Pipeline overview



# Continuous Integration:

Catch mistakes before you push your code!

34





# History of CI



(1999) Extreme Programming (XP) rule: "Integrate Often"



(2000) Martin Fowler posts "Continuous Integration" blog



Caruisecontrol. (2001) First CI tool



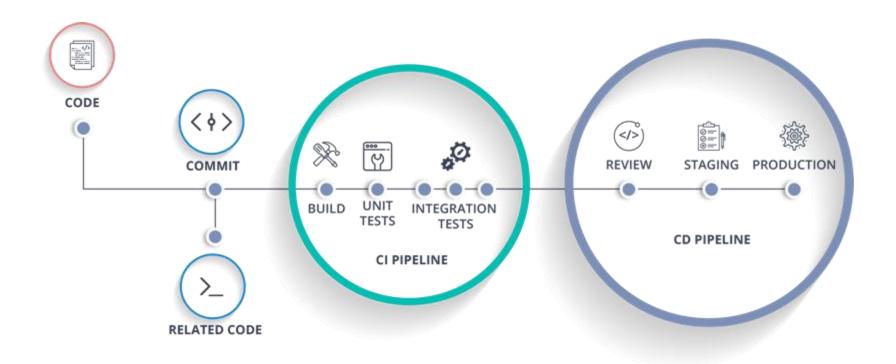
Jenkins (2005) Hudson/Jenkins





(2019) GitHub Actions

# Example CI/CD Pipeline



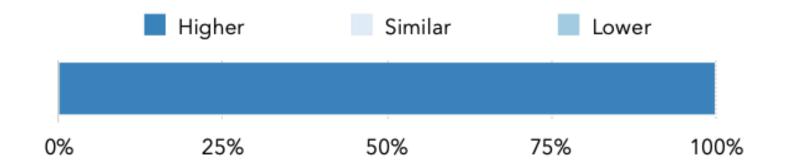
### Developers say:

CI helps us catch bugs earlier
CI makes us less worried about breaking our builds
CI lets us spend less time debugging

"[CI] does have a pretty big impact on [catching bugs]. It allows us to find issues even before they get into our main repo, ... rather than letting bugs go unnoticed, for months, and letting users catch them."

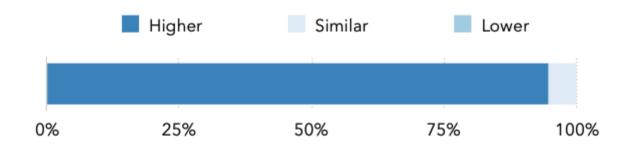


Do developers on projects with CI give (more/similar/less) value to automated tests?



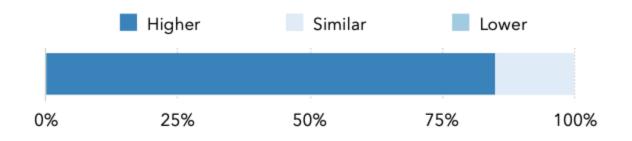
Do developers on projects with CI give (more/similar/less) value to automated tests?

Do projects with CI have (higher/similar/lower) test quality?



Do developers on projects with CI give (more/similar/less) value to automated tests?

Do projects with CI have (higher/similar/lower) test quality? Do projects with CI have (higher/similar/lower) code quality?





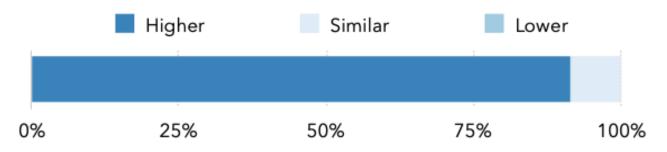


Do developers on projects with CI give (more/similar/less) value to automated tests?

Do projects with CI have (higher/similar/lower) test quality? Do projects with CI have (higher/similar/lower) code quality?

Are developers on projects with CI (more/similar/less)

productive?





### Observation

# CI helps us catch errors before others see them

44



### CI can run static and dynamic analysis

Require approval from specific reviewers before merging Add rule Rulesets ensure specific people approve pull requests before they're merged. All checks have passed Hide all checks 11 successful checks ✓ → Homework 1 Check / Homework 1 (ubuntu-latest, 16) (pull\_request) Successful in 1m **Details** ✓ ★ Lint / Lint (ubuntu-latest, 16) (pull\_request) Successful in 3m Details ✓ **(Test / Test (ubuntu-latest, 16, mongo-dev) (pull\_request)** Successful in 6m **Details** ✓ ★ Test / Test (ubuntu-latest, 16, mongo) (pull\_request) Successful in 5m Details ✓ **(pull\_request)** Successful in 5m Details ✓ **(Test / Test (ubuntu-latest, 16, postgres) (pull\_request)** Successful in 6m **Details** This branch has no conflicts with the base branch Merging can be performed automatically. Merge pull request You can also open this in GitHub Desktop or view command line instructions.

### Static Validation

- Style guides
- Compiler warnings and errors
- Static analysis
  - FindBugs
  - clang-tidy
  - Pylons Webtest
- Code review

### Style Guide

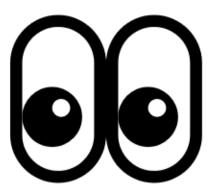
- List of environment-specific preferred practices
- Could include:
  - Libraries / idioms to use
  - Formatting

### Style Guide Examples

- https://www.python.org/dev/peps/pep-0008/
- https://github.com/airbnb/javascript
- https://subversion.apache.org/docs/communityguide/conventions.html
- https://google.github.io/styleguide/cppguide.html
- https://google.github.io/styleguide/pyguide.html
- Linux kernel style guide



### Who writes these style guides?



### Who writes these style guides?

(ad hoc ) Self-proclaimed code protectors

(wisdom) Team veteran developers

(copy-paste) Google search for blog posts by experts

(empirical study) Evidence-based analysis of code styles that

correlate with bugs



# For problems we can't easily automate, we can perform code review





## Boeing Model 299 test on October 30, 1935.

- Plane crashed because of locked elevator control surface (opposite effect of MCAS)
- 4 engines were deemed "too complex"
- Test pilots developed checklists to help them fly



### Checklists help manage complex processes







The Checklist: https://www.newyorker.com/magazine/2007/12/10/the-checklist



#### **Dr. Peter Pronovost**

- Inspired by B-17 Story
- After checklist, ten-day lineinfection rate went from eleven per cent to zero
- In 15 months, only two line infections occurred
- For one hospital, the checklist had prevented forty-three infections and eight deaths, and saved \$2M



#### The Pronovost Checklist

Central venous catheters, or lines, are used for medications, blood, fluids or nutrition and can stay in for days or weeks. But bacteria can grow in the line and spread a type of infection to the bloodstream, which causes death in one out five patients who contract it. This five-step checklist for doctors and nurses to use before inserting a line can prevent infections and death.

- Wash hands with soap and water or an alcohol cleanser
- 2. Wear sterile clothing a mask, gloves, and hair covering—and cover patient with a sterile drape, except for a very small hole where the line goes in
- 3. Clean patient's skin with chlorhexidine (a type of soap) when the line is put in

......

- 4. Avoid veins in arm and leg, which are more likely to get infected than veins in chest
- Check the line for infection each day and remove when no longer needed

Source: Dr. Peter Pronovost

https://www.wsj.com/articles/SB10001424052748704364004576131963185893084





# Difference between Pilot and Doctor error?

Which is Developer error more like?



### How to create a checklist?

- Start with problems we have seen before
  - "Safety regulations are written in blood"
- Justify why this is not automatable
- Not all checklist items need to be very specific
  - An item could be "does this team know we are proposing this change"

### Activity: Create a checklist

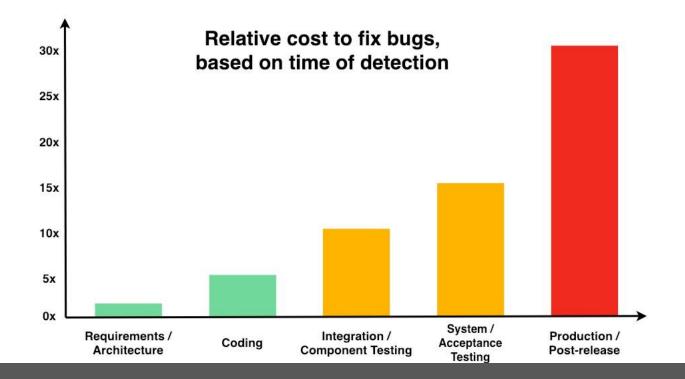
- In pairs, think about dumb mistakes your "friend" made the last time they were coding.
  - Write your names on a piece of paper.
  - Write down two checklist items that would have caught those errors.
- Divide into teams: left and right sides of the classroom.
- Which team had the most unique/good entries in their list?



# Expectations and Outcomes for code review

### Motivation

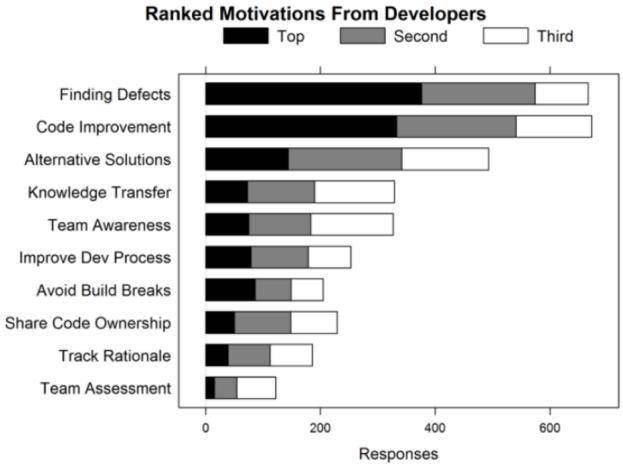
- Linus's Law: "Given enough eyeballs, all bugs are shallow."
  - - The Cathedral and the Bazaar, Eric Raymond







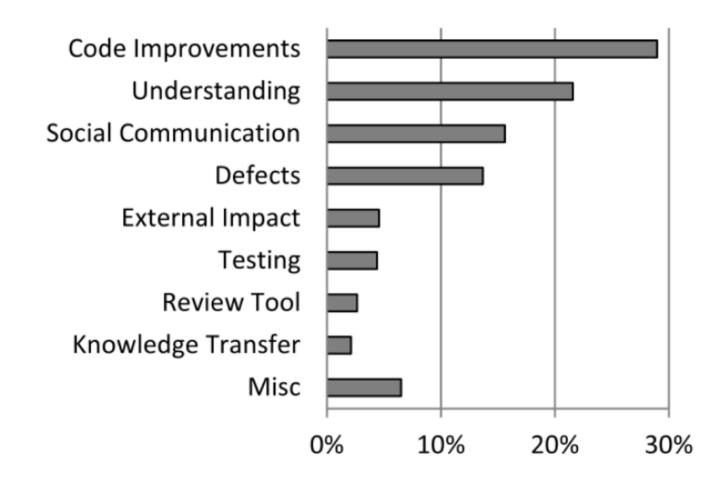
### Code Review at Microsoft



Bacchelli, Alberto and Christian Bird. "Expectations, outcomes, and challenges of modern code review." Proceedings of the 2013 International Conference on Software Engineering. IEEE Press, 2013.



### Outcomes (Analyzing Reviews)



## Mismatch of Expectations and Outcomes

- Low quality of code reviews
  - Reviewers look for easy errors, as formatting issues
  - Miss serious errors
- Understanding is the main challenge
  - Understanding the reason for a change
  - Understanding the code and its context
  - Feedback channels to ask questions often needed
- No quality assurance on the outcome

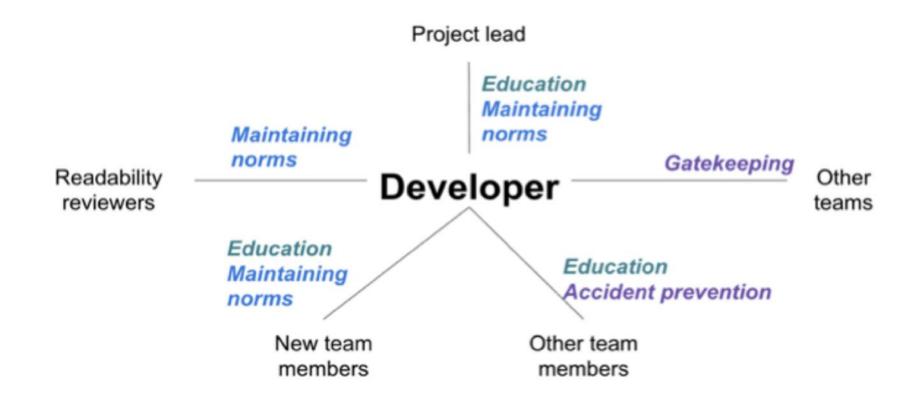
### Code Review at Google

- Introduced to "force developers to write code that other developers could understand"
- Three benefits:
  - checking the consistency of style and design
  - ensuring adequate tests
  - improving security by making sure no single developer could commit arbitrary code without oversight

Caitlin Sadowski, Emma Söderberg, Luke Church, Michal Sipko, and Alberto Bacchelli. 2018. Modern Code Review: A Case Study at Google. International Conference on Software Engineering

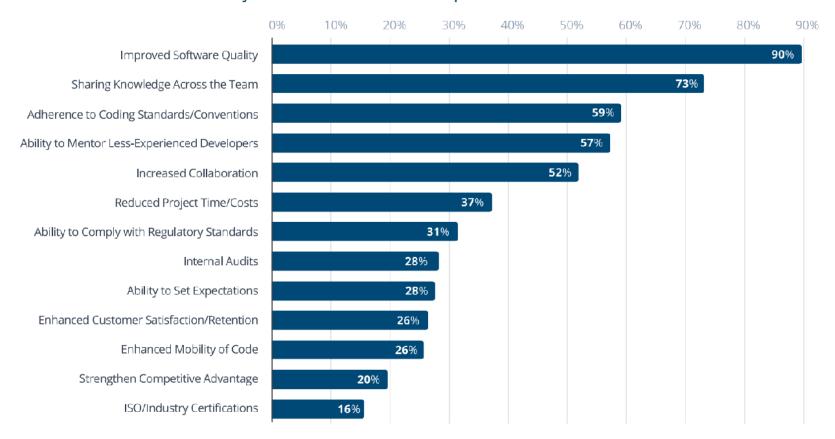


### Reviewing Relationships



### The State of Code Review survey

#### What do you believe are the most important benefits of code review?



n = 1129

### Code Review

- Start with the "big ideas"
- Automate the little things
- Focus on understanding
- Remember a person wrote the code
- Don't overwhelm the person with feedback

### Don't forget that coders are people with feelings

- A coder's self-worth is in their artifacts
- Cl can avoid embarrassment
- Identify defects, not alternatives; do not criticize coder
  - "you didn't initialize variable a" -> "I don't see where variable a is initialized"
- Avoid defending code; avoid discussions of solutions/alternatives
- Reviewers should not "show off" that they are better/smarter
- Avoid style discussions if there are no guidelines
- The coder gets to decide how to resolve fault



### Risk Analysis:

- Probability a human makes a mistake: Very Likely
- Severity: ranges, but could be extensive

#### Solution:

Use CI to catch your mistakes, make you look better, and mitigate your risks!

Use Code review to teach and learn

