CI and Developer Contribution

Omar Elazhary

Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- Data Collection
- Findings So Far
- Research Questions

Agenda

- **Motivation**
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- Data Collection
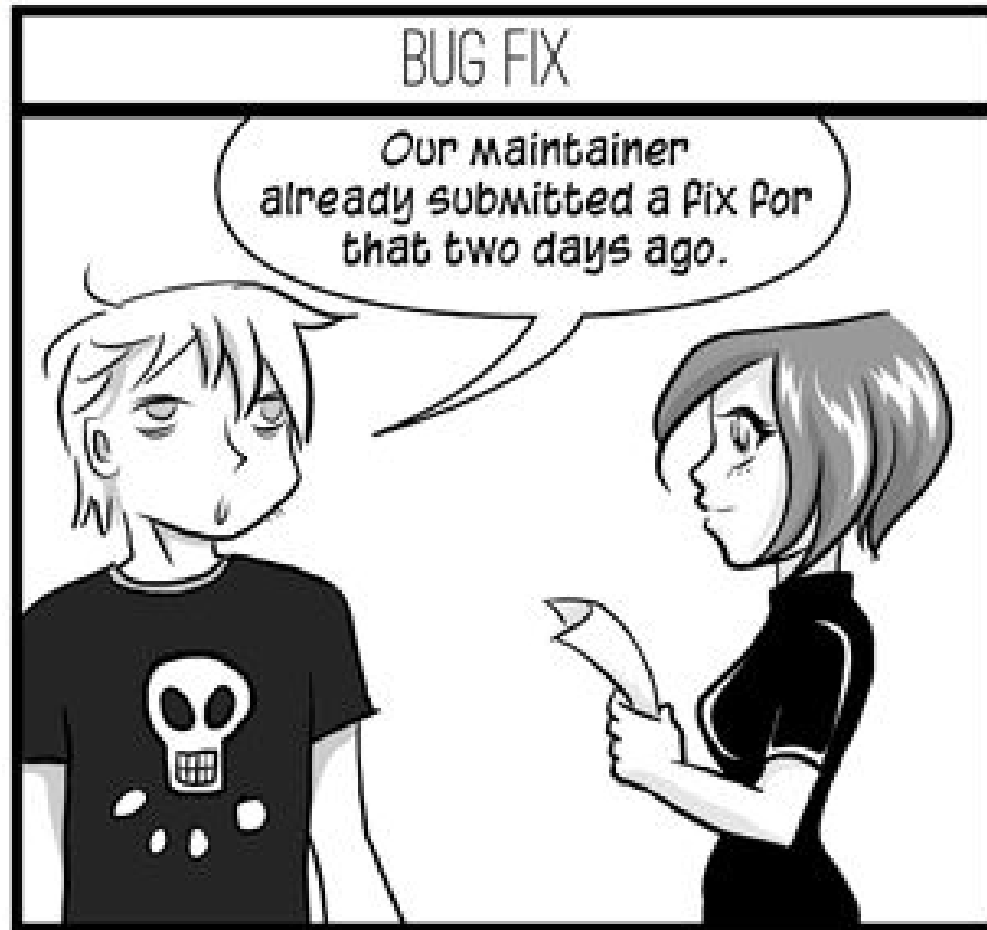- Findings So Far
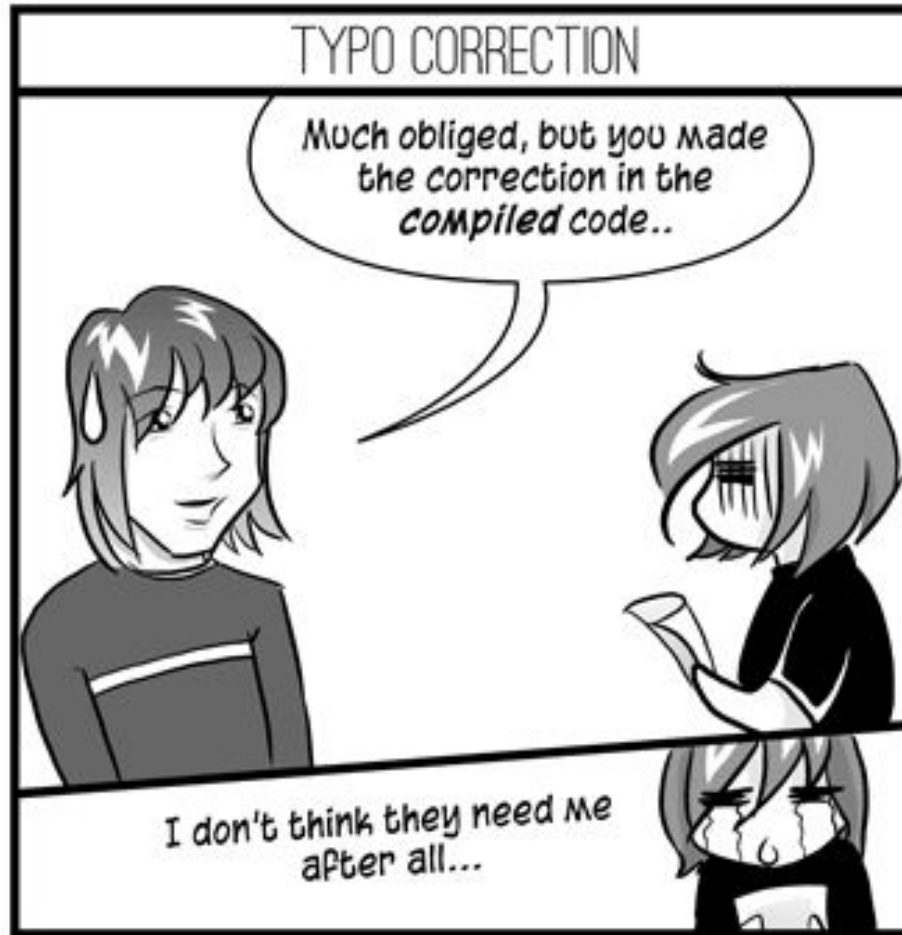- Research Questions

Motivation

Comic from http://rachelnabors.com/2012/04/of-github-and-pull-requests-and-comics/

Motivation

Comic from http://rachelnabors.com/2012/04/of-github-and-pull-requests-and-comics/

Motivation

Comic from http://rachelnabors.com/2012/04/of-github-and-pull-requests-and-comics/

Motivation

Comic from http://rachelnabors.com/2012/04/of-github-and-pull-requests-and-comics/

Motivation

Comic from http://rachelnabors.com/2012/04/of-github-and-pull-requests-and-comics/

Agenda

- Motivation
- **Background**
  - **Open Source Software Development**
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- Data Collection
- Findings So Far
- Research Questions

Open Source Software Development

- Development in large OSS communities
- Prime example: GitHub
- Not just a *meritocracy*

Agenda

- Motivation
- **Background**
  - Open Source Software Development
  - **Difference in PR Acceptance**
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- Data Collection
- Findings So Far
- Research Questions
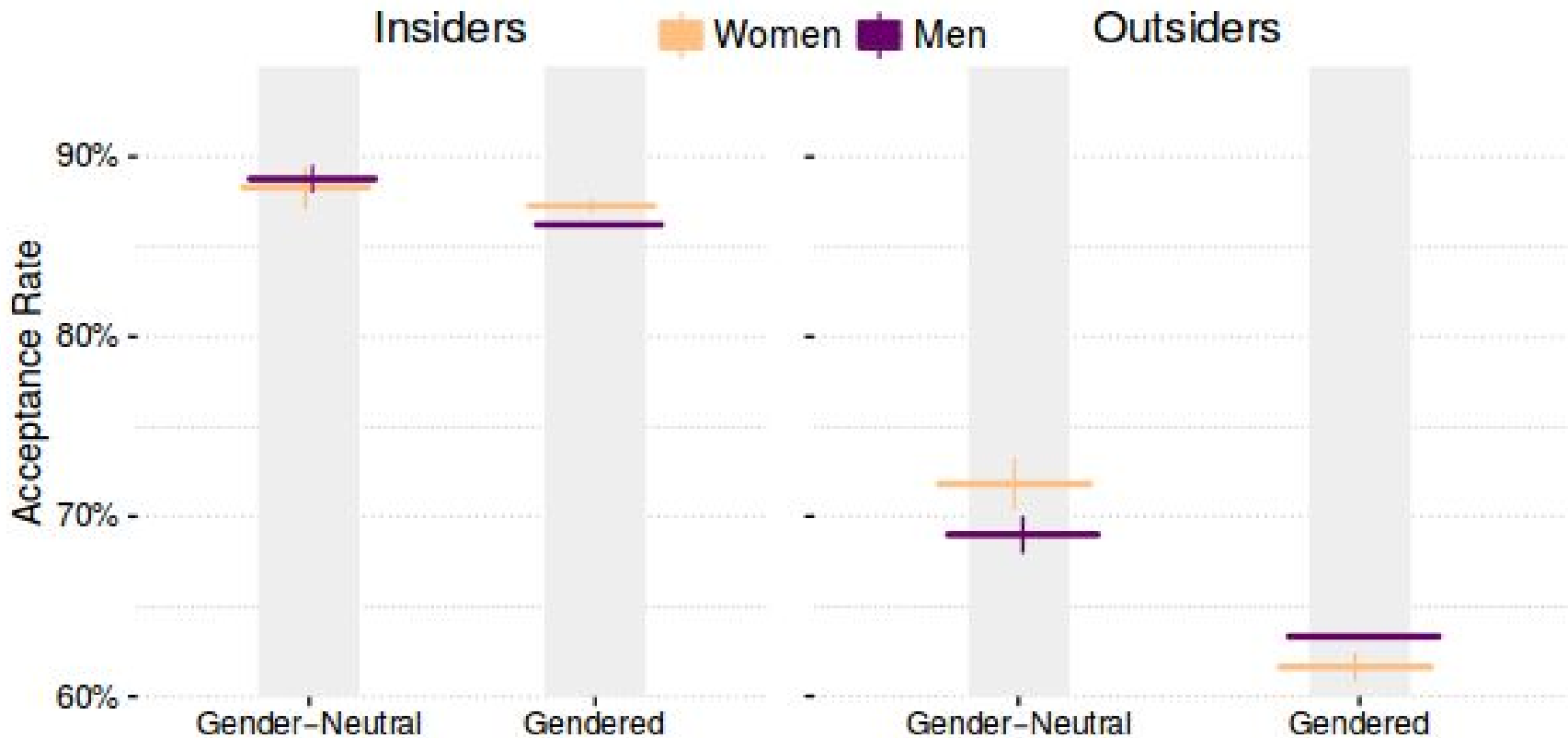
Differences in Pull Request Acceptance

Image from Terrell2016 - Gender Differences and Bias in Open Source: Pull Request Acceptance of Women Versus Men

Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - **Technical Factors in PR Acceptance**
  - Social Factors in PR Acceptance
- Theory
- Data Collection
- Findings So Far
- Research Questions

Technical Factors in PR Acceptance

- ## Objective Factors:
  - ○ Test Inclusion
  - ○ Commit Size
  - ○ Number of Files Changed
- ## Subjective Factors:
  - ○ Code Style
  - ○ Project/Technical Fit

1. J. Tsay, L. Dabbish, and J. Herbsleb, "Influence of Social and Technical Factors for Evaluating Contribution in GitHub."
2. Gousios, Georgios, et al. "Work practices and challenges in pull-based development: the integrator's perspective." *Proceedings of the 37th International Conference on Software Engineering-Volume 1*. IEEE Press, 2015.

Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - **Social Factors in PR Acceptance**
- Theory
- Data Collection
- Findings So Far
- Research Questions

Social Factors in PR Acceptance

- Social Interaction[1]
  - Social distance
  - Following
  - Prior interaction
  - Number of comments on the PR
- Social Status[1]
  - Number of followers
  - Is a core contributor

- Project Establishment[1]
  - Repo age
  - Number of collaborators
  - Stars
- Perceived Gender[2]

1. J. Tsay, L. Dabbish, and J. Herbsleb, "Influence of Social and Technical Factors for Evaluating Contribution in GitHub."
2. J. Terrell *et al.*, "Gender Differences and Bias in Open Source: Pull Request Acceptance of Women Versus Men," *PeerJ Prepr.*, no. e1733v1, pp. 1–30, 2016.

Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- **Theory**
- Data Collection
- Findings So Far
- Research Questions

Theory

- Several tools ease friction between developers
- Build automation tools (CI tools) can run *simulations* of a submitter's code before merging
- CI tools demonstrate a submitter's technical prowess
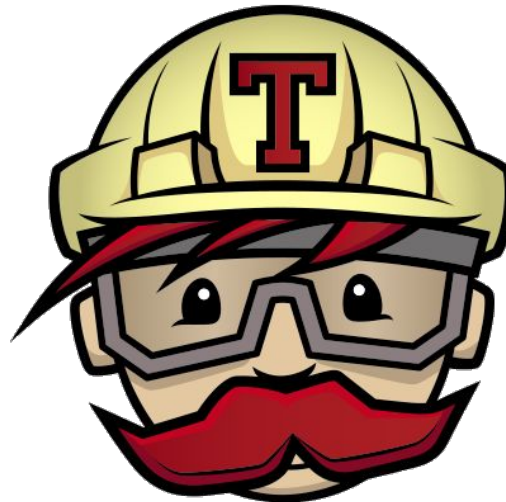- CI tools may take the focus off of social factors

Theory - Popular CI Tools
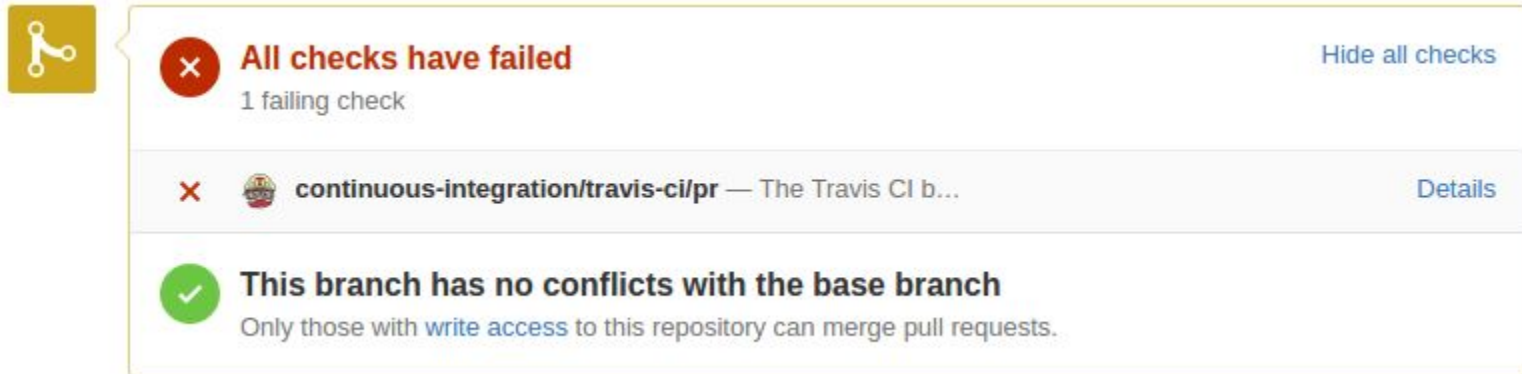
Theory - Effect of CI Tools

- Increases contribution throughput[1]
- Has a positive effect on quality[1,2]
- Effect on PR acceptance criteria??
- Effect on onboarding process??

1. B. Vasilescu, Y. Yu, H. Wang, P. Devanbu, and V. Filkov, "Quality and productivity outcomes relating to continuous integration in GitHub," in ESEC/FSE conf., 2015, pp. 805–816.
2. J. Holck and N. Jørgensen, "Continuous integration and quality assurance: A case study of two open source projects," Australas. J. Inf. Syst., vol. 11, no. 2003, pp. 40–53, 2007.
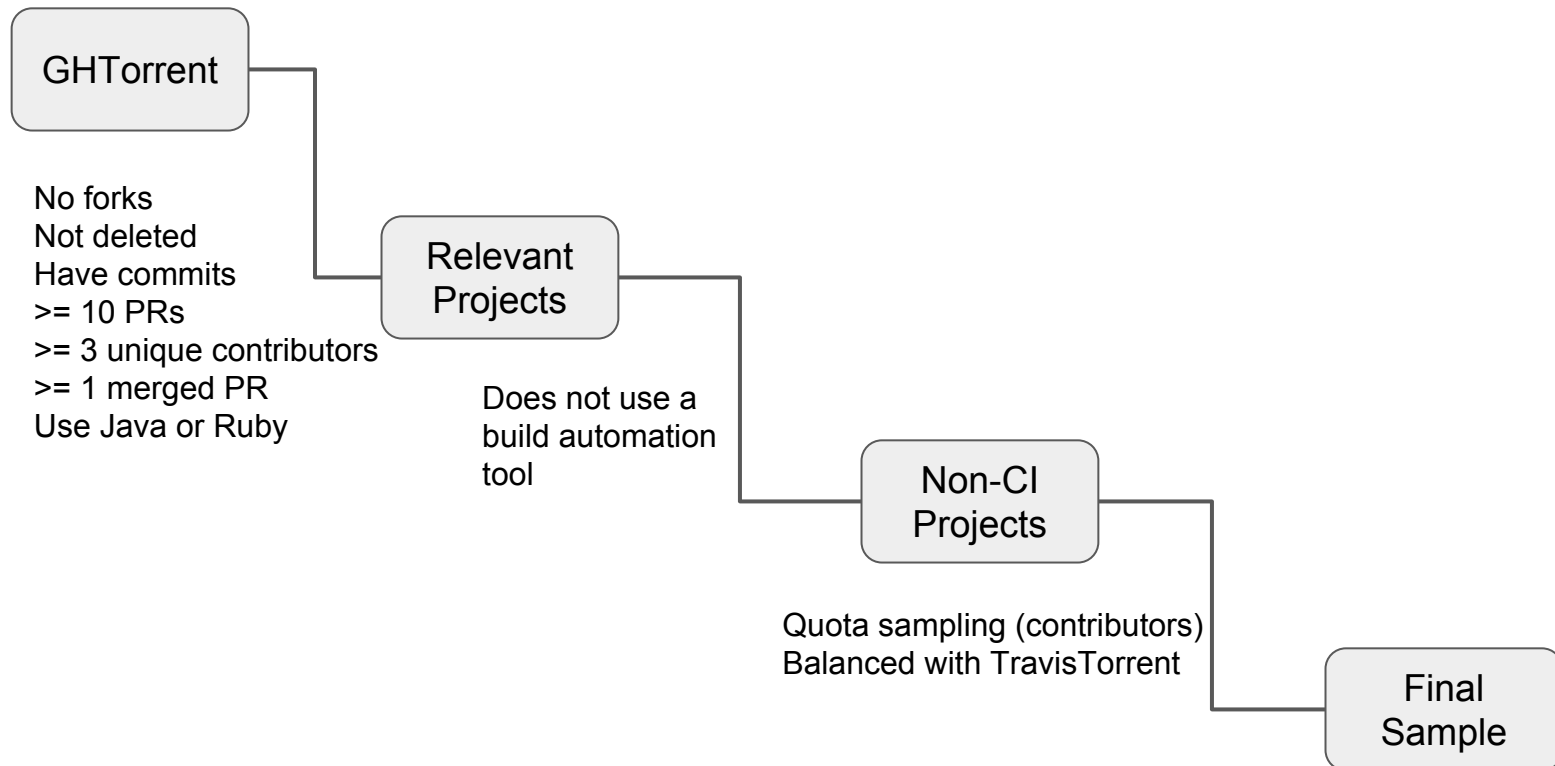
Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- **Data Collection**
- Findings So Far
- Research Questions

Data Collection

- GHTorrent
- TravisTorrent

GHTorrent

No forks
Not deleted
Have commits
>= 10 PRs
>= 3 unique contributors
>= 1 merged PR
Use Java or Ruby

Relevant Projects

Does not use a
build automation
tool

Non-CI Projects

Quota sampling (contributors)
Balanced with TravisTorrent

Final Sample

Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- Data Collection
- **Findings So Far**
- Research Questions

Findings So Far

- Very few social metrics can be **reliably** inferred from the available data sources
- GitHub is **not** the only communication channel used
- Every project uses CI build tools differently
- Quantitative methods alone are not the way to go

Agenda

- Motivation
- Background
  - Open Source Software Development
  - Difference in PR Acceptance
  - Technical Factors in PR Acceptance
  - Social Factors in PR Acceptance
- Theory
- Data Collection
- Findings So Far
- **Research Questions**

Research Questions

- **What is the effect of CI on factors considered in PR review?**
- How does CI affect the onboarding process for newcomers?