# Reproducibility, Productivity, Sustainability

**Michael A. Heroux**

# Reproducibility is essential

# Reproducibility

SCIENCE

## *Many Psychology Findings Not as Strong as Claimed*

By BENEDICT CAREY AUG. 27, 2015

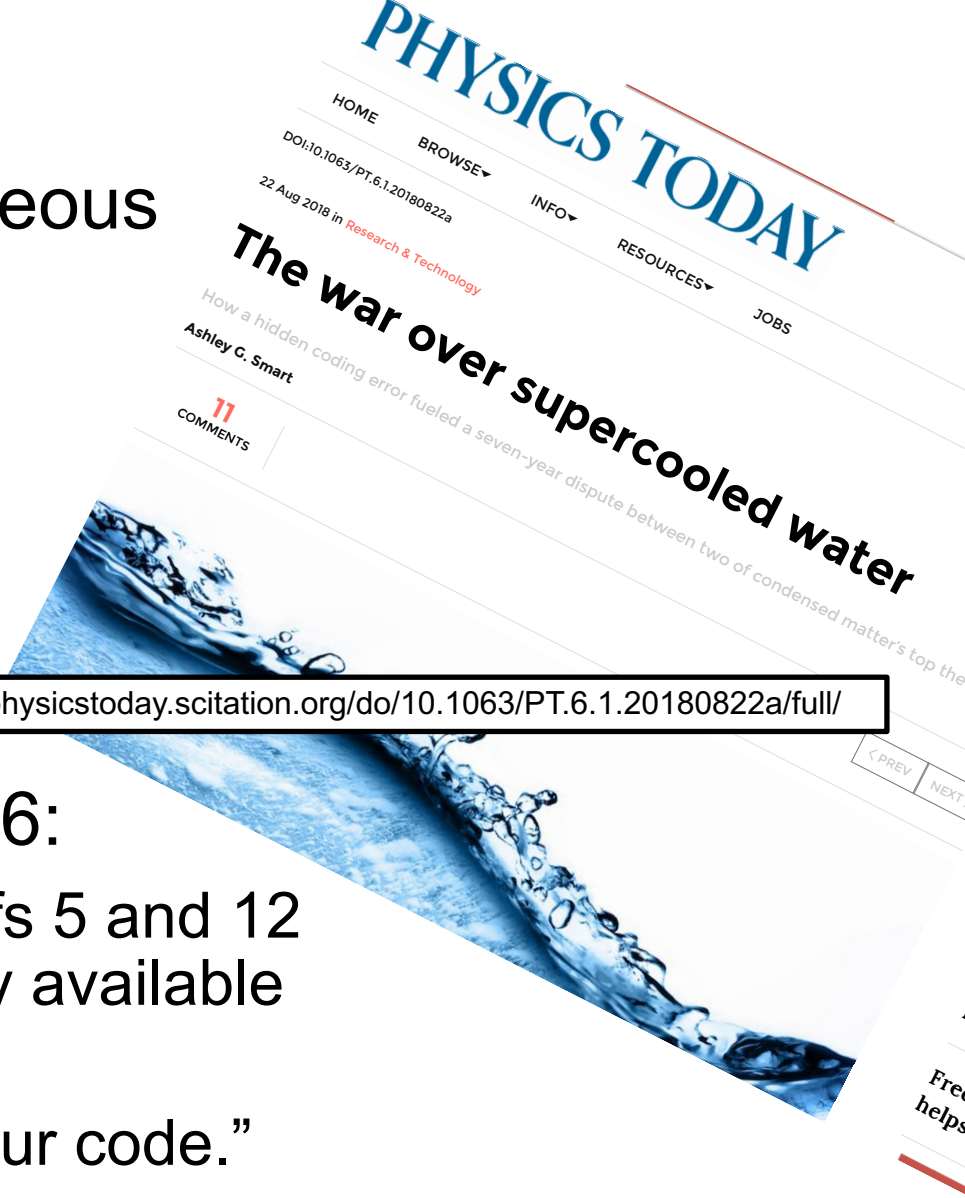
Staff of the the Reproducibility Project at the Center for Open Science in Charlottesville, Va., from left: Mallory Kidwell, Courtney Soderberg, Johanna Cohoon and Brian Nosek. Dr. Nosek and his team led an attempt to replicate the findings of 100 social science studies. Andrew Shurtleff for The New York Times

- NY Times highlights “problems”.
- Only one of many cited examples.
- Computational science ***had*** been spared this “spotlight”.

http://www.nytimes.com/2015/08/28/science/many-social-science-findings-not-as-strong-as-claimed-study-says.html?_r=0

# Computational Science Example

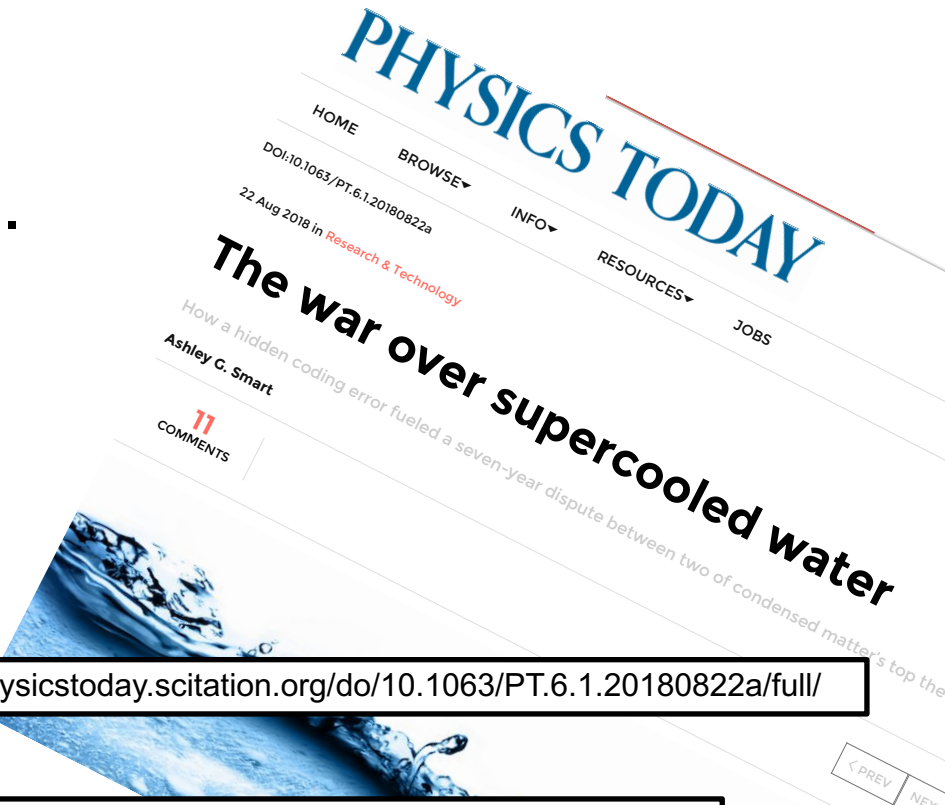
- Behavior of pure water just above homogeneous nucleation temperature (~ - 40 C/F).
- Debenedetti/Princeton (2009):
  - 2 possible phases: High or low density.
- Chandler/Berkeley (2011):
  - Only 1 phase: High density.
- No sharing of details across teams until 2016:
  - Chandler in Nature: “LAMMPS codes used in refs 5 and 12 are standard and documented, with scripts freely available upon request.”
  - Debenedetti with colleague Palmer: ”Send us your code.”
  - Received code after requests and appeal to Nature.




Source: https://physicstoday.scitation.org/do/10.1063/PT.6.1.20180822a/full/

# Computational Science Example

- Palmer located bug/feature in Berkeley code.
- Used to speed up LAMMPS execution.
- Replaced with more standard approach.
- Obtained result similar to Debenedetti 2009.
- Resolution took 7 years.




Source: https://physicstoday.scitation.org/do/10.1063/PT.6.1.20180822a/full/

*For Palmer, the ordeal exemplifies the importance of transparency in scientific research, an issue that has recently drawn heightened attention in the science community. "One of the real travesties," he says, is that "there's no way you could have reproduced [the Berkeley team's] algorithm—the way they had implemented their code—from reading their paper." Presumably, he adds, "if this had been disclosed, this saga might not have gone on for seven years."*

# Better Productivity and Sustainability

Essential for affordable reproducibility

# Tradeoffs: Better, faster, cheaper

- “Better, faster, cheaper: Pick two of the three.”
  - Scenario: (Today)
    You are behind in developing a sophisticated new model in your software that you want to use for results in an upcoming paper.
  - Which of these could be reasonable choices?
    - Develop a simpler model for the paper.
    - Set other work aside and spend more time on development.
    - Ask for an extension on the paper deadline.
    - Develop sophisticated model, but don’t test its correctness.
    - Develop sophisticated model, but don’t document it or check it in.

# Improved developer productivity

“Better, faster, cheaper: Pick all three.” – Near term.

Scenario: (6 months later)
After investing in **developer productivity improvements**, you are on time in developing a sophisticated new model in your software that you want to use for results in an upcoming paper.

Invest in developer tools, processes, practices.

# Improved software sustainability

"Better, faster, cheaper: Pick all three." – Long term.

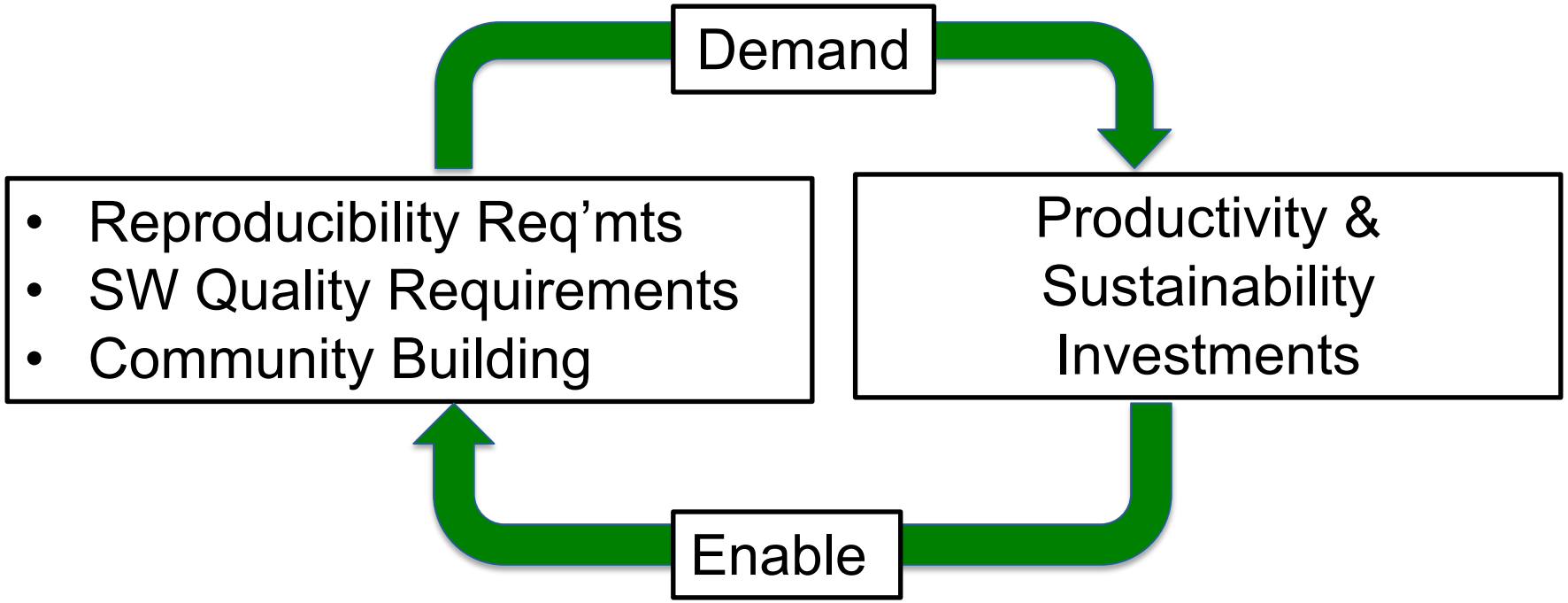
Scenario: (3 years later)
After investing in **software sustainability improvements**, you are on time in developing **several** sophisticated new models in your software that you want to use for results in upcoming papers.

Invest in testing, documentation, integration for long-term software usability.

# Which of These Enhance Reproducibility?

- Code written by first-year, untrained grad student.
- Tuning for high performance.
- Dynamic parallelism of modern processors.
- Better software testing.
- Source code and versioning management.
- Investing in developer productivity.
- Investing in software sustainability.

# Incentives Demand Investments, Enabled by Investments




Common statement: "I would love to do a better job on my software, but I need to:

- Get this paper submitted.
- Complete this project task.
- Do something my employer values more.

Goal: Change incentives to include value of better software, better science.

# Personal Expectations

Calling out the best in team members

# A Few Concrete Recommendations

*Show me the person making the most commits on an undisciplined software project and I will show you the person who is injecting the most technical debt.*

- GitHub stats: Easy to find who made the most commits.
  - Some people: Pride in their high ranking.

- Instead, be the person who ranks high in these ways:
  - Writes up requirements, analysis and design, even if simple.
  - Writes good GitHub issues, tracks their progress to completion.
  - Comments on, tests and accepts pull requests.
  - Provide good wiki, gh-pages content, responses to user issues.

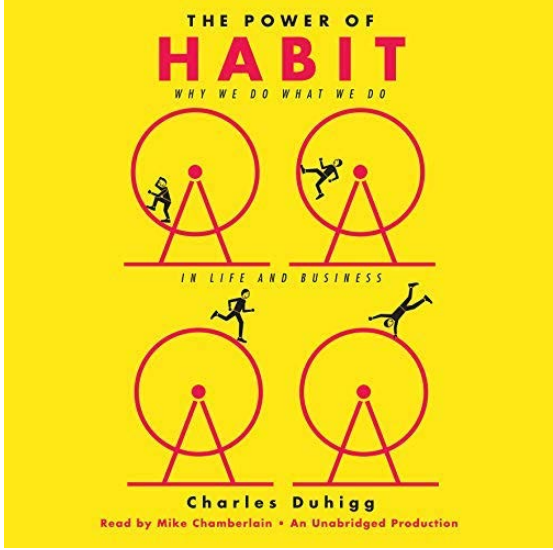
# (Personal) Productivity++ Initiative

Ask: *Is My Work _____ ?*


https://github.com/trilinos/Trilinos/wiki/Productivity---Initiative

# Reproducibility: A keystone habit

# Alcoa and Worker Safety (Duhigg)




- Year: 1987
- Investors concerned about Alcoa.
- Paul O'Neill – Selected Alcoa CEO, not well known.
- First statement: "I want to talk to you about worker safety."
- Investors panicked. But …
- Executed top-to-bottom safety focus.
- 10X injury drop, 5X revenue growth.

> "I knew I had to transform Alcoa. But you can't order people to change. So I decided I was going to start by focusing on one thing. If I could start disrupting the habits around one thing, it would spread throughout the entire company."
>
> - Paul O'Neill

# Reproducibility and Computing

- Aluminum workers:
  - 1500 degree heat, dangerous machines.
  - Safety is key.
- Reproducibility: Key for computing.
- Can we make reproducibility requirements the keystone habit?

# Questions

- What does reproducibility mean for the Solitaire and March Madness projects?
- List several ways to improve productivity on your team.
- List several ways to improve sustainability of our software products.
- What are other potential keystone habits?