

# The Secrets of High Performance: Science Edition

*Dr Nicole Forsgren & Jez Humble*

---

# The Secrets of High Performance: Science Edition

Jez Humble @jezhumble

Nicole Forsgren, PhD @nicolefv

© 2016 DevOps Research and Assessments LLC. [CC-BY-SA](#)

---

---

# Outline

How to make your data suck less

- Writing good survey questions
- Making sure the survey questions are good - with SCIENCE

What we found... that we did (AND didn't) expect

Things about Continuous Delivery

Things about Management

---

---

# Not all data is created equal

Who here thinks surveys are sh\*t?

[Nicole should probably turn around]

---

---

# Not all data is created equal

Who here thinks surveys are sh\*t?

[Nicole should probably turn around]

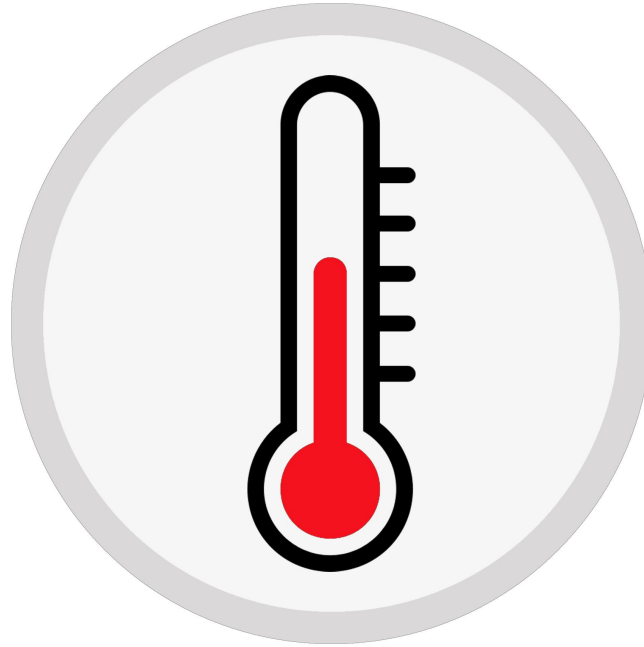
Who here LOVES the data from their log files?

And who has seen sh\*t data in a log file?

---

---

# What is a Latent Construct?



---

We use

# PSYCHOMETRICS

to make our survey data good\*

\*or give us a reasonable assurance that it's telling us what we think it's telling us (& some of this can also apply to your log data)

---

---

# Psychometrics includes:

## Construct creation (manual)

- When possible: use previously validated constructs
- Based on definitions and theory, carefully and precisely worded, card sorting task, pilot tested

## Construct evaluation (statistics)

- Establishing Validity: discriminant and convergent
  - Establishing Reliability
-



---

# Psychometrics Writing Example: Culture

- Does it matter to our study?
    - More than just intuition?
  - What KIND of culture?
    - National identity and norms
    - Adaptive culture
    - Value learning (2014 study)
    - Value information flow and trust (2014 and 2015 studies -- Westrum culture)
-

---

# Psychometrics Writing Example: Culture

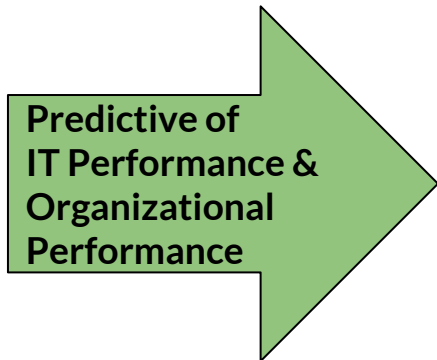
- Does it matter to our study?
    - More than just intuition?
  - What KIND of culture?
    - National identity and norms
    - Adaptive culture
    - Value learning (2014 study)
    - **Value information flow and trust (2014 and 2015 studies -- Westrum culture)**
-

Try writing  
items  
yourself!

# Westrum typology

<b>Pathological</b> <i>Power-oriented</i>	<b>Bureaucratic</b> <i>Rule-oriented</i>	<b>Generative</b> <i>Performance-oriented</i>
Low cooperation	Modest cooperation	High cooperation
Messengers shot	Messengers neglected	Messengers trained
Responsibilities shirked	Narrow responsibilities	Risks are shared
Bridging discouraged	Bridging tolerated	Bridging encouraged
Failure leads to scapegoating	Failure leads to justice	Failure leads to inquiry
Novelty crushed	Novelty leads to problems	Novelty implemented

Use strong  
statements  
with clear  
language.




---

# Westrum Culture Items

- On my team, information is actively sought.
  - On my team, failures are learning opportunities, and messengers of them are not punished.
  - On my team, responsibilities are shared.
  - On my team, cross-functional collaboration is encouraged and rewarded.
  - On my team, failure causes inquiry.
  - On my team, new ideas are welcomed.
-

---



Original in  
2014, but  
there was a  
surprise.  
Can you  
spot it?

---

# Psychometrics Analysis Example: Notification of Failure

At my organization...

- We are primarily notified of failures by reports from customers.
  - We are primarily notified of failures by the NOC.
  - We get failure alerts from logging and monitoring systems.
  - We monitor system health based on threshold warnings (ex. CPU exceeds 100%).
  - We monitor system health based on rate-of-change warnings (ex. CPU usage has increased by 25% over the last 10 minutes).
-

---


# Psychometrics Analysis Example: Notification of Failure

At my organization...

- We are primarily notified of failures by reports from customers.
- We are primarily notified of failures by the NOC.
- We get failure alerts from logging and monitoring systems.
- We monitor system health based on threshold warnings (ex. CPU exceeds 100%).
- We monitor system health based on rate-of-change warnings (ex. CPU usage has increased by 25% over the last 10 minutes).



Notification  
from  
**OUTSIDE**



Notification  
from **INSIDE**

---

# More data tests!

Plus, we test to make sure the survey doesn't have other problems.

- Common method variance (CMV) (aka CMB for Bias)
  - Early vs. late responders
  - Survey drop-off rates and bias
-

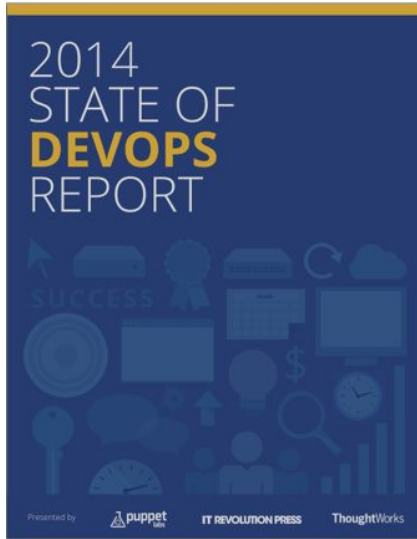
---

---

**Okay NOW we can look at the data**

---





<http://bit.ly/2014-devops-report/>

The DevOps report is a co-production between PuppetLabs and DevOps Research and Assessment LLC. Researchers: Nicole Forsgren, Jez Humble, Gene Kim, Alanna Brown, Nigel Kersten

---

# IT performance matters!

“Firms with high-performing IT organizations were twice as likely to exceed their profitability, market share and productivity goals.”



<http://bit.ly/2015-devops-report/>

---

---

# IT performance

lead time for changes

release frequency

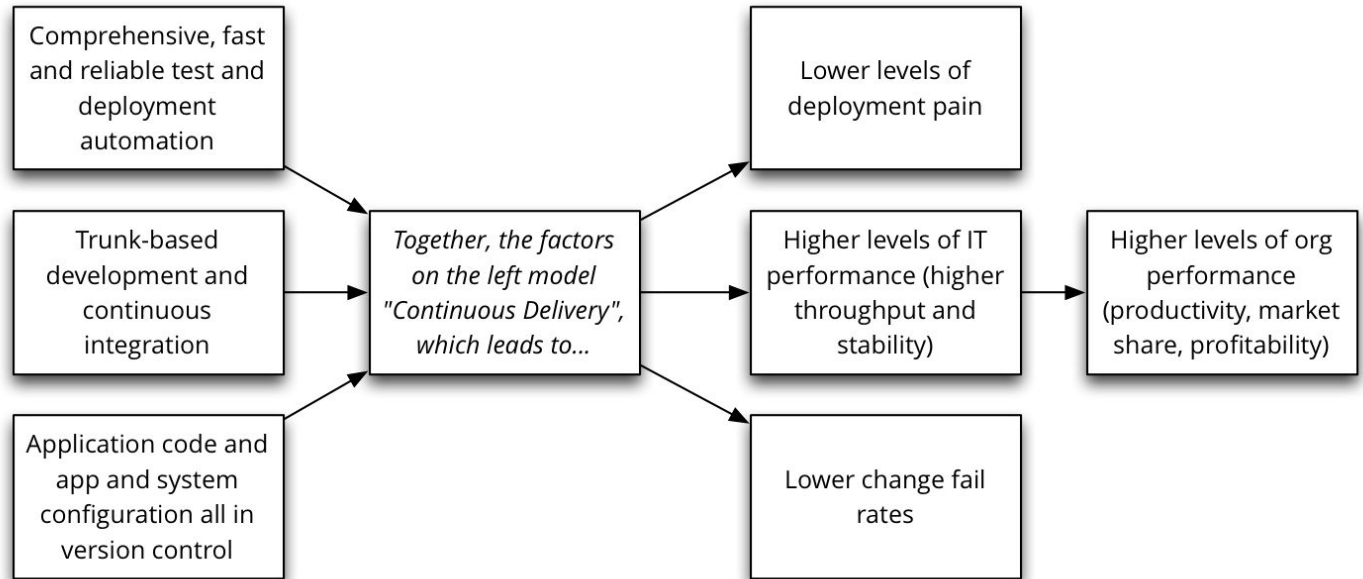
time to restore service

change fail rate

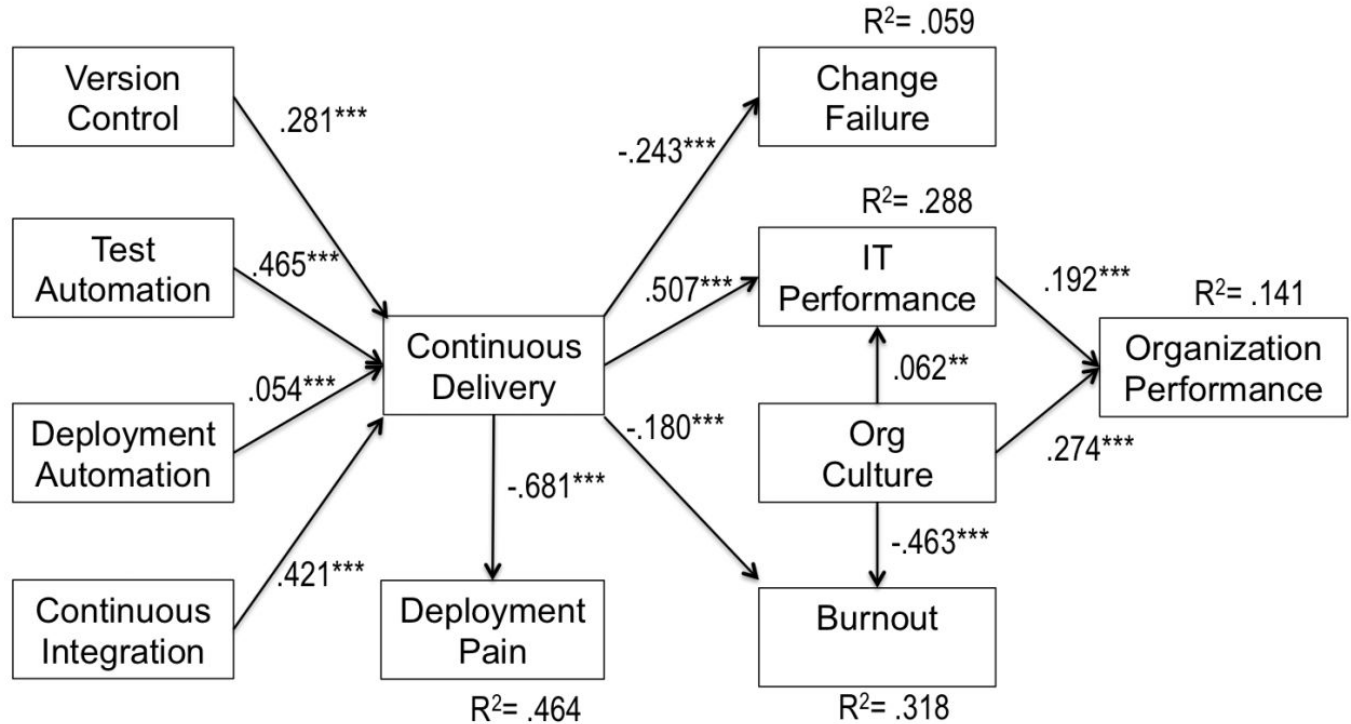
---

---

# Continuous delivery



# And now with added math



Forsgren, N., J. Humble (2016). "The Role of Continuous Delivery in IT and Organizational Performance." In the Proceedings of the Western Decision Sciences Institute (WDSI) 2016, Las Vegas, NV. Available at SSRN: <http://ssrn.com/abstract=2681909>

---

---

**some surprises**

---

---

## What's *not* strongly correlated w/ ITPerf?

Third-party scripts

Homegrown scripts

Commercial configuration management tools

Open source

Golden images

Manual configuration management

---

---

# What's *not* strongly correlated w/ ITPerf?

Third-party scripts

Homegrown scripts

Commercial configuration management tools

Open source

Golden images

Manual configuration management

---

---

## Which of these measure effective test practices?

Developers primarily create & maintain acceptance tests

QA primarily create & maintain acceptance tests

Primarily created & maintained by outsourced party

When automated tests pass, I'm confident the software is releasable

Test failures are likely to indicate a real defect

It's easy for developers to fix acceptance tests

Developers share a common pool of test servers to reproduce failures

Developers create on demand test environments

Developers use their own dev environments to reproduce failures

---



---

## Which of these measure effective test practices?

Developers primarily create & maintain acceptance tests

QA primarily create & maintain acceptance tests

Primarily created & maintained by outsourced party

When automated tests pass, I'm confident the software is releasable

Test failures are likely to indicate a real defect

It's easy for developers to fix acceptance tests

Developers share a common pool of test servers to reproduce failures

Developers create on demand test environments

Developers use their own dev environments to reproduce failures

---

---

# Change management

All production changes must be approved by an external body (e.g. change approval board, manager, etc.) before deployment or implementation (R)

Only high-risk changes, such as database changes, require approval

We have no change approval process

We rely on peer review to manage changes

---

---

# Change management

All production changes must be approved by an external body (e.g. change approval board, manager, etc.) before deployment or implementation (R)

Only high-risk changes, such as database changes, require approval

We have no change approval process

We rely on peer review to manage changes

---

---

---

# Now for management stuff

We all know managing WIP is important, right?

---

---

# Now for management stuff

We all know managing WIP is important, right?

Correlation between WIP and ITPerf is negligible

---

---

# Now for management stuff

We all know managing WIP is important, right?

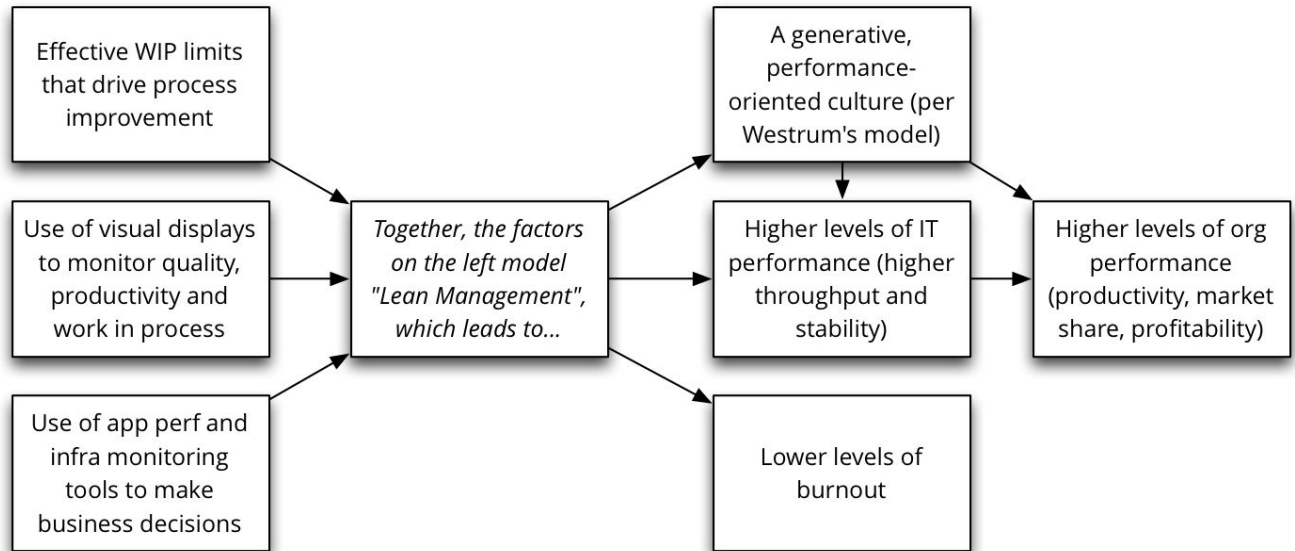
Correlation between WIP and ITPerf is negligible

What's going on?

---

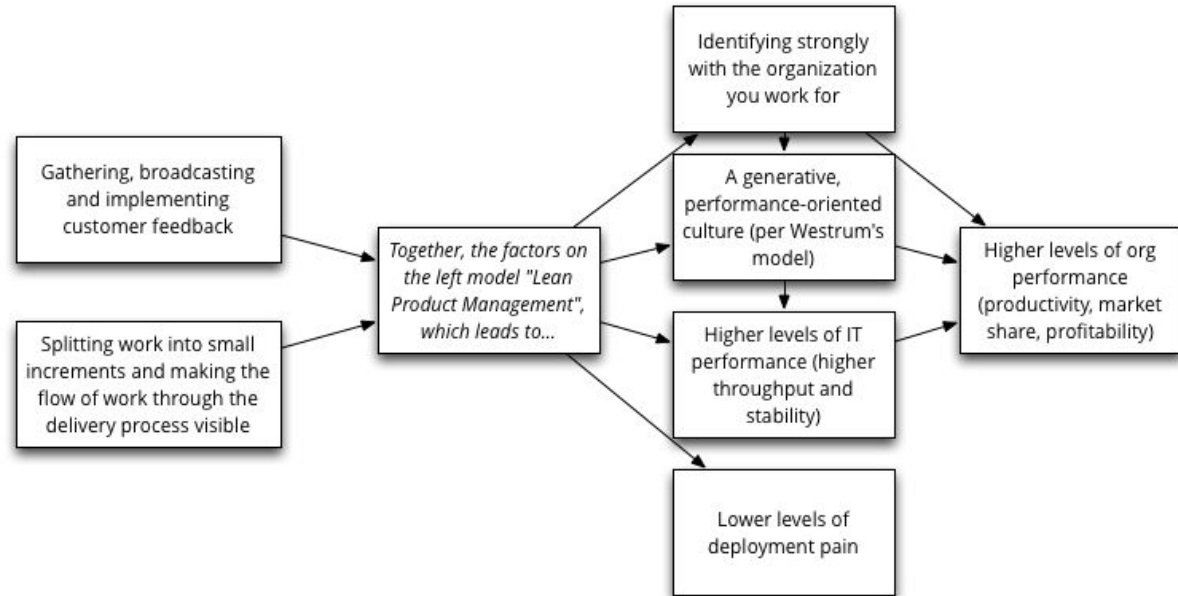
---

# Lean management SEM



---

# Lean Product Management SEM





---

# Conclusions

Even if you think it's obvious, TEST WITH DATA.

(if the results don't surprise you, you're doing it wrong)

(if you don't also confirm some things you expected, you're doing it wrong)

We CAN have it all, or at least throughput AND stability.

DevOps culture & practices have a measurable impact on IT & org perf

---

---

# Stay tuned!

Read the 2016 State of DevOps Survey results late June

Want to know how your OWN DevOps Capabilities compare?

<https://www.devops-research.com/>

just pick up your phone and send an email

to: [jezhumble@sendyourslides.com](mailto:jezhumble@sendyourslides.com)

subject: devops

---

CHICAGO

INTERNATIONAL  
SOFTWARE DEVELOPMENT  
CONFERENCE 2016

goto;  
conference



*Please*

**Remember to  
rate this session**

*Thank you!*