Software Analytics for Digital Games

Thomas Zimmermann, Microsoft Research, USA
Joint work with Nachi Nagappan and many others.

analytics is the use of analysis, data, and systematic reasoning **to make decisions**.

Definition by Thomas H. Davenport, Jeanne G. Harris
Analytics at Work – Smarter Decisions, Better Results

© Microsoft Corporation

Microsoft Research

history of software analytics

EARLY "GLOBAL" MODELS AND SOFTWARE ANALYTICS

As soon as people started programming, it became apparent that programming was an inherently buggy process. As recalled by Maurice Wilkes,[1] speaking of his programming experiences from the early 1950s: "It was on one of my journeys between the EDSAC room and the punching equipment that 'hesitating at the angles of stairs' the realization came over me with full force that a good part of the remainder of my life was going to be spent in finding errors in my own programs."

It took several decades to gather the experience required to quantify the size/defect relationship. In 1971, Fumio Akiyama[2] described the first known "size" law, saying the number of defects D was a function of the number of LOC; specifically, $D = 4.86 + 0.018 * i$. In 1976, Thomas McCabe argued that the number of LOC was less important than the complexity of that code.[3] He argued that code is more likely to be defective when his "cyclomatic complexity" measure was over 10.

Not only is programming an inherently buggy process, it's also inherently difficult. Based on data from 63 projects, Barry Boehm[4] proposed in 1981 an estimator for development effort that was exponential on program size: effort = a * KLOCb * EffortMultipliers, where $2.4 \le a \le 3$ and $1.05 \le b \le 1.2$.

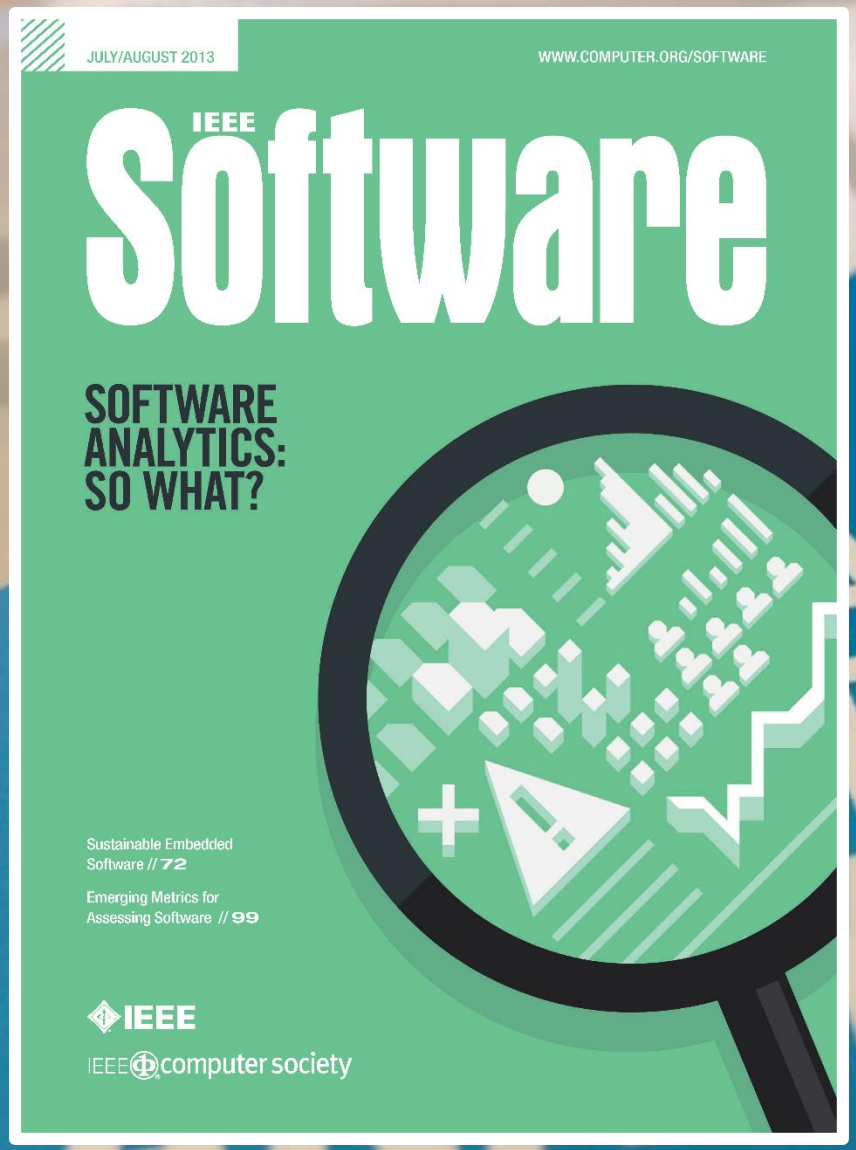
References

1. M. Wilkes, *Memoirs of a Computer Pioneer*, MIT Press, 1985.
2. F. Akiyama, "An Example of Software System Debugging," *Information Processing*, vol. 71, 1971, pp. 353–359.
3. T. McCabe, "A Complexity Measure," *IEEE Trans. Software Eng.*, vol. 2, no. 4, 1976, pp. 308–320.
4. B. Boehm, *Software Engineering Economics*, Prentice-Hall, 1981.

Tim Menzies, Thomas Zimmermann: Software Analytics: So What?
IEEE Software 30(4): 31-37 (2013)

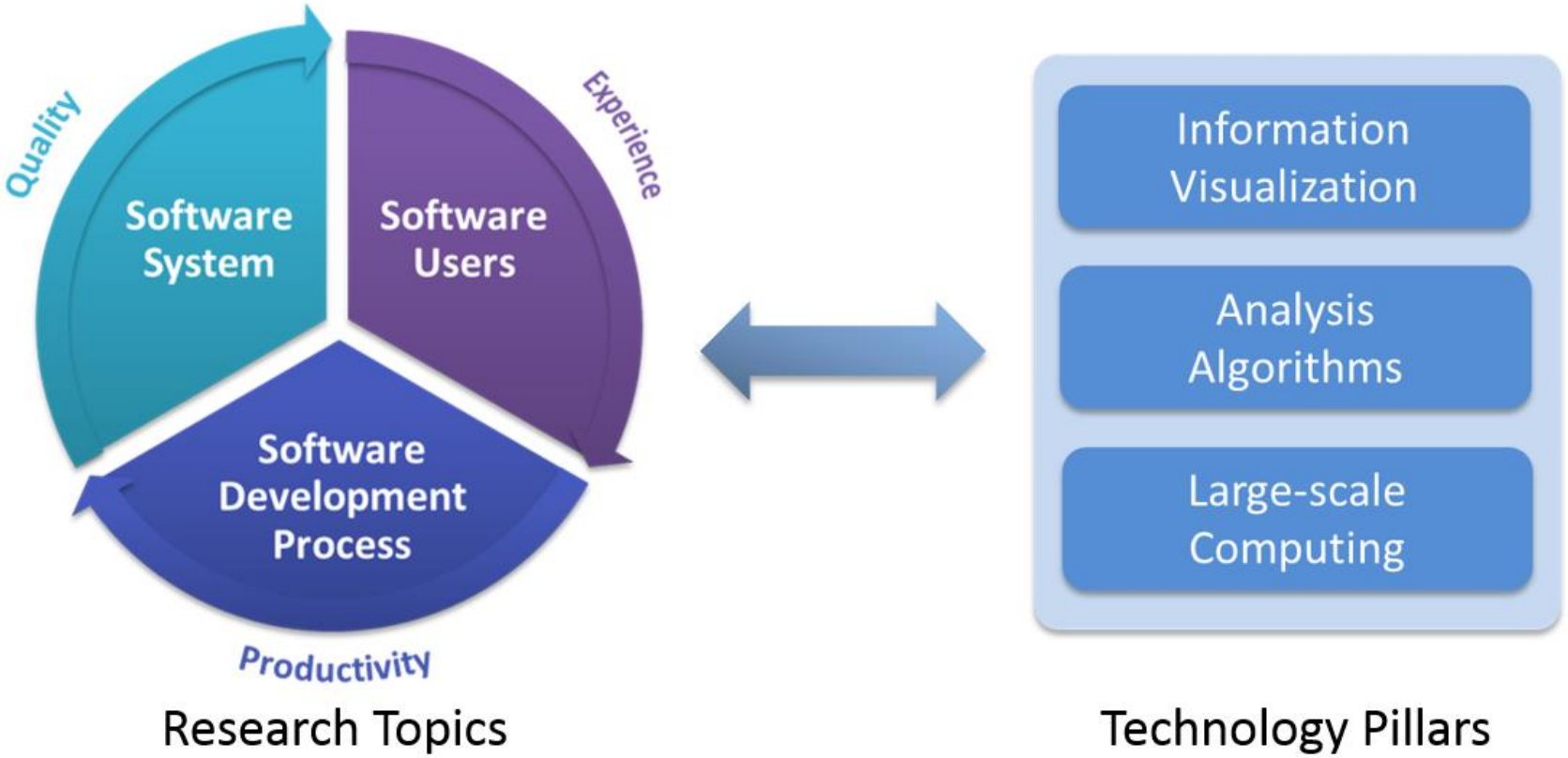
© Microsoft Corporation

Microsoft Research


JULY/AUGUST 2013
WWW.COMPUTER.ORG/SOFTWARE
IEEE
Software
SOFTWARE ANALYTICS: SO WHAT?
Sustainable Embedded Software // 72
Emerging Metrics for Assessing Software // 99
IEEE
IEEE computer society

SEPTEMBER/OCTOBER 2013
WWW.COMPUTER.ORG/SOFTWARE
IEEE
Software
THE MANY FACES OF SOFTWARE ANALYTICS
IEEE
IEEE computer society

trinity of software analytics



Dongmei Zhang, Shi Han, Yingnong Dang, Jian-Guang Lou, Haidong Zhang, Tao Xie: Software Analytics in Practice. IEEE Software 30(5): 30-37, September/October 2013.

MSR Asia Software Analytics group: http://research.microsoft.com/en-us/groups/sa/

© Microsoft Corporation

Microsoft Research

software analytics is

software analytics is

diversity

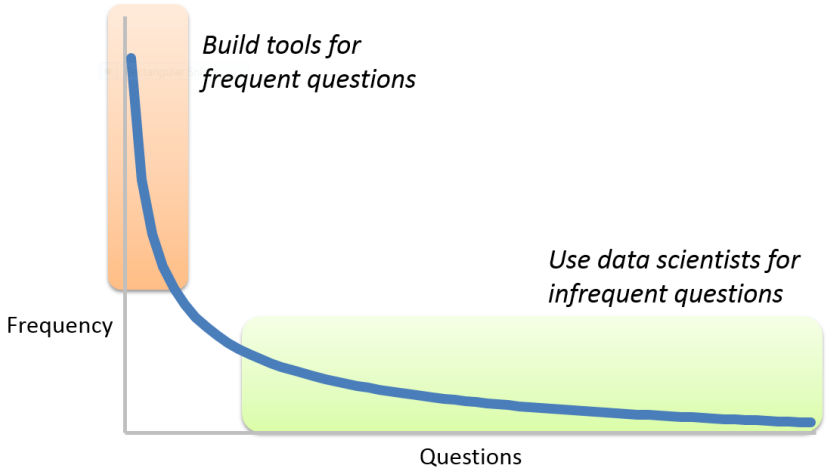
© Microsoft Corporation

Microsoft Research

The Stakeholders

The Tools

The Questions

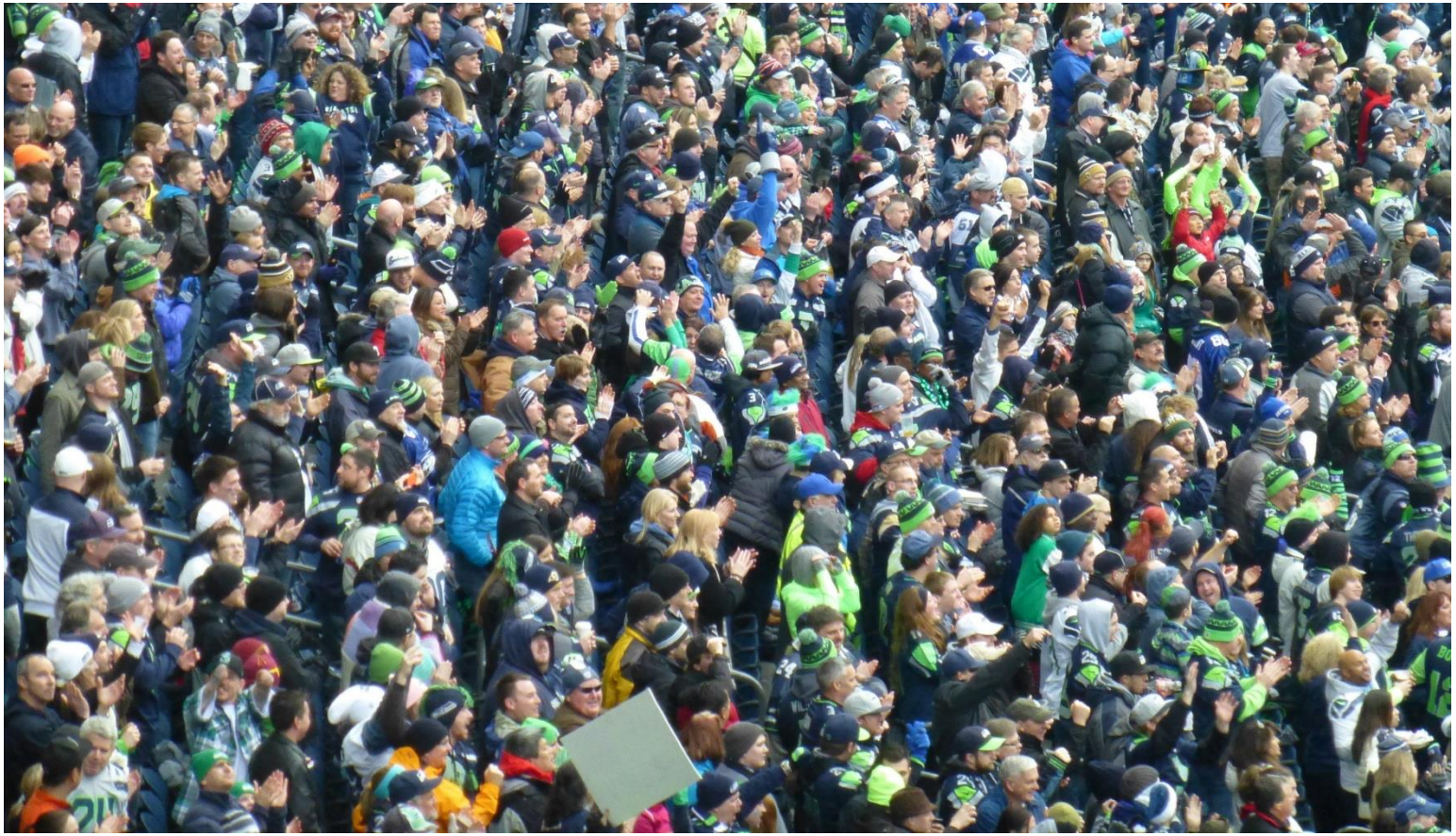


Microsoft Research

© Microsoft Corporation

http://aka.ms/**145**Questions

Andrew Begel, Thomas Zimmermann. Analyze This! 145 Questions for Data Scientists in Software Engineering. ICSE 2014



© Microsoft Corporation

Microsoft Research

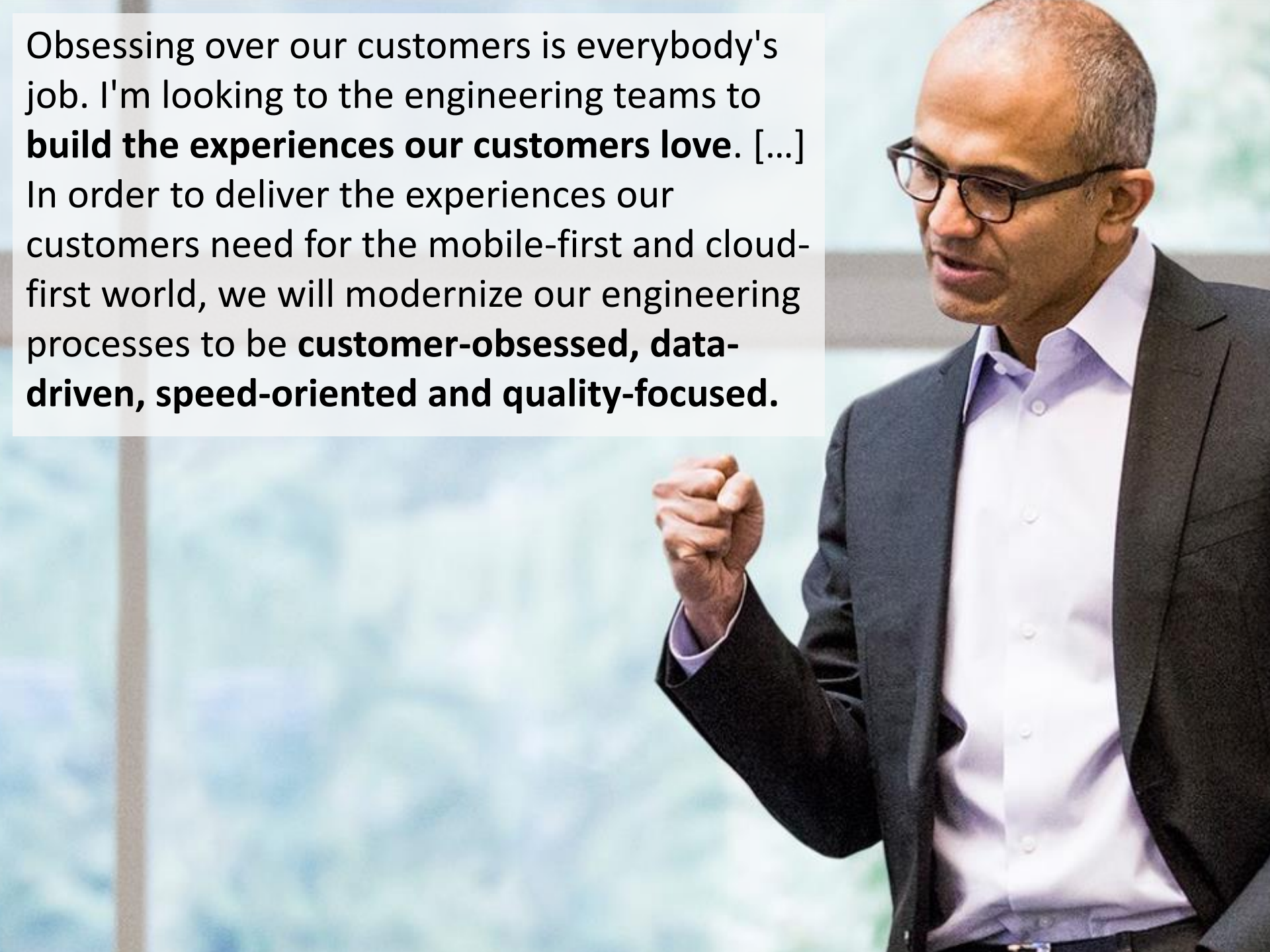
Microsoft's Top 10 Questions

	Essential	Essential + Worthwhile
How do users typically use my application?	80.0%	99.2%
What parts of a software product are most used and/or loved by customers?	72.0%	98.5%
How effective are the quality gates we run at checkin?	62.4%	96.6%
How can we improve collaboration and sharing between teams?	54.5%	96.4%
What are the best key performance indicators (KPIs) for monitoring services?	53.2%	93.6%
What is the impact of a code change or requirements change to the project and its tests?	52.1%	94.0%
What is the impact of tools on productivity?	50.5%	97.2%
How do I avoid reinventing the wheel by sharing and/or searching for code?	50.0%	90.9%
What are the common patterns of execution in my application?	48.7%	96.6%
How well does test coverage correspond to actual code usage by our customers?	48.7%	92.0%

© Microsoft Corporation

Microsoft Research

Obsessing over our customers is everybody's job. I'm looking to the engineering teams to **build the experiences our customers love**. […] In order to deliver the experiences our customers need for the mobile-first and cloud-first world, we will modernize our engineering processes to be **customer-obsessed, data-driven, speed-oriented and quality-focused.**



software analytics is

people

© Microsoft Corporation

Microsoft Research

The Decider

The Brain

The Innovator

The Researcher

Photo of MSA 2010 by Daniel M German (dmg@turingmachine.org)

© Microsoft Corporation

Microsoft Research

software analytics is

sharing

© Microsoft Corporation

Microsoft Research

Sharing Insights

Sharing Models

Sharing Methods

Sharing Data

Sharing Insights

Sharing Methods

Skill in Halo Reach

Jeff Huang, Thomas Zimmermann, Nachiappan Nagappan, Charles Harrison, Bruce C. Phillips: Mastering the art of war: how patterns of gameplay influence skill in Halo. CHI 2013: 695-704

2
E
SE
214
22
25m

How do patterns of play affect players' skill in Halo Reach?

1. General Statistics
2. Play Intensity
3. Skill after Breaks
4. Skill before Breaks
5. Skill and Other Titles
6. Skill Changes and Retention
7. Mastery and Demographics
8. Predicting Skill

The Cohort of Players

We looked at the cohort of players who started in the release week with complete set of gameplay for those players up to 7 months later (over 3 million players)

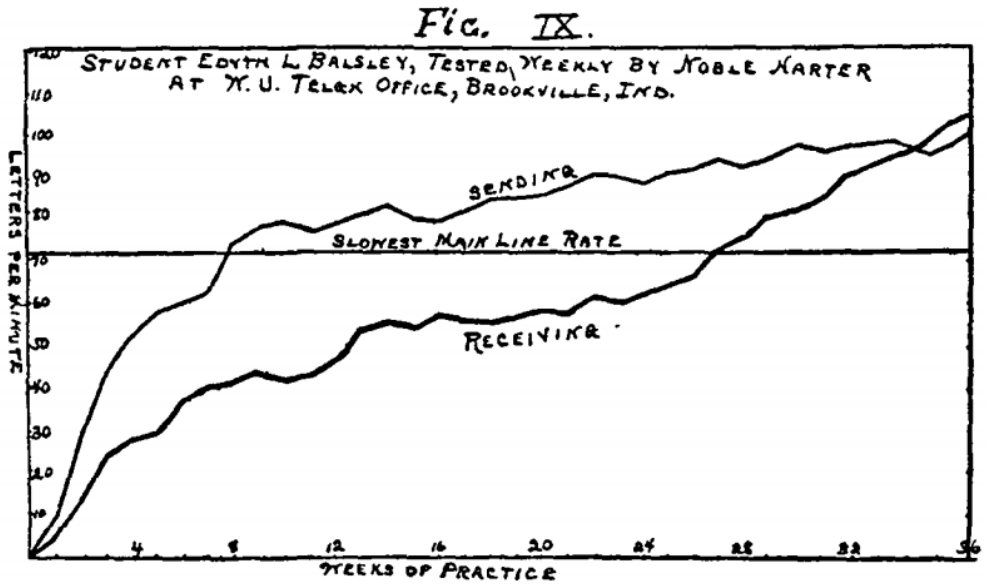
TrueSkill in Team Slayer

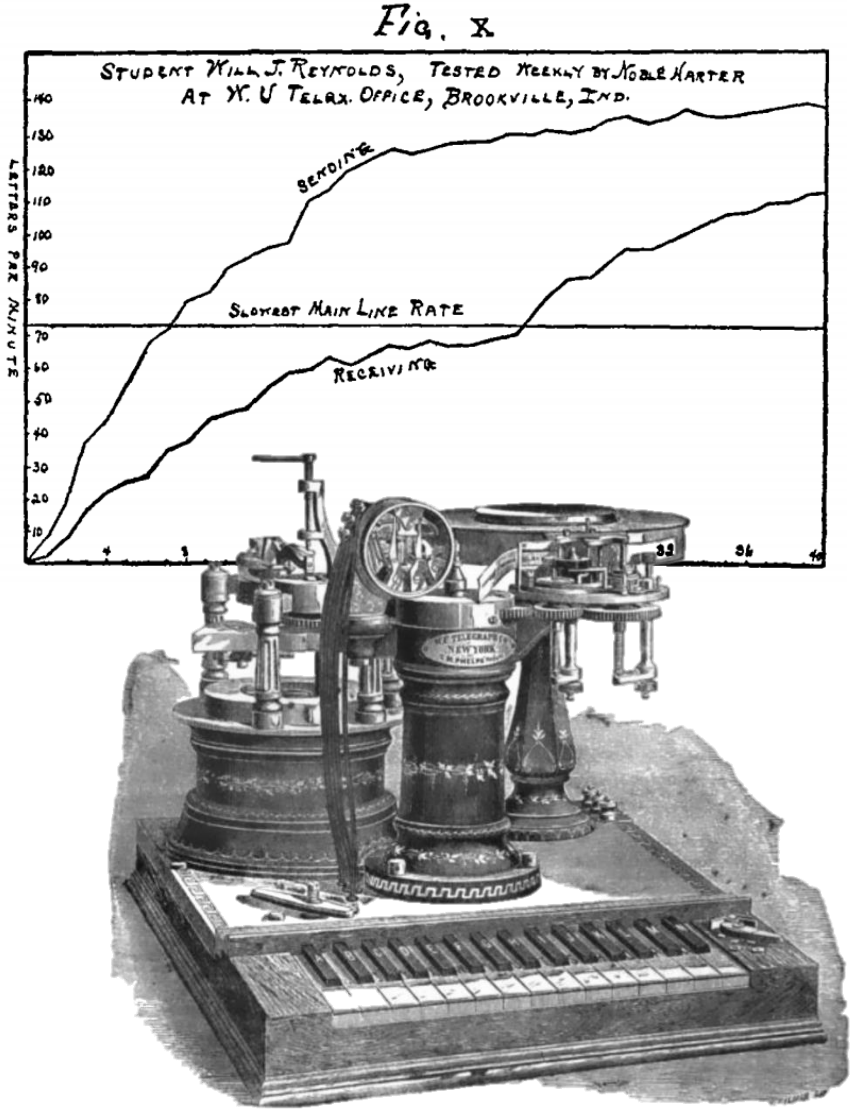
The mean skill value μ for each player after each Team Slayer match
μ ranges between 0 and 10, although 50% fall between 2.5 and 3.5
Initially μ = 3 for each player, stabilizing after a couple dozen matches

70 Person Survey about Player Experience

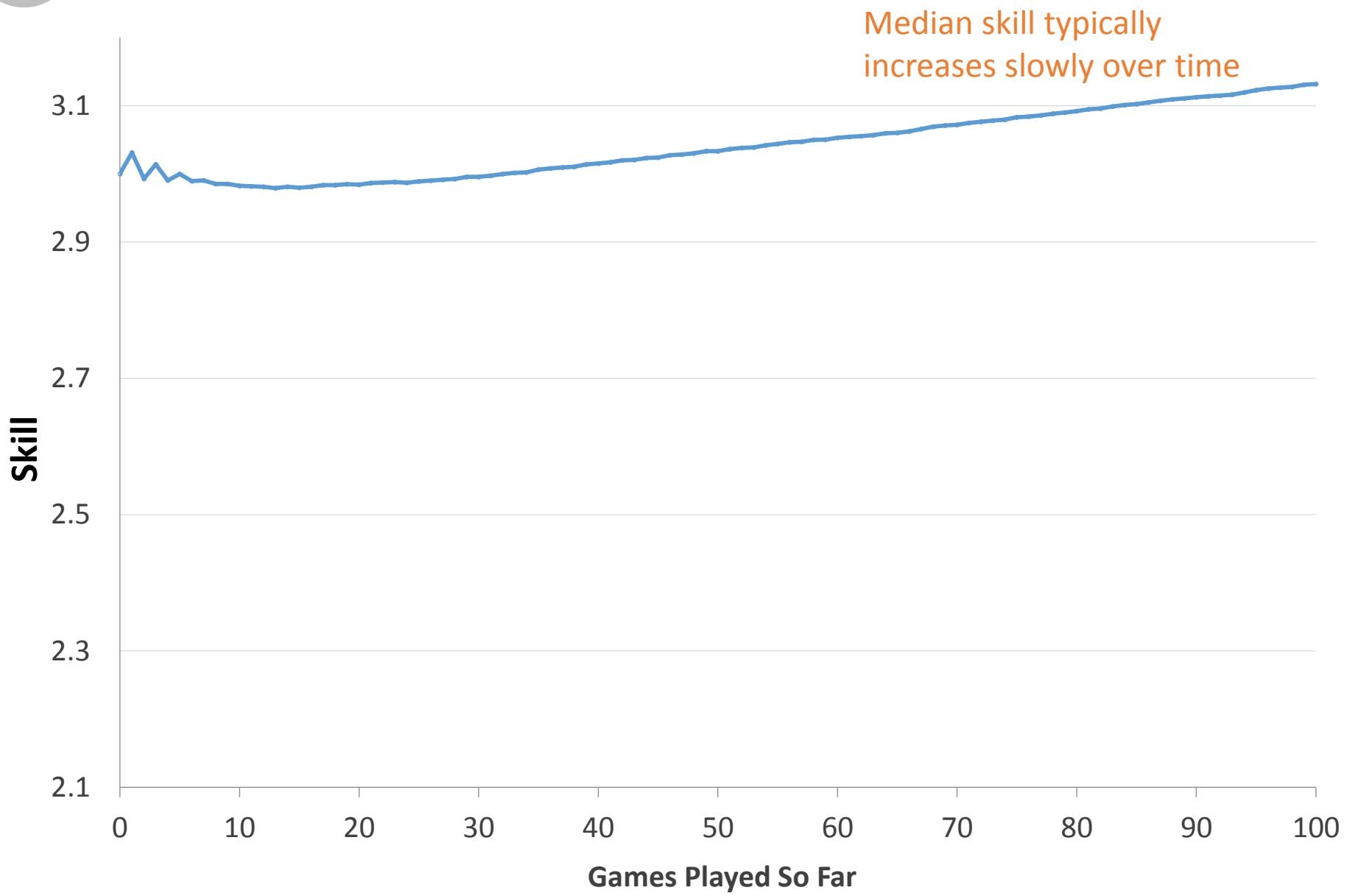
2 Play Intensity

Telegraph operators gradually increase typing speed over time

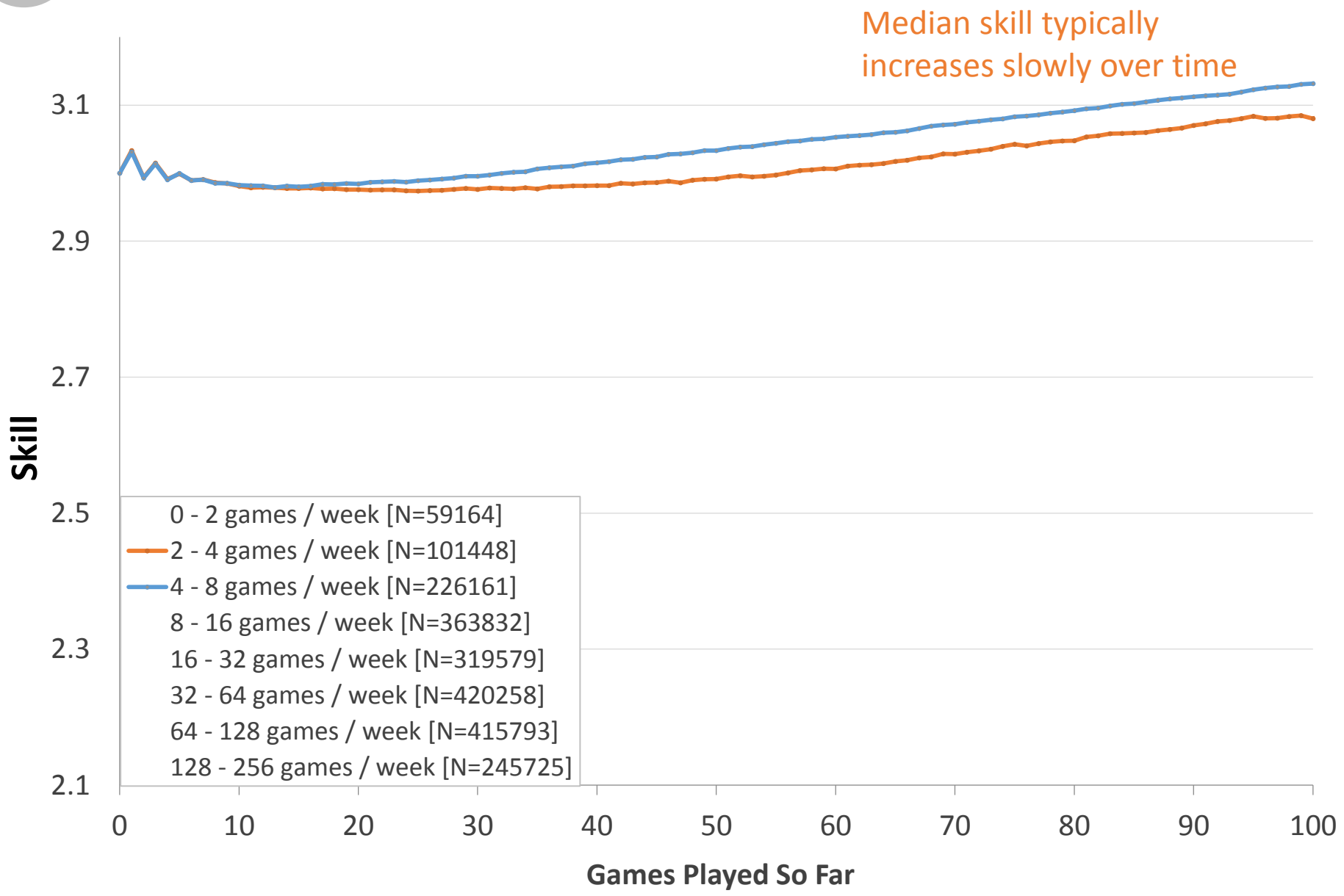




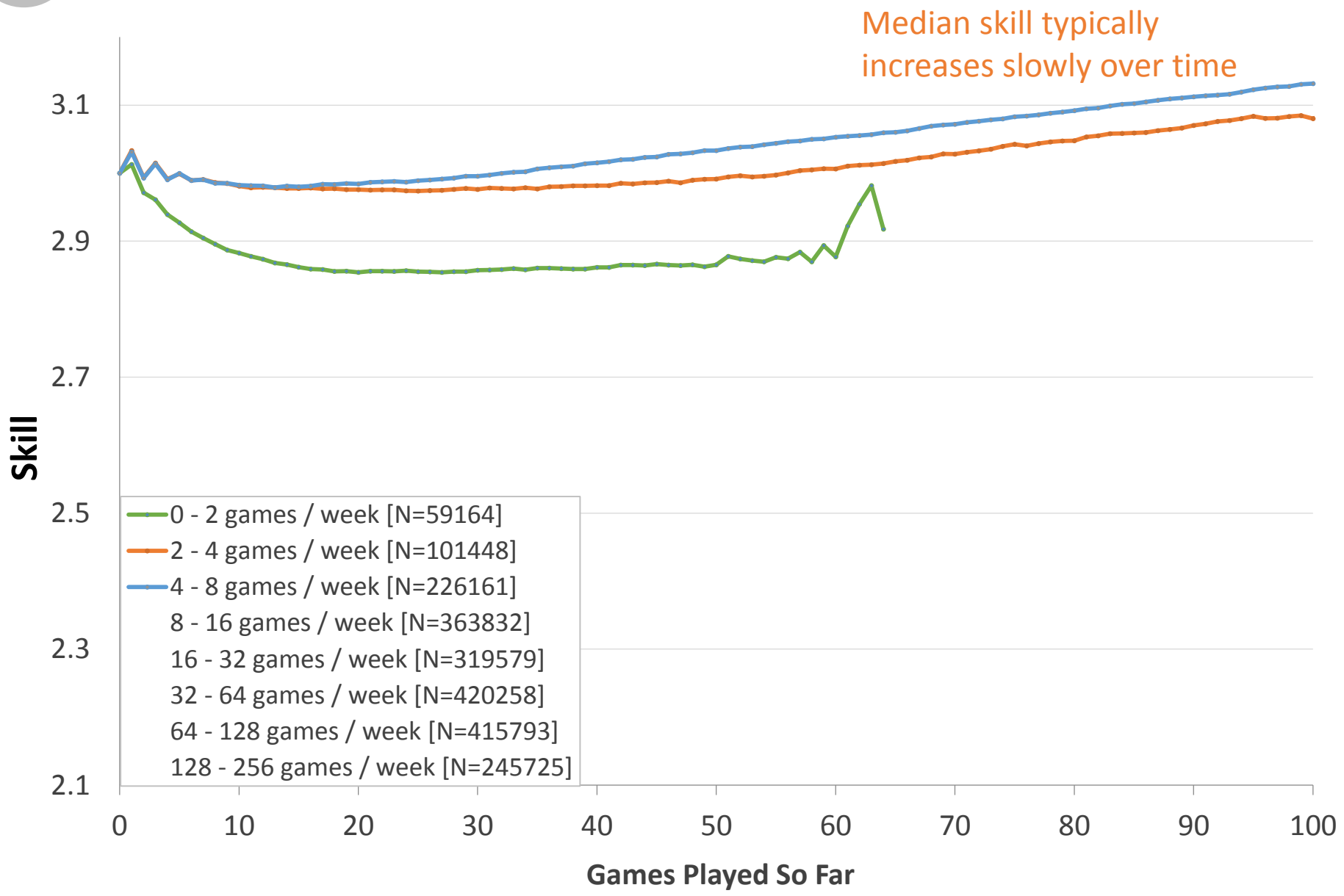
2 Play Intensity



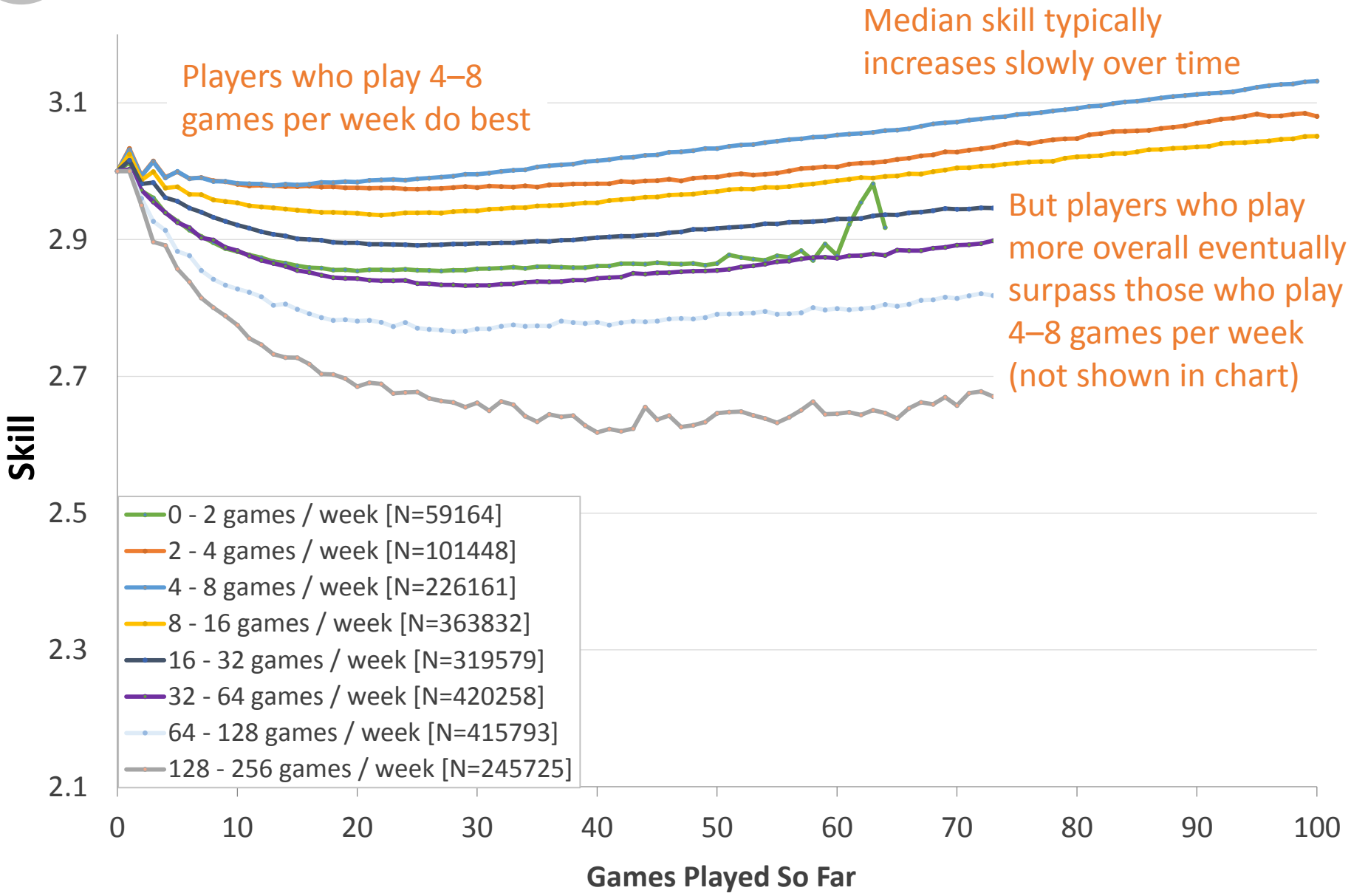
2 Play Intensity (Games per Week)



2 Play Intensity (Games per Week)



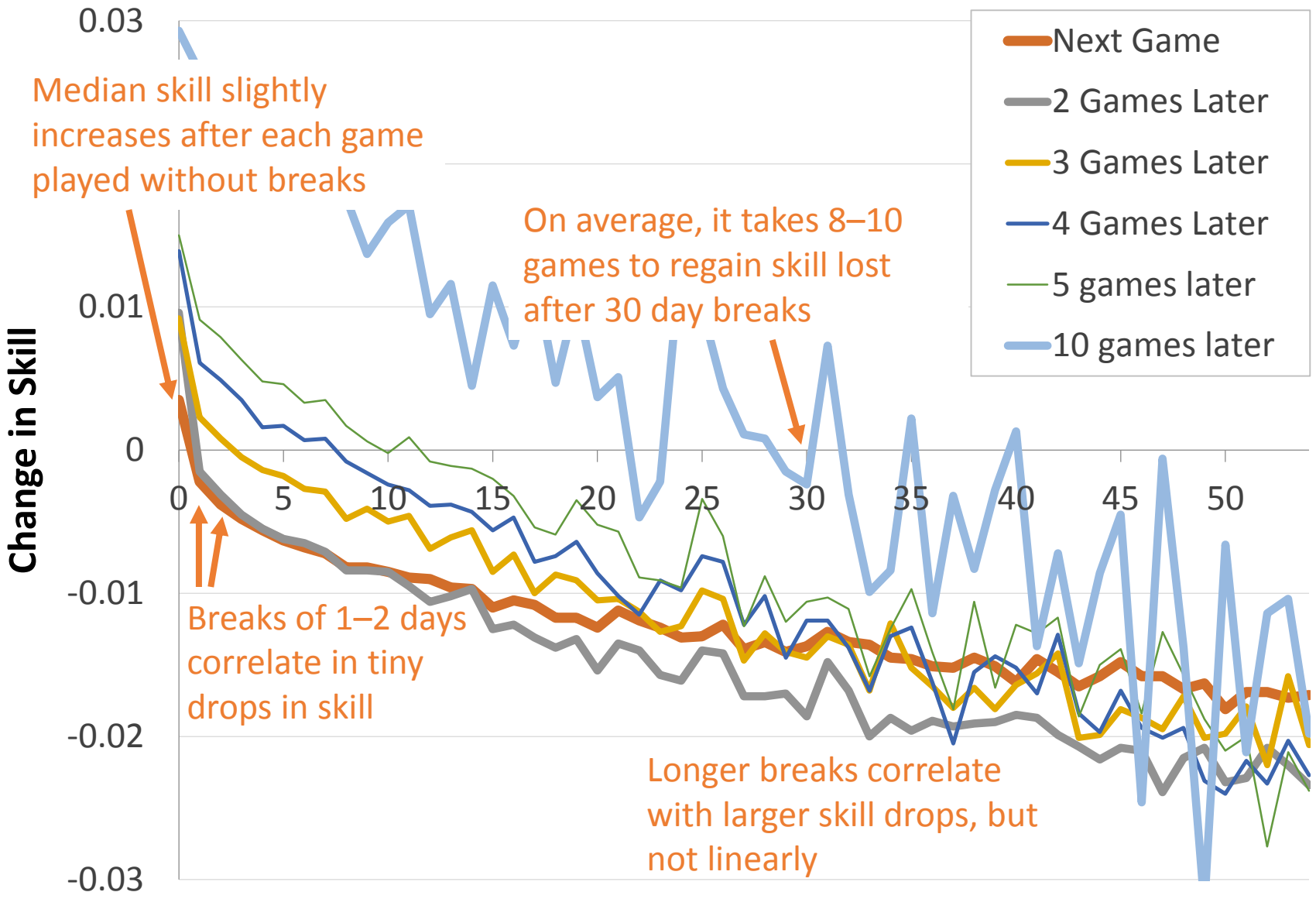
2 Play Intensity (Games per Week)



3 Change in Skill Following a Break

"In the most drastic scenario, you can lose up to 80 percent of your fitness level in as few as two weeks [of taking a break]…"

3 Change in Skill Following a Break



Analysis of Skill Data

Step 1: Select a population of players.
For our Halo study, we selected a cohort of 3.2 million Halo Reach players on Xbox Live who started playing the game in its first week of release.

Step 2: If necessary, sample the population of players and ensure that the sample is representative.
In our study we used the complete population of players in this cohort, and our dataset had every match played by that population.

Step 3: Divide the population into groups and plot the development of the dependent variable over time.
For example, when plotting the players' skill in the charts, we took the median skill at every point along the x-axis for each group in order to reduce the bias that would otherwise occur when using the mean.

Step 4: Convert the time series into a symbolic representation to correlate with other factors, for example retention.

Repeat steps 1–4 as needed for any other dependent variables of interest.

EASY mode: 1-2 buttons

EXPERT mode: 5-6 buttons

An Empirical Study of Driving Skill in Forza Motorsports 4. (FDG 2014)
Thomas Debeauvais, Thomas Zimmermann, Nachiappan Nagappan,
Kevin Carter, Ryan Cooper, Dan Greenawalt, Tyson Solberg

Cowboys,
Ankle Sprains,
and Keepers of
Quality #icse14

How Is Video Game Development Different from Software Development?

Game developers have **less clear requirements** than non-game developers.

Game developers tend to use what they perceive as an **Agile process** more than non-game developers.

Creativity is valued more in game development teams.

The ability to **communicate with non-engineers** is valued more on game development teams.

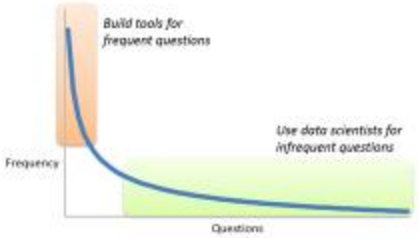
Game development requires a more **diverse team.**

People are more impressed by game developers' work.

The Stakeholders

The Tools

The Questions



© Microsoft Corporation

Microsoft Research

The Decider

The Brain

The Innovator

The Researcher

Photo of MSA 2010 by Daniel M German (dmg@turingmachine.org)

© Microsoft Corporation

Microsoft Research

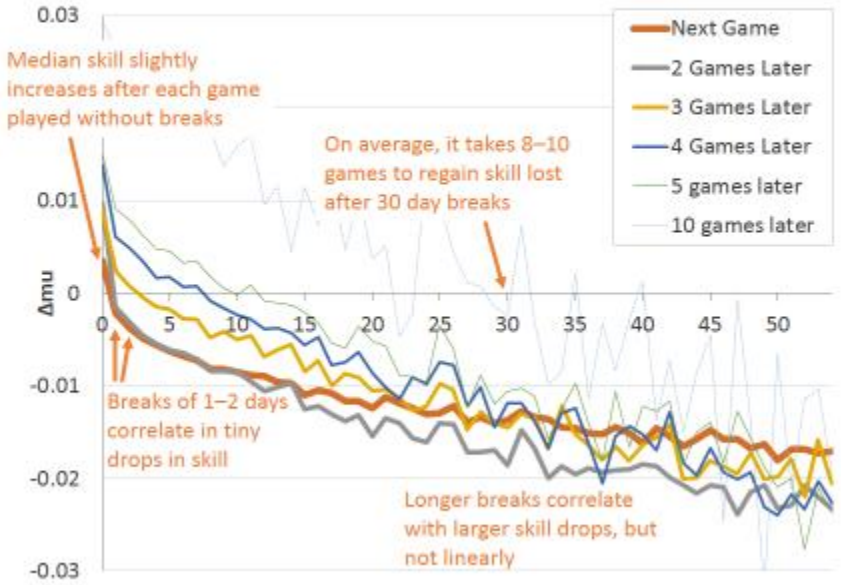
Sharing Insights

Sharing Models

Sharing Methods

Sharing Data

3 Change in Skill Following a Break



Software Analytics for Digital Games

Thomas Zimmermann, Microsoft Research, USA
Joint work with Nachi Nagappan and many others.

Thank you!