

Agile Benchmarking in 2020 - Where Are We Today?

(and why you should care)

September 4, 2020

Joe Schofield, joescho@joejr.com, <https://joejr.com>

At ISBSG's first IT Confidence Conference 2013, I presented "Using Benchmarks to Accelerate Process Improvement." This year's presentation will provide a high-level "refresh" of how organizations can benefit from benchmarking. A simple list of recommended Do's and Don'ts will be offered.

The primary focus of this presentation however, will be around experience-based and research-based agile benchmarking data. The experience-basis will incorporate client reactions and limitations to any agile data. The research-basis will trend industry-leading benchmarks including the use of agile methods, practices, and benefits since 2006 with an obvious focus on the most recent data. Insights and cautions will be shared for the attendee's consideration. Finally, recommendations on agile-related measures and metrics will be provided.

Topical Overview

Agile Trends –
Usage

Adoption Barriers
& Challenges
Trends

Noted
Improvement
Trends

Agile Practices
Trends

Success Measure

Team Value
Delivery, not
Teams'
Productivity

Benchmarking
against the best
may be the wrong
strategy

When productivity
measures are
unproductive

Benchmarking
Do's and Don'ts
for Agile
Organizations

ISBSG Agile
Guidance: Myths
& Truths

ISBSG Agile
Guidance: Myths,
Truths, and now
Reality!

Benchmarking
Suggestions for
Agile
Organizations

Fundamental
Questions around
Benchmarking

Agile
Benchmarking
Tools and Practice

Which device
'better'?

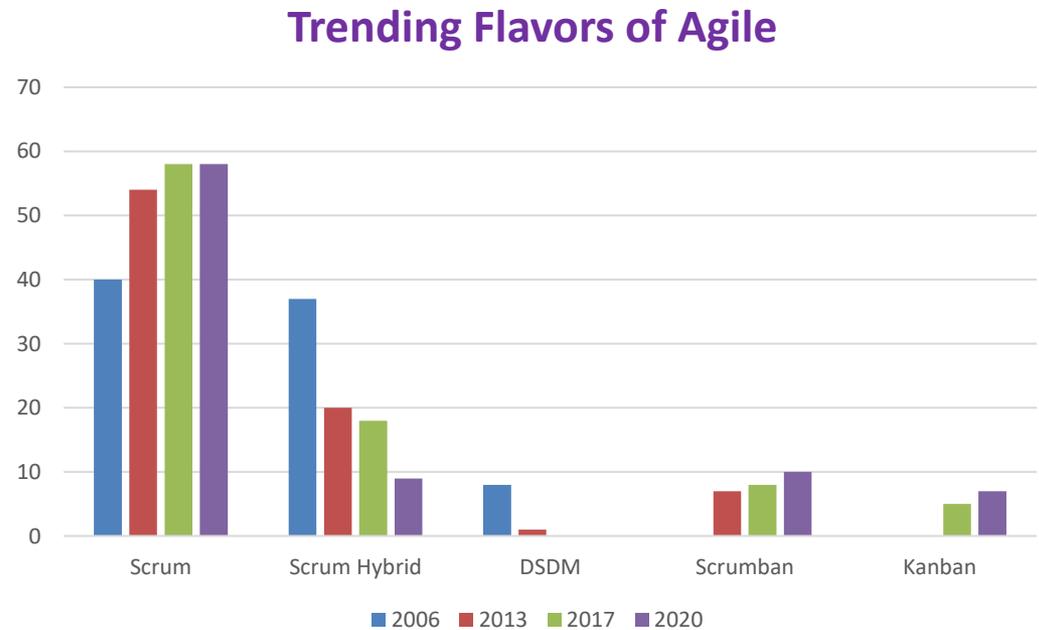
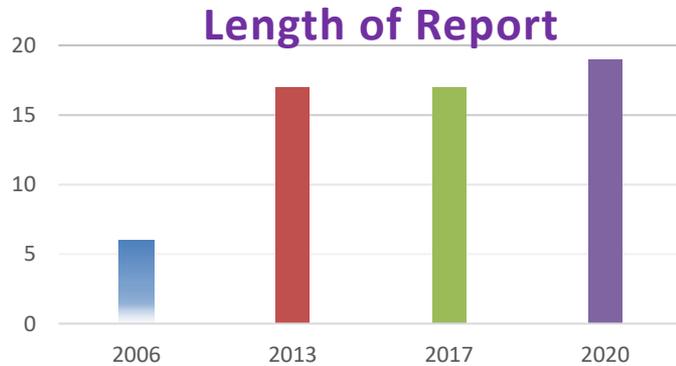
What does this
industry and ISBSG
have in common?

What do these
games have in
common?

Thank you!

Presenter Notes

Agile Trends – Usage



Adoption Barriers and Challenges Trends

2006	
%	Barriers to Success
21	Right people
20	Resistance to change
15	Customer collaboration
14	Management Support

2017	
%	Challenges to Success
47	Inexperience with agile
45	Lack of management support
43	Resistance to change
41	Lack of business support / PO

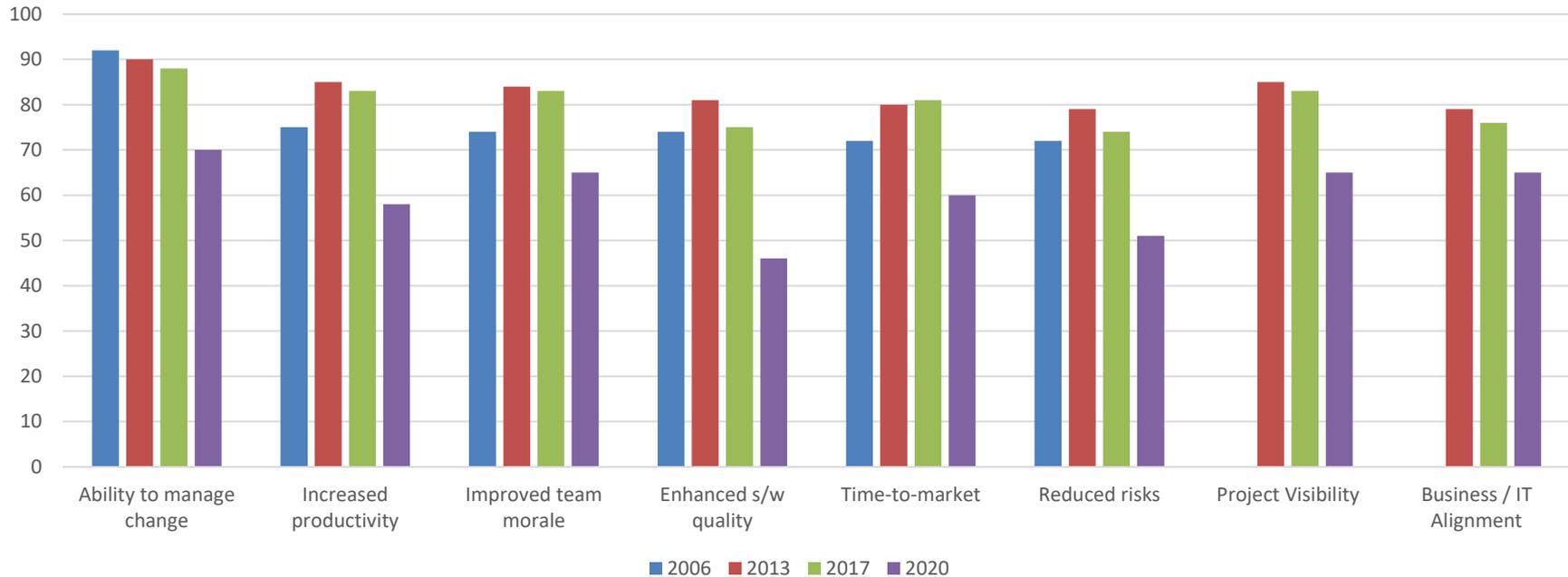
2013	
%	Causes of Failed Projects
12	Culture conflict
11	Pressure to use waterfall
11	Organization / communication
9	Inexperience

2020	
%	Challenges to Success
48	Resistance to change
46	Limited leadership participation
45	Inconsistent team practices
44	At odds with culture
43	Management support
43	Lack of skills
41	Lack of training

Annual State of Agile™ Report; VERSIONONE, digital.ai

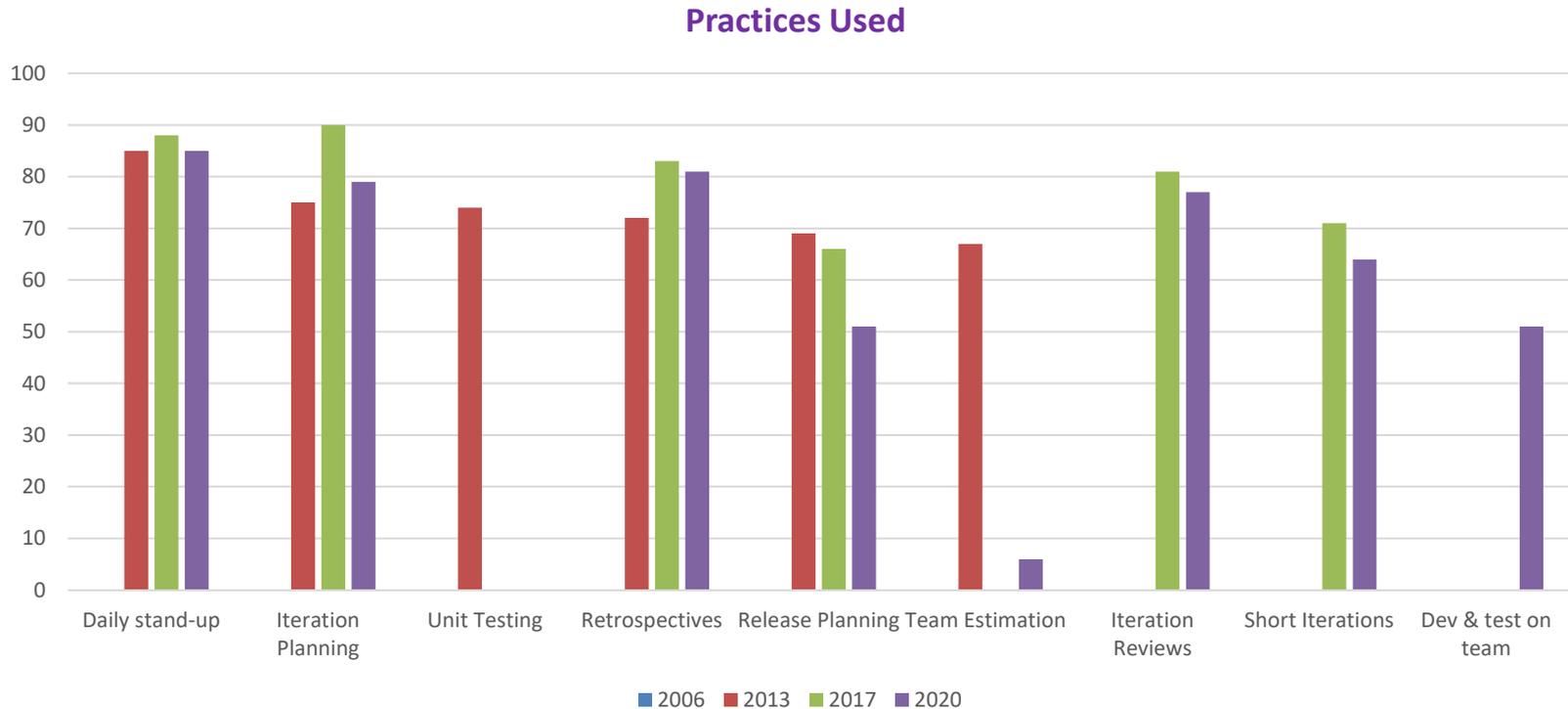
Noted Agile Improvement Trends

Observed Improvements (Percent)



*2013 these were called “top benefits”

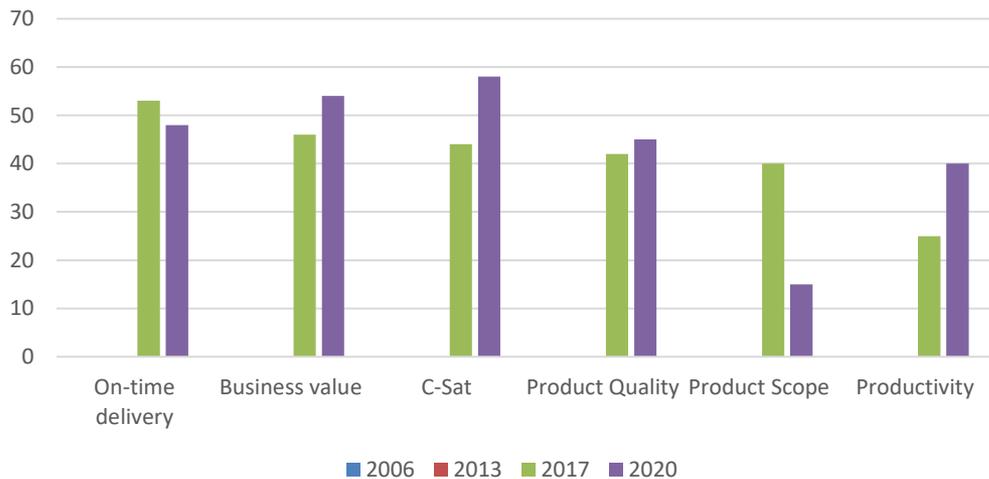
Agile Practices Trends



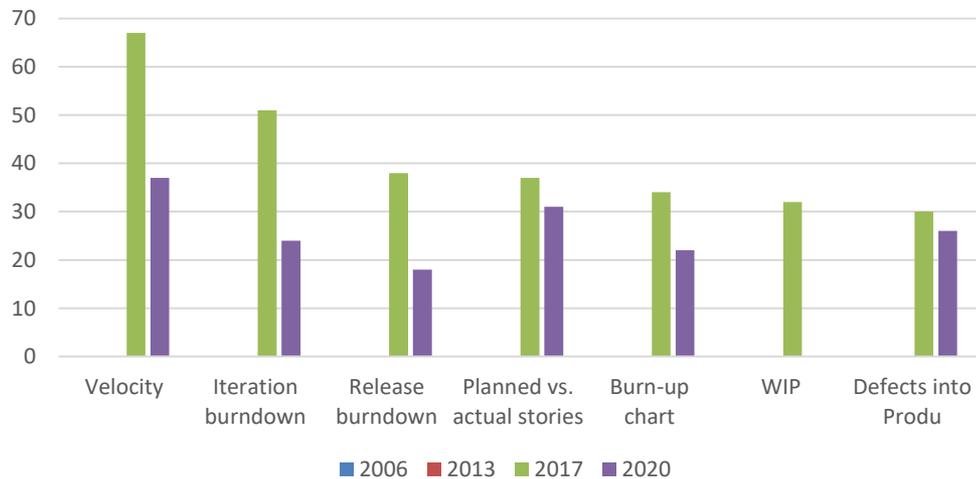
Success Measures

1,121 global respondents in 2020

Measuring Success of Agile Initiatives



Measuring Success of Agile Projects



*86% not CollabNet or VersionOne clients

It's About the Team's *Delivery*, not the Teams' *Productivity*

WHO precedes WHAT



Benchmark peer factors & considerations:

- ✓ Extent of agile use (initial or widespread)
- ✓ Which agile (Scrum, hybrid, Kanban)
- ✓ Agile maturity (getting started vs. well-established)
- ✓ DevOps teams
- ✓ Product teams (or project teams)
- ✓ Product or project funding stream
- ✓ CI / CT / CD maturity
- ✓ Scaling approach

Agile metrics:

- ✓ Value delivery
- ✓ Time to market
- ✓ T-Sat (team)
- ✓ C-Sat
- ✓ Release defects

Why Benchmarking Against the Best may be the Wrong Strategy

However, if you're considering implementing the 'SpSpSp model', please think again.

- ✓ Is *your* organization building a xxxxx player?
- ✓ Is *your* organization still trying figure out its business model?
- ✓ Is *your* organization facing hyper-growth?
- ✓ Is “move fast and break things” applicable to *your* product?

Maybe. But, probably not. When people copy the 'SpSpSp model', it often happens through a top-down directive, without taking a close look at what kind of culture and leadership is needed to make it work. Often the existing hierarchy is changed into a new, static matrix blueprint (even if the matrix is labeled with Squads, Chapters, and Tribes), instead of growing a culture of continuous participatory change.

'Don't copy the model'

“Stop trying to borrow wisdom and think for yourself. Face your difficulties and think and think and think and solve your problems yourself. Suffering and difficulties provide opportunities to become better. Success is never giving up.” — Taiichi Ohno

With all of the other cultural influences, *productivity measurement may be unproductive* (still more factors related to team and benchmarks)

Just because historic metrics are not useful in an agile environment, doesn't mean that agile measures are wrong. Nor does it mean that we need to establish standards to feed measurement systems intended for a different type of work flow.

“Those of us with any God-given sense need to resist all attempts by the organization or ourselves, to compare velocity among teams. Don't get inveigled in that thinking. First, the relative nature of estimated values renders velocity as incomparable among teams. Where your team started with an eight as a midpoint, another team could have started with a five or even a thirteen. To make matters worse, much worse, teams don't all have the **same number of members**. Teams aren't all in the **same place in their development**. Teams don't all have the same **composition of talents**. Teams have varying **rates of turnover or churnover**. Teams have more or less **experienced** scrum masters, product owners, and stakeholders. Some teams aren't **colocated**. Some teams invest more heavily in cross-functionality which typically enables their future. Do not fall victim to this mentality.” *from the forthcoming Agile novel **Aligning People and Culture for Agile Transformation**; jrs; 2020 Amazon's KDP Select (released 9/3/2020)*

Harvard research shows the ideal number of team members is 4.6

<https://www.teamgantt.com/blog/what-is-the-ideal-team-size-to-maximize-productivity>

A Leadership & Change Story

ALIGNING PEOPLE
AND CULTURE FOR
AGILE
TRANSFORMATION

JOE SCHOFIELD

Benchmarking Do and Don't Considerations for Agile Organizations

Conventional Wisdom

Use an initial *internal* benchmark to baseline practices



Agile Organizations

Capture sprint-by-iteration values for story points, capacity (available hours), number of team members, percent of commitment (100 → 1) by team members (*planned* during sprint planning, *actual* during retrospective)

Use subsequent *internal* benchmarks to trend improvement (or degradation)



Included in iteration-by-sprint data capture

Use team benchmark data to identify better / best practices among teams



Encourage teams to informally exchange improvement ideas; part of Scrum Master coaching role

Benchmark with industry **peers**



Benchmark with organizations with similar desired outcomes

Benchmark using relevant data



Benchmark with data that is three years of age or newer, only—too much change in tools, technologies, and platforms to find older data valuable. **Example:** would you track your Covid infection rate from data that is three years old? Two? One?

The more I study benchmarking, the less confident I am that I have anything to offer . . . JRS: 8/4/2020

Benchmarking Suggestions for Agile Organizations

Do

- ✓ Make benchmark objectives visible (certification intent as with the CMMI vs. internal improvement)
- ✓ Include product and process measures
- ✓ Define precisely every measure & metric
- ✓ Consider unintended consequences of collected measurements & definitions—remember that measurements influence behavior
- ✓ Apply simple statistical techniques like Capture-Recapture techniques to **predict** remaining defects in products

[Beyond Defect Removal: Latent Defect Estimation with Capture Recapture Method](#); CrossTalk, August, 2007

Doing Better

- ❑ Monitor ongoing benchmark data collection
- ❑ Use lean checklists to capture and improve repetitive activities and cross-functional growth
- ❑ Track and trend—statistically if capable—defects by release

REMEMBER:

- *Scope* is determined during sprint planning as the sprint commitment
- *Schedule* is time-boxed
- *Cost* is agreed to before the sprint and limited (for the most part) to the sprint

Agile Benchmarking Tools and Practices

For tools

- ✓ Consider Gartner's Magic for Enterprise Agile Planning Tools; Published April 21, 2020
- ✓ Planview and Atlassian are the most highly rated products
- ✓ Three other products share the Magic Quadrant for *Completeness of Vision* and *Ability to Execute*



For Scrum practices – the [Scrumbutt test](#) (aka, the Nokia test)

- ✓ From Scruminc
- ✓ 10 assessments (questions) for each team member to complete
- ✓ Perfect score is 100 per participant
- ✓ Great coverage of basics of scrum
- ✓ Good luck – you'll need it!

Which device is better?

Device A

Device B

Is known as a "smart" device

Purchase cost is $\frac{1}{2}$ of Device B

Operational cost is $\frac{1}{2}$ of Device B

Is known as a "stealth" device

Purchase cost is 2X Device A

Operational cost is 2X Device A

Is about 171 mph faster

Can carry ~30,000 lbs. of "stuff"

Is ~25 years old

Is about 171 mph slower

Can carry ~70,000 lbs. of "stuff"

Is ~5 years old

Saab Gripen

Fighter aircraft

Lockheed Martin F35

Fundamental Questions Around Benchmarking

- ✓ Seek guarantees that you are benchmarking externally with “peer” groups: industry, company type (new start-up, young and financially sustainable, mature and aged—each will have different strategies and desired outcomes) journey stage, size, maturity, education, geographic affinity, culture
- ✓ Consider that organizations *may* submit their very best projects to benchmarking studies
 - No one wants to look bad in front of their “boss”
 - “Better” projects are more likely to have available benchmark data
- ✓ “Better” projects likely received more leadership support / scrutiny / attention
- ✓ If you were an industry leader, would you provide your best practices to your competition?

ISBSG Agile 'Myths' & 'Truths'

(Agile Productivity Measures)

Myths	Truth
A. Velocity burndown charts show the progress of the team in an objective way	As story points is not a standard, they can be manipulated easily, especially when used in contract KPI's. This is usually referred to as the concept of Story Point Inflation.
B. Function Points were suitable to measure Cobol (sic) COBOL software running on mainframe computers, back in the eighties.	Function Point Analysis is an ISO/IEC standard to measure the functionality that software provides to the user, regardless the technical environment and regardless the non-functional requirements. Just like the meter was used hundreds of years ago and is still used to measure length, function points can always be used to measure the functionality provided by software to the users. The high-level variant of function points can easily be learned in one or two days by Scrum masters, requirements analysts or other team members that understand functional requirements. FPA costs less time per sprint than the sprint planning sessions and produce more accurate and reliable results.
C. It's expensive to use function points and the accuracy is low.	For the team, FPA may be considered waste, as story points metrics give them enough grip on performance on the team level. Management however needs grip on budgets, contracts, performance and need to make decisions on team size. As it is impossible to base this on story point metrics, FPA is definitely not waste on a management level.
D. Measuring function points should not be done in an agile team, as it is considered waste (in Lean terms).	
E. Even if we have the function points, there is no data available to compare to and it will take a long time to build up metrics.	The ISBSG repository has data of over 9100 Application Development projects, releases and sprints in Excel format to start with. As sprints are usually quite short (2 or 3 weeks) building up data will go quite fast. (updated 8/18/2020; thanks to Pierre Almen)

Ref: <https://www.isbsg.org/2018/11/14/productivity-measurement-in-agile-teams/>

ISBSG ‘Myths,’ ‘Truths,’ and now, Reality

(Rebuttal Considerations)

- A. JRS: Story points are not standardized, though they are often “standardized” within scaled teams. Burndown charts reflect the amount of work remaining in an iteration or release. Unless the “story” changes, the number of story points is not relevant, merely the remaining work to be completed that was “committed.” Therefore “inflation” is not a threat to making progress visible on a Burndown chart.
- B. JRS: The purpose of a software solution is to provide value to the customer. One purpose of Function Points is to provide a functional size of that solution; often that size is estimated prior to the start of the work and then extrapolated with other data to estimate cost and duration. The argument that a solution is not easily measured does not negate the legitimate value ascribed to that solution by the funding customer. User stories written at an elementary transaction level are easily counted with function points. While both practical and useful, imposing a level of detail for user stories as a standard would seem unpopular, much like telling an country that it has to measure distance with a meter.

See: [Function Points, Use Case Points, Story Points: Observations from a Case Study](#); CrossTalk; May / June, 2013

- C. JRS: Sprint planning sessions have two outcomes: the committed scope (user stories) of the sprint, and the tasks and estimated times to complete tasks for *committed* user stories. Since the team makes the commitment to complete the sprint, having a specialist—in itself an agile misalignment with generalists—perform the sizing with function points sounds as uncollaborative—another agile misalignment—as project managers ascribing time and cost to project components on behalf of the team itself. The purpose of the sprint is to define a scope of work to be completed in the time-boxed sprint, not to estimate the functional size of the product.
- D. JRS: Ironic that as business change accelerates, time-to-market shrinks, platforms and technologies expand, and a reliance on discovery for emerging products increases, that we continue to emphasize the stale and stagnate role of management and traditional tools for performance tracking. The historic and persistent reliance on cost, schedule, and scope, misleadingly reflected in red-yellow-green charts, needs to be replaced with value delivery, release frequency, and business priorities reporting. Perhaps the greater ‘waste’ is forcing 1980’s project reporting principles onto modernized solutions delivery capabilities. Standards are important, but not as important as value delivery; the business, not project management owns that determination.
- E. JRS: Rather than tout the number of projects, etc. (9100), ISBSG could display a live dashboard of agile projects, sprints, and releases, since 2018 (the three most current years (2018, 2019, and 2020)) for which they have ‘validated’ numbers to guide their subscribers in the use of agile practices and tools. Isn’t that what organizations are more interested in? **What best practices are driving results today!**

What do ISBSG and this industry have in common?

- ✓ Both capture performance data based on samples of their respective populations
- ✓ Both are interested in: **Cost** (how much), **Schedule** (how long), **Scope** (services provided)
- ✓ Both span worldwide services: billions of dollars and “users”
- ✓ Both are based on customer value delivery

But

- ✓ One is significantly older than the other
- ✓ One is significantly more dangerous than the other
- ✓ Regarding productivity, at least one may have an inverse relationship of productivity to value

Type of abuse	Outdoor workers	Indoor workers
Robbed	37%	10%
Beaten	27%	1%
Slapped, punched, kicked	47%	14%
Kidnapped	20%	2%

<http://www.havocscope.com/prostitution-statistics/>

What do the games of horseshoes, darts, and bocce ball have in common?



You can WIN even if you miss the target; *close* may be good enough



Not so much with product development; *close* may ruin your reputation and business



Thank you!



Every morning in Africa, a gazelle wakes up. It knows it must run faster than the fastest lion or it will be killed.



Every morning a lion wakes up. It knows it must outrun the slowest gazelle or it will starve to death.



It doesn't matter whether you are a lion or a gazelle. When the sun comes up, **you better start running.**

The World is Flat; Thomas Friedman; pg. 137

About the presenter . . .



SCT, SSMC, SSPOC, SMC, SPOC, SDC, SA, SAMC, CSQA, CSMS; formerly a Certified CMMI Instructor, CFPS, LSS Black Belt

Bio: <https://joejr.com/bio.htm>

Presentations: (~65)

<https://www.joejr.com/present.htm>

Publications: (~45)

<https://www.joejr.com/pub.htm>

Since 2012: Joe continues to enable enterprise-wide agile transformation through executive coaching; organization training, certification, and practice; policy and process codification; ongoing improvement; organizational alignment; collaborative teaming; and value delivery.

Selected Key Roles: Joe Schofield is a Past President of the International Function Point Users Group. He retired from Sandia National Laboratories as a Distinguished Member of the Technical Staff after a 31-year career. During twelve of those years he served as the SEPG Chair for an organization of about 400 personnel which was awarded a SW-CMM® Level 3 in 2005. He continued as the migration lead to CMMI® Level 4 until his departure.

As an enabler and educator: Joe is an Authorized Training Partner with VMedu and a Scrum Certified Trainer with SCRUMstudy™. He has facilitated ~200 teams in the areas of software specification, team building, and organizational planning using lean six sigma, business process reengineering, and JAD. Joe has taught over 100 college courses, 75 of those at graduate level. He was a certified instructor for the Introduction to the CMMI for most of the past decade. Joe has over 80 published books, papers, conference presentations and keynotes—including contributions to the books: *The IFPUG Guide to IT and Software Measurement (2012)*, *IT Measurement*, *Certified Function Point Specialist Exam Guide*, and *The Economics of Software Quality*. Joe has presented several worldwide webinars for the Software Best Practices Webinar Series sponsored by Computer Aid, Inc.

Life long learning: Joe holds eight agile-related certifications: SAFe Agilist 5.0, SCT™, SSMC™, SSPOC™, SMC™, SDC™, SPOC™, and SAMC™. He is also a Certified Software Quality Analyst and a Certified Software Measurement Specialist. Joe was a CMMI Institute certified Instructor for the Introduction to the CMMI®, a Certified Function Point Counting Specialist, and a Lockheed Martin certified Lean Six Sigma Black Belt. He completed his Master's degree in MIS at the University of Arizona in 1980.

Community & Family: Joe was a licensed girl's mid-school basketball coach in the state of NM for 21 seasons--the last five undefeated, over a span of 50 games. He served seven years volunteering in his church's children's choir; eventually called to coordinate 150 children and 20 staff. Joe is a veteran having served four years in the U.S. Air Force and six more in the Air National Guard. He was appointed to serve on the state of New Mexico's Professional Standards Commission. By "others" he is known as a husband, father, and grandfather.