Measuring Agility
A Validity Study on Tools Measuring The

Agility Level of Software Development Teams

Konstantinos Chronis

June 12, 2015

Outline

Introduction

Research Purpose

Related Work

Research Questions

Research Methodology

Results

OPS Enhancement

Answers to Research Questions

Conclusions

Future Work

Introduction

Agile methodologies adoption has increased a lot [1, 2, 3]. Many teams believe they are agile although they are not [4], while 13% of companies are at odds with core agile values [5]

The reasons for this mainly occur during the agility transition. Some of them are:

- agile methodologies are easier to misunderstand [6]

- agile methodologies are often applied to the wrong context [7]

- organizations modify agile practices before implementing them [8, 9]

- people in the same team have different perception of agile practices [10]

Based on the above, it is evident there is a huge need for tools measuring agility to verify the level of transition.

Agility measurement tools problem
For over a decade, researchers have been constantly coming up with models and frameworks in an effort to provide a solution. Unfortunately, the multiple tools have created a saturation in the field, resulting in them being used only by the organizations that participated in the empirical studies for their creation [11, 12].

Research Purpose

See if tools claiming to measure agility will yield similar results.

Related Work

Many tools for measuring agility have been created. We have separated them into two categories.

- How agile the agile methodologies are

- Agility level of software development teams

Agility of Agile Methodologies

- 4-Dimensional Analytical Tool

- XP Evaluation Framework

- Comprehensive Evaluation Framework for Agile Methodologies (CEFAM)

Agility Level of Teams

- Team Agility Assessment (TAA)

- Perceptive Agile Measurement (PAM)

- Objectives Principles Strategies Framework (OPS)

- Sidky's Agile Measurement Index (SAMI)

- Thoughtworks

- Comparative Agility

Tools Selected

- Perceptive Agility Measurement (PAM) [13] (validated with a large sample)

- Team Agility Assessment (TAA) [14] (used in industry)

- Objectives Principles Strategies (OPS) [15] (covers many agile practices)

Research Questions

1. Will PAM, TAA and OPS yield similar results?

    1.1. Does convergent validity exist among the tools?

    1.2. Will the questions that are exactly the same among the tools yield the same results?

    1.3. What is the coverage of agile practices among the tools?

2. Can the tools be combined in a way that will provide a better approach to measuring agility?

Research Methodology

Study in Company A

United States company which activates in the Point Of Sales (POS) area.
Development and QA departments (4 mixed teams).

Methodology A

The analysis made by Koch [16] was used for identifying the methodologies and
practices Company A uses.

Data Collection

Company A employees were asked to fill in the surveys created by PAM, TAA and OPS (Effectiveness part) during a 3-week period. The surveys were on a 1-7 Likert scale.Practices which were not used by Company A were excluded from the surveys.

Data Preparation

Problem

All the tools have different amount of questions and cover different amount of practices

Solution

1. Mapping of practices between tools — OPS covers more practices

2. Mapping of questions under the OPS practices

Data Analysis

Practice	Participants	PAM	TAA	OPS
Practice1	Participant1	Score1	Score1	Score1
	⋮	⋮	⋮	⋮
	ParticipantN	ScoreN	ScoreN	ScoreN

Table 1: Collected Data Structure

Correlations

- Use the practices covered by each tool and see if they correlate with the same practices from the other two tools

- Use "Spearman's rank correlation coefficient"

Same Questions Analysis

- Identify which questions were the same among the tools and form groups

- "Mann-Whitney U test" and "Kruskal-Wallis one-way analysis of variance" for statistical tests

The hypothesis is

H_0: *There is no difference between the groups of the same questions*

H_1: *There is a difference between the groups of the same questions*

Tools' Agile Practices Coverage

Check which of the tools cover more agile practices and have more questions for a practice. There are two cases:

- either the practices were the most popular according to the two case studies [1, 17].

- or they did not exist/were not popular in the case studies.

Results

Correlations

- Only 8 out of 42 plots were monotonic and allowed the use of "Spearman's rank correlation coefficient"

- Low or negative correlations

Conclusion: Inability to establish convergent validity
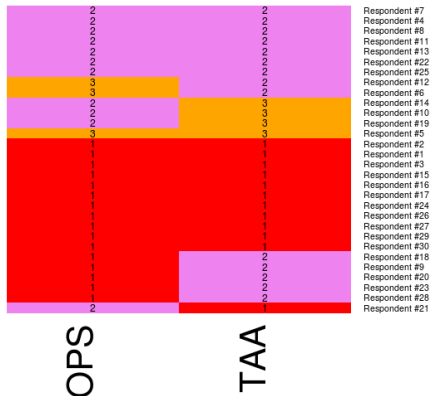
Same questions

Figure 1: Test Driven Development Heatmap

The p-values from the majority of the groups are more than the alpha level of 0.05. As a result, we cannot reject the H_0 hypothesis (*There is no difference between the groups of the same questions*).

Nevertheless, as we see in Figure 1, the frequency of questions which are the same is low, thus, we consider the results as non-significant.

Conclusion: Questions which are the same among the tools don't yield the same results

Tools' Agile Practices Coverage

Practice	TAA	PAM	OPS
Product Backlog	2		3
Smaller and Frequent Product Releases	1	1	4
Constant Velocity	1		
Iteration Progress Tracking and Reporting	17	5	5
Self-Organizing Teams	11	1	7
Daily Progress Tracking Meetings	1	5	1
Retrospective Meetings	5	6	4
Test Driven Development	3	3	4
Refactoring	2	1	4
Software Configuration Management	1		1
Adherence to Standards	3		2
Continuous Integration	5	5	10
Evolutionary Requirements			4
Total	52	26	49

Table 2: Agile Practices Coverage By Tools Based on Laurie Williams' Case Studies

Practice	TAA	PAM	OPS
Iterative and Incremental Development	2	5	3
Customer/User Acceptance Testing		5	
Appropriate Distribution of Expertise	2		5
High-Bandwidth Communication	4	8	13
Client-Driven Iterations		2	3
Minimal or Just Enough Documentation			4
Continuous Feedback		1	2
Total	8	21	30

Table 3: Agile Practices Coverage By Tools Excluding The Ones from Laurie Williams'
Case Studies

Ordered by practices	Ordered by questions
OPS (18)	TAA (52)
TAA (15)	OPS (49)
PAM (13)	PAM (26)

Table 4: Summary Of Agile Practice's Coverage

Conclusion: **The tools don't cover the same agile practices and to the same extent**

OPS Enhancement

Combine the questions of the three tools to create one tool which covers more practices and questions.

The excluded questions concerned:

- product ownership (Scrum oriented)

- iteration defects (software should be delivered to the client)

	Questions Introduced
Capability	7
Effectiveness	46

Table 5: Summary of Questions Added

Answers to Research Questions

RQ #1

1. Will PAM, TAA and OPS yield similar results?
   No. The tools measure the agility of teams in their own way

   1.1. Does convergent validity exist among the tools? **No**

   1.2. Will the questions that are exactly the same among the tools yield the same results? **No**

   1.3. What is the coverage of agile practices among the tools?
   Different practices are covered, while the number of questions is not the same for each tool

Reasons

- Few or no questions for measuring a practice

- The same practice is measured differently

- The same practice is measured by opposite questions

- Questions phrasing

- Survey answering

- Better understanding of agile concepts

- How people perceive agility

Few or no questions for measuring a practice

The *Smaller and Frequent Product Releases* practice is covered by

- 1 question in PAM

- 1 question in TAA

- 4 questions in OPS

Appropriate Distribution of Expertise practice is not covered at all by PAM

The same practice is measured differently

For the *Refactoring* practice

- PAM checks whether there are enough unit tests and automated system tests to allow the safe code refactoring

- TAA and OPS check how often the teams refactor, among other things

The same practice is measured by opposite questions

For the *Continuous Integration* practice

- TAA and PAM check if the team has synced to the latest code

- OPS checks the exact opposite

Questions phrasing

For the *Test Driven Development* practice

- PAM checks for automated code coverage

- TAA checks for 100% automated code coverage

- OPS checks for the existence of code coverage

Survey answering

- Surveys are affected by the comprehension and the assessment of a question [18]

- People misunderstand questions

- A person's judgement is subjective

- People try to look good to the survey reader [19]

Better understanding of agile concepts

In pre-post studies there is a possibility of the subjects becoming more aware of a problem in the second test as a result of the first one

How people perceive agility

- People do not seem to understand the agility concept

- Teams implement agile methodologies differently

- If the agile experts perceive agility differently, then so do the software development teams

RQ #2

2. Can the tools be combined in a way that will provide a better approach to measuring agility?
   They have been combined, but validation is required

Conclusions

- Tools claiming to measure agility do not yield similar results

- A measurement tool which satisfies the needs of one team may not be suitable for other teams

Future Work

- Conduct the study in other companies and compare the results

- Shape the data samples in another way

- Use other tools for measuring agility to conduct a similar study

- Validate the enhancements made in OPS

Questions?

References

[1] Laurie Williams. What agile teams think of agile principles. *Communications of the ACM*, 55(4):71, 2012.

[2] Kieran Conboy and Xiaofeng Wang. Understanding agility in software development from a complex adaptive systems perspective. 2009.

[3] O. Salo and P. Abrahamsson. Agile methods in european embedded software development organisations: a survey on the actual use and usefulness of extreme programming and scrum. *IET Software*, 2(1):58, 2008.

[4] Ambysoft. How agile are you?, 2013.

[5] VersionOne. 8th annual state of agile survey, 2013.

[6] V. Escobar-Sarmiento and M. Linares-Vasquez. A model for measuring agility in small and medium software development enterprises. In *Informatica (CLEI), 2012 XXXVIII Conferencia Latinoamericana En*, pages 1–10, Oct 2012.

[7] M. Taromirad and R. Ramsin. Cefam: Comprehensive evaluation framework for agile methodologies. In *Software Engineering Workshop, 2008. SEW '08. 32nd Annual IEEE*, pages 195–204, Oct 2008.

[8] C. Patel, M. Lycett, R. Macredie, and S. de Cesare. Perceptions of agility and collaboration in software development practice. In *System Sciences, 2006. HICSS '06. Proceedings of the 39th Annual Hawaii International Conference on*, volume 1, pages 10c–10c, Jan 2006.

[9] J.A. Livermore. What elements of xp are being adopted by industry practi-

tioners? In *SoutheastCon, 2006. Proceedings of the IEEE*, pages 149–152, March 2006.

[10] Scott W. Ambler. Has agile peaked?, 2011.

[11] Samireh Jalali. *Efficient Software Development through Agile Methods*. PhD thesis, Blekinge Institute of Technology, 2012.

[12] Samireh Jalali, Claes Wohlin, and Lefteris Angelis. Investigating the applicability of agility assessment surveys: A case study. *Journal of Systems and Software*, 98:172–190, 2014.

[13] Chaehan So and Wolfgang Scholl. Perceptive agile measurement: New instruments for quantitative studies in the pursuit of the social-psychological effect of agile practices. volume 31 of *Lecture Notes in Business Information Processing*, pages 83–93. Springer, 2009.

[14] Dean Leffingwell. *Scaling Software Agility: Best Practices for Large Enterprises (The Agile Software Development Series)*. Addison-Wesley Professional, 2007.

[15] Shvetha Soundararajan. *Assessing Agile Methods, Investigating Adequacy, Capability and Effectiveness*. PhD thesis, Virginia Polytechnic Institute and State University, 2013.

[16] A.S. Koch. *Agile Software Development: Evaluating The Methods For Your Organization*. Artech House, Incorporated, 2005.

[17] B. Murphy, C. Bird, T. Zimmermann, L. Williams, N. Nagappan, and A. Begel. Have agile techniques been the silver bullet for software development at microsoft? In *Empirical Software Engineering and Measurement, 2013 ACM / IEEE International Symposium on*, pages 75–84, Oct 2013.

[18] Markus Wagner and Eva Zeglovits. Survey questions about party competence: Insights from cognitive interviews. *Electoral Studies*, 34:280–290, 2014.

[19] Robert Feldt, Lefteris Angelis, Richard Torkar, and Maria Samuelsson. Links between the personalities, views and attitudes of software engineers. *Information and Software Technology*, 52(6):611–624, 2010.