The Journal of Systems and Software 107 (2015) 38–49

Contents lists available at ScienceDirect

The Journal of Systems and Software

journal homepage: www.elsevier.com/locate/jss

The prospects of a quantitative measurement of agility: A validation study on an agile maturity model

Lucas Gren [a,*], Richard Torkar [a,b], Robert Feldt [a,b]

[a] Chalmers University of Technology and the University of Gothenburg, Gothenburg SE-412 96, Sweden
[b] Blekinge Institute of Technology, Karlskrona SE-371 79, Sweden

ARTICLE INFO

Article history:
Received 11 September 2014
Revised 18 February 2015
Accepted 1 May 2015
Available online 19 May 2015

Keywords:
Agility
Empirical study
Validation

ABSTRACT

Agile development has now become a well-known approach to collaboration in professional work life. Both researchers and practitioners want validated tools to measure agility. This study sets out to validate an agile maturity measurement model with statistical tests and empirical data. First, a pretest was conducted as a case study including a survey and focus group. Second, the main study was conducted with 45 employees from two SAP customers in the US. We used internal consistency (by a Cronbach's alpha) as the main measure for reliability and analyzed construct validity by exploratory principal factor analysis (PFA). The results suggest a new categorization of a subset of items existing in the tool and provides empirical support for these new groups of factors. However, we argue that more work is needed to reach the point where a maturity models with quantitative data can be said to validly measure agility, and even then, such a measurement still needs to include some deeper analysis with cultural and contextual items.

1. Introduction

The study of agile development and management practices is a relatively new field of research. The term itself, "agile development", was first coined in the area of software development but similar concepts preceded it in the literature on manufacturing. Today it has become a general project management concept/tool, and the word "agile" is frequently used in the general business and project management literature, e.g. Miles (2013), Poolton et al. (2006), Vinodh et al. (2010).

Agile methods in software engineering evolved during the 1990s and in 2001 it became a recognized concept due to "The manifesto for agile software development" written by a group of software developers (Fowler and Highsmith, 2001). According to Cobb (2011) the background to the agile ideas was that projects in crisis sometimes took on more flexible ways of thinking and working and then were more successful. This style was named "agile", which literally means to be able to move quickly and easily (Fowler and Highsmith, 2001), and emerged in reaction to more traditional project management methods were detailed planning typically precedes any implementation work.

During the 1990s the traditional way of doing procurement, elicitation of requirements, contract negotiations and then production and, finally, delivery (e.g. what is often termed the waterfall model in software development literature), sometimes helped create computer and software systems that were obsolete before they were delivered. To try to solve these challenges the agile community thus defined a set of values that they summarized in the agile manifesto (Fowler and Highsmith, 2001):

- Individuals and interactions over processes and tools.
- Working software over comprehensive documentation.
- Customer collaboration over contract negotiation.
- Responding to change over following a plan.

Laanti et al. (2011) claim that scientific and quantitative studies on agile methods were still rare in 2011, while requesting such studies since they can give more general advice about the practices involved. Overall, if an organization wants to transition to more agile ways of working, regardless of whether they are a software organization or not, the decision-makers will benefit from measuring agility both before, during, and after such a transition. The question is if this is possible since agility is a cultural change (described in the agile manifesto above) as well as a smorgasbord of practices to support them (Ranganath, 2011; Williams, 2012; Zieris and Salinger, 2013).

There is a diversity of agile measurement tools out there, both scientific and commercial but almost none of them has been statistically validated. In order to measure agility and trust in the given results/output, both researchers and practitioners need validated tools

* Corresponding author. Tel.: +46 739 882 010.
E-mail addresses: lucas.gren@cse.gu.se (L. Gren), richard.torkar@cse.gu.se (R. Torkar), robert.feldt@bth.se (R. Feldt).

http://dx.doi.org/10.1016/j.jss.2015.05.008
0164-1212/

to guide their process. The problem is what to focus on and on what level, since the agile approach is on a diversity of levels in the organization. This empirical study will evaluate one of the agility maturity models found in research through a statistical validation process. This tool focuses a bit more on behavior and not only lists a set of practices for the research subjects to tick yes or no regarding if they are implemented or not. We also connect a Likert scale to the evaluation in order to capture more variance in connection to each item. Section 2 will outline existing agile measurement tools found in the literature, Section 3 will present how our main statistical investigation was conducted, but also describe a pretest conducted before the main study including its findings under Section 2.2, Section 4 will present main study findings, Section 5 will analyze and discuss these overall results, and, finally, Section 6 will present conclusions and suggest future work.

This study aims to contribute with the following:

1. A test to evaluate if the agile adoption framework can be used to measure current agility (instead of agile potential).
2. If practitioners think such an evaluation is relevant through a case study pretest.
3. Expand the agile adoption framework to include a Likert scale evaluation survey filled out by all the team members and not just by the assessor/researcher and connect a confidence interval to the item results.
4. Partly validate the agile adoption framework with statistical tests.
5. Suggest changes agile adoption framework and/or highlight the issues connected to agility measurement.

2. Related work

Some researchers suggest qualitative approaches like interviewing as a method for assessing agility in teams (Boehm and Turner, 2003; Pikkarainen and Huomo, 2005; Sidky et al., 2007). Hoda et al. (2012) even suggest the use of grounded theory which is an even more iterative and domain specific analysis method (Glaser and Strauss, 2006). Interviewing is a good way to deal with interviewee misinterpretations and other related biases. The work proposed by Lee and Xia (2010) compares a few agility dimensions with performance and draw conclusions about the complexity of if agile methods increase performance or not, which they do.

Datta (2009) describes an Agility Measurement Index as an indicator for determining which method of Waterfall, Unified Software Development Process (UP), or eXtreme Programming (XP) should be used. Where Waterfall is plan-driven and XP is an agile method, UP is considered to have elements of both and is a more general framework that can be adapted to specific needs but that is often used as a kind of middle ground between the other two. The author suggests that the five dimensions: duration, risk, novelty, effort, and interaction should be taken into account when selecting development method. Their method is, however, a company-specific assessment, which makes comparisons between different organizations cumbersome.

To be able to compare and guide organization in their agile implementations a diversity of agile maturity models have been suggested, as mentioned in Section 1. Leppänen (2013) presents a useful overview of these agile maturity tools selected with the following criteria: “domain” (the domains the models are targeted to), “purpose” (the purposes the models have been developed for), “conceptual and theoretical bases” (the conceptual and theoretical backgrounds upon which the models have been built), “approaches and principles” (the approaches and principles used to construct the models), “structure” (the architectures of the models), and “use and validation” (extent of deployment and validation). Based on these criteria eight tools were selected: the agile maturity model (Ambler, 2010), a road map for implementing extreme programming (Lui and Chan, 2006), toward maturity model for extreme programming (Nawrocki et al., 2001), the agile maturity map (Packlick, 2007), agile maturity model (Patel and Ramachandran, 2009), agile maturity model (Leppänen (2013)), a framework to support the evaluation, adoption and improvement of agile methods in practice (Qumer and Henderson-Sellers, 2008), and the agile adoption framework (Sidky et al., 2007). According to Leppänen (2013) some of them are merely based on conceptual studies, others are developed only in one organization, a third group has gathered more experience from organizations, and some are discussed with practitioners. However, as also Leppänen (2013) concludes, none of them are validated. He also states that higher maturity levels could partially be assessed by more lightweight methods.

A process control method often used within IT is the American CMMI (Capability Maturity Model Integration) or the European ISO/IEC 15504 SPICE (Software Process Improvement and Capability Determination). These methods also divide the organization into different maturity levels and are essentially a set of requirements for engineering processes, particularly those involved in product development. Just like stage-gate project management these older methods often co-exist with agile methods when implemented (Turner and Jain, 2002). Since agile development processes are more of a cultural change we want to use a value-driven agile maturity model connected to measuring such behavior, i.e. we want the model we use to be built on the agile principles and not on process maturity per se.

Ozcan-Top and Demirors (2013) also compared and evaluated different agile maturity models based on fitness for purpose, completeness, definition of agile levels, objectivity, correctness, and consistency. According to their analysis Sidky’s agile adoption framework was given the best assessment results. Recently, the study by Jalali et al. (2014) showed that a set of agile measurement models give different results when tested with practitioners. This further motivates our study’s scientific validation approach to such measurements (it is obvious to us that they will not show the same results since they have not been scientifically validated).

In this study we selected to focus on the Sidky’s agile adoption framework, and in order to keep the number of items as low as possible, we selected only Level 1 of this tool. We should also mention that there is a set of commercial tools available, however, their scientific foundation is hard to assess.

We would like to highlight the difficulty of measuring something that is an ambiguous construct, such as agility. Maturity is of course even harder to assess in connection to agility since maturing with a unspecific concept is even harder. However, there are some behaviors connect to “being agile” in software development and behavior connected to this way of working, which is our definition of agile maturity in this case. We do not aim to find a way to quantitatively measure agility in this study (and we neglect the agile practices’ effectiveness/quality as well), but instead to test one of the existing tools and try to understand how to proceed in measuring/dealing agility transformations in organizations.

2.1. Sidky’s agile adoption framework

In order to determine which agile methods an organization is ready to use, Sidky (2007) suggests a method called the agile adoption framework. He motivates its use by arguing that even though there are many success stories in agile development, they are not really generalizable, i.e. it is unclear how the case by case descriptions can be used to judge agility readiness for a company which has some, but not all, aspects in common with reported cases. Sidky also criticizes more general frameworks, since they address agility in its generic form and not the actual practices.

Sidky’s approach is based on a tool that has two parts. The first part is called the agile measurement index (the same name as Datta (2009) uses, but a different tool) and is: