International Journal of Information Systems and Project Management (IJISPM)
2026, 14(4), e140402, DOI: 10.12821/ijispm140402

RESEARCH ARTICLE

Adaptive metrics in agile software development projects

João Barata
University of Coimbra, CISUC/LASI, DEI, Pólo II - Pinhal de Marrocos, 3030-290 Coimbra, Portugal,
barata@dei.uc.pt

Sharon Coyle
University of Galway, J.E. Cairnes School of Business & Economics, H91 TK33 Galway, Ireland,
sharon.coyle@universityofgalway.ie

Abstract
Multiple metrics are available for agile software development (ASD), but adapting them to changing conditions of ASD is a challenge. To be truly agile, the metrics must be goal-oriented, flexible, and adhere to the principles of people interactions. This paper presents the results of a multi-year action research project conducted in different companies adopting ASD. Our contribution describes the organizational routines and a framework (3View) to incorporate meaningful metrics supported in three main pillars: comparability with past projects, relevance to measuring project expectations, and adaptability to remain valid in dynamic project conditions. The proposed framework includes (1) a reference model to build metrics and (2) a process model to guide practitioners. Measures of all types of attributes in ASD can be evaluated differently, depending on the project stakeholders and lifecycle stage. Dynamic environments require adaptive metrics that guide the interpretation and directions for project development. Failure to adopt these recommendations may lead to a risk of ceremonial conformity to measurements that do not reflect practice. Our work extends the literature on ASD metrics, expanding their role as enablers of agile project assessment and transparent communication throughout the project lifecycle. It explains how ASD metrics can be adapted to fit stakeholders' perceptions while maintaining rigor and transparency in their reasoning.

Keywords
agile software development; metrics; 3View; action-research.

Received: 16 December 2023 | Accepted: 26 September 2025

1. Introduction

Information Systems project management is constantly seeking out improvements. Becoming agile is an example, requiring an iterative process of stakeholder interaction to develop better-quality products in turbulent environments (Beck et al., 2001). Therefore, multiple metrics are available to establish and monitor agile project management quality (A. Agarwal et al., 2014; Hayes et al., 2014; Kupiainen et al., 2015). For example, some authors focus on tests and quality control (A. Agarwal et al., 2014; Janus et al., 2012), product-related metrics (Kupiainen et al., 2015; Mishra et al., 2012), defects (di Bella et al., 2013), auditing (Scharff, 2011), or stakeholder expectations (Boerman et al., 2015). Despite the plethora of metrics available, there are difficulties in adopting them in dynamic project management environments (such as agile).

Quality metrics can be defined as the degree to which a system or process holds specific quality attributes (IEEE, 1998; Kaner & Bond, 2004), and several studies have addressed this topic. However, quality metrics in the context of agile software development (ASD) projects are still evolving (Colakoglu et al., 2021; Jamieson & Fallah, 2012; Kupiainen et al., 2015; Mishra et al., 2012), and some authors suggest that additional research is necessary on "extensions of metrics and analytics to accelerate best practice" (Durbin & Niederman, 2021). The dynamic principles of ASD (Beck et al., 2001) are not aligned with the vision of ceremonial conformity, which attempts to minimize inspections that can result in complacency, whereby descriptions of a system are not aligned with ongoing practice (Biazzo, 2005; Iden, 2012; Meyer & Rowan, 1977). Moreover, project agility involves many factors besides mere practices or methods, for example, stakeholder involvement and collaborative value co-creation, empowerment, team culture, and organizational factors (di Bella et al., 2013; Li et al., 2017; Sheffield & Lemétayer, 2013).

Quality is an iterative and continuous endeavor in agile project management (Dybå & Dingsøyr, 2008), focusing on interactions between individuals and the ability to be proactive in the face of change (Beck et al., 2001; Ghobadi & Mathiassen, 2016; Kropp & Meier, 2015). By definition, agility also relates to the ability to create change (Dybå & Dingsøyr, 2008), which teams may strive to achieve in a positive, quality-centric manner. Nevertheless, current quality studies in ASD projects do not explain in detail how critical reflections of the ASD team can be conducted using metrics. Moreover, it is essential to combine quantitative evidence of ASD quality with qualitative assessments, the latter of which are critical to the core principles set out in the agile manifesto (Beck et al., 2001).

Project performance indicators are extremely popular (N. Agarwal & Rathod, 2006; Heston & Phifer, 2011). Nevertheless, Kupiainen et al. (2015) found that most metrics are non-inclusive of people which is an inherent gap given the *people over process* principle embedded in agile projects pointing out that nearly 40% of the identified metrics needed to be customized. It is surprising that "the utility and appropriateness of some of the metrics that contemporary SWQATs [software quality assessment tools] implement are highly debatable [, requiring to] support the user in defining and customizing the metrics used to measure software quality" (Pfeiffer & Aaen, 2024). Our literature review confirms that process-related metrics (e.g., velocity, story-cycle time, defects removed, test coverage) and product-related metrics (e.g., functionality, portability, reliability, usability) are well-identified (Gruschwitz & Schlosser, 2012). The central gap is *how* to use them in practice in an integrated framework that involves people in process and product measurement in a way that is both flexible and trustworthy.

In this research, we follow the description of 'metric' included in the IEEE 1061 standard (IEEE, 1998), pointing to quantitative measurements of specific quality attributes that need to be interpreted, namely a "function whose inputs are software data and whose output is a single numerical value that can be interpreted as the degree to which software possesses a given attribute that affects its quality" (IEEE, 1998). However, to be socially constructed (Berger & Luckmann, 1991), metrics must be adjusted by project stakeholders, which is a challenge in minimising the risks of opportunistic manipulation. Therefore, agile metrics are not restricted to mere observation of a fact because stakeholders must (1) have power to change their structure and construction during the entire project and (2) do not limit their role in comparing the metric to an immutable and predefined goal. Therefore, two research questions are formulated for this research:

- RQ1: How do researchers and practitioners address quality assessment and improvement in ASD projects?
- RQ2: How can project development teams develop and adapt metrics during the entire lifecycle of ASD?

To answer RQ1, we survey the state-of-the-art in agile quality and investigate the outlook of experienced ASD practitioners around quality assessment and improvement. Canonical action research (CAR) was selected to address our second research question (RQ2), building and evaluating the proposed framework (3View) to use agile metrics in a real setting. Our main objective is to adapt quality metrics that adhere to the core principles of ASD projects, incorporating approaches drawn from people, process, and outcome-related sub-factors pertaining to the project and the product (e.g., software quality and end-user satisfaction).

The following section outlines our research approach. Subsequently, we present the background literature, including a review of 69 papers on the most relevant strategies to quality management and metrics in ASD projects. It is followed by pilot findings of semi-structured interviews with ASD experts. Subsequently, we explain the proposed framework to design adaptive metrics that are socially constructed and detail two complete canonical action research cycles (Susman & Evered, 1978) conducted with a leading Information Technology (IT) supplier of healthcare information systems (HIS) and a research and technology development (RTD) institute with the mission to assist Industry 4.0 developments. The discussion follows, and our paper closes with the main conclusions, our study's limitations, and future research opportunities.

2. Methods

Our research commenced by conducting an extensive literature review (Kitchenham, 2004; Webster & Watson, 2002) to address our first research question. Kitchenham (2004) and Okoli and Schabram (2010) provide a sequence of steps to conduct comprehensive reviews, namely to (1) identify the need for the research, (2) establish a review protocol, (3) search the literature, (4) study selection and screening, (5) assess quality, (6) extract and monitor data, (7) make a synthesis, and (8) write the review.

The review pertaining to agile quality management and its metrics included journals and conference proceedings using Google Scholar, EBSCOhost, Science Direct, Web of Science, and IEEE Xplore. We started with broad search terms and progressively refined the results (Boell & Cecez-Kecmanovic, 2014), screening the titles and abstracts obtained with the search terms "agile quality" and "quality in agile". Using different databases is advantageous: Google Scholar has a broader scope but presents the most extensive lists. Additionally, this provided an opportunity to iterate the "search and acquisition circle and the wider analysis and interpretation circle" (Boell & Cecez-Kecmanovic, 2014) by comparing databases. Subsequently, we tested a combination of related search terms, for example, "agile development" + "quality metric"; "agile project" + "quality metric"; "agile development" + "quality monitoring"; and "quality improvement" + "agile practice". We also performed citation analysis (Webster & Watson, 2002) to identify related works and tested keywords related to our study, namely "agile software development" + "composite metrics". Papers that did not address quality assessment or improvement within ASD contexts and practices were excluded. We did not include books, non-English papers, tool presentations, editorials, posters, patents, keynotes, or panel conclusions. A total of 69 papers were reviewed. The most essential concepts in the literature include quality assessment in agile and agile practices for quality improvement. We subsequently identified units of analysis within each concept: (1) quality assessment in agile included process, outcome, and metrics in the form of evidence, goals, and quality indices; and (2) agile practices for quality improvement included stakeholders' interaction within agile projects and the benefits and pitfalls of agile practices for quality improvement.

Secondly, exploratory qualitative interviews (Myers & Newman, 2007) were completed in two geographically dispersed organizations (Portugal and Australia). The goal was to identify how metrics are used in practice, ascertain quality assessment difficulties, and obtain initial insights for improving metrics. We created a script for semi-structured interview sessions, including questions for (1) company identification, (2) experience and practices of agile, (3) forms of measuring quality, (4) metrics, (5) improving agile quality, and (6) expectations for a new framework for agile quality. The authors conducted in-depth interviews, averaging two hours. Afterward, answers were discussed with interviewees and subsequently compared with findings across literature to help generate rich data (Schultze & Avital, 2011) and insights for further development.

Thirdly, the canonical form of action research was followed with a dual aim (McKay & Marshall, 2001; Susman & Evered, 1978) of creating the 3View framework to implement agile quality management using adaptive metrics and solving concrete organizational problems in two organizations. Action research is conducted "collaboratively in an immediate situation using data feedback in a cyclical process" (Hult & Lennung, 1980). Canonical action research (CAR) is one of the most popular and well-documented (Davison et al., 2004), evolving in cycles with five phases (Davison et al., 2004; Susman & Evered, 1978):

- Diagnosing, identifying, or defining the situation in collaboration between researchers and practitioners. The project participants evaluate the phenomenon and formulate a hypothesis for the subsequent phases;
- Action planning, deciding possible courses of action to improve the situation;
- Action taking, implementing the selected course of action, and trying to generate improvements in the organization;
- Evaluating actions involving a critical joint reflection by the research team;
- Specifying learning and establishing the contribution to the body of knowledge.

According to Eden and Huxham (1996), action research results must be transferable to other contexts and not be restricted to solving the case company's problem. Moreover, theory must be produced and guide the development of possible tools or methods, as planned in our work. The principles suggested by Davison et al. (2004) were used to evaluate our research: Principle of the Researcher–Client Agreement; Principle of the Cyclical Process Model; Principle of Theory; Principle of Change through Action; and Principle of Learning through Reflection.

The research team considered alternatives to CAR, namely design science research (Hevner & Chatterjee, 2010), which is particularly suited for developing innovative artifacts. However, the primary focus of our work was the analysis of the social setting. Moreover, this approach has also recently been used and verified in the context of agile software development (Kuciapski & Marcinkowski, 2023). The opportunity to collaborate closely with the practitioners in ongoing ASD projects, studying changes introduced by the 3View, was an additional motivation to select CAR.

Two organizations participated in our action research. The HIS company puts quality management at the top of its priorities. Their global presence increases pressure for short development cycles and immediate feedback to their customers and partners. The RTD institute is the second case that we address. It is an interface institution created to support the ceramic and glass industries. In recent years, Industry 4.0 has been one of the sector's priorities, requiring innovative software developments for production and quality. The institute has private and public associates, and the software development team adopts agile. These organizations were selected due to the need to comply with quality standards and receive regular audits from customers and third-party assessors, their ASD practices focusing on several projects evolving in parallel, and their motivation to test a new framework developed in collaboration by researchers and practitioners.

CAR cycles may be conducted in parallel, with benefits for case comparison, or as a sequence, when one cycle complements or extends the previous one, as in our study. The first cycle started due to the challenges of the HIS quality manager in their audits and the commitment of their team to ASD practices. The lack of documentation was a problem, and KPIs were usually available later in the development lifecycle. It was also the starting point for the literature review to understand the research landscape of ASD metrics and approaches for their adoption in practice. However, the team's access to details about their projects was limited due to confidentiality reasons. The first cycle was sufficient to propose a reference model. Still, it lacked the details of continuous work with the ASD practitioners, which was possible in the second cycle with the RTD institute. The initial findings of our work were published in top conference publications, ICIS and ACIS (Barata & Coyle, 2016; Coyle & Barata, 2016), and substantially extended and improved for journal publication. Additional details about the settings in both companies are presented in section 5.1, explaining the diagnosis phase of our CAR.

3. Literature review

3.1. Quality assessment in agile

There is a long tradition of studying metrics for ASD, but it is also recognized that practitioners' guidance is still underdeveloped (Maddox & Walker, 2021). Different approaches have been used for their selection. It is possible to import metrics from traditional software development (TSD) approaches (Concas et al., 2012). A recent study based on four case studies suggests templates and guidelines for KPI development, although recommending a top-down approach for the implementation (Fatima et al., 2024). Still, some authors argue that TSD metrics are insufficient to evaluate progress in an iteration, releases, or feature correction rates (Mishra et al., 2012). Other authors propose a combination of TSD and agile-specific metrics (Padmini et al., 2015). The number of metrics in the literature exceeds 130 (Maddox & Walker, 2021). For example, Kupiainen et al. (2015) found 89, most quantitative and related to lead time, defects, velocity, cycle times, and burndown/burnup, among others. They are used for "Sprint and Project Planning, Sprint and Project Progress Tracking, Understanding and Improving Quality, Fixing Software Process Problems, and Motivating People" (Kupiainen et al., 2015). Interestingly, these authors found that many ASD teams customize metrics (almost 40%), and most relate to outcomes and process. More recently, Natarajan and Pichai (2024) suggest a tailored framework with eight ASD metrics, particularly aiming at the process and product levels of analysis, suggesting a continuous refinement. Surprisingly, people-centric metrics are not a priority in the literature.

Several factors are crucial to evaluate the quality of an ASD project: people (Begier, 2010; Fernández-Sanz & Misra, 2011; Ghobadi & Mathiassen, 2016), process (Hayes et al., 2014; Kupiainen et al., 2015), product (Kupiainen et al., 2015), and outcome (N. Agarwal & Rathod, 2006). The Capability Maturity Model Integration (CMMI) embodied the argument that "the quality of a system or product is highly influenced by the quality of the process used to develop and maintain it" (CMMI, 2010). Metrics in ASD have also been studied for product evaluation (Kupiainen et al., 2015; Mishra et al., 2012), testing (A. Agarwal et al., 2014; Janus et al., 2012), and software defects (di Bella et al., 2013). Nevertheless, meeting stakeholder expectations is also essential, and research confirms that a comprehensive framework to evaluate quality in ASD supported by meaningful metrics is necessary. While many studies present the metrics that can be used (Kupiainen et al., 2015; Maddox & Walker, 2021) and their integration as a pillar of evidence-based quality awareness in ASD teams (Karhapaa et al., 2024), the gap found in the literature is the question of "how to use them" in adherence to agile principles.

There are important quality measures for the outcome and process of ASD. For example, one of the few examples that address the transition from TSD to ASD is presented by Olszewska et al. (2016), who explain significant changes when shifting from a "plan-driven" way of working to agile. Product and process quality were the focus of Rauch et al. (2008) and Gruschwitz and Schlosser (2012), while Bansiya and Davis (2002) proposed an integrated index. Their model includes product-related attributes such as (among others) effectiveness, extendibility, reusability, or flexibility (Bansiya & Davis, 2002) and defines properties that can affect the attributes (e.g., complexity), defining weights for each one and its influence on each attribute. For example, higher complexity reduces understandability. Additionally, the model includes design metrics (e.g., number of methods) and required components. As before, these metrics are also quantitative, and it is now known that agile practices have social implications for the team (Koch et al., 2023). For example, the works of Alami et

2026, 14(4), e140402, DOI: 10.12821/ijispm140402

al. (2022) and Machuca-Villegas et al. (2021) conclude that commitment is one of the most relevant human factors in productivity assessment and the achievement of technical excellence, the ninth principle of the agile manifesto.

Agile quality has several challenges (Alsaqaf et al., 2019); metric selection is only part of the problem (Kārkliņa & Pirta, 2018). Is it reliable even when data is available, and are we using it correctly? A study by Ram et al. (2018) identified 132 metrics, most of which were for process and product—for example, velocity, development speed, testing performance, or estimation accuracy. The authors found three main challenges for metric adoption, namely, (1) lack of data or tools to produce it, (2) changes in the process to produce required data for metrics, and (3) "lack of actionable input" or the difficulty in using metrics for decision making. Moreover, "for highly iterative or agile systems, the metrics will not remain effective forever [... with practical limitations] throughout several software iterations as the software matures and the dynamic nature of the software development process subsides". Challenges in information quality, team, or practices are well-known causes of failure in ASD estimations, and few studies exist to support practitioners in the estimation efforts. New solutions for the effective use of metrics are therefore needed.

Stakeholder interaction in ASD is important to understand the impact of improvement actions and the acquisition and sharing of tacit knowledge (Ryan & O'Connor, 2013) during communication and the implementation of agile practices (Hummel et al., 2015). More "social metrics" like "team capabilities" or "user involvement" are needed to assess and improve quality in ASD. These metrics can address people's behavior and assess communications or aspects of interactions (Dorairaj et al., 2012; Gren et al., 2015; Ibrahim et al., 2010; Wiese et al., 2014). Examples include technology awareness and customer collaboration, empowerment, personal skills, or motivation (di Bella et al., 2013; Tseng & Lin, 2011). Several years ago it was stated that "[t]here is an obvious need for more scientific evidence and further research to understand the requirements of agile quality [operating] in different cultural and social contexts" (Siakas & Siakas, 2007), but companies are still looking for an integrated solution for agile quality management (Jain et al., 2016).

Quality in ASD is a continuous endeavor from the early phases (such as requirements identification and risk analysis), during the development, and after the software delivery. However, metrics do not have the same importance in each phase of the development process (Almeida & Carneiro, 2023), and "more work is needed to reach the point where a maturity model 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" (Gren et al., 2015).

3.2. Agile practices for quality management and improvement

Quality management includes social and technical aspects. A vigorous quality culture is based on customer focus, evidence-based decisions, the involvement of people, and continuous improvement, which is aligned with agile (Stålhane & Hanssen, 2008). For example, the belief in continuous improvement is rooted in retrospective meetings (Babb et al., 2014; McHugh et al., 2012). In highly regulated projects, however, different quality standards and improvement frameworks (like ISO 9001 or ITIL) are adopted in agile environments (Torrecilla-Salinas et al., 2019). For example, Stålhane and Hanssen (2008) highlight documentation difficulties when combining agile and ISO 9001. There exists an inherent tension between growing pressures to adopt agile while simultaneously and harmoniously trying to align such explorative agile practices with highly regulated, often quantitatively focused practices.

There are essential approaches for integrating quality in agile practices. For example, the model presented by Hongying and Cheng (2011) includes 20 key areas for quality assurance in ASD. They identify best practices for each area and a maturity model. The 3C proposal combines metrics and continuous integration (Janus et al., 2012), while Sidky et al. (2007) propose two components for agile quality, namely, (1) an agile adoption index for the principles of "Embrace change to deliver customer value", "Plan and deliver software frequently", "Human centric", "Technical excellence", and "Customer collaboration", and (2) an agile adoption process starting with the identification of factors that can prevent agile success, conduct a project and organizational readiness assessment, and reconciliation to ensure that the required practices are implemented for the project.

End-user feedback can be obtained in different phases of the ASD lifecycle (Hayes et al., 2014). In addition, Baxter and Sommerville (2011) support that the product owner "needs to be extended to consider a broader set of system stakeholders". Nevertheless, ASD and traditional approaches are dissimilar: "the traditional approach of tracking progress against a pre-made plan and measurable goals conflicts with the Agile value of embracing the change [... and its] rather comprehensive set of metrics, which does not align well with the Agile principle of simplicity" (Kupiainen et al., 2015). Research has indicated that when comparing a "waterfall" process to "agile", development performance and product quality are far superior in agile (Tarhan & Yilmaz, 2014), highlighting attempts by agile methodologies to embed quality efforts into processes that are iteratively evaluated.

Agile methodologies may improve quality (Boerman et al., 2015; Jamieson & Fallah, 2012; Torrecilla-Salinas et al., 2019). However, there are different views of quality to consider: (1) quality as excellence, usually requiring a standard; (2) quality as value, including the cost-benefit analysis of quality; (3) quality as conformance with specifications; and (4) quality as meeting or exceeding the expectations of stakeholders. In ASD, "quality as conformance" addresses software quality metrics and requirements satisfaction. However, there are also quality principles and standards ("quality as excellence") and the potential benefits provided by utilizing agile techniques ("quality as value"). Nevertheless, as Kaner and Bond (2004) stated, "there are too many simplistic metrics that do not capture the essence of whatever it is that they are supposed to measure." They are also used "after the fact" and are difficult to apply in iterative and adaptive environments.

Retrospective meetings are an essential ASD practice, pushing continuous improvement. However, development pressures can make this task difficult in practice (Babb et al., 2014; McHugh et al., 2011). Artifacts and guiding steps for retrospective meetings have advantages (Péraire & Sedano, 2014), supporting the reflection about previous iterations and future practices. For example, dashboards can be a valuable tool to present metrics to ASD teams (Boon et al., 2023). Hayes et al. (2014) emphasize quality practices in the early project phases and the potential for the later stages to include documenting and user stories. Quality concerns appear in the early stages of agile projects, proceed in the complete documentation of user stories, and "can be supplemented with a more direct measure of customer-perceived value—using customer satisfaction feedback" (Hayes et al., 2014). There is a need to incorporate adaptive quality practices with traditional techniques that are relevant and suitable for ASD projects (Salo & Abrahamsson, 2007). Fig. 1 presents the main quality touchpoints in ASD.

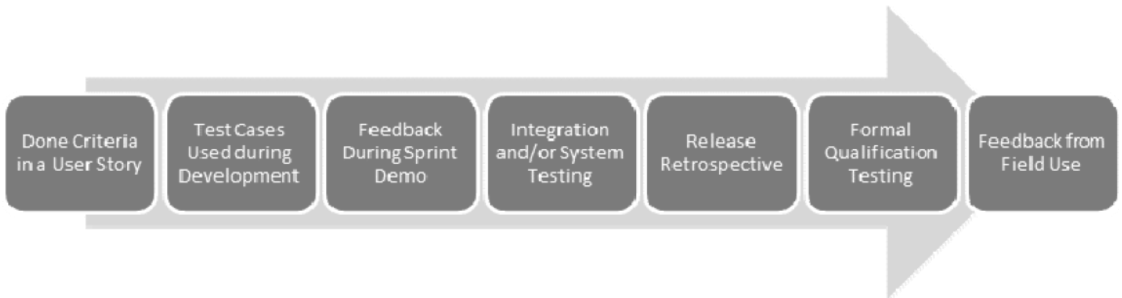


Fig. 1. Important touch-points for quality in agile projects (adapted from Hayes et al. (2014))

As Baxter and Sommerville (2011) put it, it is necessary to extend the notion of requirements ownership to multiple stakeholders in ASD. Moreover, constant pressures can make reflection and analysis difficult (Babb et al., 2014; McHugh et al., 2011), and ASD projects present different challenges when compared to traditional approaches, namely, "the traditional approach of tracking progress against a pre-made plan and measurable goals conflicts with the Agile value of embracing the change [... and its] rather comprehensive set of metrics, which does not align well with the Agile principle of simplicity" (Kupiainen et al., 2015). Measuring defects or functionality issues is insufficient in agile projects (Hayes et al., 2014).

Backlogs and quantitative metrics are essential for assessing past projects or tasks, but are limited in supporting a proactive stance or efforts for continuous improvements. In fact, in "an ever-changing world, such an approach [methods that only focus on analysing and evaluating past actions] bears the risk of producing pictures of a moving target without providing a substantial orientation for change" (van der Aalst et al., 2018). Moreover, agile teams must be able to autonomously interpret metrics and use them in ongoing decision-making processes. In ASD projects, there are daily opportunities to do this due to intense stakeholder interactions among developers, facilitators, testers, and end-user representatives. Ghobadi and Mathiassen (2016) suggest this approach when addressing knowledge-sharing barriers in agile projects and stress that different roles must be considered to overcome communication gaps and promote a common understanding in the team. Despite this, we could not find any literature guidelines for fostering quality assessment and evaluation techniques in ASD projects that incorporate ongoing qualitative criteria with quantitative measures such as defects, velocity, or cycle time.

4. Preparing a frame of reference for action research: pilot interviews

The research team conducted semi-structured interviews (Myers & Newman, 2007) in two organizations (anonymized) adopting ASD, selected from our contacts list. The HIS company indicated an interest in improving its quality metrics, and the consulting company had extensive experience in delivering agile projects. The case companies are presented in Table 1.

Table 1. Qualitative interviews

	HIS company	*Consulting company*
Industry	Healthcare Information Systems	Software Consultancy
Employees	60	230
Interviewee	Quality manager	Senior Developer
Agile practices	Scrum	Scrum; Kanban
Quality Standards	ISO 9001:2015; NP 4457 for innovation management; multiple healthcare standards	None specified

The HIS company is a European software provider of healthcare for hospitals and clinics, founded more than 30 years ago. They serve over 100.000 users and 25 million clinical processes. When we interviewed their quality manager, she was preparing the external audit for ISO 9001. According to the quality manager, the company "has numerous indicators; however, only a few are valid for agile quality". Sometimes, "the numbers are highly dependent on the context and must be carefully interpreted". In other cases, "[she] does not think it fair to establish goals, for example, regarding the number of defects or features implemented; these metrics depend on multiple factors". Agile quality is challenging because "40% of our major customers [representing 80% of the income] require quality indicators and evidence for each iteration, due to the critical nature of healthcare IT".

A lack of suitable implementation of retrospectives contributed to "difficulties in creating improvement on our project and without appropriate communications we are not sharing the knowledge which is a critical aspect of our business due to the complexity of product lines". The interviewee also pointed to the importance of being able to change metrics for each project or team in an "agile way". Quality metrics can be used to build appealing dashboards, "but we need to assess and improve quality [...not] metrics that do not correspond with practice”. Even worse, “template” metrics “and unrealistic goals can reduce the team’s commitment to quality during agile projects”.

Interestingly, when discussing social metrics, the answer was that: “in my previous work, our motivation was evaluated by top managers on a quantitative scale ranging from 1 to 5… How could they know my motivation or the factors affecting it? I hated to be evaluated that way”. When we asked how user intervention might assist in constructing metrics, she stated how “this would be a very useful, inclusive approach and has the potential to address our main issues of knowledge sharing, obtaining quality evidence for our team and external audits, re-invigorating our retrospectives, providing support for weekly meetings and customer request, and “provide meaning” to our agile numbers, according to the team’s perspective”.

Founded fifteen years ago, the Consulting company is a geographically dispersed firm that helps clients develop innovative software projects. Our interviewee has over eight years of experience working on ASD projects across many sectors in the consulting industry and stated that when it comes to measuring quality in agile, “the biggest and most important test is what the user thinks”, highlighting the importance of user involvement in quality evaluation. In particular, this research participant described three quality metrics. The first relates to setting benchmarks for quality at the start of the project, which is “extremely difficult to do when you have nothing to compare it to”, but in his opinion, “mature teams (in performing stages) are much better at doing this”. The second relates to “comparison exercises that we do; so we compare the latest increment against the previous one”, and finally, “our centralized continuous integration server runs automated tests, which is a very good way of measuring software quality”. However, the interviewee stressed that “even when each of those metrics shows us very positive results, if the user is not happy, then we still have not reached our quality pinnacle”. This shows that quantitative-driven assessments alone are not complete quality indicators for agile projects.

When we asked about social metrics, they “could see this making a lot of sense and a huge difference for our agile projects, but then...how we apply those metrics in the context of teams during storming or norming stages is the challenge because these teams often have misaligned observations for quality assessment”. This highlights the challenges of our research effort, but also the importance and benefits of including adaptive metrics for assessing quality in ASD teams while simultaneously raising quality awareness.

5. Implementing agile quality management: insights from action research

This section follows the steps proposed by Susman and Evered (1978). Next, we describe the starting point in each case and the action plan, followed by field results. The action research evaluation in section 5.3 alludes to the guiding principles proposed by Davison et al. (2004).

5.1. Diagnosing and action planning

The diagnosis in the HIS company included the interview with the quality manager described in Section 4, document collection, and meetings with the IT infrastructure manager. Our literature review proceeded throughout the research to identify best practices for ascertaining quality and metrics (Section 3).

The cycle in the RTD institute started one year later, and the first version of the framework was already available. We interviewed a small team of three software developers and a manager of the software development unit. The context of this case is radically different because software development is only one of the priority areas in the case organization (which also includes laboratories, management consulting, and research and development, with several ongoing European co-funded projects). Only a few metrics exist, and according to the most senior software developer, “we have a vast background in using the balanced scorecard, risk-based thinking, and agile metrics; however, measurements are only one part of the equation (…) how do you use them to improve our work? Who evaluates what is good or not – in some cases, she adds, the people who look at the metrics do not understand the numbers. What are the consequences – especially if only some of us are contributing to the results?". Moreover, the team told us that it was challenging to keep their quality platforms updated (project documentation, version control, customers' meetings), so they prefer only a few metrics but "with a script that allows others to read them properly and produce real results".

The first CAR cycle (CAR#1) with the HIS company included the activities of model creation for metrics adaptation, identification of measures that could be applied for each metric type (people, process, outcome), and defining how metrics should be calculated. According to three dimensions of (1) evidence from practice, (2) stakeholders' expectations, and (3) stakeholders' evaluation, a valuable tool for daily meetings, retrospectives, and audits can be developed.

The limitations found in the first cycle were a starting point for planning the next. It was also an opportunity to test and refine the proposed framework in a different setting. Secondly, we wanted to include other stakeholders in our study (as the first cycle only involved the researchers and the selected company managers), namely, auditors, partners, and customers. Therefore, the second CAR cycle (CAR#2) with the RTD institute started one year later, when the first version of the 3View framework had already been created. Our action plan in this case was to:

- Evaluate the 3View framework in different settings with other types of software products and organizational structures;
- Produce data that organizations can use to (self-) evaluate their quality in ASD and the efficacy of improvement actions by comparing how the indicators change over time;
- Investigate the changes (e.g., knowledge transfer, team motivation) involved in adopting the metric in daily meetings, retrospectives, and audits.

5.2. Action taking: a collaborative framework for agile quality management

The next section presents the final version of the proposed 3View framework, followed by a description of its development in CAR#1 (5.2.2) and CAR#2 (5.2.3).

5.2.1 The 3View framework for agile quality management

For our research, a metric is socially constructed when stakeholders can alter its structure and critically evaluate its results. In this context, stakeholders' perceptions are innate to metric construction. Unlike traditional approaches, which compare against predefined goals (e.g., number of defects below X% when compared to lines of code), stakeholders are not extracted from the process or only included at the end. They are involved in metric construction.

Two main artifacts were created to support the practitioners in 3View adoption, namely, (1) a reference model explaining the quantitative assessment and (2) a process model. The 3View reference model includes three dimensions in adaptive metrics for ASD, namely (1) evidence from practice, (2) stakeholders' expectations, and (3) stakeholders' evaluation. It is non-prescriptive about the types of measures that can be socially constructed. Therefore, it may potentially apply to any measurement selected by agile teams. Quantitative data must be enriched with contextual and cultural issues (Gren et al., 2015). On the one hand, we follow numerous authors (Bansiya & Davis, 2002; Janus et al., 2012; Sidky et al., 2007) in the quest to obtain metrics for quality assessment as an essential step of ASD quality improvement. On the other hand, we show how to assess and discuss improvements in ASD projects simultaneously. Metric construction should be a moment of critical reflection in ASD teams (Hodgson & Briand, 2013) and can potentially increase team commitment in ascertaining quality.

The insights obtained in our literature review, pilot interviews, and the CAR action suggest that a comprehensive assessment of agile quality requires three main types of metrics, represented to the left of Fig. 2, people-oriented, process-oriented, and outcome-oriented, summarizing the 3View reference model. Fig. 3 depicts the process.

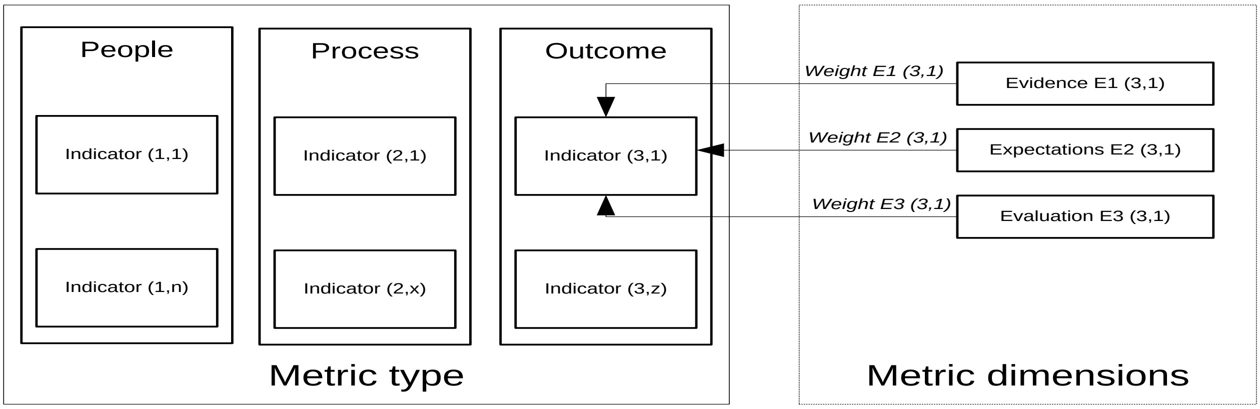


Fig. 2. The 3View reference model

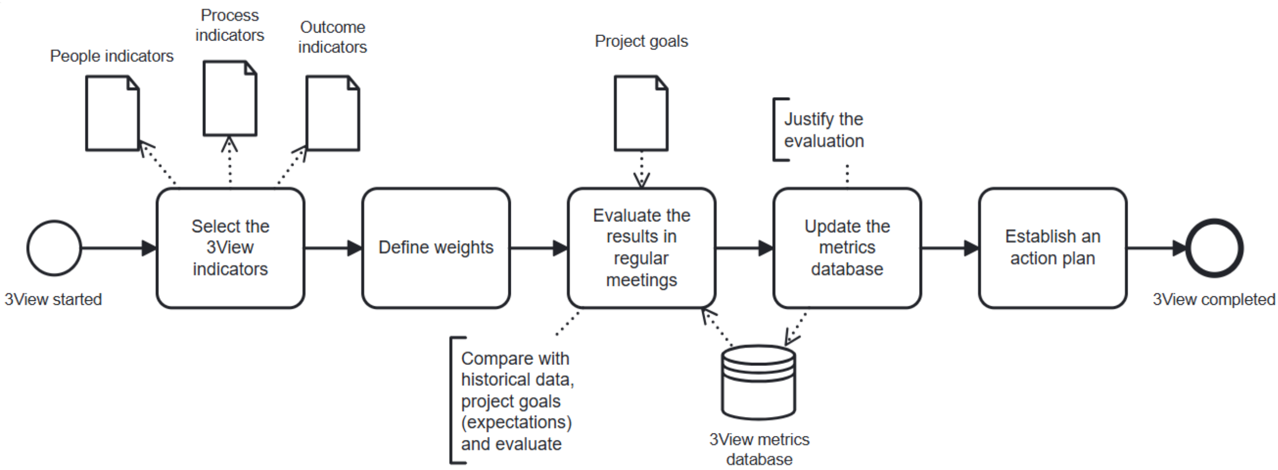


Fig. 3. The 3View process model

The process starts with stakeholders selecting the indicators (step 1). Identifying specific indicators for use in each type is outside the scope of this research. The creation of adaptive metrics occurs in step 2. Each one is weighted according to the three dimensions of (E1) evidence, (E2) expectations, and (E3) evaluation. The resulting value for the indicator is a weighted average (step 3). It will include a comparison with past results to identify if improvement occurred (E1), with the expected result according to the stakeholders' initial plan (E2), and a critical analysis performed by the agile team (E3). The reasons for grading need to be recorded, and a database of metrics, or "metadata," about the assessment and the decisions taken needs to be created.

These steps aim to sketch a pattern for action to encourage a reference framework for a metrics program (Oza & Korkala, 2012). Nevertheless, "it is not sufficient to merely collect all possible metrics, but driving the culture of continuous measurement is imperative" (Oza & Korkala, 2012). Socially-constructed metrics integrate project evidence or a comparison with predetermined goals while adding context-bound interpretation, critical reflection, and evaluation provided by the stakeholders. The project participants decide the weights of each dimension (W1 to W3). Each weight can vary from 0 to 1, and the three weights total (W1+W2+W3) must be calculated.

Fig. 4 presents an example of grading an indicator of "customer complaints" during an ASD project.

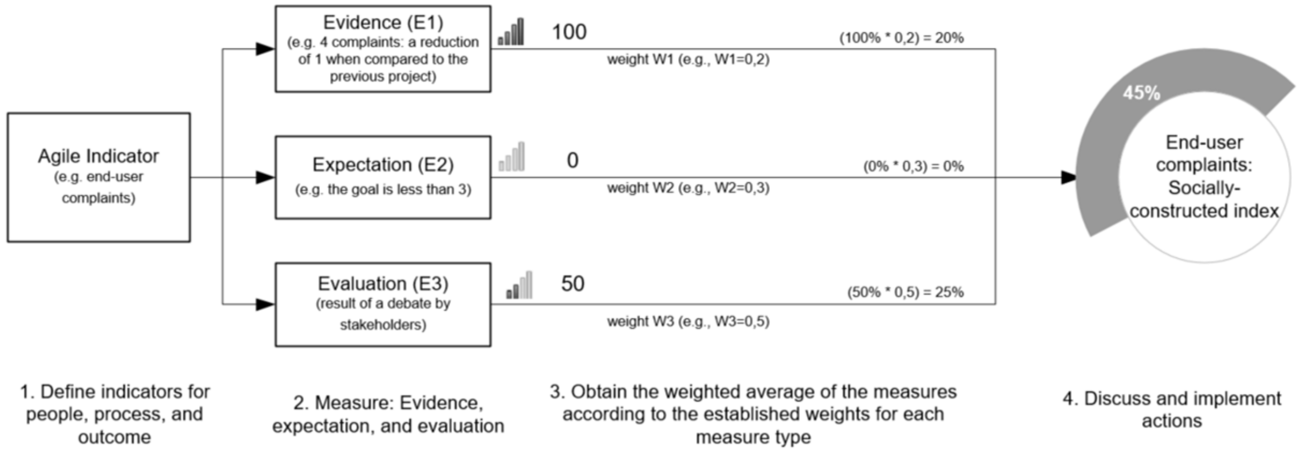


Fig. 4. Example of an adaptive metric for customer complaints

In the case we illustrate in Fig. 4, the team considered that complaints are highly variable according to multiple factors, including the project type. Therefore, they decided that for the weights W1 (evidence) = 0.2; W2 (expectation) = 0.3; and W3 (evaluation) = 0.5, the result would be evidence (100%*0.2) + expectation (0%*0.3) + evaluation (50%*0.5) = 45% as represented. We used the suggestions in Table 2 to weigh the dimensions of each selected indicator, while Table 3 describes evaluation guidelines.

Table 2. Three dimensions of adaptive metrics for ASD and guidelines for weighting

	Definition	**Potential ways to consider weightings**	**How it was calculated in Fig. 4**
Evidence	Quality is based on facts. Evidence represents the indicator’s practical improvement compared to the backlog.	The weight can be higher if the indicator is not significantly affected by uncontrolled aspects.	According to the indicator's history, if the current project has 4 complaints and the previous 5, then the indicator is 100% (improvement occurred).
Expectations	In agile development, there are goals to achieve. These include technical goals (e.g., reducing defects), social goals (e.g., improving motivation), and others.	The weight can be higher if stakeholders’ decisions mainly influence the indicator.	The team defines less than 3 complaints as a goal. In this case, it was above (4), so the expectation indicator would have the value “0%” (did not improve). We use 3 grades (success 100%, more or less 50%, unsuccessful 0%), but it could be a continuous scale.
Evaluation	Agile quality requires reflection and debate (e.g., about the meaning of the data) and the identification of lessons learned.	The weight can be higher if the indicator is not consensual or highly variable according to external factors.	The team agrees that the end-user was difficult, so 4 complaints in this case are a success (the reverse could also occur). However, they agree that improvement is necessary. They allocate 50% to evaluation.

Table 3. How to calculate the dimensions of each Indicator

	0 (regression)	50 (no improvement)	100 (clear improvement)
Evidence	Worse than the last measurement	Similar to the previous result	Better than the previous measurement
Expectations	Below expectations	Within expectations	Better than expected
Evaluation	Negative opinion	Neutral opinion	Positive opinion

Combining the three forms of grading that need to be justified by the team, we reach a balanced index that includes evidence comparing past results of similar projects, conformance to expectations (goals), and ensures debate and critical evaluation of the results. Moreover, it is participative, aiming to build adaptive metrics that can be tailored in the early phases of ASD. These are helpful during project execution and provide valuable insights for retrospectives.

5.3. Results in CAR#1

The 3View framework introduced in Section 5.2.1 was initially tested during this cycle. Therefore, our priorities were designing the weighting reference model, building and creating tools to adopt it in practice, and evaluating its applicability. Figs. 5-7 present the tool developed for adopting 3View in practice.

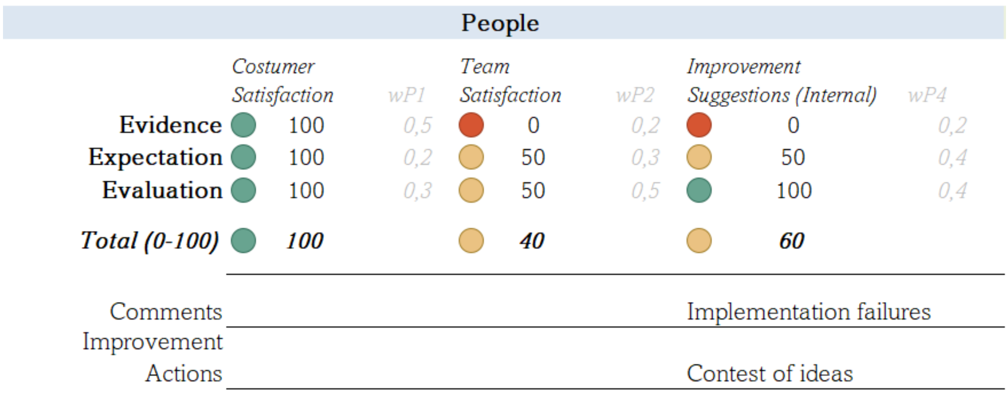


Fig. 5. Agile metrics – template for people (corrective actions omitted for the sake of simplicity)

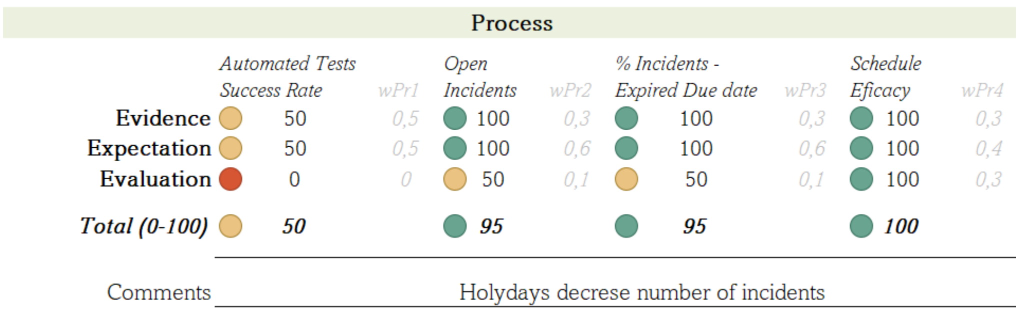


Fig. 6. Agile metrics – template for process

Outcome								
	Implemented Features	*wO1*	*Failed Features*	*wO2*	*Critical defects sent by customer*	*wO3*	*% Improvement Features*	*wO4*
Evidence	↑ 50	*0,2*	↓ 50	*0,2*	↑ 100	*0,3*	↑ 100	*0,5*
Expectation	↑ 50	*0,6*	↓ 50	*0,4*	↑ 100	*0,5*	→ 50	*0,3*
Evaluation	↑ 50	*0,2*	↑ 100	*0,4*	↓ 50	*0,2*	→ 50	*0,2*
Total (0-100)	! ***50***		✓ ***70***		✓ ***90***		✓ ***75***	
Comments			Lacking complete information		Few updates sent to customers		Holydays decrese requests (increase %	

Fig. 7. Agile metrics – template for outcome

The initial problems appeared when we needed to select indicators for agile quality. The research team established the rule that each type of metric should have at least one indicator. Then, the team proposed weights for the dimensions of evidence, expectation, and evaluation for each one.

Figs. 5-7 include tables created to assess our agile metrics. We selected three for people (the team provided the columns: customer satisfaction, team satisfaction improvement, suggestions (internal)), four indicators for the process (Fig. 6), and another four concerning outcome (Fig. 7). The aggregated result of each adaptive metric (line Total ranging from 0 to 100) in each figure is a weighted average. For example, for "customer satisfaction" in Fig. 5 (column 1), evidence is weighted 0.5, expectation 0.2, and evaluation 0.3. We decided not to use a continuous scale to simplify evaluation. Therefore, each dimension can have a measure of 100 (clear improvement), 50 (no improvement), and 0 (regression). The project participants can comment on the results and propose actions that remain in the table if they are active. Proposing new metrics is out of the scope of our study; our proposal is a new way to evaluate them in a three-dimensional space.

On analyzing the "process" metric in Fig. 6, indicators relating to "Open Incidients" and "% Incidents - expired due date"; initially, there was a decrease in both indicators for the last week ("100" for evidence), and it was clearly below their established target ("100" for expectations). Still, the team highlighted that customer holidays usually have fewer incidents, so their number and percentage allocations are not justifiably comparable with other periods. They considered this as "normal" but not excellent, the latter of which would be interpreted if we only looked at the quantitative value compared to a pre-determined target. Outcome-related metrics (Fig. 7) are also insightful. For example, on initial inspection, "failed features" presented worrying results, but the reason attributed to this was external to the team (problems in information completeness).

Another example is "critical defects sent by customers" that improved compared to the target (expectation) and past values (evidence), but the team attributed this to a reduction in system updates. Finally, the "% of improvement features" increased compared to previous periods ("100" in evidence) while still not on target ("50" in expectations). The percentage increased because the total number of features decreased, making the number of improvements more significant in that situation.

This cycle was a starting point, but understanding the changes in different settings was necessary. The company auditors were enthusiastic about this proposal because it could provide evidence of improvement and individuals' commitment to improvement. The next cycle allowed tool testing over several months with a deeper evaluation of the internal and external effects on the organization.

5.3.1 Results in CAR#2

The artifacts created in CAR#1 were used from the beginning of our intervention in the RTD institute (CAR#2). The company selected two ASD projects. The team decided to adopt the same indicators and weighting formula of CAR#1 to simplify

2026, 14(4), e140402, DOI: 10.12821/ijispm140402

the cycle initiation in practice, and were eager to ascertain whether the selected indicators were also applicable to them. One of the projects was the implementation of mobile platforms in a food company (A), and the other (B), also mobile, in a paper manufacturing company (apps for manufacturing). The projects are summarized in Table 4.

Table 4. ASD projects

	Project A	Project B
Industry	Food	Paper
Goal	Mobile applications for document management and quality control in food production lines: mobile audit tool for IFS and BRC standard (BRC, 2015; IFS, 2012)	Mobile applications for production, maintenance, and quality data
Team	3 (6 months)	4 (12 months)
Agile practices	Scrum	Scrum

Fig. 8 presents an excerpt of the results obtained for "Project A" during four sprints.

		People			Process				Outcome			
		Costumer Satisfaction	*Team Satisfaction*	*Improvement Suggestions (Internal)*	*Automated Tests Success Rate*	*Open Incidents*	*% Incidents - Expired Due date*	*Schedule Eficacy*	*Implemented Features*	*Failed Features*	*Critical defects sent by customer*	*% Improvement Features*
S1	*Evidence*											
	Expectation	50	100	100	50	50	50	0	100	100	50	100
	Evaluation	100	100	100		0	50	50	100	100	100	50
	Total (0-100)	*80*	*100*	*100*	*50*	*40*	*50*	*20*	*100*	*100*	*65*	*80*
S2	*Evidence*	50	50	0	50	100	50	50	100	0	0	0
	Expectation	0	50	0	50	50	50	50	100	50	0	50
	Evaluation	100	100	50		50	100	50	100	0	50	0
	Total (0-100)	*55*	*75*	*20*	*50*	*65*	*55*	*50*	*100*	*20*	*10*	*15*
S3	*Evidence*	50	0	50	100	50	100	50	0	50	100	100
	Expectation	100	0	100	100	50	100	50	50	50	50	50
	Evaluation	100	50	50		0	100	50	100	50	100	50
	Total (0-100)	*75*	[illegible]	*70*	*100*	*45*	*100*	*50*	*50*	*50*	*75*	*75*
S4	*Evidence*	0	50	50	0	50	0	50	100	100	50	50
	Expectation	50	100	50	50	100	50	100	50	50	50	100
	Evaluation	50	50	100		50	100	100	100	50	50	100
	Total (0-100)	[illegible]	*65*	*70*	[illegible]	*80*	*40*	*85*	*70*	*60*	*50*	*75*

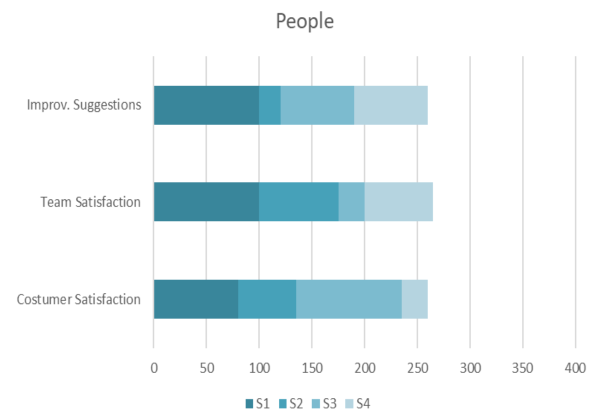
Fig. 8. 3View Applied to four sprints – project A (excerpt)

The first column in Fig. 8 presents the four sprints with the following goals: S1 - Audit Management, S2 - Audit Checklist, S3 - Actions System, and S4 - Audit Analytics. The four elements delivered in Project A are part of the mobile audit system for the food industry. The evidence grades for S1 were not considered in this case due to the lack of previous values. The other grades follow the procedure described in the earlier sections. Fig. 9 presents a sample of useful analytic reports based on the adaptive metrics approach.

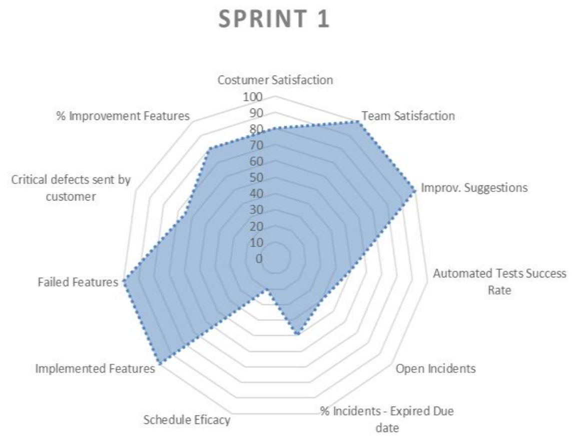
CAR#2 focused on using adaptive metrics to make social and organizational changes. One of the researchers participated regularly in the project meetings to evaluate the debate and understand how the metrics evolved in each sprint. The team considered the tools simple to use, and the results were generally positive. According to a food auditor in Project A, the tables provide a guide to interpreting raw data for improvement, and (...) the metadata is a possible way to assess the employee's commitment to quality principles. Also, the developers considered the framework accessible in their projects

2026, 14(4), e140402, DOI: 10.12821/ijispm140402

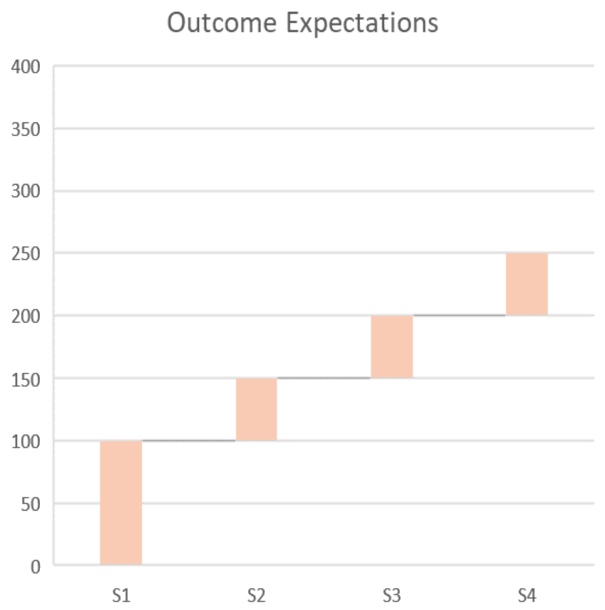
– it does not increase bureaucracy and is flexible to change. The major shortcoming, as expected, was the necessity to have metadata generated during the artifacts' evaluation. However, the research team, the scrum master, and the customer considered this an advantage to improve project documentation. During this cycle, we also asked the customers in both projects to fill in the tables presented in Fig. 8, allowing for contrasting perspectives (only accessing a part of the indicators) with the team. A deeper discussion of these findings is presented in Section 6.



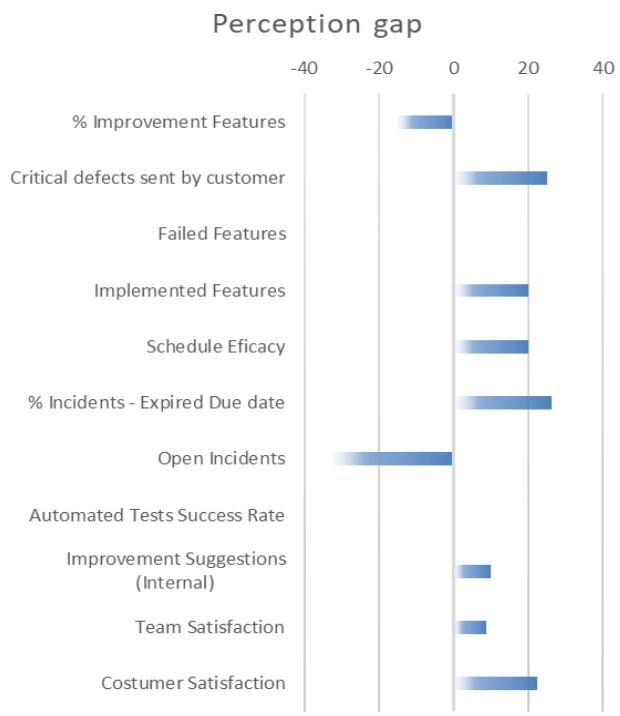
Agile quality – People dimension evolution in four sprints (S1-S4, maximum 400)



Radar chart of the results in S1 (totals)



Evolution of outcome expectations in four sprints – burnup chart for a maximum of 400 (100 for each sprint)



Average difference between grades of evaluation and total in four sprints (the team evaluated higher in most cases)

Fig. 9. 3View analytics (sample)

5.4. Evaluating the two CAR cycles

This section is structured according to the principles presented by Davison et al. (2004):

- Principle of the Researcher–Client Agreement

Researchers and practitioners agreed that CAR was an appropriate approach to studying socially constructed metrics in practice. The participating organizations explicitly committed to the project and adopted our proposed solutions within their teams. Their objective was to improve quality assessment using meaningful metrics that they could apply to their project. Interviews, observation, and document collection were used.

- Principle of the Cyclical Process Model

This research followed the five stages of CAR (Susman & Evered, 1978). We created our frame of reference with a literature review and semi-structured interviews that were discussed at top-tiered information systems conferences (Barata & Coyle, 2016; Coyle & Barata, 2016). Next, we made a diagnosis of the situation in each company. The researchers and the quality manager developed an action plan and evaluated it according to CAR principles. The framework outlined in CAR#1 served as the starting point for CAR#2.

- Principle of Theory

Our theoretical frame of reference included studies in agile metrics and models for quality improvement. The academic contribution of this work is an agile framework for using metrics that introduce flexibility in their selection and weightings (Li et al., 2017). Moreover, we follow a human-centric approach to metric adaptation, fostering continuous interaction and debate to accommodate variable factors.

- Principle of Change through Action

First, we created an innovative way of using and calculating metrics, including self-evaluation within a metric structure. We created artifacts and promoted routines to guide the development team and other stakeholders. The organizational situation was evaluated before, during, and after the intervention, ensuring the research team documented and evaluated change. Changes occurred internally with the use of the 3View artifacts and externally when we asked the customer to complete the tables at the end of each sprint. The customer found the proposal interesting to give feedback to the team and allow them to evaluate the possible interpretation gaps. In particular, customer satisfaction, schedule efficacy, and critical defects were the customer-included comments regarding the interpretation of metrics.

- Principle of Learning through Reflection

This principle was applied in different CAR stages. Joint reflection ensured that our results would be relevant, helped to improve the client's situation, and ensured project success. Progress reports were provided to the participating organizations. We learned about the benefits and the challenges of our framework, emerging from critical analysis and adaptive agile metrics that required justification. New questions were also addressed in future work, more restricted to each organization's strategy. For example, "Should we create rules to enforce corrective actions below value X? Should we enforce explanations if the evaluation differs from the other two dimensions (e.g., evaluation 0 when the other dimensions receive 100)?"

6. Discussion

We can conclude that traditional singular perspectives on quality are neither sufficient nor practical for ASD. We found that it is necessary to obtain synergies from *evidence* that emerges from ASD practices, *expectations* expressed in the form of goals, and *evaluation* resulting from a critical reflection. Combining different perspectives can promote sensemaking (Weick et al., 2005) and a debate about improvement actions supported by more meaningful metrics. Our work extends recent research focusing on the methods and tools of quality evaluation in ASD (Fatima et al., 2024; Pfeiffer & Aaen, 2024), studying the co-creation of metrics that include a subjective component in their structure and measurement process.

The first cycle, aiming at unraveling socially constructed metrics for agile, has encountered some challenges. First, regarding selecting measures for people, process, and outcome, the team looked across the literature and within the organization. For simplicity, we reduced the number of indicators to four for each metric type and selected indicators relevant to the project priorities. However, the set of indicators can be adapted according to each company and project's characteristics, which are aligned with the needs of ASD. Allocating weight to each dimension of the indicator and its grading was not straightforward. In this case, the project managers selected the weights. Still, it is possible to use a more inclusive approach, like a workshop, to define the indicators and weights according to the guidelines presented in Tables 1 and 2. Moreover, the first CAR cycle allowed us to identify a possible solution for one of the main challenges put forward by Ram et al. (2018), namely, when “a company is uncertain about that metric’s potential in providing actionable inputs”. 3View actionable inputs emerge from the deeper evaluation of the three metric components, two based on evidence, and the other obtained by subjective evaluation.

Allocating values to each indicator later opened up discussions that were considered invaluable to the teams. We concluded that the weighted value of the indicator (60) in Fig. 5 is the least important compared to the debate that included the search for solutions and open communication among team members and management. This conclusion is aligned with the importance of commitment (Alami et al., 2022; Machuca-Villegas et al., 2021). The participants also felt more empowered, taking part in the decisions based on project evidence (Karhapaa et al., 2024), and were flexible enough to adapt to the context. Additionally, the practitioners in CAR#2 stated that one of the most important aspects of 3View is to offer an opportunity to share different viewpoints and conflicting perspectives on the development performance.

We address additional key challenges for ASD metrics in the second CAR cycle, namely, the availability of data sources and process changes (Ram et al., 2018), but also to make templates useful for practice (Durbin & Niederman, 2021). It includes (1) how agile teams define the relative importance of each dimension of the metric; (2) the potential advantages of our adaptive metrics when compared with “traditional” metrics; (3) the potential conflicts that can emerge in the debate and construction of metrics, and how to solve them; (4) the potential difficulties of metrics that require metadata – because there is a combination of evidence, expectations and critical evaluation of stakeholders; and finally, (5) the benefits of socially-constructed metrics for improving communication between different stakeholders, for example, in daily meetings, retrospectives or quality audits.

The main advantages of adopting the framework in CAR#2 were (1) to guide the regular meetings during project development, (2) to provide concrete guidance to interpret quantitative measures, (3) to produce metadata to the selected metrics (e.g., justification of evaluations), (4) assist in the interaction with customers, offering a detailed description of each sprint, (5) identify preventive actions to discuss with the client, and (6) improve ASD reporting, minimizing the effort. The latter is particularly relevant when studies confirm reporting as one of the critical challenges in ASD (Schüll et al., 2023). Curiously, one of the advantages of 3View framework is also its main difficulty, namely, metrics metadata. Adaptive metrics, as we propose, are inherently subjective and specific to the project context. Although it is possible to identify trends (e.g., positive evolution of evidence, expectations, or evaluations) and global assessment (e.g., if most metrics have satisfactory results), the rationale behind each evaluation is not clear from the data alone. Each number (e.g., 50 or 100) has a particular meaning (e.g., comparing the expected goal for that metric and the actual result). The interpretation must contrast the metric metadata and the results.

The customers of projects A and B (CAR#2) found the approach interesting and complementary to specific approaches to quality (e.g., regulations and standards adoption). For example, the test documentation is critical. Still, the 3View framework lets the customer know the test compliance and the actions produced in each case. Therefore, it is possible to trace project decisions at each point (which may be necessary for auditing purposes), the evidence for measuring quality (Karhapaa et al., 2024), and the continuous improvement effort. The RTD institute that participated in CAR#2 considered that not all the indicators should be shared with the customer. So, they selected the indicators with more potential to "extract" customer perception of the sprint and contrast with the developer's perspective, promoting debate. The justification is that they would feel more comfortable identifying specific weaknesses (e.g., team satisfaction or failed features) and keeping them private to the team while sharing their vision of the project with the customer. For example, the team faced a problem in sprint S2 when the most experienced developer was off duty, creating difficulties for some development tasks. This problem was shared with the customer, and the necessary adjustments were made in sprint 3, prioritizing improvements and unexpected requests.

The 3View framework can be used as a guide to interpret raw data, fostering interaction, customer collaboration, and change management. The 3View reference model proposes three types of metrics but does not prescribe specific indicators for each one. Similarly, project stakeholders should agree on weighting the three dimensions (evidence, expectations, and evaluation). Although not solving the complex interplay of metrics and actions (Bouwers et al., 2012), 3View can reduce the bias of metrics that do not adhere to the dynamics of ASD and the undesirable effect of "obtaining what is measured".

7. Conclusions

This paper proposes the 3View framework to create adaptive metrics in ASD. In our proposal, various stakeholders' perceptions, perspectives, and priorities influence the meaning of quality, which can change from one release (or project) to the next. There are several metrics available in the form of quality goals or quantitative indicators emerging from ASD practices; however, (1) few metrics are related to people, (2) they do not comprehensively provide contextualized information within agile projects, and (3) do not result from a reflection made by project teams, which is essential in the context of both being 'agile' and continuously striving for quality.

Three main types of adaptive metrics are crucial in ASD: people-related, process-related, and outcome-related. We suggest that a small set of indicators should be used; however, in compliance with adaptive project management, companies should allow these to evolve over time. Moreover, agile metrics must include three interrelated dimensions: evidence from practice, stakeholder expectations, and stakeholder evaluation. Our dashboards include accessible indicators and allocation rules for improvement and critical analysis. This combination, however, needs to be cautiously evaluated for inaccuracies or bias, particularly for teams during the "storming" and "norming" stages, as indicated in our findings.

7.1. Implications

Our theoretical contribution is aligned with the tenets of action research. It includes new artifacts, lessons learned, and changes observed in the client-system infrastructure that justify a flexible approach to using metrics in agile. The literature is fertile with ASD metrics that are useful in different software development phases. These metrics focus on various aspects, including people, process, and outcome. However, the dynamic context of ASD may lead to different interpretations of metrics. For example, comparison with past results may reveal excellent results when past project performance was poor, and comparison with project goals may also provide a false sense of success if the goals were not well established. People's interpretation of each metric, or what can be called its social construction (Berger & Luckmann,

1991) allows to make a reflection about the current context of the ASD project, promotes a sense of fairness in metrics analysis, and puts metrics in the center of ASD project meetings.

The popular notion that "you get what you measure" is true in software development projects, because metrics steer people and must be compared with the past results and project goals (Bouwers et al., 2012). We could not agree more, but it is also well known that "culture eats strategy for breakfast", as stated by Peter Drucker. Teams can be tempted to change the system as a way to improve the metric value (Bouwers et al., 2012). Based on our findings, we would add that metrics in ASD must be interpreted, socially constructed, and used to communicate a shared vision of project performance. We argue that the traditional view of imposed metrics "is in contrast to most accounts of agile which involve voluntary, bottom up adoption on small co-located teams developing systems deemed to be suitable for agile development" (Conboy et al., 2011).

Several implications for practice can be summarized. First, accessible artifacts are proposed to guide practitioners in using meaningful metrics, namely, how to weight them and the steps to do it within the ASD project lifecycle. Second, the proposed framework can integrate any quantitative or qualitative metric in different facets of project management, being metrics-agnostic. Third, discrepancies in the three pillars of 3View (comparability with past projects, project expectations, and adaptability to dynamic conditions) are not a problem, but an opportunity to gain a clearer understanding of each metric's meaning and its value in informing improvement actions.

7.2. Limitations

Several limitations must be stated. First, the selection of databases and keywords for the literature review. Our work has evolved over several years, and we can evaluate the changes over time, but the ASD literature is constantly evolving. Second, action research allows the transferability of the findings, but only two complete CAR cycles were conducted. Third, social interventions are complex, and there are risks of the Hawthorne effect, suggesting that the observed participants' behavior and feedback could be affected by the attention they receive (French, 1950). To minimize threats to the validity of our study, we followed existing guidelines for literature reviews (Webster & Watson, 2002), elaborated an interview protocol (McLellan et al., 2003), and compared the interviews with other sources in the literature. Two authors of the paper proceeded in parallel during the entire research project, contrasting different data sources and constantly challenging the results at each stage of the research process.

The practitioners consider this model an improvement in team interaction. However, some limitations are inherent in our use of adaptive metrics. First, it is necessary to see the values of the three dimensions to understand the result. Second, it is a contextualized evaluation and cannot be used to compare different companies, although it may be used to compare various in-house projects. Third, it includes a subjective part of the evaluation that makes the value representative of the team's reality.

7.3. Future Work

Several opportunities for future research are identified. First, the approach could be tested in different development contexts and software development methodologies. Second, it was interesting to identify profiles of users based on historical data, for example, evaluating optimistic/pessimistic views on the project and understanding the motivations and interpretations for the subjective evaluation components. Third, it is possible to contrast the impact of adaptive metrics (or each of its dimensions) on project success. Fourth, although we present a way of calculating metrics, we do not propose specific indicators for each dimension. Suggesting a list of indicators for each case and recommendations for weighting in different project typologies is also possible. More studies are necessary to understand which indicators affect motivation, compliance with project schedules, team interaction, and client communication. Fifth, the metrics metadata can be used to create a knowledge base for training and decision support systems for ASD. For example, they identify the actions with the most effect in each metric improvement. Sixth, the customer of project B (CAR#2) suggested using the

3View framework for employee reward and recognition. According to them, it has the potential to merge direct measures with the feedback from the development team, potentially improving the accuracy of rewards and a better understanding of the reward selection, according to evidence (comparison with the past), expectation (compare to the goals for each employee) and evaluation (balancing contextual factors and reflection about practices).

Quality procedures are challenging to incorporate into the daily practice of ASD. Still, we hope this work may inspire future contributions to improve the situation and build a socially constructed vision of quality, perhaps less inspective and prescriptive but certainly more agile.

Acknowledgments

This work is financed through national funds by FCT - Fundação para a Ciência e a Tecnologia, I.P., in the framework of the Project UIDB/00326/2025 and UIDP/00326/2025. Thank you to Ali Karabey, Erdinc Akkaya, and Iftrue for their guidance in developer experience studies.

References

Agarwal, A., Garg, N. K., & Jain, A. (2014). Quality assurance for Product development using Agile. International Conference on Reliability Optimization and Information Technology (ICROIT), 44–47.

Agarwal, N., & Rathod, U. (2006). Defining 'success' for software projects: An exploratory revelation. International Journal of Project Management, 24(4), 358–370. https://doi.org/10.1016/j.ijproman.2005.11.009

Alami, A., Krancher, O., & Paasivaara, M. (2022). The journey to technical excellence in agile software development. Information and Software Technology, 150, 106959. https://doi.org/10.1016/j.infsof.2022.106959

Almeida, F., & Carneiro, P. (2023). Perceived Importance of Metrics for Agile Scrum Environments. Information, 14(6), 327. https://doi.org/10.3390/info14060327

Alsaqaf, W., Daneva, M., & Wieringa, R. (2019). Quality requirements challenges in the context of large-scale distributed agile: An empirical study. Information and Software Technology, 110, 39–55. https://doi.org/10.1016/j.infsof.2019.01.009

Babb, J., Hoda, R., & Nørbjerg, J. (2014). Embedding Reflection and Learning into Agile Software Development. IEEE Software, 31(4), 51–57. https://doi.org/10.1109/MS.2014.54

Bansiya, J., & Davis, C. G. (2002). A hierarchical model for object-oriented design quality assessment. IEEE Transactions on Software Engineering, 28(1), 4–17. https://doi.org/10.1109/32.979986

Barata, J., & Coyle, S. (2016). Developing Socially-Constructed Quality Metrics in Agile: A Multi-Faceted Perspective. ICIS - International Conference on Information Systems, 1–12. http://aisel.aisnet.org/icis2016/ManagingIS/Presentations/5/

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. Interacting with Computers, 23(1), 4–17. https://doi.org/10.1016/j.intcom.2010.07.003

Beck, K., Beedle, M., Bennekum, A. van, Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). Manifesto for Agile Software Development. In http://agilemanifesto.org/.

Begier, B. (2010). Users' involvement may help respect social and ethical values and improve software quality. Information Systems Frontiers, 12(4), 389–397. https://doi.org/10.1007/s10796-009-9202-z

Berger, P. L., & Luckmann, T. (1991). The social construction of reality: A treatise in the sociology of knowledge. Penguin Books.

Biazzo, S. (2005). The new ISO 9001 and the problem of ceremonial conformity: How have audit methods evolved? Total Quality Management & Business Excellence, 16(3), 381–399. https://doi.org/10.1080/14783360500054145

Boell, S. K., & Cecez-Kecmanovic, D. (2014). A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches. Communications of the Association for Information Systems, 34(1), 257–286. https://doi.org/10.17705/1CAIS.03412

Boerman, M. P., Lubsen, Z., Tamburri, D. A., & Visser, J. (2015). Measuring and Monitoring Agile Development Status. 6th International Workshop on Emerging Trends in Software Metrics (WETSoM), 54–62.

Boon, G. C., Stettina, C. J., Visser, J., & El-Baz, Y. (2023). Beyond Dashboards: Operationalising a Measurement Framework for Agile Teams. International Conference on the Quality of Information and Communications Technology. https://link.springer.com/10.1007/978-3-031-43703-8_10

Bouwers, E., Visser, J., & van Deursen, A. (2012). Getting what you measure. Communications of the ACM, 55(7), 54–59. https://doi.org/10.1145/2209249.2209266

BRC. (2015). BRC Global Standard (Issue 7). British Retail Consortium.

CMMI. (2010). CMMI for Development, Version 1.3. In Carnegie Mellon University (Issue November). https://doi.org/CMU/SEI-2010-TR-033 ESC-TR-2010-033

Colakoglu, F. N., Yazici, A., & Mishra, A. (2021). Software Product Quality Metrics: A Systematic Mapping Study. IEEE Access, 9, 44647–44670. https://doi.org/10.1109/ACCESS.2021.3054730

Conboy, K., Coyle, S., Wang, X., & Pikkarainen, M. (2011). People over Process: Key Challenges in Agile Development. IEEE Software, 28(4), 48–57.

Concas, G., Marchesi, M., Destefanis, G., & Tonelli, R. (2012). An empirical study of software metrics for assessing the phases of an agile project. International Journal of Software Engineering and Knowledge Engineering, 22(04), 525–548. https://doi.org/10.1142/S0218194012500131

Coyle, S., & Barata, J. (2016). Socially-Constructed Metrics for Agile Quality: An Action-Research Study. ACIS - Australasian Conference on Information Systems, 1–11.

Davison, R., Martinsons, M. G., & Kock, N. (2004). Principles of canonical action research. Information Systems Journal, 14(1), 65–86. https://doi.org/10.1111/j.1365-2575.2004.00162.x

di Bella, E., Fronza, I., Phaphoom, N., Sillitti, A., Succi, G., & Vlasenko, J. (2013). Pair Programming and Software Defects--A Large, Industrial Case Study. IEEE Transactions on Software Engineering, 39(7), 930–953. https://doi.org/10.1109/TSE.2012.68

Dorairaj, S., Noble, J., & Malik, P. (2012). Understanding Team Dynamics in Distributed Agile Software Development. In Lecture Notes in Business Information Processing (Vol. 111, pp. 47–61). https://doi.org/10.1007/978-3-642-30350-0_4

Durbin, M., & Niederman, F. (2021). Bringing templates to life: Overcoming obstacles to the organizational implementation of Agile methods. International Journal of Information Systems and Project Management, 9(3), 5–18. https://doi.org/10.12821/ijispm090301

Dybå, T., & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. Information and Software Technology, 50(9–10), 833–859. https://doi.org/10.1016/j.infsof.2008.01.006

Eden, C., & Huxham, C. (1996). Action Research for Management Research. British Journal of Management, 7(1), 75–86. https://doi.org/10.1111/j.1467-8551.1996.tb00107.x

Fatima, I., Funke, M., & Lago, P. (2024). Providing Guidance to Software Practitioners: A Framework for Creating KPIs. IEEE Software, 1–9. https://doi.org/10.1109/MS.2024.3456446

Fernández-Sanz, L., & Misra, S. (2011). Influence of Human Factors in Software Quality and Productivity. In Lecture Notes in Computer Science (pp. 257–269). https://doi.org/10.1007/978-3-642-21934-4_22

French, J. R. P. (1950). Field Experiments: Changing Group Productivity. In J. G. Miller (Ed.), Experiments in Social Process: A Symposium on Social Psychology (pp. 81–96). McGraw-Hill.

Ghobadi, S., & Mathiassen, L. (2016). Perceived barriers to effective knowledge sharing in agile software teams. Information Systems Journal, 26(2), 95–125. https://doi.org/10.1111/isj.12053

Gren, L., Torkar, R., & Feldt, R. (2015). The prospects of a quantitative measurement of agility: A validation study on an agile maturity model. Journal of Systems and Software, 107(2015), 38–49. https://doi.org/10.1016/j.jss.2015.05.008

Gruschwitz, S., & Schlosser, F. (2012). Towards an integrated model for managing product and process quality in agile software projects. International Research Workshop on IT Project Management, 147–155.

Hayes, W., Miller, S., Lapham, M. A., Wrubel, E., & Chick, T. (2014). Agile Metrics: Progress Monitoring of Agile Contractors. CMU/SEI-2013-TN-029.

Heston, K. M., & Phifer, W. (2011). The multiple quality models paradox: how much 'best practice' is just enough? Journal of Software Maintenance and Evolution: Research and Practice, 23(8), 517–531. https://doi.org/10.1002/smr.481

Hevner, A., & Chatterjee, S. (2010). Design Science Research in Information Systems. In Information Systems (pp. 9–22). https://doi.org/10.1007/978-1-4419-5653-8_2

Hodgson, D., & Briand, L. (2013). Controlling the uncontrollable: 'Agile' teams and illusions of autonomy in creative work. Work, Employment and Society, 27(2), 308–325. https://doi.org/10.1177/0950017012460315

Hongying, G., & Cheng, Y. (2011). A customizable agile software Quality Assurance model. The 5th International Conference on New Trends in Information Science and Service Science, 382–387.

Hult, M., & Lennung, S.-Å. (1980). Towards a definition of action research: a note and bibliography. Journal of Management Studies, 17(2), 241–250. https://doi.org/10.1111/j.1467-6486.1980.tb00087.x

Hummel, M., Rosenkranz, C., & Holten, R. (2015). The Role of Social Agile Practices for Direct and Indirect Communication in Information Systems Development Teams. Communications of the Association for Information Systems, 36(2015), 273–300. https://doi.org/10.17705/1CAIS.03615

Ibrahim, W. M., Bettenburg, N., Shihab, E., Adams, B., & Hassan, A. E. (2010). Should I contribute to this discussion? 2010 7th IEEE Working Conference on Mining Software Repositories (MSR 2010), 181–190.

Iden, J. (2012). Investigating process management in firms with quality systems: a multi-case study. Business Process Management Journal, 18(1), 104–121. https://doi.org/10.1108/14637151211215037

IEEE. (1998). IEEE Standard for a Software Quality Metrics Methodology. Institute of Electrical and Electronics Engineers.

IFS. (2012). IFS Food (version 6) - Standard for auditing quality and food safety of food products. IFS Management GmbH.

Jain, P., Ahuja, L., & Sharma, A. (2016). Current state of the research in agile quality development. 3rd International Conference on Computing for Sustainable Global Development, 1177–1179.

Jamieson, M. J., & Fallah, M. H. (2012). Agile Quality Management Techniques. Software Quality Professional, 14(2), 12–21.

Janus, A., Dumke, R., Schmietendorf, A., & Jager, J. (2012). The 3C approach for Agile Quality Assurance. 3rd International Workshop on Emerging Trends in Software Metrics, 9–13.

Kaner, C., & Bond, W. P. (2004). Software Engineering Metrics: What Do They Measure and How Do We Know? 10th International Software Metrics Symposium, 1–12.

Karhapaa, P., Behutiye, W., Seppanen, P., Rodriguez, P., Oivo, M., Franch, X., Martinez-Fernandez, S., Lopez, L., Choras, M., Bagnato, A., Aaramaa, S., & Partanen, J. (2024). Evidence-Based Quality-Aware Agile Software Development Process: Design and Evaluation. IEEE Access, 12, 86487–86512. https://doi.org/10.1109/ACCESS.2024.3414614

Kārkliņa, K., & Pirta, R. (2018). Quality metrics in Agile Software Development Projects. Information Technology and Management Science, 21, 54–59. https://doi.org/10.7250/itms-2018-0008

Kitchenham, B. (2004). Procedures for Performing Systematic Reviews. Technical Report TR/SE-0401 and NICTA Technical Report 0400011T.1. (Keele University Technical Report).

Koch, J., Drazic, I., & Schermuly, C. C. (2023). The affective, behavioural and cognitive outcomes of agile project management: A preliminary meta-analysis. Journal of Occupational and Organizational Psychology, 96(3), 678–706. https://doi.org/10.1111/joop.12429

Kropp, M., & Meier, A. (2015). Agile Success Factors A qualitative study about what makes agile projects successful. Retrieved from University of Applied Sciences and Arts Northwestern Switzerland.

Kuciapski, M., & Marcinkowski, B. (2023). Agile software development approach for “ad-hoc” IT projects. International Journal of Information Systems and Project Management, 11(4), 28–51. https://doi.org/10.12821/ijispm110402

Kupiainen, E., Mäntylä, M. V., & Itkonen, J. (2015). Using metrics in Agile and Lean Software Development – A systematic literature review of industrial studies. Information and Software Technology, 62(1), 143–163. https://doi.org/10.1016/j.infsof.2015.02.005

Li, Y., Shepherd, M., Liu, J. Y.-C., & Klein, G. (2017). Enhancing development team flexibility in IS projects. Information Technology and Management, 18(1), 83–96. https://doi.org/10.1007/s10799-016-0258-4

Machuca-Villegas, L., Gasca-Hurtado, G. P., & Muñoz, M. (2021). Measures related to social and human factors that influence productivity in software development teams. International Journal of Information Systems and Project Management, 9(3), 43–67. https://doi.org/10.12821/ijispm090303

Maddox, M., & Walker, S. (2021). Agile Software Quality Metrics. 2021 IEEE MetroCon. https://doi.org/10.1109/MetroCon54219.2021.9666049

McHugh, O., Conboy, K., & Lang, M. (2011). Using agile practices to influence motivation within IT project teams. Scandinavian Journal of Information Systems, 23(2), 59–84.

McHugh, O., Conboy, K., & Lang, M. (2012). Agile Practices: The Impact on Trust in Software Project Teams. IEEE Software, 29(3), 71–76. https://doi.org/10.1109/MS.2011.118

McKay, J., & Marshall, P. (2001). The dual imperatives of action research. Information Technology & People, 14(1), 46–59. https://doi.org/10.1108/09593840110384771

McLellan, E., MacQueen, K. M., & Neidig, J. L. (2003). Beyond the Qualitative Interview: Data Preparation and Transcription. Field Methods, 15(1), 63–84. https://doi.org/10.1177/1525822X02239573

Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. American Journal of Sociology, 83(2), 340–363. https://doi.org/10.1086/226550

Mishra, D., Balcioglu, E., & Mishra, A. (2012). Measuring project and quality aspects in agile software development. Technics Technologies Education Management, 7(1), 122–127.

Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. Information and Organization, 17(1), 2–26. https://doi.org/10.1016/j.infoandorg.2006.11.001

Natarajan, T., & Pichai, S. (2024). Behaviour-driven development and metrics framework for enhanced agile practices in scrum teams. Information and Software Technology, 170, 107435. https://doi.org/10.1016/j.infsof.2024.107435

Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. Sprouts Working Papers on Information Systems, 10(26), 1–49.

Olszewska (née Pląska), M., Heidenberg, J., Weijola, M., Mikkonen, K., & Porres, I. (2016). Quantitatively measuring a large-scale agile transformation. Journal of Systems and Software, 117, 258–273. https://doi.org/10.1016/j.jss.2016.03.029

Oza, N., & Korkala, M. (2012). Lessons Learned in Implementing Agile Software Development Metrics. UK Academy for Information Systems Conference (UKAIS).

Padmini, K. V, Dilum Bandara, H. M. N., & Perera, I. (2015). Use of software metrics in agile software development process. Moratuwa Engineering Research Conference (MERCon), 312–317.

Péraire, C., & Sedano, T. (2014). Essence reflection meetings: Field study. 18th International Conference on Evaluation and Assessment in Software Engineering (EASE), 1–4.

Pfeiffer, R. H., & Aaen, J. (2024). Tools for monitoring software quality in information systems development and maintenance: five key challenges and a design proposal. International Journal of Information Systems and Project Management, 12(1), 19–40. https://doi.org/10.12821/ijispm120102

Ram, P., Rodriguez, P., & Oivo, M. (2018). Software Process Measurement and Related Challenges in Agile Software Development: A Multiple Case Study. In Lecture Notes in Computer Science (Vol. 11271, pp. 272–287). https://doi.org/10.1007/978-3-030-03673-7_20

Rauch, N., Kuhn, E., & Friedrich, H. (2008). Index-based Process and Software Quality Control in Agile Development Projects. CompArch2008, 1–4.

Ryan, S., & O'Connor, R. V. (2013). Acquiring and sharing tacit knowledge in software development teams: An empirical study. Information and Software Technology, 55(9), 1614–1624. https://doi.org/10.1016/j.infsof.2013.02.013

Salo, O., & Abrahamsson, P. (2007). An iterative improvement process for agile software development. Software Process: Improvement and Practice, 12(1), 81–100. https://doi.org/10.1002/spip.305

Scharff, C. (2011). Guiding global software development projects using Scrum and Agile with quality assurance. Software Engineering Education and Training (CSEE&T), 274–283.

Schüll, M., Hofmann, P., Philipp, P., & Urbach, N. (2023). Reporting in large-scale agile organizations: insights and recommendations from a case study in software development. Information Systems and E-Business Management, 21(3), 571–601. https://doi.org/10.1007/s10257-023-00643-1

Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research. Information and Organization, 21(1), 1–16. https://doi.org/10.1016/j.infoandorg.2010.11.001

Sheffield, J., & Lemétayer, J. (2013). Factors associated with the software development agility of successful projects. International Journal of Project Management, 31(3), 459–472. https://doi.org/10.1016/j.ijproman.2012.09.011

Siakas, K. V., & Siakas, E. (2007). The agile professional culture: A source of agile quality. Software Process: Improvement and Practice, 12(6), 597–610. https://doi.org/10.1002/spip.344

Sidky, A., Arthur, J., & Bohner, S. (2007). A disciplined approach to adopting agile practices: the agile adoption framework. Innovations in Systems and Software Engineering, 3(3), 203–216. https://doi.org/10.1007/s11334-007-0026-z

Stålhane, T., & Hanssen, G. (2008). The Application of ISO 9001 to Agile Software Development. In Product-Focused Software Process Improvement. Lecture Notes in Computer Science 4589 (pp. 371–385). Springer.

Susman, G. I., & Evered, R. D. (1978). An Assessment of the Scientific Merits of Action Research. Administrative Science Quarterly, 23(4), 582–603.

Tarhan, A., & Yilmaz, S. G. (2014). Systematic analyses and comparison of development performance and product quality of Incremental Process and Agile Process. Information and Software Technology, 56(5), 477–494. https://doi.org/10.1016/j.infsof.2013.12.002

Torrecilla-Salinas, C. J., De Troyer, O., Escalona, M. J., & Mejías, M. (2019). A Delphi-based expert judgment method applied to the validation of a mature Agile framework for Web development projects. Information Technology and Management, 20(1), 9–40. https://doi.org/10.1007/s10799-018-0290-7

Tseng, Y.-H., & Lin, C.-T. (2011). Enhancing enterprise agility by deploying agile drivers, capabilities and providers. Information Sciences, 181(17), 3693–3708. https://doi.org/10.1016/j.ins.2011.04.034

van der Aalst, W. M. P., Becker, J., Bichler, M., Buhl, H. U., Dibbern, J., Frank, U., Hasenkamp, U., Heinzl, A., Hinz, O., Hui, K.-L., Jarke, M., Karagiannis, D., Kliewer, N., König, W., Mendling, J., Mertens, P., Rossi, M., Voss, S., Weinhardt, C., ... Zdravkovic, J. (2018). Views on the Past, Present, and Future of Business and Information Systems Engineering. Business & Information Systems Engineering, 60(6), 443–477. https://doi.org/10.1007/s12599-018-0561-1

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare the future. MIS Quarterly, 26(2), xiii–xxiii.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. Organization Science, 16(4), 409–421. https://doi.org/10.1287/orsc.1050.0133

Wiese, I. S., Côgo, F. R., Ré, R., Steinmacher, I., & Gerosa, M. A. (2014). Social metrics included in prediction models on software engineering. 10th International Conference on Predictive Models in Software Engineering (PROMISE), 72–81.

2026, 14(4), e140402, DOI: 10.12821/ijispm140402

Biographical notes

João Barata is an Assistant Professor at the Department of Informatics Engineering, Faculty of Science and Technology of the University of Coimbra, Portugal. He is also a researcher at CISUC/LASI - Centre for Informatics and Systems of the University of Coimbra and was invited to join the Polytechnic Institute of Coimbra, in 2016, as Assistant Professor of Information Systems (IS). Prior to this, João was the IS unit manager of a technological center. João has an experience of 22 years as an IS consultant and has worked with over 300 public and private organizations in Europe, America, and Africa. João holds a Ph.D. in Information Systems, participated in 23 co-funded research and innovation projects, and published over 90 papers, including leading IS conferences (ICIS, ECIS, AMCIS, PACIS) and journals such as Technovation, Business & Information Systems Engineering, Applied Soft Computing, Journal of Industrial Information Integration, and Electronic Markets. His main research interests are information systems, digital transformation, digital twins, supply chain management, and Industry 4.0/5.0.

ORCID: 0000-0002-7456-594X

Sharon Coyle is Director of Executive Education and an Assistant Professor in Digital Business at the University of Galway. She previously worked as a Senior Lecturer with the School of Information Systems and Technology Management (ISTM) at the University of New South Wales (UNSW), Sydney, and as a Lecturer at the University of Sydney Business School. She is an award-winning educator and holds a Ph.D, a Masters of Business Studies, and a Bachelor of Commerce from the University of Galway. Sharon has published and presented her research in leading international IS journals, conferences and industry forums. Prior to joining academia, Sharon gained extensive industry experience in her work as a Management Consultant for Accenture, working on a variety of high-profile projects for government bodies and financial services institutes.

ORCID: 0000-0003-1875-4719