Why Do Organizations Adopt Agile Scaling Frameworks? — A Survey of Practitioners

Abheeshta Putta
abheeshta.putta@aalto.fi
Aalto University
Espoo, Finland

Ömer Uludağ
oemer.uludag@tum.de
Technische Universität München
München, Germany

Shun-Long Hong
shunlong.hong@tum.de
Technische Universität München
München, Germany

Maria Paasivaara
maria.paasivaara@lut.fi
LUT University
Lahti, Finland

Casper Lassenius
casper.lassenius@aalto.fi
Aalto University
Espoo, Finland

ABSTRACT

Background: The benefits of agile methods in small, co-located projects have inspired their adoption in large firms and projects. Scaling frameworks, such as Large-Scale Scrum (LeSS) and the Scaled Agile Framework (SAFe), have been proposed by practitioners to scale agile to larger contexts, and become rather widely adopted in the industry. Despite the popularity of the frameworks, the knowledge on the reasons, expected benefits, and satisfaction of organizations adopting them is still limited.

Aims: This paper presents a study of practitioners who have adopted an agile scaling framework in their organization and investigates the reasons for, expected benefits of, and the satisfaction level with the adoption of the selected framework.

Method: We conducted a survey of software practitioners. We received data from 204 respondents representing ten frameworks adopted in 26 countries and located in six continents.

Results: The results show that SAFe is the most widely adopted framework among our respondents. The two most commonly mentioned reasons for adopting agile scaling frameworks are *to scale to more people* and *to remain competitive in the market*. The most common expected benefits are *improving the collaboration and dependency management between teams*. We also found some unique reasons and expected benefits for the framework adoption, such as *inculcating an agile mindset, addressing the needs of regulated environments, dissolving silos, and technical excellence*. Our findings indicated statistically significant differences for reasons, expected benefits, and satisfaction between different frameworks. Most of our respondents report that the selected framework met their expectations.

Conclusions: This paper offers the first quantitative assessment of reasons, expected benefits, and satisfaction of firms for adopting

agile scaling frameworks. Future studies comparing scaling frameworks could help firms in selecting the most suitable framework fitting their needs.

CCS CONCEPTS

• **Software and its engineering** → **Agile software development**.

KEYWORDS

Agile scaling frameworks, reasons, expected benefits, satisfaction level

ACM Reference Format:

Abheeshta Putta, Ömer Uludağ, Shun-Long Hong, Maria Paasivaara, and Casper Lassenius. 2021. Why Do Organizations Adopt Agile Scaling Frameworks? — A Survey of Practitioners. In *ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '21)*, October 11–15, 2021, Bari, Italy. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3475716.3475788>

1 INTRODUCTION

Since the creation of the Agile Manifesto in 2001, agile software development methods have become mainstream in software projects that are continuously affected by external drivers, such as changing customer demands, shifting regulatory requirements, and new technological advances [18, 28]. Since then, academics and practitioners have devoted a great deal of attention to agile methods [17]. Agile methodologies, such as Extreme Programming and Scrum, were initially created for small, co-located, and self-organizing teams developing software in close collaboration with customers applying short iterations [28]. Given the successful adoption of agile methods in small organizations and projects, many software practitioners are increasingly engaged in using these methods in large projects and organizations [16].

Some custodians of existing agile methods and software practitioners, who have worked with companies to scale agile approaches to their settings, have promoted various agile scaling frameworks, such as Large Scale Scrum (LeSS)¹ and Scaled Agile Framework (SAFe)² to address issues related to the adoption of agile methods in large companies and projects [8, 16, 58]. As large organizations

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ESEM '21, October 11–15, 2021, Bari, Italy

© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-8665-4/21/10...\$15.00
<https://doi.org/10.1145/3475716.3475788>

¹<https://less.works/>, last accessed on: 05-11-2021.

²<https://www.scaledagileframework.com/>, last accessed on: 05-11-2021.

face increasing pressure to become more agile, the adoption of agile scaling frameworks in the industry has grown, as they claim to provide off-the-shelf solutions for scaling agile methods [8, 58, 62].

Not only is there a growing interest in the adoption of agile scaling frameworks from a practical perspective [62], but there is also a growing interest from academia to study their adoption in large organizations to provide companies with valuable knowledge to support their agile adoption [58]. A systematic mapping study [58] uncovered the topic of agile scaling frameworks as a vital research stream in the large-scale agile development research field.

The current literature on agile scaling frameworks covers the following four main topics: (i) the investigation on how individual frameworks are adopted in organizations (cf. [21], [37], [40]), (ii) the qualitative comparison of agile scaling frameworks based on their characteristics (cf. [1], [53], [59]), (iii) the identification of challenges and success factors related to the adoption of agile scaling frameworks in various organizations (cf. [9]), and (iv) the quantitative investigation related to the benefits and challenges of SAFe adoption in organizations (cf. [31], [45], [47]).

Even though there is a growing body of knowledge on agile scaling frameworks, the existing literature neglects to investigate why companies select a specific framework, what benefits they expect, and how satisfied they are with their selected framework. We start filling this gap in the literature by presenting our findings on the reasons, expected benefits, and satisfaction of organizations adopting agile scaling frameworks based on a survey of industry participants worldwide. We aim to answer the following three research questions (RQs):

- *RQ1: What are the reasons organizations adopt agile scaling frameworks?*
- *RQ2: What are the expected benefits of adopting agile scaling frameworks?*
- *RQ3: How satisfied are software practitioners with the agile scaling framework their organization adopted?*

The remainder of this paper is structured as follows. In Section 2, we describe the background and related work. In Section 3, we portray the procedure of our survey. In Section 4, we present our results, provide a discussion of our findings, and discuss the limitations of our study. We conclude our paper with a summary of the results and remarks on future research in Section 5.

2 BACKGROUND AND RELATED WORK

The successful adoption of agile methods in small teams inspired large organizations to start adopting agile methods in large projects, even beyond software development, across the enterprise [36]. This phenomenon is often denoted as “*large-scale agile development*” in the literature [15]. Similar to previous studies (cf. [14], [20]), we understand the term ‘*large-scale agile development*’, as the application of agile methods in large multi-team settings consisting of 50 persons or more, or at least six teams. The application of agile methods for larger projects or larger companies involves two unprecedented challenges. First, additional complexity factors, such as a large number of teams, geographical distribution, or formal governance structures, challenge the optimal conditions, i.e., small and co-located teams of less than 50 persons with easy access to the user and business experts and that develop non-life-critical

software [5], in which agile methods can be applied successfully [2]. Second, agile methods do not provide sufficient guidance on dealing with these complexity factors [33]. Motivated by this backdrop, custodians of existing agile methods and consultants that have worked with firms in tailoring agile approaches to their settings have proposed various agile scaling frameworks to address the limitations of the agile methods in large organizations and projects [1, 8, 59]. These frameworks include predefined best practices to tackle challenges related to a large number of teams, inter-team coordination, and customer involvement [1, 9].

Due to the high practical relevance of agile scaling frameworks for companies introducing agile methods at scale, researchers have increasingly begun to study the adoption of these frameworks in organizations [59]. For example, [56] compare twenty agile scaling frameworks, such as LeSS, Nexus, and SAFe, and assess their maturity levels based on the number of case studies reporting them as well as available documentation, training courses or certifications. [1] compares agile scaling frameworks, such as LeSS, Nexus, and Spotify, based on their characteristics. [53] provides a subway map that visualizes which agile practices different agile scaling frameworks are supporting. Apart from comparing multiple frameworks based on their characteristics, Conboy and Carroll [9], and Uludağ et al. [59] investigate the benefits, success factors, and challenges of adopting agile scaling frameworks. While extant literature provides a thorough description of the benefits and challenges of adopting agile scaling frameworks (cf. [45, 59]), researchers also highlight serious critics about the adoption of agile scaling frameworks. For example, Uludağ and Matthes [57] criticize that companies tend to use scaling frameworks as recipes resulting in various problems, e.g., companies become pseudo agile, people do not develop the aspired agile mindset, and companies spend a lot of money on training people in roles and practices they do not need.

Besides the previous qualitative studies of agile scaling frameworks, a few quantitative studies also exist. Laanti and Kettunen [31] surveyed companies in Finland to study the state of SAFe adoption by investigating applied SAFe practices, as well as reported benefits and challenges. Gustavsson and Bergkvist [22] surveyed three companies from the automotive, financial, and public sectors in Sweden to investigate their perceptions of the benefits and drawbacks of adopting SAFe. Salikhov et al. [48] surveyed employees of sixteen organizations to study factors that ease SAFe adoption, and the challenges organizations face during the SAFe adoption process. Finally, Putta et al. [45] surveyed the benefits and challenges of adopting SAFe with respondents from all over the world.

Although academics have devoted a considerable interest to study the adoption of agile scaling frameworks in organizations [58], to the best of our knowledge, there is no other work that quantitatively assesses the satisfaction, expected benefits, and reasons of organizations for adopting different agile scaling frameworks.

3 RESEARCH METHODOLOGY

In this section, we describe our research method, data collection process, and data analysis.

3.1 Survey Design

Questionnaire Design. The questionnaire consisted of five sections. The first section described the research goals and who should answer the survey (only practitioners were asked to respond). The second section included questions on the organization's transformation background, such as the development methods used before adopting a scaling framework. In the next three sections, we asked the participants to assess their agreement regarding their own organization on a list of reasons, expected and realized benefits, and challenges, which were consolidated based on three previous studies on agile and large-scale agile development [54, 56, 60]. Additionally, in each of the three sections, we included an open-ended question to write-in other reasons, benefits, and challenges experienced/witnessed respectively to limit the anchoring effect. In the fifth section, we captured the participants' background information, such as respondent's primary role in the organization, and the location of the organization.

Survey Validation. The questionnaire was initially carefully reviewed by the authors. Then, we consulted an academic expert on large-scale agile. Next, two survey experts assisted in checking that the questionnaire adhered to the best practices of survey research. Finally, an industry expert on large-scale agile reviewed the questionnaire and suggested a few changes, e.g., reducing the length. After incorporating all suggestions, we conducted a pilot survey with three respondents from our target audience and asked for their feedback on the questionnaire's length, understandability, and readability. After final modifications, the survey was sent to the target audience.

Sampling and Target Audience. We used non-probabilistic convenience sampling that involves "getting responses from the individuals who are willing and are available" [29]. The target audience for our survey included software professionals from various roles, e.g., managers, agile coaches, product owners, and developers, who use agile scaling frameworks in their organizations.

Data Collection. Data collection took place between May and September 2019 using the online tool "LimeSurvey"³. To reach our target population, we used the following media to promote the survey: (1) international, agile conferences, (2) agile meetup groups, (3) social media groups, and (4) personal networks.

We promoted the survey at three conferences: the 20th International Conference on Agile Software Development (XP 2019)⁴, the 14th International Conference on Global Software Engineering (ICGSE 2019)⁵, and the Agile 2019⁶.

In addition, the survey link was published on selected social media platforms, promoted at agile meetups, and distributed to software professionals. In June 2019, the survey link was posted in the worldwide LinkedIn group "Lean and Agile Software Development"⁷, which is the largest online community consisting of Lean and Agile practitioners from all over of the world. We promoted the survey in two agile meetups, where practitioners share their experiences, ideas, and knowledge on issues regarding agile software development. One was held in Helsinki, Finland, in August 2019 with 32

participants and one in Copenhagen, Denmark, in June 2019 with 30 participants. The survey leaflets, containing the link and QR code, were distributed to all meetup participants. Finally, professionals from different organizations worldwide were approached via email, LinkedIn, and other social media channels and asked to fill in the survey. By snowballing of contacts, we aimed to find new contacts, i.e., the personal networks of existing contacts were leveraged.

3.2 Data Analysis

We conducted a two-day workshop for cleaning the data, in which all the authors were involved. During the workshop, the authors removed the partially filled or incomplete responses. We received a total of 4044 responses, of which 204 were complete responses. Next, the completed responses from LimeSurvey were imported to SPSS Statistics⁸. Later on, we created variables and ran few basic statistical tests, e.g., checking the normality of data.

Descriptive Statistics. We started the data analysis by running basic descriptive statistics for contextual information, reasons, and expectations for adopting agile scaling frameworks, such as frequencies, to get an overview of the data and insights on how to proceed with inferential statistics. Then, we calculated the mean values for the reasons, expected benefits, and satisfaction.

Inferential Statistics. We conducted the *Kolmogorov-Smirnov test* [32], which showed that our data had a non-normal distribution. Thus, we adopted non-parametric tests to perform inferential statistics. We used the *Kruskal-Wallis H test* [10] to compare the differences between more than two independent groups, e.g., primary framework, industry sector, when the dependent variable is either ordinal or interval/ratio, e.g., reasons, expected benefits, and satisfaction level.

The *Kruskal-Wallis H test* will only find if there exists a statistical difference between several independent groups (e.g., SAFe, LeSS, S@S (Scrum at scale)). To find the statistically significant difference between two independent groups, we conducted a post-hoc analysis, which consists of pair-wise comparisons between the individual frameworks (e.g., SAFe and LeSS) using the *Dunn-Bonferroni test* [19]. We verified the results of the pair-wise comparisons to that of *Mann-Whitney U* [10] test, which is used to identify the statistically significant differences between two independent groups.

Qualitative Analysis. Data for two open-ended questions, *other reasons and other expected benefits*, was imported into Excel for qualitative coding. We coded the data by following the guidelines from [11]. We started by open coding, which included breaking down the data into meaningful labels. After that, we grouped the open codes into axial codes based on the similarities and differences.

4 RESULTS AND DISCUSSION

In this section, we present our results. We start by describing the demographics, followed by the answers to the research questions. We also compare our findings to the previous literature and discuss the limitations of our study.

4.1 Demographics

Roles of the Respondents. The majority, 56.4%, of our respondents had a process-related role, e.g., Scrum Master or agile coach

³<https://www.limesurvey.org/>

⁴<https://www.agilealliance.org/xp2019/>, last accessed on: 05-11-2021.

⁵<https://conf.researchr.org/home/icgse-2019>, last accessed on: 05-11-2021.

⁶<https://www.agilealliance.org/agile2019/>, last accessed on: 05-11-2021.

⁷<https://www.linkedin.com/groups/37631>, last accessed on: 05-11-2021.

⁸<https://www.ibm.com/products/spss-statistics>, last accessed on: 05-11-2021.

(see Figure 1). 25.5% were in management roles, such as project or line manager. Roles from the development team-level (10.3%) included developers and team leaders, and roles from the product-level included Product Owners and requirement engineers (6.4%).

Figure 1: Roles of the respondents

Geographic Location of Respondents. Our respondents were distributed across all continents. As shown in Figure 2, the highest number of respondents were from the USA (22.1%), followed by Germany (19.6%), and Denmark (5.9%). A probable explanation that a large part of our respondents are from the USA is that we actively recruited participants at conferences that took place in the USA (Agile 2019) and Canada (XP 2019, ICGSE 2019). The high response percentages from Germany and Denmark can be explained by a higher geographical focus during the sampling process, as three authors of this paper were located in these countries. It is worth noting that the geographical distribution of our respondents matches the one reported in the State of Agile Survey, where most responding organizations also were located in North America, and Europe [62].

Primary Adopted Framework. The highest number of our respondents ($n=100$, 49%) had adopted SAFe as their primary framework⁹. The second most popular framework was LeSS ($n=28$, 13.7%), followed by internally created methods for scaling ($n=12$, 5.9%), as shown in Figure 3. Spotify ($n=10$) and S@S ($n=10$) were represented each by 4.9% of respondents. This finding confirms the results of other surveys [30, 62], which also reported SAFe being the most popular framework.

Duration of Usage. The highest number, 41.2%, of our respondents, had started using the agile scaling framework 1–2 years before answering the survey. Only 10.8% had used the framework for more than five years, as illustrated in Figure 4. This indicates that agile scaling frameworks have become popular only in the recent past. The increase in the usage of the agile scaling frameworks has been observed only after 2014 [61].

Number of Teams. 36.8% of our respondents had 10–50 agile teams, as illustrated in Figure 5, 34.3% had 1–9 teams, and 28.9% had more than 50 teams.

Previous Development Methodology. We asked the respondents to indicate the methodology used before adopting a scaling framework as plan-driven, hybrid, or agile. 44.6% of the respondents had used plan-driven methods before adopting an agile scaling

Table 1: Previous Development Methodology

Framework	Agile	Hybrid	Plan-driven
SAFe	14.0 %	23.0 %	52.0 %
LeSS	17.9 %	7.10%	53.6%
Spotify	60.0%	10.0%	10.0%
S@S	30.0%	20.0%	40.0%
Internally created methods for scaling	33.3%	8.3%	50.0%

framework, see Figure 6. We also calculated the percentages for individual frameworks¹⁰ as shown in Table 1. While the majority of companies adopting SAFe and LeSS used them to introduce agile methods in the organization, the Spotify approach was mostly adopted in settings that already had experience with agile methods.

Industry Sector. The organizations in our sample represent different domains. As illustrated in Figure 7, the highest percentage, 33%, was from the financial sector, followed by the technology sector (13%), and the public sector (12%). The distribution is similar to [62], which reported that the majority of organizations using agile (of which 76% of respondents used agile scaling frameworks) were from the finance and technology sectors.

4.2 Reasons for Adopting Scaling Frameworks

We asked the respondents to indicate, on a five-point Likert scale (strongly agree – strongly disagree) to what degree they agreed with the statement: *The following reason was important in your decision to adopt the agile scaling framework* for the six reasons presented in Figure 8. The reason *to scale to more people* (4.07) received the highest mean value. This was followed by *to remain competitive in the market* (3.87), *because the framework is well defined and clearly documented* (3.83), and *because it is well supported by coaching, training, and guidance* (3.8). These reasons were reported in previous literature on SAFe and LeSS, e.g., scaling to more people in [12, 34] and remain competitive in [3]. Similarly, *well-defined and clearly documented*, and *well supported by training and guidance* were reported as significant reasons for adopting SAFe [24, 26, 27, 44, 51] and LeSS [55].

The open question identified 34 additional reasons, presented in Table 2. The most mentioned reason was *the framework helps in dealing with the organization's pain points/needs/current challenges*, which unfortunately does not actually explain what the pain points were and would probably hold true for any improvement action. Other common reasons include *the framework is the most popular and proven successful* and *the framework provides flexibility*. The popularity and claimed success of SAFe were reported as significant reasons in [44]. Similarly, the success of LeSS as a motivator for adoption was reported in [55]. In the literature, SAFe has been considered the least flexible and most prescriptive framework [4, 13]. Interestingly, some of our survey respondents considered its flexibility important when deciding to adopt it. Other write-in reasons with several mentions, such as *addressing complexity, de-scaling organizations, improving customer value*, and *scaling to more teams*, were reported as the key reasons behind the creation of agile scaling frameworks according to the framework founders [58].

⁹Some respondents also used more than one framework, in such cases, they were asked to choose the framework that is predominately used in their organization, which we called the primary framework.

¹⁰We took the five agile scaling frameworks with the most responses in our survey.

Figure 2: Geographical distribution of the respondents

Figure 3: Primary adopted agile scaling frameworks

Figure 4: Number of years of experience since agile scaling framework adoption

Other reasons, such as *product focus*, *light-weight process model*, and *people with experience to guide* were also mentioned for LeSS adoptions [55].

We identified some unique reasons for adopting scaling frameworks, e.g., *inculcating an agile mindset*, *addressing the needs of regulatory environments*, *fostering innovation*, *increasing attractiveness as an employer*, and *having cross-functional/feature teams*. While these

Figure 5: Number of agile teams

Figure 6: Previously used software development approaches before agile scaling framework adoption

reasons were not found in previous literature on scaling frameworks, they might reflect general goals of agile adoption that are, per se, not tied to any particular framework.

Each agile scaling framework provides organizations with various rules, principles, practices, roles and structures. Organizations are likely to select a framework based on a mapping between their needs and the attributes of the frameworks [13]. Thus, we were interested in determining whether the reasons to adopt a respective agile scaling framework differ between different frameworks.

Table 2: Write-in Adoption Reasons

Framework Specific Reasons	Framework(s)	n
Popular and proven successful	SAFe	5
Provides flexibility	SAFe, LeSS, S@S	5
Helps in de-scaling the organization	LeSS	4
Provides a light-weight process model	LeSS, Nexus	3
Provides a common language/structure	SAFe	2
Maps well for traditional organizations, e.g., provides an easier transition for managers	SAFe	2
Helps in mixing practices from other methods, e.g., Scrum/DevOps	SAFe, SoS	2
Has a focus on lean portfolio management and system thinking	SAFe	1
Well documented approaches	Nexus	1
Good materials and people with experience to guide	SAFe	1
Organizational/Business Reasons	Framework(s)	n
Helps in dealing with the organization's pain points/needs/current challenges	SAFe, LeSS, Nexus, SoS, DAD, Internal methods	8
To inculcate an agile mindset	SAFe, Spotify	5
To address growth and complexity	SAFe, Spotify, DAD, Internal methods	4
To have product focus	LeSS, CAF (Continuous Agile Framework)	3
To improve or provide customer value	SAFe, LeSS, CAF	3
To scale agile to more teams or other units	SAFe, LeSS, Spotify	3
To improve employee satisfaction/morale/empowerment	SAFe, LeSS, S@S	3
To improve productivity	SAFe	2
To address the needs of regulatory environments/compliance	SAFe	2
Adopted to support client transition	LeSS	2
Improve agile process and business agility	DAD	2
Increase profits/market share	LeSS, Nexus	2
To be more effective in managing work and teams	SoS	2
To have cross-functional/feature teams	CAF	1
To improve alignment	SAFe	1
Improve performance of the organization	SAFe	1
Foster innovation	LeSS	1
To have secured delivery	SAFe	1
To support new digital ways of working	SAFe	1
To become more flexible	LeSS	1
To generate value faster	SAFe	1
To increase attractiveness as an employer	SAFe	1
Other reasons	Framework	n
Consultants recommended the framework	Spotify	1
Association with founders of the framework	Nexus	1

Figure 7: Industry sectors of the respondents' organizations

To this end, we first calculated mean values for all reasons (see Table 3¹¹) for the individual frameworks. We observed that the means differed. Next, we conducted the *Kruskal-Wallis H test* by taking the primary adopted framework as the independent variable (SAFe, LeSS, S@S, Spotify, Internal created methods for scaling¹²) and all six reasons as the dependent variables (e.g., to remain competitive).

¹¹Lower mean values are coloured red and higher values green.

¹²We took five most frequently adopted frameworks based on our survey results, i.e., for n more than or equal to 10.

Table 3: Mean values for Adoption Reasons

Reason	SAFe	LeSS	Internal	S@S	Spotify
Widely adopted	3.69	2.5	2	2.9	3.1
Well defined and documented	4.26	3.8	2.27	3.9	2.3
Addresses architectural challenges	3.52	3.3	3.27	3.9	3
Well supported	4.12	3.7	2.55	4.6	2.7
To remain competitive	3.91	3.9	3.45	4.4	3.4
To scale to more people	4.12	4	4.18	4.5	3.7

We found statistically significant differences for the reasons: (i) *because it is widely adopted* ($p=0.000$), (ii) *because the framework is well defined and clearly documented* ($p=0.000$), (iii) *because it is well supported by coaching, training, and guidance* ($p=0.000$). We conducted post-hoc analysis to identify statistically significant differences between individual frameworks for the three reasons mentioned above, see Table 4. The table contains the mean ranks¹³ for each of the frameworks along with the frameworks pairs which had statistically significant differences for the respective reasons.

For reason, (i) *because it is widely adopted*, SAFe had the highest mean rank and internal methods the lowest, an expected result. We also identified a significant difference between SAFe and internally created methods, which shows the perceived popularity of SAFe as an important adoption reason among SAFe users. Conversely, this was not an important decision factor for organizations opting to develop their scaling method internally.

¹³The higher the mean rank, the higher the agreement among the respondents.

Figure 8: Adoption reasons arranged by mean values

Table 4: Post-hoc Analysis for Adoption Reasons

Reason for adoption	Mean Ranks	Statistically Significant Pairs
Because it is widely adopted	SAFe=88.81	Internal-Spotify (p=0.039) Internal-SAFE (p=0) LeSS-SAFE (p=0)
	LeSS = 47.38	
	Spotify=71.10	
	S@S= 62.40	
Because the framework is well defined and clearly documented	Internal= 32.70	Internal-LeSS (p=0.001) Internal-S@S (p=0.003) Internal-SAFE (p=0) Spotify-LeSS (p=0.003) Spotify-S@S (p=0.005) Spotify-SAFE (p=0) LeSS-SAFE (p=0.039)
	SAFe = 90.71	
	LeSS = 71.93	
	Spotify = 26.00	
Because it is well supported by coaching, training, and guidance	S@S = 79.45	Internal-LeSS (p=0.017) Internal-SAFE (p=0) Internal-S@S (p=0) Spotify-LeSS (p=0.06) Spotify-SAFE (0.001) Spotify-S@S (p=0) LeSS-SAFE (p=0.038) LeSS-S@S (p=0.016)
	Internal = 22.30	
	SAFe = 86.46	
	LeSS = 67.77	

SAFe [50], LeSS[6], and S@S [52] are clearly defined and document their practices on their official websites, unlike Spotify [4] and internally created methods. Thus, it is no surprise that the reason, (ii) *because the framework is well defined and clearly documented*, seems to have been a more important decision criterion for adopters of the three first mentioned frameworks compared to Spotify and internally created methods.

SAFe [49], LeSS [7], and S@S [23] are well supported with coaching and guidance in contrast to internally created methods and Spotify [4]. For instance, their official websites list many training and consulting partners to support their adoption [56]. Therefore, SAFe, LeSS, and S@S had statistically higher agreement for the reason, (iii) *because it is well supported by coaching, training, and guidance* over internally created methods and Spotify. Similar reasons related to support and guidance were mentioned for SAFe and LeSS adoption in [44] and in [55] respectively.

SAFe respondents had a statistically higher agreement for reasons (i), (ii) and (iii) when compared with LeSS respondents. SAFe is known to be more popular, highly defined, documented, and supported compared to LeSS [4, 13]. Previous literature also mentioned that SAFe was chosen for its well-defined structure [25] and popularity [44] when compared to other frameworks.

Table 5: Write-in Expected Benefits

Expected benefits	Framework(s)	n
To improve employee engagement/satisfaction	SAFe, LeSS	3
To address most common scaling challenges	LeSS	1
To be customer-centric and product focused	LeSS	1
To foster innovation	SAFe	1
To improve productivity	SAFe	1
To manage multiple requirements	SAFe	1
On-time delivery	CAF	1
To have scalable products	SAFe	1
To have technical excellence	LeSS	1
To dissolve silos	Internal	1
Better understanding of Cost of Delay, queuing theory	SAFe	1
To reduce cost	LeSS	1

4.3 Expected Adoption Benefits

We asked our respondents, on a five-point Likert scale (strongly agree – strongly disagree), to indicate their agreement with eleven expected benefits of framework adoption. Figure 9 shows the results, ordered from highest to lowest mean values. The expected benefit *to improve collaboration between teams* (4.08) received the highest mean value, followed by *to improve dependency management between teams* (3.92).

To catch expected benefits not covered by our list, we had an open question, which identified twelve additional items, presented in Table 5. These largely matched with the reasons discussed earlier. We identified a few new expected benefits, e.g., *scalable products*, *dissolve silos*, and *technical excellence*.

The expected benefits are similar to common reasons for framework adoption. For instance, managing dependencies between agile teams were mentioned as a common reason for LeSS adoption [55]. Similarly, improving collaboration was mentioned as a reason for SAFe adoption in [35]. Other expected benefits such as improve transparency, faster time to market, improve predictability were also previously found as reasons for SAFe adoption [34, 41, 51]. The aforementioned expected benefits were also mentioned as realized benefits of adopting SAFe in [43].

Furthermore, we were interested in determining if there was any difference in expected benefits based on which agile scaling framework an organization had adopted. We calculated the mean values, see Table 6. We tested for differences between the groups using the *Kruskal Wallis H* test, which indicated statistically significant differences for the following expected benefits: (i) *to increase delivery predictability* (p= 0.038), (ii) *to increase responsiveness* (p=

Figure 9: Expected benefits arranged by mean values

Table 6: Mean values for expected benefits per framework

Expected benefit	SAFe	LeSS	Internal	S@S	Spotify
To improve dependency management between teams	4.09	3.36	3.92	3.9	4.1
To improve collaboration between teams	4.14	4.36	4.08	3.7	4
To improve customer satisfaction	3.82	4.04	3.83	3.7	3.8
To improve team autonomy	3.34	4.18	3.25	3.7	4.2
To have more frequent deliveries	3.89	4.14	3	3.5	3.5
To have shorter time to market	3.9	4.25	3.5	3.6	3.5
To enable faster feedback	3.82	4.21	3.08	3.6	3.3
To increase responsiveness	3.68	4.21	3.17	3.6	3.5
To increase delivery predictability	3.76	3.86	3.83	3.6	2.3
To improve software quality	3.45	4	3.17	2.4	2.6
To improve transparency	3.84	4.46	3.33	3.7	3.5

Table 7: Post-hoc Analysis for Expected Benefits

Expected benefits	Mean Ranks	Statistically Significant Pairs
To improve team autonomy	SAFe = 72.52 LeSS = 99.23 Spotify=102.85 S@S = 98.15 Internal = 61.77	Internal-LeSS (p=0.017) Internal-Spotify (p=0.033) SAFe-LeSS (p=0.005) SAFe-Spotify (p=0.038)
To enable faster feedback	SAFe = 80.87 LeSS = 94.21 Spotify = 59.25 S@S = 86.30 Internal = 49.05	Internal-SAFe (p=0.022) Internal-LeSS (p=0.004) Spotify-LeSS (p=0.030)
To increase responsiveness	SAFe = 77.86 LeSS = 96.71 Spotify = 67.65 S@S = 89.95 Internal = 51.68	Internal-S@S (p=0.042) Internal-LeSS (p=0.003) SAFe-LeSS (p=0.041)
To increase delivery predictability	SAFe = 83.45 LeSS = 79.11 Spotify = 38.25 S@S = 88.50 Internal = 81.14	Spotify-LeSS (p=0.011) Spotify-Internal (p=0.024) Spotify-SAFe (p=0.002) Spotify-S@S (p=0.010)
To improve software quality	SAFe = 81.56 LeSS = 96.39 Spotify = 55.05 S@S = 60.90 Internal = 64.18	Spotify-LeSS (p=0.011) S@S-LeSS (p=0.029) Internal-LeSS (p=0.040)

0.032), (iii) *to improve software quality* ($p = 0.033$), (iv) *to improve team autonomy* ($p = 0.006$), and (v) *to enable faster feedback* ($p = 0.027$). Next we conducted a post-hoc analysis to find the individual differences between frameworks, see Table 7.

Spotify adopters had the highest mean rank for the item *to improve team autonomy* among all frameworks. Statistically significant differences were found between SAFe and Spotify, and Spotify and

internally created for this item. This indicates that improving autonomy is an important expected benefit for Spotify adoption, as mentioned in [46]. LeSS respondents also had a statistically higher agreement for autonomy compared to SAFe, showing that team autonomy is considered more critical to LeSS than SAFe adopters.

For the item *to enable faster feedback*, LeSS and SAFe adopters had statistically higher agreement than those who adopted internally created methods. SAFe users have mentioned this as a common benefit of SAFe usage [43]; therefore, many respondents would expect this benefit from SAFe.

For the expected benefit, *to increase responsiveness*, LeSS respondents had a statistically higher agreement over internally created methods and SAFe. This indicates that improving responsiveness is a significant expected benefit for LeSS adoption over SAFe and internally created methods. The aforesaid expected benefit was a key reason for LeSS adoption discussed in [38].

SAFe adopters had a statistically higher agreement for the item *to increase delivery predictability* than Spotify adopters. An improvement in delivery predictability is mentioned as a common benefit among SAFe users, a fact that could have motivated several respondents to think SAFe brings this benefit [43].

For the item *to improve software quality*, LeSS respondents had the highest mean rank among all the frameworks. A statistically significant difference was also found for LeSS with all the frameworks except SAFe. This shows the importance of improving quality among the LeSS respondents compared to Spotify, S@S, and internally created methods.

4.4 Satisfaction with the Frameworks

We measured the satisfaction with the selected framework by asking the respondents to indicate their agreement on a five-point Likert scale with four statements. Figure 10 shows the satisfaction level of our respondents using agile scaling frameworks arranged based on highest to lowest mean values. The satisfaction level was measured based on four different statements in our survey: (i) meeting the expectations of the organization, (ii) willingness to move back to old ways of working, (iii) willingness to shift to another framework, and (iv) willingness to recommend the framework to other similar organizations. We have reverse coded the statements (ii) and (iii). The higher mean value represents higher disagreement for (ii) and

Table 8: Satisfaction level means per framework

Satisfaction Level of the Practitioners	SAFe	LeSS	Internal	S@S	Spotify
The framework met the expectations of my organization	3.81	4.12	3.83	4.6	4
I would not* like to move back to the old ways of working	4.57	4.67	4.6	4.7	3.4
I would not* like to shift to another framework	3.63	4.17	3.4	4.5	3.5
I would like to recommend this framework to other similar organizations	3.95	4.21	3.36	4.5	2.6

*) For readability reasons, both statements have been reverse coded and negatively worded by a "not".

(iii), and converse for (i) and (iv). For statements (i) and (iv) majority of the respondents agreed, and the majority disagreed for (ii) and (iii). By analyzing the results for these four statements, we can conclude that majority of our respondents are satisfied with using their chosen scaling frameworks in their respective organizations.

We did not find any literature explicitly measuring satisfaction concerning the usage of scaling frameworks. However, we aimed to compare our results with employee satisfaction as the nearest match to our research question. We found that employee satisfaction was a popular benefit among the SAFe users [43] and Spotify users [39]. We did not get information from other frameworks, e.g., LeSS or S@S, to compare our results due to a lack of research on scaling frameworks.

Using agile scaling frameworks in large organizations gives rise to several challenges and can lead to different results based on the context and needs of the organizations [9, 13]. Thus, we were interested in determining whether the satisfaction level differed based on the respondents' choice of agile scaling framework. First, we started by calculating mean values for all the four statements, see Table 8. We observed mean values differed between different frameworks for statements that measured satisfaction level. Next, we conducted the *Kruskal-Wallis H test* by taking the primary adopted framework as the independent variable (SAFe, LeSS, S@S, Spotify, Internal) and the satisfaction level as the dependent variable (e.g., the willingness to recommend the framework to other similar organizations).

We found statistically significant differences for the following expectations: (i) *the framework met the expectations of my organization* ($p = 0.032$), (ii) *I would like to shift to another framework* ($p = 0.020$), and (iii) *I would like to recommend this framework to other similar organizations* ($p = 0.000$). Next, we did a post-hoc analysis to find out statistically significant differences between the individual frameworks. The results of the post-hoc analysis for these three statements are presented in Table 9.

For the statement *the framework met the expectations of my organization*, S@S respondents had the highest mean rank followed by LeSS. A statistical significance was found between internally created methods and S@S, and SAFe and S@S. Thus, S@S respondents had higher satisfaction when compared to the other two afore-said frameworks. S@S and LeSS are light-weight process models, which were derived directly from Scrum [63]. In contrast, SAFe is a heavy-weight process model and often criticized for its complexity, and overhead [43]. This might be one reason for higher satisfaction among the S@S and LeSS respondents compared to SAFe respondents.

For statement, *I would like to shift to another framework*, SAFe respondents had least mean rank amongst all other frameworks, and a statistical significance was found between S@S and SAFe, and LeSS and SAFe (we have reverse coded this statement in the

Table 9: Post-hoc Analysis for Framework Satisfaction

Satisfaction aspect	Mean Ranks	Statistically Significant Pairs
The framework met the expectations of my organization	SAFe = 73.69 LeSS = 86.37 Spotify = 76.94 S@S = 113.15 Internal = 65.95	SAFe-S@S ($p=0.004$) Internal-S@S ($p=0.008$)
I would like to shift to another framework ¹⁴	SAFe = 73.09 LeSS = 98.04 Spotify = 76.70 S@S = 108.75 Internal = 73.77	S@S-SAFe ($p=0.015$) LeSS-SAFe ($p=0.008$)
I would like to recommend this framework to other organizations	SAFe = 80.51 LeSS = 91.61 Spotify = 32.35 S@S = 103.50 Internal = 50.90	Spotify-SAFe ($p=0.001$) Spotify-LeSS ($p=0$) Spotify-S@S ($p=0$) Internal-SAFe ($p=0.039$) Internal-LeSS ($p=0.011$) Internal-S@S ($p=0.007$)

data). This means shifting to another framework is more likely to happen for respondents using SAFe when compared to LeSS and S@S. Literature has indicated SAFe to be the most criticized framework when compared to all other frameworks for its non-agile practices, e.g., 8-12 weeks Program Increments (PIs) in SAFe felt like moving back to plan-driven methods [45]. The aforesaid drawbacks of SAFe might be a reason for many respondents to consider shifting to another framework from SAFe.

For statement, *I would like to recommend this framework to other similar organizations* S@S respondents had the highest mean rank among all the frameworks. The respondents of S@S, SAFe, and LeSS have statistically higher agreement towards recommending their respective frameworks when compared to Spotify and internally created methods respondents.

Based on the above analysis, we conclude that respondents using Scrum derived frameworks, e.g., S@S, are more satisfied when compared to SAFe. We could also infer that SAFe is more likely to be adopted for its popularity and claimed benefits, rather than actual benefits of its adoption [43, 59]. Thus, there is more likeliness among SAFe respondents to shift to another framework. We conclude that software practitioners should always look for empirical evidence of the framework usage, e.g., benefits, rather than just getting drawn by the marketing and popularity. However, the current literature lacks empirical evidence of framework usage, and the need for research-based studies on scaling frameworks seems to be highly at the moment [42].

Our results show that our respondents' satisfaction differed based on the choice of the frameworks; deeper analysis on why they differ between frameworks is an exciting area for future research.

4.5 Threats to Validity

In this section, we describe the limitations of our study [64].

Internal Validity. This threat relates to factors that influence the relationship between the research process and the obtained results, e.g., respondent bias. We mitigated the respondent bias by collecting data from reliable sources, i.e., the majority of our responses (more than 90%) came either from people we met during the conferences and agile Meetups or via personal contacts, and we knew that they were using scaling frameworks, which helped us to avoid unreliable responses.

Figure 10: Framework satisfaction arranged by means

External Validity. This threat is concerned with the generalizability of the survey results. Due to our opportunistic sampling strategy, we are unable to identify in what way our sample might be skewed. We had respondents with different roles, using different scaling frameworks, from various domains and geographical locations. Our basic demographics are similar to other surveys in the field, giving us some degree of confidence that our sample at least does not completely misrepresent the population. Our respondents are likely biased towards positive answers through their roles, e.g., agile coaches. Further, our respondents likely represent firms and people who are active agile community participants.

Construct Validity. This threat is concerned whether the questions asked in the questionnaire represent the attributes being measured. We formulated the survey statements on the reasons, expected benefits, and satisfaction based on earlier findings in the realm of agile and large-scale agile. We had to limit the questionnaire length, and we could not include all the possible reasons or statements that measure satisfaction that we identified from the literature. We compensated for this through open questions, which we think helped get the most probable reasons and expected benefits for agile scaling framework adoption. We validated the questionnaire with a domain expert, survey experts and tested it by conducting a pilot study, which helped ensure that the questionnaire was clear and understandable to the respondents.

Conclusion Validity. This threat is concerned with the ability to draw the proper conclusion from the collected data. The survey data was mainly Likert-scale, and we conducted appropriate non-parametric tests to identify differences between independent groups. We also verified our post-hoc analysis with the Mann-Whitney U test that confirmed that we did the proper tests to determine the significant differences for individual frameworks. The qualitative codes were thoroughly discussed among all the authors, which mitigated the misinterpretation of data. We also compared the results with the existing literature for validating our results.

5 CONCLUSIONS AND FUTURE WORK

Driven by pressure to respond quickly to changing business environments, software practitioners are increasingly interested in benefiting from the advantages of agile methods by adopting agile scaling frameworks in large organizations and projects [8, 62]. Along with this, researchers show a growing scientific interest in studying the adoption of these frameworks [59].

Although there is a considerable body of knowledge on agile scaling frameworks, less research has been conducted to show the

state-of-practice of agile scaling framework adoptions employing surveys of companies that have adopted these frameworks. For this reason, we surveyed 204 software practitioners worldwide to quantitatively assess the reasons and satisfaction of organizations for adopting agile scaling frameworks.

Our findings show that the top reasons for adopting agile scaling frameworks are *scale to more people* and *to remain competitive in the market*. The top expected benefits are *to improve dependency management and collaboration between teams*. We identified 34 reasons and 12 expected benefits from the write-in question. Both the reasons and expected benefits had similarities. We also identified some unique reasons for adopting scaling frameworks, e.g., *to inculcate agile mindset*, and *to address the needs of the regulatory environments*. Moreover, some unique expected benefits were identified, e.g., *to dissolve silos*, and *technical excellence*. These were not reported in the previous literature related to scaling frameworks, however, they seem to be similar to the goals or reasons for agile adoption in general.

The inferential analysis of the five most frequently adopted frameworks, i.e., SAFe, LeSS, S@S, Spotify, and internally created methods, showed statistically significant differences between the reasons and expected benefits for adopting these frameworks. Notably, we revealed that the *popularity* and the *clear definition and documentation* of SAFe were significant adoption reasons when compared to internally created methods and Spotify. In addition, we identified that the *improvement of autonomy and quality* were important expected benefits for adopting Spotify and LeSS respectively among the survey respondents.

Our results further revealed that most of our respondents were satisfied with their current framework selections. We performed an inferential analysis to determine whether the satisfaction level differs between the choice of agile scaling frameworks. We found out that *our respondents were the most satisfied using S@S. Our respondents using SAFe showed more likeliness towards shifting to another framework*, while it is criticized among agile software development practitioners for its non-agile practices [43].

Finally, this paper leaves room for further research. We encourage researchers to conduct further studies to quantitatively assess the benefits and challenges of adopting the different agile scaling frameworks. Further studies on the impact of contextual factors such as agile maturity, complexity, or multi-product development on the suitability of different frameworks could provide organizations with a decision framework to aid in framework selection.

REFERENCES

- [1] Mashal Alqudah and Rozilawati Razali. 2016. A Review of Scaling Agile Methods in Large Software Development. *International Journal on Advanced Science, Engineering and Information Technology* 6, 6 (2016), 828–837.
- [2] Scott W Ambler. 2007. Agile software development at scale. In *IFIP Central and East European Conference on Software Engineering Techniques*. Springer, 1–12.
- [3] Arpitha Badanahatti and Sapna Pillutla. 2020. Interleaving Software Craftsmanship Practices in Medical Device Agile Development. In *Proceedings of the 13th Innovations in Software Engineering Conference on Formerly known as India Software Engineering Conference*. 1–5.
- [4] Julian M Bass. 2019. Future Trends in Agile at Scale: A Summary of the 7 th International Workshop on Large-Scale Agile Development. In *International Conference on Agile Software Development*. Springer, 75–80.
- [5] Barry Boehm. 2002. Get ready for agile methods, with care. *Computer* 35, 1 (2002), 64–69.
- [6] The LeSS Company B.V. [n.d.]. Large Scale Scrum. <https://less.works/case-studies/index.html>.
- [7] The LeSS Company B.V. [n.d.]. LeSS Framework Training and Courses. <https://less.works/courses/less-courses>. [Online; accessed 10-MAY-2021].
- [8] Noel Carroll and Kieran Conboy. 2019. Applying Normalization Process Theory to Explain Large-Scale Agile Transformations. In *Proceedings of the 14th International Research Workshop on IT Project Management*.
- [9] Kieran Conboy and Noel Carroll. 2019. Implementing Large-Scale Agile Frameworks: Challenges and Recommendations. *IEEE Software* 36, 2 (2019), 44–50.
- [10] William Jay Conover. 1998. *Practical nonparametric statistics*. Vol. 350. John Wiley & Sons.
- [11] Juliet M. Corbin and Anselm L. Strauss. 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3rd ed ed.). Sage Publications, Inc., Los Angeles, Calif.
- [12] Steph Adolph Curtis Michelson. 2019. Experience reports from agile alliance Bias From The Bottom: A Different Way to Bootstrap A SAFe Train. In *International Conference on Agile Software Development*. AgileAlliance.
- [13] Philipp Diebold, Anna Schmitt, and Sven Theobald. 2018. Scaling agile: how to select the most appropriate framework. In *Proceedings of the 19th International Conference on Agile Software Development: Companion*. 1–4.
- [14] Kim Dikert, Maria Paasivaara, and Casper Lassenius. 2016. Challenges and success factors for large-scale agile transformations: A systematic literature review. *Journal of Systems and Software* 119 (2016), 87–108. <https://doi.org/10.1016/j.jss.2016.06.013>
- [15] Torgeir Dingsøy, Tor Erlend Fægri, and Juha Itkonen. 2014. What Is Large in Large-Scale? A Taxonomy of Scale for Agile Software Development. In *Proceedings of the 15th International Conference on Product-Focused Software Process Improvement*. Springer, Cham.
- [16] Torgeir Dingsøy, Nils Brede Moe, Tor Erlend Fægri, and Eva Amdahl Seim. 2018. Exploring software development at the very large-scale: a revelatory case study and research agenda for agile method adaptation. *Empirical Software Engineering* 23, 1 (2018), 490–520.
- [17] Torgeir Dingsøy, Sridhar Nerur, VenuGopal Balijepally, and Nils Brede Moe. 2012. A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software* 85, 6 (2012), 1213–1221. Special Issue: Agile Development.
- [18] Tore Dybå and Torgeir Dingsøy. 2008. Empirical studies of agile software development: A systematic review. *Information and software technology* 50, 9-10 (2008), 833–859.
- [19] Alan C Elliott and Linda S Hynan. 2011. A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal–Wallis analysis. *Computer methods and programs in biomedicine* 102, 1 (2011), 75–80.
- [20] Christoph Fuchs and Thomas Hess. 2018. Becoming agile in the digital transformation: The process of a large-scale agile transformation. In *Proceedings of the 39th International Conference on Information Systems (ICIS)*.
- [21] Tomas Gustavsson. 2019. Dynamics of Inter-Team Coordination Routines in Large-Scale Agile Software Development. In *Proceedings of the 27th European Conference on Information Systems (ECIS)*.
- [22] Tomas Gustavsson and Linda Bergkvist. 2019. Perceived Impacts of Using the Scaled Agile Framework for Large-Scale Agile Software Development. (2019).
- [23] Scrum Inc. [n.d.]. Scrum at Scale Courses. <https://www.scrumatscale.com/scrum-at-scale-practitioner-course/>. [Online; accessed 10-APR-2021].
- [24] Scaled Agile Inc. [n.d.]. Intel Case study. <http://www.scaledagileframework.com/case-study-intel/>. [Online; accessed 06-APR-2020].
- [25] Scaled Agile Inc. 2016. Amdocs Case Study. <http://www.scaledagileframework.com/amdocs-case-study/>. [Online; accessed 06-APR-2020].
- [26] Scaled Agile Inc. 2016. HPE Case Study. <http://www.scaledagileframework.com/hpe-case-study/>. [Online; accessed 06-APR-2020].
- [27] Scaled Agile Inc. 2017. Capital One Case Study. <http://www.scaledagileframework.com/capital-one-case-study/>. [Online; accessed 06-APR-2020].
- [28] Petri Kettunen. 2007. Extending Software Project Agility with New Product Development Enterprise Agility. *Software Process: Improvement and Practice* 12, 6 (2007), 541–548.
- [29] Barbara Kitchenham and Shari Lawrence Pfleeger. 2002. Principles of survey research: part 5: populations and samples. *ACM SIGSOFT Software Engineering Notes* 27, 5 (2002), 17–20. <https://doi.org/10.1145/571681.571686>
- [30] Marco Kuhrmann, Philipp Diebold, Jürgen Münch, Paolo Tell, Vahid Garousi, Michael Felderer, Kitija Trekter, Fergal McCaffery, Oliver Linssen, Eckhart Hanser, et al. 2017. Hybrid software and system development in practice: waterfall, scrum, and beyond. In *Proceedings of the 2017 International Conference on Software and System Process*. 30–39.
- [31] Maarit Laanti and Petri Kettunen. 2019. SAFe adoptions in Finland: a survey research. In *Proceedings of the 20th International Conference on Agile Software Development: Agile Processes in Software Engineering and Extreme Programming – Workshops*. Springer, Cham, 81–87. https://doi.org/10.1007/978-3-030-30126-2_10
- [32] Hubert W Lilliefors. 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Amer. Statist. Assoc.* 62, 318 (1967), 399–402.
- [33] Chuck Maples. 2009. Enterprise Agile Transformation: The Two-Year Wall. In *Proceedings of the 2009 Agile Conference*. IEEE, 90–95.
- [34] David McMunn and Phillip Manketo. 2017. Building Strong Foundations... Underwriting Fannie Mae's Agile Transformation. In *International Conference on Agile Software Development*. AgileAlliance.
- [35] Maria Paasivaara. 2017. Adopting SAFe to Scale Agile in a Globally Distributed Organization. In *Proceedings of the 2017 IEEE 12th International Conference on Global Software Engineering*. IEEE, 36–40. <https://doi.org/10.1109/ICGSE.2017.15>
- [36] Maria Paasivaara, Benjamin Behm, Casper Lassenius, and Minna Hallikainen. 2018. Large-scale agile transformation at Ericsson: a case study. *Empirical Software Engineering* 23, 5 (2018), 2550–2596.
- [37] Maria Paasivaara, Ville Heikkilä, and Casper Lassenius. 2012. Experiences in Scaling the Product Owner Role in Large-Scale Globally Distributed Scrum. In *Proceedings of the 7th International Conference on Global Software Engineering (ICGSE)*. Institute of Electrical and Electronics Engineers (IEEE), 174–178. <https://doi.org/10.1109/ICGSE.2012.41>
- [38] Maria Paasivaara and Casper Lassenius. 2016. Scaling Scrum in a Large Globally Distributed Organization: A Case Study. In *2016 IEEE 11th International Conference on Global Software Engineering*. IEEE, 74–83. <https://doi.org/10.1109/ICGSE.2016.34>
- [39] Vaishnavi Patil, S Panicker, and M Kv. 2016. Use of agile methodology for mobile applications. *Int. J. Latest Technol. Eng. Manag. Appl. Sci.* 5, 10 (2016), 73–77.
- [40] Jan Pries-Heje and Malene M Krohn. 2017. The SAFe Way to the Agile Organization. In *Proceedings of the XP2017 Scientific Workshops*. ACM, 1–4.
- [41] Jan Pries-Heje and Malene M Krohn. 2017. The SAFe way to the agile organization. In *Proceedings of the XP2017 Scientific Workshops*. ACM, 18.
- [42] Abheeshta Putta. 2018. Scaling Agile Software Development to Large and Globally Distributed Large-Scale Organizations. In *Proceedings of the 13th International Conference on Global Software Engineering*. ACM, 141–144. <https://doi.org/10.1145/3196369.3196386>
- [43] Abheeshta Putta, Maria Paasivaara, and Casper Lassenius. 2018. Benefits and Challenges of Adopting the Scaled Agile Framework (SAFe): Preliminary Results from a Multivocal Literature Review. In *Proceedings of the 19th International Conference on Product-Focused Software Process Improvement*. Springer, Cham, 334–351. https://doi.org/10.1007/978-3-030-03673-7_24
- [44] Abheeshta Putta, Maria Paasivaara, and Casper Lassenius. 2019. How Are Agile Release Trains Formed in Practice? A Case Study in a Large Financial Corporation. In *Proceedings of the 20th International Conference on Agile Software Development: Agile Processes in Software Engineering and Extreme Programming*. Springer, Cham, 154–170. https://doi.org/10.1007/978-3-030-19034-7_10
- [45] Abheeshta Putta, Ömer Uludağ, Maria Paasivaara, and Shun-Long Hong. 2021. Benefits and Challenges of Adopting SAFe-An Empirical Survey. In *International Conference on Agile Software Development*. Springer, Cham, 172–187.
- [46] Abdallah Salameh and Julian M Bass. 2019. Spotify tailoring for promoting effectiveness in cross-functional autonomous squads. In *International Conference on Agile Software Development*. Springer, 20–28.
- [47] Dilshat Salikhov, Giancarlo Succi, and Alexander Tormasov. 2020. An empirical analysis of success factors in the adoption of the scaled agile framework—first outcomes from an empirical study. *arXiv preprint arXiv:2012.11144* (2020).
- [48] Dilshat Salikhov, Giancarlo Succi, and Alexander Tormasov. 2020. An empirical analysis of success factors in the adoption of the scaled agile framework—first outcomes from an empirical study. *arXiv preprint arXiv:2012.11144* (2020).
- [49] Inc. Scaled Agile. [n.d.]. SAFe Courses. <https://www.scaledagile.com/certifications/which-course-is-right-for-me/>. [Online; accessed 10-APR-2021].
- [50] Scaled Agile Inc. [n.d.]. SAFe Home Page. <https://www.scaledagileframework.com/>.
- [51] J. Schongot and Y. Man. 2018. Lean engineering at elbit systems aerospace division - Implementation journey. In *Annual Forum Proceedings - AHS International*, Vol. 2018-May. www.scopus.com
- [52] Scrum Inc. [n.d.]. Scrum at Scale. <https://www.scrumatscale.com/>.

- [53] Sven Theobald, Anna Schmitt, and Philipp Diebold. 2019. Comparing scaling agile frameworks based on underlying practices. In *Proceedings of the 7th International Workshop on Large-Scale Agile (XP)*. Springer, 88–96. https://doi.org/10.1007/978-3-030-30126-2_11
- [54] Ömer Uludağ, Martin Kleehaus, Christoph Caprano, and Florian Matthes. 2018. Identifying and Structuring Challenges in Large-Scale Agile Development Based on a Structured Literature Review. In *2018 IEEE 22nd International Enterprise Distributed Object Computing Conference*. IEEE, 191–197. <https://doi.org/10.1109/EDOC.2018.00032>
- [55] Ömer Uludağ, Martin Kleehaus, Niklas Dreymann, Christian Kabelin, and Florian Matthes. 2019. Investigating the Adoption and Application of Large-Scale Scrum at a German Automobile Manufacturer. In *Proceedings of the 2019 ACM/IEEE 14th International Conference on Global Software Engineering*. IEEE, 22–29. <https://doi.org/10.1109/ICGSE.2019.00019>
- [56] Ömer Uludağ, Martin Kleehaus, Xian Xu, and Florian Matthes. 2017. Investigating the Role of Architects in Scaling Agile Frameworks. In *2017 IEEE 21st International Enterprise Distributed Object Computing Conference (EDOC)*. IEEE, 123–132. <https://doi.org/10.1109/EDOC.2017.25>
- [57] Ömer Uludağ and Florian Matthes. 2019. Identifying and Documenting Recurring Concerns and Best Practices of Agile Coaches and Scrum Masters in Large-Scale Agile Development. In *Proceedings of the 26th International Conference on Pattern Languages of Programs*. ACM.
- [58] Ömer Uludağ, Pascal Philipp, Abheeshta Putta, Maria Paasivaara, Casper Lassenius, and Florian Matthes. 2021. Revealing the State-of-the-Art in Large-Scale Agile Development: A Systematic Mapping Study. *arXiv preprint arXiv:2007.05578* (2021).
- [59] Ömer Uludağ, Abheeshta Putta, Maria Paasivaara, and Florian Matthes. 2021. Evolution of the Agile Scaling Frameworks. In *International Conference on Agile Software Development*. Springer, Cham, 123–139.
- [60] Version One. [n.d.]. 13th State of Agile Survey. <https://bit.ly/3sadydS>. last accessed on 03-11-2021.
- [61] Version One. [n.d.]. State of Agile Survey. <https://stateofagile.com/#ufh-i-613554519-9th-annual-state-of-agile-report/7027494>.
- [62] VersionOne. 2020. 14th Annual State of Agile Survey. <https://stateofagile.com/#ufh-i-615706098-14th-annual-state-of-agile-report/7027494>. last accessed on: 05-11-2021.
- [63] Ewelina Wińska and Włodzimierz Dąbrowski. 2020. Software development artifacts in large agile organizations: a comparison of scaling agile methods. In *Data-Centric Business and Applications*. Springer, 101–116.
- [64] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. 2012. *Experimentation in Software Engineering*. Springer Science & Business Media.