Perceived Impacts of Using the Scaled Agile Framework for Large-Scale Agile Software Development

Tomas Gustavsson
Karlstad Business School, Karlstad University
Karlstad, Sweden   tomas.gustavsson@kau.se

Linda Bergkvist
Karlstad Business School, Karlstad University
Karlstad, Sweden   linda.bergkvist@kau.se

Abstract

Software development organizations are adopting values, principles, and frameworks to implement agile ways of working today, even in larger organizations. When several teams need to cooperate, and development needs to scale, many organizations are implementing the Scaled Agile Framework. At the same time, both researchers and practitioners have raised critical voices towards this framework, and the impacts are not much studied. This study aims to fill that gap by providing perceptions of benefits and drawbacks experienced in three different organizations: one in the automotive industry, one government agency, and one bank in Sweden. The analysis of survey answers from 154 respondents showed that the most commonly perceived benefits from implementing the Scaled Agile Framework were increased visibility, overview, and transparency. Authors of the framework claim massive productivity gains in every single team but, regarding perceived drawbacks, a lack of productivity, focus, and efficiency were most commonly reported.

Keywords: Scaled Agile Framework, Agile Software Development, Large-scale, Project management.

1. Introduction

The importance of self-organizing teams is put forth in the Agile Manifesto [11], which stresses the importance of allowing autonomy and trust in the team's ability to make wise decisions, solve problems and deliver results. Granting autonomy to the team and team members is a primary reason for success in agile development projects [4], and research in other industries also confirm that autonomous and empowered teams are more productive and proactive [7]. Today there is an industry trend towards adopting agile methodologies in-the-large [1] and although research into the agile approach to software development has matured in the past years, agile in large-scale settings is not much explored [12]. In larger environments, with several teams working towards a common goal, new problems arise such as dependencies between teams and individuals. These dependencies might limit the amount of autonomy and empowerment to the single team, as well as the individual team member. To manage dependencies, there is a need for additional coordination both between teams but also to a great deal towards the rest of the organization [14]. The balance between the benefits of autonomous teams versus alignment towards a common goal is, therefore, an essential problem for the software industry [12]. To maintain this balance, large-scale practices for coordinating teams have been proposed to reduce negative impacts while maintaining the positive impacts of agile ways of working in the teams. But do the large-scale coordination practices really have the intended impacts on the team members and the teams? Or, perhaps the coordinating practices cause problems to achieve the benefits from the agile ways of working in autonomous teams?

According to an annually recurring industry survey [1], the most commonly adopted framework today for large-scale agile ways of working is, by far, the Scaled Agile Framework (SAFe). Leffingwell et al. [10] make several claims regarding expected

beneficial impacts by implementing SAFe, such as employee engagement, productivity increases of 20 to 50 percent and even shorten the time to market with 30 to 75 percent and no drawbacks of implementing SAFe are mentioned. SAFe has been criticized by researchers [3,16] as well as agile practitioners such as Ken Schwaber [18], one of the originators of Scrum, regarding the risk of delimiting autonomy to the individual team which could cause negative impacts to the software development process. Still, no research has been conducted to investigate these possible negative impacts.

In 2008, Laanti et al. [8] conducted a study at Nokia who implemented a predecessor of SAFe, based on Leffingwell's [9] earliest suggested practices. The research did not specifically focus on the large-scale practices but on the overall perceived benefits and challenges of implementing agile ways of working at Nokia. The result from the study was a number of benefits but also many challenges in the implementation. The question is, are these problems present in organizations implementing SAFe today, ten years later? This study aims to investigate the positive and negative impacts of SAFe practices implemented in the organization. It will contribute to further understanding of the impacts of large-scale coordination practices prescribed by large-scale frameworks in organizations working with agile software development. Three organizations were studied to investigate the following research question: *What are the perceived benefits and drawbacks of working in a large-scale agile setting using the SAFe framework?*

The paper is structured as follows: Section 2 discusses large-scale agile software development, and Section 3 presents the data collection and method for data analysis. Section 4 contains results, and Section 5 covers the discussion of the results and its limitations. Finally, Section 6 presents conclusions and future work.

2.   Frameworks for Large-Scale Agile Software Development

The most commonly spread methods for agile ways of working such as Scrum and eXtreme Programming [1] describes concepts and recommendations, like using small, self-managing teams, which works well for small projects. Nevertheless, guidance is lacking when it comes to coordinating multiple teams working in a large project or a large-scale setting. Practices for coordinating several teams are neither addressed in the values or the principles of the Agile Manifesto [11]. However, in recent years, several frameworks have been proposed suggesting several coordination practices when many teams and managers are cooperating towards a common goal.

SAFe is one of the frameworks that seek to address these coordination problems. There are other large-scale agile frameworks such as Large-Scale Scrum (LeSS) and Disciplined Agile Delivery (DAD), but according to the annually recurring industry survey conducted by VersionOne Inc. [1], SAFe is by far the most commonly implemented framework in large organizations. In their study, 29 percent of the respondents who scaled up their agile way of working chose SAFe. The second and third most commonly adopted framework were LeSS and DAD, both at only 5 percent. SAFe is intended to guide enterprises in scaling lean and agile practices by suggesting principles, organizational structure, and workflow patterns promoting alignment and collaboration. The framework is based on three primary bodies of knowledge: agile software development, lean product development, and systems thinking [10]. It was developed by practitioners and is freely available for anyone to use but is a registered trademark of Scaled Agile, Inc.

The first versions of the framework were aimed at visualizing a big picture view of how work flowed from stakeholders and product management, through programs or project portfolios to development teams and finally out to customers. Leffingwell [9], in collaboration with others in the agile community, refined the framework, and formally described it in 2011. SAFe continues to be developed and the latest edition, version 4.6, was released in October 2018 [15].

SAFe describes practices for different hierarchical levels of implementation in an enterprise and suggests a distinction between three levels: portfolio, program, and team level. Since this study is focused on perceived benefits and drawbacks reported mainly by team members, Scrum Masters and Product Owners, we will here focus on describing

practices on the team level. Four practices are prescribed on the team level to achieve inter-team coordination: Program Increment (PI) planning, Scrum of Scrums (SoS), Communities of Practice (CoP) and the use of a program board.

A PI is prescribed to span from eight to twelve weeks, consisting of four timeboxed iterations (often called sprints) of work followed by an Innovation and Planning (IP) iteration. The teams demonstrate results at the end of every sprint to get feedback and thereby to understand the need for further coordination. The PI starts with a two-day PI planning workshop to create a new release plan for the whole PI. PI planning is debated [2,13] whether it is a step backward towards the traditional waterfall model approach or if it is a step forward toward better predictability in agile projects. Critics claim that eight to twelve weeks is a too long time horizon to make detailed plans while supporters claim to perceive a better understanding of the big picture [13]. The two-day workshop has a detailed standardized schedule created to involve all team members as well as stakeholders in performing and committing to a comprehensive plan.

SoS is a face-to-face coordination meeting between representatives from every team and SAFe prescribes that the Scrum Master should be the team representative. Critics [13] claim that with only Scrum Masters attending, there is a risk of "whispering game" problems and that anyone with a dependency issue should be allowed to participate. SoS is suggested to be performed one or two times per week and to be divided into two parts: 1) SoS for everybody attending and 2) Meet-After/Problem solving for representatives that need to sort out dependency issues not resolved during the first part of the meeting. SAFe suggests following a meeting agenda to focus on what problems that are of importance to the other teams regarding what has been done, what will be done and obstacles to reach the current goal of the iteration.

The CoP is described in the SAFe framework as organized groups of people who have a common interest in a specific technical or business domain who regularly collaborate to share information, improve skills, and work to advance the general knowledge of the domain [10]. Several examples of organizing CoPs are also described, such as role-based communities (CoP for architects or testers) or topic-based communities (CoP for Java development or automated testing). CoPs are supposed to both plan for coordination, such as prescribing architectural changes in advance, but also manage emerging dependency issues raised within the specific domain of the CoP.

SAFe prescribes a visual representation of dependencies between team backlogs called a program board. A program board shows the different teams involved in the upcoming release on the Y-axis and the team backlogs along the X-axis. Dependencies between features are demonstrated by using red strings as displayed in Fig. 1.
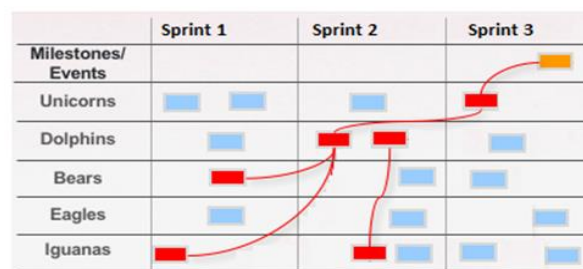

Fig. 1. Program board.

The program board is updated during the PI planning workshop and may or may not be updated when new dependency issues arise during the iterations [15].

Besides the four coordination practices, SAFe prescribes a new role called Release Train Engineer. The word "train" comes from the SAFe terminology where the word for all the teams working together led by a Release Train Engineer is called an Agile Release Train. Since one Product Owner and Scrum Master per team is prescribed, the Release Train Engineer has the responsibility to facilitate meetings to synchronize Product Owners as well as to facilitate SoS meetings.

On the SAFe website [15], several claimed benefits are expressed together with quotes from case studies written by SAFe end users. The benefits are expressed within four areas: 20 to 50 percent increase in productivity, 30 to 75 percent faster time to market, 10 to 50 percent happier, more motivated employees and 25 to 75 percent defect reduction.

There are, of course, no drawbacks mentioned on the website but the framework has been criticized in several ways such as being too top-down and inflexible [3,16,18] and is aggregating too many different practices [5]. SAFe suggests an organizational structure of a product manager coordinating multiple team backlogs and Product Owners. Although SAFe assumes that Product Owners work closely with product management, it has nonetheless been criticized for separating Product Owners into the development organization [2].

3.   Method

This research was conducted as a questionnaire-based survey in three organizations. In this section, the cases are described as well as the method used for data collection and analysis.

3.1.   The case organizations

The first case, called Auto, is a product development department within the Automotive industry. Auto started using Scrum and agile ways of working six years prior to the study, in 2012, and consists of 24 development teams divided into three areas, consisting mostly of software developers but some employees are hardware developers. Roughly 80 percent of the product development is software development, 20 percent hardware development. Some of the teams are rather small, so the total amount of employees is 141 people, which means an average of 5,9 people per team. The average number of years working at Auto is 9,5 for the respondents ($SD = 7,3$). The department is organized in three Agile Release Trains with one Scrum Master per team, almost one Product Owner per team (some are responsible for two teams) and three Release Train Engineers, one per Agile Release Train.

Auto implemented all four described coordination practices on team-level (PI planning, SoS, CoP, and the program board) with some tailoring. The PI follows a strict ten week period consisting of three sprints of three weeks each and a final one-week IP sprint for reflections, experimenting, and planning the next PI. SoS meetings are only held once a week and sometimes even less than so. The program board is only used for planned coordination during a PI and is not updated continuously during the PI.

The second case, called Gov, is a project started in a software development department within a large Government Agency in Sweden (more than 13 000 employees). The project was started as a pilot project for implementing SAFe with the ambition to get best-practice knowledge for further SAFe implementations in the rest of the organization. Gov started using Scrum and agile ways of working five years before the study, in 2013, and consists of 8 teams developing software to be used both in-house, for administrators, and citizens in need of their services. The total amount of employees in the Gov project is 70 people, which means that the average team consisted of 8,75 people, the largest average team size of the three investigated organizations. The average number of years working at Gov is 10,5 for the respondents ($SD = 10,2$). Gov is organized as one Agile Release Train in the same manner as Auto with one Scrum Master per team, one Product Owner per team and one Release Train Engineer.

Gov only implemented three of the described coordination practices on team-level: PI planning, SoS, and the program board. Management did not allow CoPs to be formed since they believed it to be of low value and rather a waste of time. The PI at Gov is a twelve to sixteen week period consisting of four sprints of where the last sprint ends with one day of reflections and a two-day PI planning workshop. SoS meetings are only held once a week and are not organized by Scrum Masters only. Several managers and stakeholders are also invited, which means that the meeting is an informational status meeting rather than a problem-solving session for emerging dependency issues. The program board is only used for planned coordination during a PI and is not updated continuously during the PI.

The third case, called Bank, is a software development department in one of the largest business banks in Sweden (more than 15 000 employees). The studied department in the Bank introduced agile ways of working in dedicated self-organizing teams during 2014. Today, Bank consists of 7 teams with 42 employed developers and testers, which means that the average team consists of 6 people. The average number of years working at Bank is 9,6 for the respondents ($SD = 7,5$). Bank is also organized as one Agile Release Train in the same manner as both Auto and Gov with one Scrum Master per team, one Product Owner per team and one Release Train Engineer.

Bank implemented all four described coordination practices on team-level (PI planning, SoS, CoP, and the program board) but also with some tailoring. The length of the PI has ranged between six and twelve weeks consisting of two to four sprints. Bank deliberately started with shorter PI:s, with only two sprints, to teach everyone involved in how it worked and ended up in four sprint PI:s after nine months of implementation. SoS meetings are held daily, and an extra mid-sprint review meeting has been implemented. The program board is updated twice per sprint: once at the start of the sprint and once during the mid-sprint review meeting.

All three organizations implemented the SAFe way of working one year before the conducted study only a couple of months apart. Auto started to implement SAFe at the beginning of January 2017; Gov started at the end of March 2017, and Bank began in April 2017. Before implementing SAFe, all three organizations worked according to Scrum, in self-organized teams without any formal practices to coordinate between teams. Instead, all dependency issues were dealt with in an ad-hoc manner by, what Thompson [17] calls, "mutual adjustments."

3.2.   Data collection and analysis

The method chosen for this study is to use qualitative findings to reveal the perceptions of the benefits and drawbacks of large-scale agile software development. To provide a view on this topic, respondents are categorized as having a negative, neutral, or positive attitudes towards the adoption of SAFe in the organization. Then, we analyze differences in perceptions between the two extreme groups, i.e., the negative and positive groups of respondents, to see if perceptions differ between the groups.

To do this, paper-based questionnaires were used at all organizations. They were handed out and collected during a two-day PI planning workshop. The survey was conducted in February 2018 for Auto, in March 2018 for Gov and in April 2018 for Bank. The questionnaire consisted of multiple sections: (1) background (e.g., the number of years worked in the organization), (2) agile role, team, and department, (3) opinions of working in a large-scale agile context, (4) effects on teamwork. This study focuses only on Section 3.

Section 3 consisted of one statement and the following two open questions:
1) "What do you consider as the main benefit of using SAFe in your organization?"
2) "What do you consider as the main drawback of using SAFe in your organization?"
Since the question 1) and 2) were open questions, some respondents gave several answers to these questions respectively. The statement in the questionnaire Section 3 was:
3) "I would like to go back to the old way of working (instead of working according to SAFe)."
This statement was followed by a five-point Likert-scale for the respondent to use, reaching from "Not true" (1) to "Neutral" (3) to "Very true" (5).

Laanti et al. [8] used a similar questionnaire in their article on opinions on agile transformation in a large-scale setting at Nokia who, as mentioned previously, did not implement SAFe but a predecessor to SAFe. Nokia was inspired by the early books from the author Leffingwell [9] who later created SAFe, which makes a comparison between these two studies suitable. Instead of using a paper-based questionnaire, they used a web-based survey which was open for two weeks in April 2008. In their study, instead of asking "…benefit of using SAFe…", they asked "…the main benefits of agile methods…" and since the focus of the study was the actual transformation, instead of asking for "drawbacks

of using…", they asked for "the main challenges" of the transformation. Although there is a slight difference of focus between Laanti et al. [8] and this study, they are comparable since several of the departments at Nokia had implemented the large-scale practices more than a year before the survey.

The questionnaire answers in the study by Laanti et al. [8] were grouped into key areas, based on the "agility evaluation grid outline" presented in Kettunen and Laanti [6], and some key areas were added which ended up in 17 key areas in their study. In this study, the questionnaire answers are analyzed by using the same 17 key areas. This was chosen for two reasons: first, to be able to compare the results of Laanti et al. [8] with this study since they are based on the same method originator [9] and, second, because the key areas are based on a framework for measuring agile maturity in different areas important for successful implementations of large-scale agile practices. The framework is, therefore, suitable for the analysis of this study.

Two researchers conducted the analysis in the following way: all data were coded based on the 17 key areas, and the unclear cases were discussed whether the answers were interpreted in the right way and that the true meaning of the answer was captured. Only eight of these cases led to changing the key area. The qualitative responses from Auto were first read through in several increments to code all data. Two examples of coding responses to key areas are presented in Table 1.

Table 1. Coding examples of benefits reported from survey respondents at Auto.

Group	Survey response	Key area
Positive	"Better planning of upcoming work"	(1) Requirements/goals/planning/prioritizing
Negative	"Overview"	(4) Visibility/overview/transparency

Then, the same procedure took place by going through responses from Gov and Bank. During each increment, new findings were added to the already identified areas. Two key areas were renamed for increased clarity. In Laanti et al. [8], the first key area was named "(1) Requirements management/planning/changing requirements/flexibility". Several responses from all three organizations explicitly stated "goals and requirements" and rather than talking about flexibility or changing requirements, "prioritizing" of requirements were expressed. Therefore, the key area was renamed into "(1) Requirements/goals/planning/ prioritizing". Key area number 4 was, in Laanti et al. [8] named "(4) Visibility/ transparency" but, as exemplified in Table 1, several responses simply stated "overview" which can be seen as an effect of visibility or transparency. Since several answers only contained this single word, the name of the key area was extended to become "(4) Visibility/overview/transparency/".

Also, just as in Laanti et al. [8], cluster analysis was used to discover differences between groups, primarily consisting of two different attitude shifts: positive and negative. Statement 3) "I would like to go back to the old way of working (instead of working according to SAFe)" in the questionnaire was used for grouping as it was considered to distinguish the positive group (responses of 1 and 2) and the negative group (responses of 4 and 5) from the survey population. As in the study by Laanti et al. [8], neutral responses (responses of 3) were not included in the analysis.

4.   Results

This section presents the attitudes towards working in a large-scale agile setting using SAFe compared to the previous way of working agile and the identified benefits and drawbacks as perceived by the respondents.

4.1.   Working according to SAFe compared to the previous way of working

The questionnaires were distributed and collected during a PI planning workshop, and since attendance was mandatory in all three organizations, the response rate was very high. At Auto, 125 answers from survey respondents were collected (a response rate of 88,7

percent) of which 103 could be used. The discarded answers were due to lack of response to the question regarding if they would like to go back to the old way of working. Out of the remaining 103 survey respondents, 74 belonged to the Positive group, 12 to the Negative group, and 17 people in the Neutral group. After removing the answers from the Neutral group from the analysis, the responses from 86 survey respondents remained. Since this was an open question, some respondents wrote more than one benefit or drawback, and some only wrote a benefit but no drawback and vice versa, which ended up in 157 answers, 117 from the Positive group and 40 from the Negative group.

At Gov, answers from 65 survey respondents were collected (a response rate of 92,9 percent) of which 57 could be used since eight answers did not state if the respondent would like to go back to the old way of working. Of the remaining 57 survey respondents, 29 belonged to the Positive group, 16 to the Negative group and 12 in the Neutral group. After removing the answers from the Neutral group, answers from 45 respondents remained. As explained previously, since some respondents wrote more than one response, 86 answers were put forth in the survey at Gov, 51 from the Positive group and 35 from the Negative group.

Bank had the lowest response rate where 36 filled out questionnaires from survey respondents were collected (a response rate of 85,7 percent). Answers from 30 survey respondents could be used since six answers did not state if the respondent would like to go back to the old way of working. Of the remaining respondents, 21 belonged to the Positive group, only 2 to the Negative group and 7 in the Neutral group. Remaining answers after removing answers from the neutral group was from 23 survey respondents. In total, 50 responses were put forth in the survey at Bank, 46 from the Positive group, and only four from the Negative group.

All in all, a total of 293 answers from 154 survey respondents were used for classification and analysis from the three organizations. Of the 154 respondents, 16 of them worked as Product Owners, 17 as Scrum Masters, 88 as team members and 33 were various kind of stakeholders such as managers from various departments, project managers, architects, and business analysts.
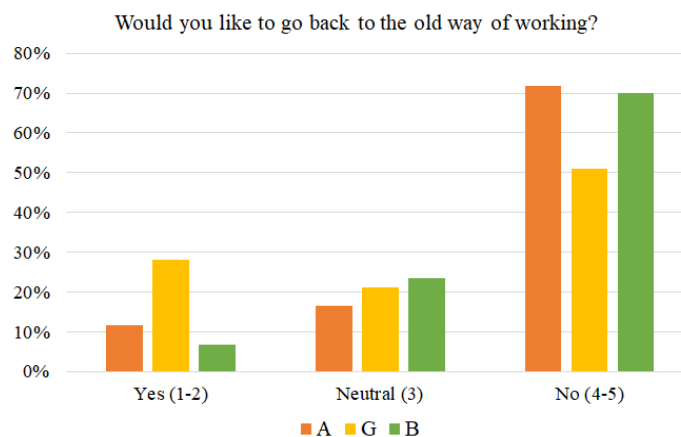

Fig. 2. Working according to SAFe versus the old way of working at Auto (A), Gov (G) and Bank (B).

Fig. 2 shows the overall distribution of responses to the statement "I would like to go back to the old way of working (instead of working according to SAFe)" and how the responses were distributed between respondents from Auto (A), Gov (G), and Bank (B). The overall attitude towards working according to SAFe can be considered as positive, as 72% from Auto, 51% from Gov and 70% from Bank wanted to continue to work according to SAFe. Only 12% from Auto, 28% from Gov, and only 7% from Bank wanted to go back to work in an agile way without using the SAFe practices. However, as many as 17% at Auto, 21% at Gov and 23% of the respondents at Bank were neutral. In the study at Nokia in 2008 [8] 60% wanted to continue working in a large-scale agile way, 31% were neutral, and 9% wanted to go back to the former way of working.

4.2.   Perceived benefits and drawbacks

Fig. 3 presents the distribution of findings, in percentage, in each key area for Auto, Gov, Bank, and the average percentage of answers from the three organizations combined. In the figure, the top three key areas are in bold letters for every group.

Key area	Negative group								Positive group							
	Benefits % (A: #18, G: #14, B: #2)				Drawbacks % (A: #22, G: #21, B: #2)				Benefits % (A: #63, G: #29, B: #27)				Drawbacks % (A: #54, G: #22, B: #19)			
	Avg	A	G	B	Avg	A	G	B	Avg	A	G	B	Avg	A	G	B
(1) Requirements/goals/planning/ prioritizing	**18**	17	21	-	**11**	23	-	-	**14**	16	21	7	**19**	28	9	5
(2) (Product) integration	3	-	7	-	2	5	-	-	-	-	-	-	-	-	-	-
(3) Resourcing/effort management	3	6	-	-	2	5	-	-	1	-	3	-	6	2	18	5
(4) Visibility/overview/ transparency/	**24**	33	7	50	2	5	-	-	**26**	24	24	33	-	-	-	-
(5) Deployment/tailoring/process improvement	-	-	-	-	**18**	14	9	50	-	-	-	-	**25**	30	23	16
(6) Self-organization/teamwork	6	11	-	-	4	5	5	-	8	11	7	-	4	2	5	11
(7) Competence	6	-	14	-	-	-	-	-	2	-	7	-	1	-	5	-
(8) Management	3	-	7	-	7	9	5	-	2	2	3	-	5	4	14	-
(9) Quality/testing/tools	-	-	-	-	4	5	5	-	3	5	-	-	2	2	-	5
(10) Team size/distribution	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(11) Other	3	6	-	-	4	5	5	-	3	3	3	-	5	6	5	5
(12) No benefits or challenges or I do not know	9	11	7	-	-	-	-	-	-	-	-	-	1	-	-	5
(13) Frequent delivery/speed	3	-	7	-	2	-	5	-	5	3	3	11	3	6	-	-
(14) Productivity/focus/efficiency	6	6	7	-	**29**	18	38	50	13	17	3	11	**21**	19	14	37
(15) People/communication/ collaboration	**15**	6	21	50	4	9	-	-	**14**	10	24	15	3	2	-	11
(16) Dependencies/co-operation	3	6	-	-	9	-	19	-	11	10	3	22	3	2	9	-
(17) Error fixing/maintenance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Fig. 3. Benefits and drawbacks based on the key areas as presented by Laanti et al. [8].

SAFe can be implemented and tailored in different ways in different organizations. Even so, the reported perceived impacts do not differ much between Auto, Gov, and Bank, as displayed in Fig. 3. Two differences are apparent, however: many respondents at Auto put forth drawbacks regarding *(1) Requirements/goals/planning/prioritizing* and few at Gov and Bank mentioned it as a benefit while respondents at Gov perceived it in a reversed way. At Gov, many respondents put forth the benefits of *(15) People/communication/ collaboration* which few did in the Auto organization.

In total, for the negative group, the top three perceived drawbacks of working according to SAFe were: *(14) Productivity/focus/efficiency*, *(5) Deployment/tailoring/process improvement*, and *(1) Requirements/ goals/planning/prioritizing*. The top three challenges for the negative group in the study by Laanti et al. [8], were (5), (1), and (3).

For the positive group, the top three perceived drawbacks of working according to SAFe were the same as for the negative group but in another order: *(5) Deployment/tailoring/process improvement*, *(14) Productivity/focus/efficiency*, and *(1) Requirements/goals/planning/prioritizing*. The top three challenges for the positive group in the study by Laanti et al. [8], were (5), (1), and (16).

Fig. 4 shows this summary and comparison of drawbacks in the same manner as presented in Laanti et al. [8] and Fig 5. shows the summary and comparison of perceived benefits in the same way. To get a better overview in both Fig. 4 and Fig. 5, challenges and drawbacks that occur in more than one place are color-coded in the same manner.

Top three perceived drawbacks of large-scale work according to SAFe in 2018	
Positive group	Negative group
1. Deployment/tailoring (5)	1. Productivity/focus/efficiency (14)
2. Productivity/focus/efficiency (14)	2. Deployment/tailoring (5)
3. Requirements/goals/planning/prioritizing (1)	3.Requirements/goals/planning/prioritizing (1)
Top three perceived challenges of agile development in Laanti et al. [8] in 2008	
Positive group	Negative group
1. Deployment/tailoring (5)	1. Deployment/tailoring (5)
2. Requirements/goals/planning/ prioritizing (1)	2. Requirements/goals/planning/ prioritizing (1)
3. Dependencies, co-operation (16)	3. Resourcing/effort management (3)

Fig 4. Top three drawbacks at Auto, Gov, and Bank versus challenges in Laanti et al. [8].

The most reported perceived benefits were the same in both the positive and the negative group in the study. They were, in the following falling order: *(4) Visibility/overview/transparency*, *(1) Requirements/goals/planning/prioritizing*, and *(15) People/communication/collaboration*.

Top three perceived benefits of large-scale work according to SAFe in 2018	
Positive group	Negative group
1. Visibility/overview/transparency (4)	1. Visibility/overview/transparency (4)
2. Requirements/goals/planning/prioritizing (1)	2. Requirements/goals/planning/prioritizing (1)
3. People/communication/collaboration (15)	3. People/communication/collaboration (15)
Top three perceived benefits of agile development in Laanti et al. [8] in 2008	
Positive group	Negative group
1. Visibility/overview/transparency (4)	1. Visibility/overview/transparency (4)
2. Requirements/goals/planning/prioritizing (1)	2. Requirements/goals/planning/prioritizing (1)
3. Productivity/focus/efficiency (14)	3. Frequent delivery/speed (13)

Fig 5. Top three benefits at Auto, Gov, and Bank versus benefits in Laanti et al.[8].

The top three key areas from the positive group in the study by Laanti et al. [8] were (4), (1) and (14) and the top three for the negative group were (4), (1) and (13).

5.   Discussion

There are some interesting differences between the study conducted at Nokia [8], who implemented a predecessor of SAFe, and this study. First of all, Laanti et al. [8] report that, surprisingly, the single largest area of drawbacks in the negative group with 22% of the answers was "(12) No benefits or challenges or I do not know". Some of their comments showed much frustration and even anger towards the implementation. That has not been the case in this survey where only a total of 9% (11% in Auto and 7 % in Gov) saw no benefits at all. The rest of this section contains a discussion on the top three benefits and drawbacks and limitations to the study.

5.1.   Top three benefits and drawbacks

Regarding the top three drawbacks and challenges, for both the positive and negative groups in this study and Laanti et al. [8], two key areas were within the top three: *(5) Deployment/tailoring*, and *(1) Requirement management/planning*. In this study, however, the third key area within the top three was *(14) Productivity/focus/efficiency*. For the negative group, this was the number one most reported drawback and for the positive group as the second most reported drawback. This indicates that many employees might perceive working in large-scale agile development according to SAFe as not being very productive or efficient.

This is an important contribution of this study since Leffingwell et al. [10] claim that implementing SAFe increases employee engagement, motivation and productivity and even state that the increase of productivity in every single team is as high as between 20 and 50 percent. In this study, the actual productivity was not measured, so there are,

unfortunately, no numbers to verify whether productivity has increased or decreased. But regardless of the actual effect, it is important to understand that while some respondents perceive an increase in productivity (13% in the positive group, 6% in the negative), many will actually perceive that productivity is lacking (29% in the positive group, 21% in the negative). For a practitioner, this is an important takeaway from this study: by implementing the coordination practices of the SAFe framework, there might be resistance from employees due to a perceived lack of productivity.

Regarding the top three benefits, the first and second place are identical between this study and Laanti et al. [8] for both the positive and the negative groups. In third place in Laanti et al. [8] for the negative group is *(13) Frequent delivery/speed* which is not even close to being among the top three benefits in any of the three investigated organizations. Although this is a small study with only three organizations involved, it seems possible that working according to SAFe is not perceived as increasing the speed of delivery. This also is an important contribution of this study since Leffingwell et al. [10] claim speed to be one of the expected benefits from implementing SAFe, with an increased time to market with 30 to 75 percent. As noted above, this is also an essential takeaway for practitioners as it is in line with the lacking sense of productivity. Many employees in an organization where the SAFe framework is implemented will not perceive increased speed as a benefit. No matter if speed has increased or not.

As was described regarding the practice called PI planning: critics view this as a step towards the waterfall model while supporters claim to perceive a better understanding of the big picture [13]. Looking at reported benefits, a better overview is reported from many respondents who might indicate that PI planning does give the suggested benefit, as supporters have claimed. However, one key area noted as the third most reported of the top three benefits by both negative and positive respondents in this study is *(15) People/communication/collaboration* which is not within top three of any respondent group in Laanti et al. [8]. These findings suggest working according to SAFe might increase collaboration and communication between people; at least it is perceived to do so. This benefit is not explicitly highlighted as an expected outcome of implementing SAFe in Leffingwell et al. [10].

One important contribution of this study is that it confirms several benefits and challenges/drawbacks that will be perceived when large-scale coordination practices are implemented in an organization. The results of this survey conducted in three organizations and the study in Laanti et al. [8] in another organization have many similarities and confirm several key areas, even though they are conducted ten years apart in different organizations.

Many employees will most likely perceive the benefits of *(4) visibility/overview/ transparency,* and *(1) requirements/goals/planning/prioritizing* when large-scale practices are implemented. They were in first and second place in both studies and thereby further confirms the evidence of these beneficial impacts. But many employees will also struggle with *(5) deployment/tailoring*, and *(1) requirements/goals/planning/prioritizing* since they are both within the top three perceived drawbacks/challenges in both studies.

Looking at benefits which were not within the top three, several respondents in the positive group saw benefits of *(14) Productivity/focus/efficiency* (13 %) compared to the negative group (6 %). The same goes for *(16) Dependencies/co-operation*, where 11% of the positive group perceived benefits versus 3% of the negative group.

Regarding drawbacks not within the top three, more respondents in the negative group viewed *(16) Dependencies/co-operation* as problematic (9%) compared to the positive group (3%).

5.2.   Limitations

The study is scoped into the respondents' opinions and their perceived benefits and drawbacks of working according to the SAFe framework. This study has not taken the success, or lack of success, of the implementation of SAFe into account whether it might have had an impact on the results. This would have required a comparative study over a more extended period.

The survey was conducted in three organization so that it might be biased by the organizational culture and the specific interpretation and implementation of SAFe. Another possible bias that also may be caused by the organizational culture is whether the respondents answered truthfully or not. That means that there remains a question of external validity, i.e., to what extent these results can be transferred. External reliability deals with the issue of whether or not independent researchers can replicate studies in the same or similar settings. There is always a question of how much the results were affected by the response rate, i.e., would it be possible that only those respondents who are more satisfied with SAFe answered the survey. Fortunately, the response rate was very high in the studied organizations, 88,7 percent in Auto, 92,9 percent in Gov, and 85,7 percent at Bank. Although, since discarding respondents who were neutral to the newly implemented way of working (or did not answer the question) actually means that the data used from each organization represents the view of 61 percent of the employees from Auto, 64,3 percent of employees from Gov and 54,8 percent of employees at Bank. These are still high response rates, so this possible bias is likely out of the question.

The negative group was much smaller than the positive group. This, in itself, is an interesting finding, but could also impact the robustness of the negative response answers. Notably, comparing differences in answers from the negative group between each organization is not feasible since the negative group at Bank consisted of only two respondents. The perceived benefits and drawbacks based on the respondents' role are also scoped out from this paper, even though the impact on some roles have probably been more thorough than to other roles. This impact could be further analyzed in future research.

6.   Conclusions and future research

The motivation of this study was to provide empirical evidence of benefits and drawbacks of the most common [1] large-scale agile framework adopted today: SAFe. Since the originators of SAFe claim several benefits but no drawbacks to being expected [10], it is important to investigate if the practices of SAFe will provide any real improvement from the viewpoint of adopters and if there are any potential drawbacks of scaling up according to SAFe.

By answering the research question "What are the perceived benefits and drawbacks of working according to the SAFe framework?" the research goal of this study was to contribute to the scientific and industrial body of knowledge on the perceived benefits and drawbacks of large-scale agile development. As a result, the three main drawbacks of working according to SAFe in three studied organizations were: *(14) Productivity/focus/ efficiency*, *(5) Deployment/tailoring*, and *(1) Requirements management/planning*. This is in contrasts with the study by Laanti et al. [8] where very few respondents saw *(14) Productivity/focus/efficiency* as a challenge. This is an important contribution since this implies that several of the practices prescribed by the SAFe framework is perceived as unproductive and not helpful for coordination purposes. These results confirm previous voices of critics [3,16,18] regarding SAFe as being too bureaucratic and having too much administrative overhead.

The three main benefits perceived in this study were: *(4) Visibility/overview/ transparency*, *(1) Requirements management/planning*, and *(15) People/communication/ collaboration*. Comparing these results, it is interesting to see that very few in this study put forth one of the main three benefits in Laanti et al.[8]: *(13) Frequent delivery/speed*. This is especially interesting since SAFe claims implementation of the framework will increase the time to market with 30 to 75 percent [10]. It is important to remember that these are the perceived effects of SAFe and no real measurements on the actual increase (or decrease) of speed has been conducted. However, it is an important contribution for practitioners to understand that high delivery speed is not perceived by many employees, whether there is an increase or not. To speculate, this might be due to a common perception of practices as being unproductive (overhead), as mentioned in the reported drawbacks.

Further studies are needed to confirm these preliminary results with only 293 responses from 154 people in three organizations. The study was carried out at one point in time one

year after the implementation of SAFe. A suggestion of further work is not only to conduct studies in more settings but also to investigate the perceived impacts of SAFe at a later place in time in the same organizations to study if the perceived benefits and drawbacks will change over time and especially to investigate differences between roles. Another important area of further work is to conduct actual measurements of the effects of SAFe implementations. As many respondents reported lack of productivity as a drawback and very few reported the increased speed of delivery as a benefit, this is in contrast with the claimed benefits of increased productivity and shorter time to market.

References

1. 12th Annual State of Agile™ Survey (2018), http://stateofagile.versionone.com. Accessed March 1, 2019
2. A critical view on SAFe (2015), http://scrumorakel.de/blog. Accessed March 9, 2019
3. Alqudah, M., Razali, R.: A review of scaling agile methods in large software development. International Journal on Advanced Science, Engineering and Information Technology. 6(6), 828-837 (2016)
4. Chow, T., Cao, D. B.: A survey study of critical success factors in agile software projects. Journal of systems and software. 81(6), 961-971 (2008)
5. Has SAFe Cracked the Large Agile Adoption Nut? (2013) http://www.infoq.com/news/2013/08/safe. Accessed February 27, 2019
6. Kettunen, P., Laanti, M.: Combining agile software projects and large-scale organizational agility. Software Process: Improvement and Practice. 13 (1), 183–193 (2008)
7. Kirkman, B. L., Rosen, B.: Beyond self-management: Antecedents and consequences of team empowerment. Academy of Management Journal. 42 (1), 58-74 (1999)
8. Laanti, M., Salo, O., Abrahamsson, P.: Agile methods rapidly replacing traditional methods at Nokia: A survey of opinions on agile transformation. Information and Software Technology. 53 (3), 276-290 (2011)
9. Leffingwell, D.: Scaling Software Agility: Best Practices for Large Enterprises. Addison-Wesley, Willard (2007)
10. Leffingwell, D., Yakyama, A., Knaster, R., Jemilo, D., Oren, I.: SAFe reference guide: Scaled Agile Framework for Lean Software and Systems Engineering. Addison Wesley, Willard (2017)
11. Manifesto for Agile Software Development (2001), http://www.agilemanifesto.org. Accessed March 1, 2019
12. Moe, N. B., Olsson, H. H., Dingsøyr, T.: Trends in Large-Scale Agile Development: A Summary of the 4th Workshop at XP2016. In: Proceedings of the Scientific Workshop Proceedings of XP2016, pp. 1-4. ACM, XP2016 (2016)
13. Paasivaara, M., Lassenius, C., Heikkilä, V.T.: Inter-team coordination in large-scale glob-ally distributed scrum: Do Scrum-of-Scrums really work? In: Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement, pp. 235-238, Lund (2012)
14. Rolland, K. H., Fitzgerald, B., Dingsøyr, T., Stol, K.-J.: Problematizing Agile in the Large: Alternative Assumptions for Large-Scale Agile Development. In: Proceedings of the International Conference on Information Systems, pp. 1-21. Dublin, Ireland (2016)
15. Scaled Agile Framework version 4.6. (2018), http://www.scaledagileframework.com. Accesses February 6, 2018
16. Stojanov, I., Turetken, O., Trienekens, J. J.: "A maturity model for scaling agile development", In: Proceedings of 41st Euromicro Conference on Software Engineering and Advanced Applications, pp. 446-453. Funchal (2015)
17. Thompson, J. D.: Organizations in Action. McGraw-Hill, Chicago (1967)
18. UnSAFe at any speed (2013), http://kenschwaber.wordpress.com/2013/08/06/unsafe-at-any-speed/. Accessed March 8, 2019