# Performance Alignment Work: How software developers experience the continuous adaptation of team performance in Lean and Agile environments

Fabian Fagerholm[a,*], Marko Ikonen[a], Petri Kettunen[a], Jürgen Münch[a], Virpi Roto[b], Pekka Abrahamsson[c]

[a]*Department of Computer Science, University of Helsinki, P.O. Box 68, FI-00014 University of Helsinki, Finland*
[b]*School of Arts, Design and Architecture, Aalto University, P.O. Box 31000, 00076 Aalto University, Finland*
[c]*Faculty of Computer Science, Free University of Bozen-Bolzano, Piazza Domenicani 3, Bolzano, Italy*

## Abstract

**Context:** Companies increasingly strive to adapt to market and ecosystem changes in real time. Gauging and understanding team performance in such changing environments present a major challenge. **Objective:** This paper aims to understand how software developers experience the continuous adaptation of performance in a modern, highly volatile environment using Lean and Agile software development methodology. This understanding can be used as a basis for guiding formation and maintenance of high-performing teams, to inform performance improvement initiatives, and to improve working conditions for software developers. **Method:** A qualitative multiple-case study using thematic interviews was conducted with 16 experienced practitioners in five organisations. **Results:** We generated a grounded theory, Performance Alignment Work, showing how software developers experience performance. We found 33 major categories of performance factors and relationships between the factors. A cross-case comparison revealed similarities and differences between different kinds and different sizes of organisations. **Conclusions:** Based on our study, software teams are engaged in a constant cycle of interpreting their own performance and negotiating its alignment with

*Corresponding author
Email addresses:* `fabian.fagerholm@helsinki.fi` (Fabian Fagerholm),
`marko.ikonen@alumni.helsinki.fi` (Marko Ikonen), `petri.kettunen@cs.helsinki.fi` (Petri Kettunen), `juergen.muench@cs.helsinki.fi` (Jürgen Münch), `virpi.roto@aalto.fi` (Virpi Roto), `pekka.abrahamsson@unibz.it` (Pekka Abrahamsson)

other stakeholders. While differences across organisational sizes exist, a common set of performance experiences is present despite differences in context variables. Enhancing performance experiences requires integration of soft factors, such as communication, team spirit, team identity, and values, into the overall development process. Our findings suggest a view of software development and software team performance that centres around behavioural and social sciences.

## 1. Introduction

Performance is a multi-faceted concept that is used on several levels of an organisation to mean different things [30]. The desired outcome, a successful and well-performing software product or service, is contingent on a complex combination of factors that can be found in projects, processes, organisations, teams, and individuals (e.g. [50, 54, 58, 59]). Within these categories, there are multiple characterisations of performance that are relevant in different contexts and for different purposes. Even the performance of the end result, the software itself, can be viewed in different ways; e.g. in terms of technical quality, fitness for purpose, or generated profits. Many of today's software development organisations operate in highly volatile environments in which different elements of performance can change rapidly. As corporate strategy changes, performance targets may sometimes change implicitly, sliding continuously to meet the updated understanding of conditions in the business milieu. Some organisations aim to improve performance by being more responsive to changing market needs, e.g. by treating R&D as a continuous experimentation system [40]. However, propagating goal changes to all levels of the organisation in a comprehensive and timely manner may be hampered by communication and transparency problems. Also, if goals change too quickly and frequently, organisational activity may become erratic and self-defeating.

When the objective is to analyse and understand teams, human factors are brought to the forefront. A team may be evaluated, e.g., in terms of its productivity [59], speed [7], or ability to produce novel and innovative results [45]. It may also be evaluated in terms

of process control [54], or the knowledge it produces [55]. Many factors influence performance within these areas and time is frequently an important factor to consider. However, since software development is largely a human-based activity, most types of outcome depend on human factors. Motivation, skill, satisfaction, values, and personality are factors to consider when forming teams, creating and designing processes and development environments, and structuring organisations and communication. The importance of such human aspects on performance in software development is well known [6, 8, 26, 37, 59]. However, there is a lack of understanding in many software development environments of how software practitioners themselves experience the pursuit of high performance, and how striving for performance could simultaneously be a meaningful and positive experience.

In a previous paper [27], we studied how professional software developers experience performance in a Lean and Agile context. Drawing on an earlier conceptualisation of Developer Experience [28], we approached the issue through a cognitive, affective, and conative lens. We viewed team performance from the perspective of individual software practitioners, gaining insights that may be of use in evaluating teams from an internal perspective. The study showed why it is not sufficient to consider performance only as meeting predefined objectives. It also showed how practitioners reason as they attempt to perform in their work, and what they perceive as beneficial and detrimental for those attempts.

The present article is an extension of the previous study that adds additional analysis. We aim to cast further light on the similarities and differences in performance experiences among professional software developers in different types of companies. We augment our previous results with findings that show reasoning appearing consistently across companies of different types, and reasoning that emerges when moving between types: from smaller to larger companies, between companies in different fields of industry, and different degrees of globalisation. We also show that understanding how individual software developers experience the striving for performance in their teams can help formulate hypotheses of how and why the company is currently performing in its software development activities. Such hypotheses may be of use in performance

3

improvement efforts, such as software process improvement initiatives. Our specific research questions are:

**RQ1** How do software practitioners experience team performance in Lean and Agile environments?

**RQ2** How do software practitioners reason about the relationships between perceived performance factors?

**RQ3** How do performance factors experienced by software practitioners differ between different types of companies?

The remainder of the article is organised as follows. In Section 2, we discuss the concept of performance in software engineering, with particular focus on human factors on the team and individual levels. In Section 3, we describe our research approach: the data collection and analysis methods used. In Section 4, we present the empirical results. We discuss the implications and limitations of our findings in Section 5. Finally, we conclude the paper in Section 6 and briefly outline possible future work.

## 2. Theoretical background

One of the foremost practical objectives of team performance research is the pursuit of ways to improve the work outcome of teams. It is interesting to note that teams were once considered an improvement over individual work: teams can potentially offer greater adaptability, productivity, and creativity than any single individual [31, 35, 61]. However, gaining the potential benefits of teams is not easy. For example, it is not enough to merely group skilled individuals together [36]. In this section, we briefly discuss how to define performance, and shortly review some previous research on performance factors and models of team performance.

### 2.1. Definition of performance

One definition of high-performing teams is that they outperform "all reasonable expectations as well as all other similarly situated teams" [43]. While this definition

proceeds to say that the performance of these teams surprises even themselves, organisations find high-performing teams highly desirable and wish to replicate their success. However, reports describing such high-performing teams are typically on an anecdotal level, based more on assumptions than on a valid causal analysis. Part of the problem may stem from the lack of a sound measure for "success" in software engineering, although it is a central dependent variable [57].

Performance is often divided into efficiency and effectiveness. Efficiency means accomplishing objectives quickly and with minimal resource usage. Effectiveness refers to accomplishing the right objectives, e.g. those that have the greatest value. However, the terms can be used differently; e.g. Salas et al. [61] use them as follows. Team performance refers to "the outcomes of the team's actions regardless of how the team may have accomplished the task". Team effectiveness considers "not only whether the team performed" (e.g. completed a task), but also "how the team interacted to achieve the team outcome" (e.g. team processes, teamwork). The distinction is important since many factors may influence the outcome, and confound the causal reasoning assumed in team performance measures. This may result in an incorrect understanding of the team and the group processes which govern its performance [61]. In this work, we use "performance" as an umbrella term for all the meanings described above and use more specific terms as needed.

## 2.2. Performance influence factors

Sudhakar et al. [65] list four classes of factors which influence team performance: (i) technical, (ii) non-technical (soft), (iii) organisational, and (iv) environmental. The technical factors include project-specific traits such as size, complexity, and processes, as well as product characteristics. There are numerous reported soft factors, and fully explaining them is beyond the scope of this paper. However, some examples can be mentioned.

On the individual level, cognitive factors include skill [9, 10, 65], knowledge [49], competence [37], and logical reasoning [13]. Motivation is a conative factor that has received much attention in software engineering research [8, 29]. Personal values [49], beliefs [23, 56], and personality [8, 65] have also been investigated as direct or indirect

performance factors. In addition, affective factors have been examined, showing that developers do experience several emotions in their work, and that these change over time [64]. Moods can influence programming tasks such as debugging [46]. Enthusiasm [67], and emotional valence and dominance [34], can have a positive effect on performance, while frustration is a negative risk factor for performance [67].

On the group or team level, some of the reported factors include cohesion [22, 41, 69], trust [1, 22, 41], clarity of purpose and goal-setting [1], group structure and communication [18, 59, 68], knowledge sharing [37], team relationships, diversity, and leadership [65], and coordination processes [47, 68]. The organisational factors include organisational culture, climate, structure, and values [65]. Finally, the environmental factors include industry characteristics and volatility, and also factors relating to customers and competitors [65].

### 2.3. Team performance models

Many researchers have shown that team effectiveness is the result of the team's processes (e.g. [31, 35, 52]). However, it is less clear what those processes are and how they result in improved outcomes. Salas et al. [61] note that teams "do more than simply interact with tools; they require the ability to coordinate and cooperatively interact with each other to facilitate task objectives through a shared understanding [of available resources, goals and objectives, and constraints]". Research has shown that different types of teams manifest teamwork processes differently [61].

Team performance models aim to describe causal relationships between variables that result in performance outcomes or at least provide actionable advice for managing performance. Dingsøyr and Dybå [20] discuss three teamwork models concerned with team effectiveness from an internal perspective. In the Salas model [61], five components (team leadership, mutual performance monitoring, backup behaviour, adaptability, and team orientation) interact to produce performance. Three coordinating mechanisms (shared mental models, closed-loop communication, and mutual trust) are proposed as means to raise the level of performance. The Dickinson and McIntyre model [19] is similar to the Salas model. It adds feedback and coordination, and is intended for self-managed teams. The Hoegl model [38] has six facets: communication, coordination,

balance of member contributions, mutual support, effort, and cohesion. The model has been shown to have a significant association with team performance (divided into effectiveness and efficiency) and team members' personal success (work satisfaction and learning).

Dingsøyr and Dybå [20] assert that although several team performance (or effectiveness) models exist in other disciplines, there are many open questions regarding their use in software engineering. The relationship between team performance and project success also remains an open question. Success includes not only meeting schedules and making profits, but also encompasses employee well-being and public impact [58]. The notion of performance must then be considered dynamic, to include the activity of defining its meaning. In this expanded definition, performance can be understood in many different ways depending on the viewpoint [30, 58], and viewpoints may conflict [42, 45, 58]. An open question is therefore how software development practitioners experience the pursuit of high performance in an uncertain environment. Could the pursuit of high performance be more than improving the work outcome of teams?

## 3. Research approach

Due to the nature of our research questions, we chose an exploratory, embedded multiple-case study method [70]. Case studies aim at investigating contemporary phenomena in their context [60] and are suitable for research questions of an exploratory and explanatory nature [70]. Our aim is to generate, not test, theory. There are several ways in which case studies can be used to inductively build theory [24, 25]. We used an analysis strategy based on grounded theory coding methods [16].

### 3.1. Sample and context

Being a multiple-case study, this study aims to understand the dynamics of performance in software development teams by viewing it from the perspective of multiple practitioners in multiple organisations. On both company and participant levels, we used maximum variation sampling [24, 32], where the focus is on finding variants on a

7

common theme [53]. Widely varying instances are of particular interest since they allow capturing the core experiences and common dimensions of a setting or phenomenon [32]. Such variations expand the range of applications in which the results can be used [53].

We approached five companies with offices in Finland (Table 1), varying in terms of size, industry, and market. They were selected because they used Lean and Agile software development approaches and because they operated in volatile markets. All five agreed to participate in a research project on team performance. The companies vary in size, ranging from around 50 employees to almost 1000 employees in the Finnish location and tens of thousands worldwide. All are at least 10 years old; the oldest traces its roots back more than 100 years. One is a Fortune 500 company, two are publicly traded on Nordic stock exchanges, and two are privately owned. The companies operate in several different application domains, including telecommunications, embedded and wireless systems, data and network security, and general software and business development services. Some of them provide consultation services and product and service development to third parties, while others market their own products directly to businesses and consumers. All companies had offices in or near the Finnish capital of Helsinki at the time of the study, the large ones with several offices in different parts of Finland. All companies have adopted Lean and Agile development principles and use some variant of Lean and Agile software development in their development process. In the older and larger companies, multi-year organisational transformations have been conducted to replace earlier software development approaches with more modern ones.

Following our instructions, contact persons within each company purposively selected participants with sufficient experience to give relevant information regarding the research topic. We sought participants from different parts of the development organisation, to cover a wide range of perspectives. At the time of the study, the participants worked in teams of approximately 5–15 persons. However, all of them had worked in several teams of different sizes during their career, and thus had experience with many different team conditions to draw from. In total, our sample consists of 16 practitioners, including managers on the company and department levels (3), coaches/team leaders (11), and team members (2); 13 were male and 3 female. All except two participants were native Finns.

Table 1: Key Demographic Facts of Case Companies. Company sizes are reported according to European Commission Recommendation 2003/361 [15], which classifies companies according to headcount and turnover as follows. Micro: < 10, ≤ 2 M€; Small: < 50, ≤ 10 M€; Medium: < 250, ≤ 50 M€ (both criteria must be fulfilled). In addition, we separate large companies, which exceed the criteria for medium company in the EC recommendation, and very large companies, which we define by headcount ≥ 5000 and turnover ≥ 500 000 M€. We consider the headcount and turnover of the entire business group, not only the national subsidiary.

| Company | Field of Industry | Size | Degree of Globalisation |
|---------|-------------------|------|-------------------------|
| A | Business development, consulting | Small | National |
| B | Telecom networks and services | Very large | Multinational |
| C | Computer security | Large | Multinational |
| D | Embedded and wireless systems | Large | Multinational |
| E | Software design and development | Medium | National |

## 3.2. Data collection

We used thematic, semi-structured interviews [5, 53] for data collection. In thematic interviews, participants are interviewed about issues directly related to the object of study – the theme [5]. They allow quick access to a wide and deep range of practitioner expertise and are particularly useful for aspects that the interviewee is not accustomed to speak about on a daily basis, such as values, intentions, or ideals. The amount of structure in such interviews may vary. A semi-structured interview, used here, is a mix of more and less structured questions, but with flexible wording and question order. A base set of questions is always covered, but there is room for open-ended, exploratory conversation. Multiple perspectives on the same issues can thus be examined, resulting in triangulated data both within and between participants [5, 53].

An interview guide is a list of questions to be asked or topical areas to be covered by the interviewer, possibly including their order and other instructions [53]. We designed an interview guide for discussing performance from several perspectives in order to gain data to answer our research questions. The guide was designed in a chronological fashion to help recollect experiences from participants' entire careers. It was kept flexible enough to allow constant analysis of interviews to affect the direction of subsequent interviews, supporting the grounded theory approach of constant comparison [17, 32]. The guide is shown in Appendix A.

Data collection was carried out during two months in 2011. Two researchers carried out 16 semi-structured thematic interviews of about 1.5 hours each with participants from the five companies under study. Each interview was recorded for later reference. Notes were primarily taken by one of the researchers while the other primarily concentrated on interviewing the participant. After each interview, the researchers discussed the interviews, wrote supporting field notes, and constantly compared the interview material with previous interviews. When participants spontaneously raised a topic or told a story, we allowed them to do so freely, while making sure that the themes in the guide were covered. Where applicable, we specifically asked follow-up questions about personal opinions, reasoning, and feelings. Interviews were carried out in English or Finnish, according to the participant's preference.

### 3.3. Data analysis

Our analysis procedure consists of two parts. In the first part, the aim was to make sense of the collected interview data and inductively construct a theoretical structure which would explain the information given by the interview participants, providing answers to the first two research questions. This part was closely linked with the data collection phase and the two informed each other. In the second part, the aim was to increase the depth and breadth of the theoretical structure and provide an answer to the third research question. In the latter part, we re-analysed the entire data set and results from the first part in order to make a cross-case synthesis on the company level.

In the first part of the analysis, we employed coding strategies from grounded theory method [16, 32] to analyse the interview data. Grounded theory can be thought to proceed in three phases [16]: open, axial, and selective coding. During open coding, we identified categories in the data by grouping related interview fragments. Fragments that mentioned, e.g., a similar performance factor, or reasoned similarly about causes and effects, were grouped together. During open coding, related groups may also be clustered into higher-order categories. In axial coding, we related categories to each other, creating a refined category scheme with links between the categories.

To enable multiple researchers to collaboratively perform open and axial coding, we used the Affinity wall method. The method originates in quality management

research [44] but has been transferred into system design as a tool for consolidating large amounts of seemingly disparate information into a coherent picture [11, 12]. The Affinity wall method ensures that several researchers consider and discuss each and every piece of data, bringing researcher triangulation into a central position in the research process. The method involves multiple participants iteratively categorising pieces of information written on paper notes onto a flat surface, usually a wall. In our case, the notes were self-contained pieces of interview data. The result is an Affinity diagram, a three-level hierarchical diagram which structures field data (notes) into conceptually similar groups, which are then further organised into higher-level categories. The diagram is often referred to simply as an Affinity wall.

In the final phase, we used selective coding. Selective coding develops a core category, propositions, or a hypothesis. Here, analysis does not seek to summarise material without losing detail as in the Affinity wall method, nor to expand to generate new perspectives, but aims to proceed quickly and selectively towards a coherent, integrated theory [32, 62]. In our case, the overall understanding in the data pointed towards a core category that described a sense-making and negotiation process. The details of the core category and the relationships between it and the other categories are given in Section 4. The categories generated during selective coding form an additional hierarchy above that of the Affinity wall diagram and organises the top-level Affinity wall categories into a meaningful theory.

In the second part of the analysis, we performed a cross-case synthesis on the company level by revisiting all the interview data and results from the first part of the analysis. In a cross-case synthesis, each individual case is treated as a separate study, and findings are aggregated across the studies [70]. Comparing similarities and differences across the cases can potentially result in knowledge about the general types to which the cases belong, thus increasing the generality of the findings. In this study, cross-case comparison also helps refine the theoretical contribution, as it adds information about the contextual conditions in which the theory is relevant. The results of the cross-case analysis are given in Section 4.5.
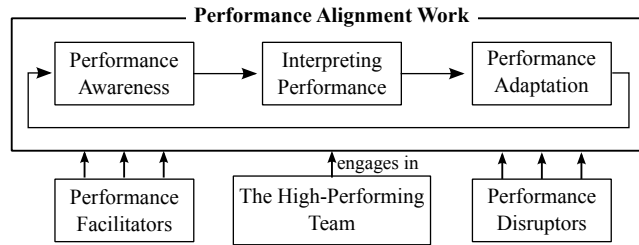
Figure 1: Performance Alignment Work is a continuous process in which people work to negotiate the meaning of performance in different situations and adapt their performance to changing conditions. Several factors can facilitate or disrupt this work. High-performing teams engage actively in Performance Alignment Work and are particularly good at it, resulting in superior performance. The parts of this figure are broken down in greater detail in figures 2–9.

## 4. Results

The core category *Performance Alignment Work* (see Figure 1) was constructed to summarise the entire data set. This category refers to the continuous process that all participants were engaged in to negotiate the meaning of performance in different situations, interpret their current performance, and adapt it to changing circumstances. In this section, we introduce the categories around the core category and relationships among them that arose during the analysis. We illustrate the results using diagrams which combine hierarchies and effect relationships in order to show the reasoning described in the interview material. We include interview quotes as examples of the data behind the categories. Each category is emphasised in the text. In Table 2, we also show the 33 categories that emerged from the Affinity wall analysis, to facilitate traceability into the interview material. These top-level categories represent sets of interview contents that share a common theme. Below them is a second level of more detailed categories which in turn consist of a third level of individual interview fragments. For each top-level category, the number of supporting second- and third-level items are given, as well as an indication of which categories were found in which companies. It is thus possible to assess the relative strength of the categories. The table also links the top-level Affinity wall categories to the visual diagrams that depict the components of the final grounded theory. Finally, we present a cross-case synthesis of the entire data

| Understanding job roles | individual | **Performance Awareness** | systemic | Seeing the big picture |
|---|---|---|---|---|

Figure 2: *Performance Awareness* refers to becoming aware of performance concerns, with perceptions ranging from individuals and their job roles to a larger, more systemic understanding which includes teams, units, the organisation, and the entire market.

set which shows differences and commonalities between the studied organisations and provides more contextual detail to deepen the theoretical contribution of this paper.

*4.1. The meaning of performance*

Participants expressed different views regarding the meaning of performance. Two top-level categories form orthogonal axes that summarise these differing views. The first axis, *Performance awareness*, describes the level of self- or other-orientation in participants' perception of performance (see Figure 2). This axis forms a continuum ranging from the individual (*Understanding job roles*) to the team, unit, organisation, and market (*Seeing the big picture*).

> "[The way I was trained], the project manager had to know everything about everything. He had to think about everything. I realised that in Agile, the team has the power to decide, and is trusted." (*Coach, Company E*)

The second axis, *Interpreting performance*, describes the desirable level of performance: meeting or exceeding predefined objectives – or transcending them altogether by participating in their definition and assessment (see Figure 3). These interpretations stem from different *Ways to see success*, different understandings of the development philosophy (*Understanding Agile*), different views on the desirability of being involved in social interaction (*Need for communication*) and different notions of how goals should be set and pursued (*Planning of work*).

> "Good performance is such that it fulfils expectations and the expectations come from some kind of conception about the end customer." (*Manager, Company D*)

13

Table 2: Categories from the Affinity wall. The number of supporting second- and third-level subcategories, and the companies from which the supporting interview data originated, are shown. The table is sorted by number of third-level subcategories, followed by second-level subcategories and last by the number of source companies. The crosses ($\times$) indicate which companies' interview data are included in each category, and thus which categories were present in which companies. The number of second- and third-level categories, and the number of top-level categories per company (number of crosses) are shown in the bottom row.

| Top-level category | Shown in Figure | Number of subcategories | | Companies | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $2^{nd}$ level | $3^{rd}$ level | A | B | C | D | E |
| Need for communication | 3 | 6 | 36 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Team spirit | 9 | 8 | 36 | $\times$ | $\times$ | $\times$ | | $\times$ |
| Improving the process | 8 | 7 | 33 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Re-organisation | 5 | 7 | 31 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Team setup | 9 | 8 | 30 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Tools | 7 | 8 | 30 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Decision power | 4 | 6 | 29 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Facilitating communication | 4 | 3 | 29 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Organisational learning | 5 | 5 | 27 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Organisational support | 4 | 8 | 26 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Planning of work | 3 | 6 | 26 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Ways to see success | 3 | 6 | 25 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Reward | 9 | 6 | 24 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Atmosphere | 4 | 5 | 24 | $\times$ | $\times$ | $\times$ | | $\times$ |
| Time investment | 8 | 6 | 23 | $\times$ | $\times$ | $\times$ | $\times$ | |
| Seeing the big picture | 2 | 5 | 23 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Collaboration and cooperation | 6 | 3 | 23 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Distributed work | 5 | 5 | 21 | $\times$ | $\times$ | $\times$ | $\times$ | |
| Personal development | 9 | 4 | 21 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Understanding job roles | 2 | 4 | 21 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Goal setting | 6 | 6 | 20 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Team identity | 9 | 5 | 20 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Understanding Agile | 3 | 4 | 19 | $\times$ | $\times$ | $\times$ | | $\times$ |
| Adapting to change | 8 | 4 | 18 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Pride | 9 | 4 | 17 | | $\times$ | $\times$ | | $\times$ |
| Social skills | 9 | 5 | 16 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Control of my own work | 4 | 4 | 16 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Intrinsic motivation to perform | 9 | 4 | 15 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Learning from failures | 8 | 4 | 14 | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
| Prioritisation | 8 | 4 | 14 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Testing | 7 | 4 | 14 | | $\times$ | $\times$ | $\times$ | $\times$ |
| Open office | 6 | 3 | 11 | | $\times$ | $\times$ | | $\times$ |
| Getting buy-in | 8 | 3 | 10 | | | $\times$ | $\times$ | $\times$ |
| Total | | 170 | 742 | 15 | 32 | 33 | 28 | 31 |

14

```
            notions    ┌──────────────┐
  ┌──────────┐of value │ Interpreting │ goals  ┌──────────┐
  │ Ways to  │─────────│ Performance  │────────│ Planning │
  │see success│        └──────────────┘        │ of work  │
  └──────────┘   development  │    │  interaction└──────────┘
               philosophy    /      \
           ┌──────────────┐      ┌──────────────┐
           │Understanding │      │   Need for   │
           │    Agile     │      │communication │
           └──────────────┘      └──────────────┘
```
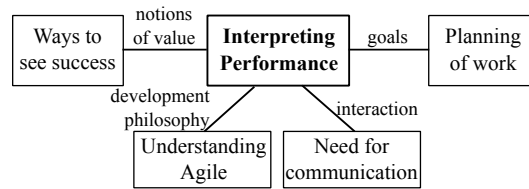
Figure 3: Interpreting Performance describes the activity of defining whether results match the desirable level of performance: meeting, exceeding or transcending predefined objectives. Several factors influence this interpretation, from differing notions of value and understandings of the development philosophy to differences in willingness to engage in social interaction, goal setting and planning.

*4.2. Factors impacting performance*

Participants reported on factors that they perceived to facilitate and disrupt performance. In addition, they reported on factors that they saw as having both a positive and negative effect. Two types of *Performance Facilitators* were reported: factors that the organisation could concretely influence (*Organisational support*), and soft environmental factors (*Atmosphere*) (see Figure 4). *Decision power* placed within teams, and individual autonomy in how tasks are carried out (*Control of my own work*) were seen as two positive factors that the organisation should implement. *Facilitating communication*, e.g. by removing barriers, was also seen as the responsibility of the organisation. This factor was linked to *Atmosphere*, since open communication requires a supportive environment and vice versa. Creating a good atmosphere by facilitating communication was seen as important. Practitioners indicated both that the presence of these factors were beneficial for performance, but also that their absence was detrimental.

> "Last summer our team worked really well, everyone knew what everyone else was doing. We didn't need any formal meetings, communication was natural and direct. Whenever someone from the outside asked us what we were doing, we were able to give a direct answer about our current status and give predictions on when different things would be ready. It was fun to come to work because everything just worked without any extra challenge or effort. [Then] a colleague and I went to work abroad for half a year. When we returned, something had happened to the team. It didn't work any
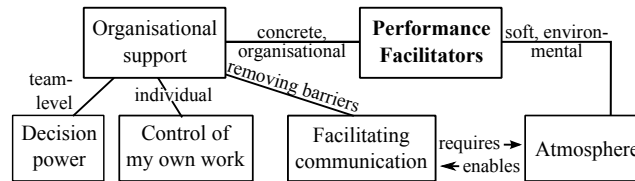
15

Figure 4: *Performance Facilitators* have a positive impact on Performance Alignment Work. Generally, a good *Atmosphere* is beneficial. Organisations can support individuals and teams in different ways. *Facilitating Communication* requires a good *Atmosphere*, which in turn enables facilitation.

> more. [People] didn't talk to each other in the same way as before." (*Team leader, Company B*)

Two strong subcategories emerged as *Performance Disruptors* (see Figure 5). *Distributed work* was seen as having a continuous negative influence on performance, but one which is manageable through increased emphasis on communication. *Reorganisations* were seen as events with a major negative performance impact that requires both time and effort to recover from.

> "When the team changes, all forecasts on team performance like velocity estimates become invalid. I am trying to figure out all the time how much work we can do when the skill set and motivation changes in the team. The [layoff negotiations] have been kind of a trauma in the team. We had to find ways to motivate people to take the work that was left over from the team which was shut down." (*Coach, Company C*)

These were not the only factors that could be construed as negative impact factors, but they were the only ones that were clearly indicated as such in the material. Other negative factors were not as clearly distinguished and their connections to other concepts meant that they were interpreted as being part of other categories.

Three *Double-edged factors* were described (see Figure 6). The potential of an *Open office* to facilitate communication was seen as positive. However, the amount of communication could become disruptive, and participants reported on the need for norms and behavioural signals to avoid the drawbacks. The category *Goal setting*

16

| Distributed work | continuous | **Performance Disruptors** | single events | Re-organi-sation |

Figure 5: *Performance Disruptors* have a negative impact on Performance Alignment Work. Disruptors may be continuous, such as *Distributed Work*, or they can be single events, such as a *Re-organisation*.
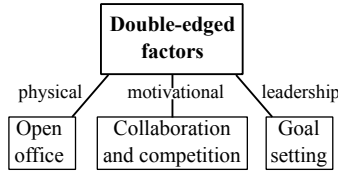


Figure 6: *Double-edged factors* can have both a positive and negative impact on Performance Alignment Work, depending on context. Physical, motivational, and leadership factors can both enhance and worsen performance experiences.

refers to the balance between goals being set outside and within the team. Some participants were firmly against goals being set outside the team because the team had the information needed to set them realistically. Others felt that the team did not always have the expertise or information required to set goals, and that outside guidance in these cases could be beneficial. *Collaboration and competition* included comments regarding competition between teams. While some reported temporary higher performance as a result, cross-team collaboration was seen as more motivational.

Two interrelated categories relate to using automation to facilitate performance (see Figure 7). The category *Testing* reflects the primary means by which participants approached technical quality. *Tools* were frequently mentioned in relation to software development, testing, and communication. Participants reported that they deeply integrate tools into their development process, to the extent that their process-related discourse contains terminology and jargon borrowed from the tools themselves.

> "We have information radiators that show the condition of the code in the version control system. Before we started using git and other related tools, we didn't have very strict control over our code. During the last months, we have made stricter rules. The code in the master branch has to be in
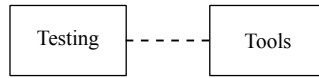
Figure 7: *Testing* and *Tools* were mentioned as two aspects of automation that participants used to facilitate performance. These were interlinked, since tools for automated testing were seen as desirable for continuous quality control. However, participants emphasised that tools alone were not sufficient; to achieve performance gains, tools must fit into the work procedures.
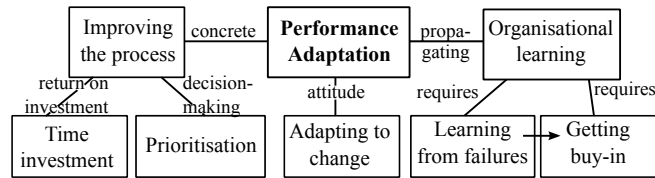


Figure 8: *Performance Adaptation* refers to activities by which people adapt their performance as conditions change. Adaptation may occur through process improvement, whether in a formal sense or in the sense of rapid micro-level changes. Propagating adaptation across an organisation can be seen as an organisational learning activity. Continually *Adapting to change* requires a certain attitude, which is influenced by the perceived source of change – internal or external to the organisation.

> very good shape, so that merges and other version control operations work
> properly." (*Team leader, Company B*)

However, the tools do not necessarily need to be sophisticated; a pragmatic approach was often favoured by the participants.

> "We used an electronic tool for planning, but it didn't really work. Then we
> just started putting notes on the wall for everyone to see. [...] Technology is
> seldom the solution, but it can add efficiency." (*Team leader, Company C*)

### 4.3. Adapting performance

Three main categories emerged that described how participants adapted their performance – in terms of changes to work procedures – when conditions changed (see Figure 8). First, the whole notion of adaptation emerged as an attitude towards *Adapting to change*. Practitioners differed in whether they viewed adaptation as necessary or

18

desirable at all, depending on whether they perceived the source of change as external to the organisation or not. If they did, their attitude was more favourable towards adaptation.

> "When the world is changing, there is no such thing as the old, familiar, and safe. You need to change the way of working." (*Coach, Company D*)

Concrete ways of adaptation centred around *Improving the process*, but the notion of what the process is and how it should be improved differed. One view was that processes are beneficial because they formalise and capture procedures that can be reliably replicated across large parts of the organisation. They were seen as a *Time investment* that must be pragmatically balanced against the gained benefits. Another view was that processes formalise and improve decision-making, helping to avoid overwork and biased decisions through *Prioritisation* procedures.

> "[In the past], there was a constant fight [among project managers on] whose task gets priority. The project managers fighting over team resources used to say that it is impossible to set priorities for new development tasks and bug fixing tasks. We can laugh at it now that we have clear priority order on the Kanban board." (*Coach, Company B*)

A third view was that processes need to be more fine-grained and that separate processes are needed for different parts of the development cycle. This reflected a preference for teams owning their own processes, selecting and developing them for their own particular needs. Participants that expressed this view also viewed processes as ephemeral: they felt that Lean and Agile values and the spirit of a continuous search for better performance were more important than following methods to the letter.

> "Steering mechanisms are completely different in the starting end of an innovation funnel than in the end. In the beginning of the innovation funnel, the impact of a single person is large. A firm process at the beginning of the innovation funnel will kill innovation. At the end of an innovation funnel, the cost and impact of choices grow, and therefore decision-making and

steering mechanisms are different at different stages." (*Management team leader, Company A*)

Another concern for the participants was how to propagate adaptive actions across the organisation (*Organisational learning*). For example, a team could seek to propagate a successful change to work processes or procedures, or they could seek change in another part of the organisation in order to enable further improvement for themselves.

"A good organisation does not change things just for the sake of change. [It] really learns from its mistakes, and does not make them again, [but] the larger an organisation, the harder it is to get it to work well." (*Team leader, Company B*)

*Learning from failures* was seen as occurring first locally in teams, and the challenge was then to convince other teams and the rest of the organisation to adopt the solution (*Getting buy-in*).

"I think the Scrum master needs to let the team fail – once. After the failure, we should discuss together how to avoid such a failure in the future. In practice, it is hard to let the team fail and learn from the failure. Nobody wants to take the [temporary decrease in performance]." (*Coach, Company C*)

"The best way to improve the work of the surrounding organisation is by example. Showing how we do Agile work is more effective than lecturing or forcing. When we increase the visibility of our way of working, little by little elements of it sneak into other teams as well." (*Developer, Company E*)

## 4.4. The high-performing software team

Views, opinions, and experiences of high-performing software teams formed a distinct category in the interview material (see Figure 9). During our analysis, we grouped data fragments and categories related to this theme into a structure that explains how our participants reasoned about high-performing teams. Descriptions of these teams often depicted them as self-directed and resourceful:
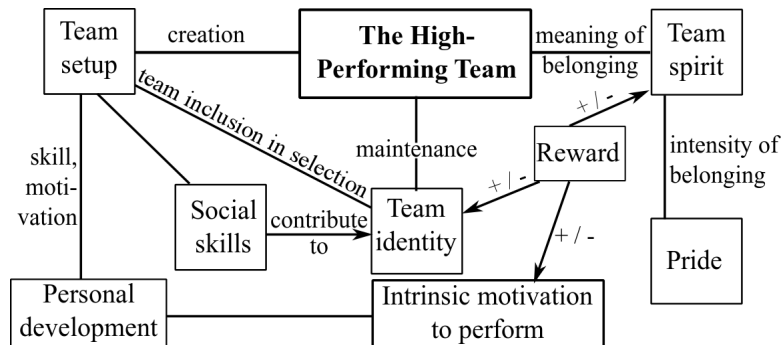
20

Figure 9: *The High-Performing Team* is particularly capable of doing Performance Alignment Work. Such a team has a particular identity and its members have a feeling of *Team spirit* and *Pride*. *Social skills*, an *Intrinsic motivation to perform*, and a desire for *Personal development* are key traits of members in such teams. The *Team setup* phase, as well as changes to the team composition, have critical impact on the team. Also, *Reward*s can influence the team in both positive and negative ways, depending on how they are given.

> "[Based on] my experience, good teams usually have a do-it spirit like a small company. They have established a flow through which people commit to their work and see it as interesting, important and the right thing to do. They put in more than just office hours. [These teams] were formed in an unusual way. There was no line organisation, really not much of an organisation at all. The organisation had a largely fractal structure that changed all the time, like a kaleidoscope. The core is an individual or a team which gets an idea that is larger than life. They combine competencies inside and outside the company and produce a product or service. They find investors for their project, customers, and resources to implement their idea." (*Management team leader, Company A*)

We found some reasoning on the teamwork processes of high-performing software teams. A first category concerned the creation of high-performing teams (*Team setup*). Participants expressed that *Social skills* should be a selection factor for such teams.

> "There are some people who do not fit into a Scrum team and its active way of working. I think that if we are looking for very efficient teams, then their members will be in one end of a scale of social skills. Multitalented people

are usually socially skilled. [...] People who do not have the social skills to work in a Scrum team could work in some other kind of team that does not develop itself or have these scary social practices." (*Coach, Company C*)

Social skills contribute to the formation and maintenance of a powerful *Team identity*. Participants expressed that the team should be included in the selection process to preserve the identity. They expressed that this applies also in the case of adding a new member into the team. Participants indicated that on the individual level, *Intrinsic motivation to perform* spurs a desire for *Personal development*.

"[I'm motivated because] I'm solving my problems for the next two years. [That] motivates to start learning a new area." (*Coach, Company C, discussing setting up a new continuous integration system*)

"We are a company that provides you a chance to learn and develop [yourself] better than anywhere else. [...] In the new organisation, everyone knows that it is important to develop oneself in order to be competitive." (*Manager, Company B*)

Thus participants were motivated to join high-performing teams to both express and develop their own skills, and to gain opportunities for achieving and experiencing high performance. According to some participants, existing teams will respond to that kind of motivation and welcome individuals who are prepared to work hard for exceptional results.

"The team I worked in previously decided to take as our fundamental mode of work to be the best team in the company. When a new guy joined our team, who had been in the company for several years, he said he had never had to work as hard as in our team." (*Team leader, Company C*)

Some participants expressed strong reluctance against letting individuals without these traits join their teams, reflecting a maintenance strategy of the *Team identity*.

"We have chosen people who have an excellent track record, and not tolerated anyone who is not excellent. It might be that it would be better

for [a person] who can't work in an Agile way to work elsewhere." (*Team leader, Company C*)

Other participants took a more inclusive view, reflecting a different kind of identity:

"I feel that having seen the world a little bit, nobody is really that much more [competent] than anyone else. It's always a case of 'personal chemistry'. It is sometimes forgotten that you need someone to facilitate inclusion into the team. Someone who says 'hey, come and join us'." (*Coach, Company E*)

While becoming part of a high-performing team appears to be associated with some challenges, belonging to one can be a meaningful experience. Participants expressed that such teams have a special *Team spirit*: a feeling of belonging to a group of like-minded individuals with a high degree of respect for each others' talents.

"More than a company, this is a coalition of intelligent people who think alike." (*Management team leader, Company A*)

"A good way of working has more to do with the the work spirit than with the methods themselves. If you have a good work spirit, then you are automatically considerate towards others, both in what the team does and how it does it." (*Team leader, Company B*)

When this feeling of belonging was intense, participants expressed a feeling of *Pride*.

"I feel good when I know I've done good quality work. If I cannot be proud of the outcome of my work, I get frustrated. The possibility of having an influence on my own work drives me." (*Developer, Company E*)

Finally, participants expressed how *Reward* could influence the dynamics of the high-performing team. The process of obtaining a reward was seen as more important than the reward itself. It should be linked to an actual episode of good performance, preferably one that the team can influence.

"Promising rewards like money, free trips, or a car to use, in return for meeting a deadline is the wrong place to start. If the team itself wants some

kind of rewards in the form of doing things together, then that is something

to encourage." (*Team leader, Company B*)

A badly motivated or timed reward, given without consideration of how the team itself has perceived its performance, may result in dissonance and emotional rejection of the reward, impacting *Intrinsic motivation to perform*, *Team identity*, and *Team spirit* negatively. Also, the reward should be valued by the team as a group – a social experience could have a better effect than individual monetary rewards.

"It was almost impossible to reach the given target in the given schedule,

but we worked night and day because we had a common dream about

enjoying the sunset in Mexico with the whole team as a reward. And we

did it." (*Manager, Company B, discussing a successful project experience*)

*4.5. Similarities and differences across organisations*

We found several important similarities and differences across the organisations under study. We traced each category which emerged in the first part of analysis back to its source, and examined which categories were represented in the different organisations (see Tables 2 and 3). The number of categories in the small Company A (15) is roughly half of the medium Company E (31), the large Company C (33) and Company D (28), and the very large Company B (32) (see Table 1). We now examine the commonalities and differences across the case organisations.

*4.5.1. Common categories*

Ten important categories were present in all companies: *Need for communication*, *Improving the process*, *Team setup*, *Organisational learning*, *Organisational support*, *Ways to see success*, *Reward*, *Team identity*, *Adapting to change*, and *Learning from failures*. Of these, *Need for communication* and *Ways to see success* belong to the axis *Interpreting performance* (see Section 4.1). This indicates that the main concerns when team members attempt to negotiate an acceptable level of performance among themselves and with other stakeholders are the level of desirability to be involved in social interaction, and the differing options and opinions on how to define success.

Table 3: Core categories present in all case companies, marked using a greyed background, and categories present in all but Company A. The table is a subset of Table 2 and retains the order in that table.

| Top-level category | Companies | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | E |
| Need for communication | × | × | × | × | × |
| Improving the process | × | × | × | × | × |
| Re-organisation | | × | × | × | × |
| Team setup | × | × | × | × | × |
| Tools | | × | × | × | × |
| Decision power | | × | × | × | × |
| Facilitating communication | | × | × | × | × |
| Organisational learning | × | × | × | × | × |
| Organisational support | × | × | × | × | × |
| Planning of work | | × | × | × | × |
| Ways to see success | × | × | × | × | × |
| Reward | × | × | × | × | × |
| Seeing the big picture | | × | × | × | × |
| Collaboration and cooperation | | × | × | × | × |
| Personal development | | × | × | × | × |
| Understanding job roles | | × | × | × | × |
| Goal setting | | × | × | × | × |
| Team identity | × | × | × | × | × |
| Adapting to change | × | × | × | × | × |
| Social skills | | × | × | × | × |
| Control of my own work | | × | × | × | × |
| Intrinsic motivation to perform | | × | × | × | × |
| Learning from failures | × | × | × | × | × |
| Prioritisation | | × | × | × | × |
| Testing | | × | × | × | × |

Only one category from the *Performance Facilitators* axis (see Section 4.2) was present in all companies: *Organisational support*. It is perhaps not surprising that this category was present while the other, more specific categories were not. It appears plausible that support from the organisation is perceived as a facilitating factor for performance in most situations, but that the specifics will vary depending on the type and size of the organisation. The medium and larger companies had many similar facilitating factors, however.

Four common categories, *Improving the process*, *Organisational learning*, *Adapting to change*, and *Learning from failures*, belong to the *Performance Adaptation* axis (see Section 4.3). The reality of many software companies, regardless of size, is a focus on, and belief in, constant change. It appears that among the companies in our sample, all understood the importance of concrete process improvement for facilitating performance adaptation. It is interesting that *Organisational learning* was present in all companies together with the requirement to learn from failures. This indicates that the attitude towards adapting performance includes a tolerance for temporary failure and an understanding of process improvement as learning rather reduction of variability. In other words, we interpret the presence of these four categories in all the case companies as meaning that the aim of performance adaptation is not limited to improvement of technical product quality, but rather an improvement of the ability to search for and develop products that customers want. This "adaptive performance" is one central, common element in all our case companies.

*Team setup*, *Reward*, and *Team identity* were present from our characterisation of *The High-Performing Team* (see Section 4.4). This is significant because it highlights three important factors that may be of general relevance for software development teams. The emphasis on *Team setup* indicates a strong belief in the importance of investing in the creation of teams. Teams are social groups, and their performance depends on group dynamic processes which in turn depend on individual factors of all group members. The data indicate that the conditions for high-performing teams are created during team setup and revisited when teams are changed. *Team identity* further emphasises the insight that regardless of company size, team performance depends on human factors. The identity of the team is its uniqueness compared to other teams. Members of a

high-performing team experience a difference between the in-group that is their team and the out-group, which is other teams and the rest of the organisation. The data indicate that this sense of identity is important for maintaining the team. It also implies that forming teams and selecting new team members is a social process in which team members should be involved.

Finally, and perhaps most significantly, is the presence of *Reward* in all companies under study. In our characterisation of *The High-Performing Team*, this category is an important influencing factor on *Team spirit*, *Team identity*, and *Intrinsic motivation to perform*. As noted in Section 4.4, *Reward* can influence team dynamics, and may be an important social factor that influences performance. Together, the prevalence of these three categories across companies reminds us that the same social psychological mechanisms are present in all teams, regardless of company type or size. It is tempting to consider such soft team factors more in small companies than in large ones, because people are more depersonalised in the latter due to a higher level of abstraction. However, high performance must be built on the level of teams regardless of organisational size.

### 4.5.2. Dissimilarities between organisations

When moving from the small Company A to the medium Company E, the number of categories increases by 18, from 15 to 31, thus encompassing nearly all of the 33 categories in the total set (see Tables 2 and 3): *Performance Awareness* emerges as an axis, with the two categories *Seeing the big picture* and *Understanding job roles*. This reflects a greater degree of ambiguity in the larger organisations. As the company size increases, it appears to become more difficult for individual employees and teams to understand how their work fits into organisational activity on the systemic level. On the individual level, it becomes less clear what is expected given a certain job role. While *Interpreting Performance* includes the definition and communication of success in all companies, *Planning of work* emerges as an additional interpretive activity in the larger companies. Planning was seen as related to performance goals: how goals should be set and pursued.

*Performance Facilitators* (see Section 4.2) become more complex as the organisational size increases. While only *Organisational support* was present in Company A,

data from the other companies also include *Decision power*, *Facilitating communication*, and *Control of my own work*. The larger social structures of these companies raise the importance of factors which may not be as visible in smaller companies. *Decision power* is a team-level support issue where the organisation can delegate authority and give the team a degree of autonomy on decisions which has an impact on its work. The same issue is reflected on the individual level in the category *Control of my own work*. By *Facilitating communication*, the organisation can remove barriers to performance which arise because of the larger numbers of people involved, and the more complex communication hierarchies or networks. In addition to these facilitating factors, the medium- and larger-size organisations include the categories *Tools* and *Testing*. These categories are examples of using automation to facilitate performance. Large investments in automating work may be beyond the capabilities of small organisations, either because of costs or due to a lack of time to properly analyse work practices and implement automation that actually improves performance.

*Re-organisation* was reported as a *Performance Disruptor* (see Section 4.2) in all the medium and larger companies. This should be understood against the contextual background of the organisations. Restructuring, including layoffs, is part of the managerial tactics and strategies for influencing financial and operational performance in these organisations. In contrast, small organisations such as Company A may be dependent on all their employees, and there is little to be gained in terms of performance by altering an already very simple organisational structure. If faced with conditions that would suggest a re-organisation in a larger company, a small company may have to acquire additional funding, seek to be bought by a larger company, or it may go out of business. For these reasons, it is less likely that re-organisation would emerge as a disrupting factor in a small organisation. While re-organisation can be understood as occurring regularly in larger organisations, the question remains if Company E is a typical example of a medium company. This company is special in the sense that its teams operate within other companies on software development projects. Therefore, the presence of re-organisation in the experiences of Company E's participants should be understood more as a reflection of re-organisation within its client companies. Other medium-sized software companies may not have this concern.

28

All the *Double-edged factors* except *Open office* emerge in the medium-, large-, and very large-size companies. The importance of *Collaboration and cooperation*, referring to the balance between cross-team competition and cooperation, seems to increase with company size. The balance of *Goal setting* within and outside the team also enters the experience of the participants, likely because of the increased sources of goals and pressures to meet externally set goals that stem from and depend on remote parts of a larger organisation. Opinions on having an *Open office* also emerged in the larger companies. It is natural that such a physical office arrangement has a smaller impact with fewer employees.

Some of the details in our characterisation of *The High-Performing Team* differed between the companies. *Personal development* was not present in the small Company A. In the other companies, it was linked to an *Intrinsic motivation* to perform and to the set-up of a new team. Being part of a high-performing team was thus often a question of personal development for the participants in the medium and larger companies. In our interpretation, this does not mean that participants in the small company would not aspire to develop personally or be intrinsically motivated. Rather, the emphasis on personal development could stem from the career path requirements in medium and larger companies. Employees may wish to emphasise skills and experience which they have accumulated, in order to motivate a salary raise or transfer to another role with more responsibility or more interesting work tasks. In a smaller company, the possibilities for career advancement in the form of formal titles or roles are more limited. Two other categories emerged in all but one of the medium and larger companies: *Team spirit* and *Pride*. *Team spirit* describes the meaning of belonging to a team: the definition of what sets the team apart from other teams in terms of beliefs, opinions, values, and behaviour. In contrast to *Team identity*, which is the sense that there is a team which is distinct from other social structures, *Team spirit* defines what it means to belong to the team. The intensity of that sense of belonging manifests itself as *Pride* among its members. *Social skills* also emerged in the medium and larger companies as a factor which contributes to the formation of a *Team identity* and which is at play when forming and re-forming a team.

Three categories emerged in the medium and larger organisations under the *Performance Adaptation* axis (see Figure 8). *Prioritisation* refers to improving the decision-making process. Focusing on tasks which have been deemed most important, and setting other tasks aside, seemed to be an important concern in these companies. The reason why this was not reported in the small Company A may be that there is no time in a small company to even consider more than the absolutely essential tasks. In larger companies, it may be easier to create redundant ideas which must then be pruned. Keeping employees in those companies focused on the essentials can be a continuing concern for adapting performance. Another adaptive facet is how to propagate adaptation across the organisation. When a team has found an adaptive strategy, they face more resistance in larger organisations for *Getting buy-in* than in a smaller one.

### 4.5.3. Case company-specific observations

Company D, while being a large organisation, displayed some differences in the details compared to the otherwise similar medium- and larger-size organisations. The categories *Team Spirit*, *Atmosphere*, *Understanding Agile*, *Pride*, and *Open Office* were not present in this company. It is interesting to note that Company D was undergoing heavy restructuring at the time of the interviews. One of the offices of this company was closed shortly after the interviews were completed, and employees were either laid off or forced to relocate to another city. We hypothesise that team members may have tried to cope with the changes by withdrawing from their teams and decreasing their perceived importance. This would explain why several categories related to social factors were not emphasised by participants from this company. Lack of mention of such factors may be a weak signal that can be used to detect a potential dysfunction or stressor. Company C had already completed a restructuring, including layoffs. The presence of the aforementioned categories in that company may indicate that it had mostly recovered from the negative effects. Comments by the interview participants indicate that this was indeed the case. For example:

"I think this was the darkest period of the team and my career in this company. The team talks about the dark period from time to time to

remember that we changed. Not survived, but evolved." (*Coach, Company C*)

Company B did not share the category *Getting buy-in* with Company C, although both have all other categories in common. Based on comments and explanations from the interview participants, we hypothesise that this may be partly due to company culture and differences in how the companies have executed their transitions to Lean and Agile methods. Company B has a history of rigorous management practises and measurement-based process control. Their transformation was well grounded in Lean thinking and management was heavily involved in and led the transformation. Although the implementation of the transformation was carried out in a very short time, it was preceded by an inclusive process which established its reasons and demonstrated management commitment to carrying it through. It was accompanied by an understanding that it was a long-term effort which would require several attempts and learning to reach good results. There was thus less concern among teams of how to convince other teams and other parts of the organisation to adopt insights they had reached. In contrast, the transformation in Company C was based more on Agile principles than Lean thinking. While the transformation in Company C was encouraged by management, it was driven primarily by former project managers who had taken the new role of Scrum Master, and by software development teams themselves. Other parts of the organisation, such as product development, usability and user experience, and sales teams, did not participate in the transformation. The transformation was seen solely as concerning the performance of the software development units. Teams experienced resistance to change in other parts of the organisation, did not perceive long-term commitment by management, and thus struggled to persuade other teams and other internal stakeholders to act according to what they had learned during their own attempts at enhancing performance through Lean and Agile methods. This difference may explain why *Getting buy-in* is not present in Company B while it is present in Company C. Our interpretation is that significant performance improvements require wide involvement from all parts of the organisation. Performance improvement targeting only one function or unit can only have limited results.

## 5. Discussion and limitations

The main finding of our study is that software development practitioners experience team performance in social and behavioural terms. Regardless of organisational size and despite the use of advanced measurement programs in some of the case organisations, the concept of performance is elusive. Practitioners are well aware that that ultimately, performance is judged in the marketplace when customers make purchasing decisions. At the same time, they realise that there are many layers of performance and that good performance in one area does not translate directly into good performance in another area. Our study participants experience performance as a continuous process of negotiation within their teams and with external stakeholders. They are aware of performance aspects on multiple levels of the organisation. They perceive high-performing teams in terms of group processes that link skilled and motivated developers to a powerful team identity. The experience of performance varies with the size and type of organisation, although an important set of core factors are common to all organisations.

Our results indicate that understanding how software practitioners experience the striving for performance in their teams can help formulate hypotheses of how and why their company is currently performing in its software development activities. These can in turn be used in performance improvement initiatives, such as software process improvement, and in improving work-related well-being by giving structure and meaning to the striving for high performance and to foster a culture of sustainable high performance. In this section, we discuss our research questions in light of the results and consider the potential wider implications of the findings for practice.

### 5.1. Addressing the research questions

To address the research questions established in Section 1, we draw on the results presented in Section 4, compare with previous research, and discuss the implications of the results for the research questions. Each research question is addressed in turn.

*5.1.1. RQ1: How do software practitioners experience team performance in Lean and Agile environments?*

Considering RQ1, our results point to the close connection in practitioners' experience between performance and success. They support the claim that practitioners experience both as multifaceted, socially negotiated, changing over time, and sometimes as conflicted between different stakeholders (cf. "success" in Ralph et al. [58]). In our study, Performance Alignment Work is the activity by which practitioners deal with the fluidity of the performance concept (see Figure 1). It contributes to the body of knowledge by describing a particular type of teamwork that specifically addresses performance goals and the process by which software teams attempt to reach them (see e.g. Salas et al. [61] for a discussion of teamwork).

The reported experiences have similarities to earlier findings on team performance (e.g. [19, 21, 38, 61]): as would be expected, many of the concerns expressed by our participants revolve around communication, coordination, and group dynamics. A detailed comparison would be interesting, but is not within the scope of this paper. We instead make a comparison to some of the most similar studies we have found. Whitworth and Biddle [66] qualitatively study how Agile practices mediate the experience of individuals developing software. They find several factors that are uncovered in our study, such as motivation, pride, and social (team) identity. Cedergren and Larsson [14] investigated how software product managers perceive and evaluate performance in large organisations. Their findings indicate that managers are dissatisfied with their current way of evaluating performance. The main focus was found to be on the easily measurable but not necessarily important factors cost, time, and quality. Also, measurements were found to be result oriented rather than process oriented, which means they are difficult to integrate into the management practice. They suggest that in order to improve the current situation, managers should focus on "how the organisation perceives performance and how important performance criteria are being developed". They also observe that the currently used measurements affect the manager's perception of performance, leading to the conclusion that their perception must be changed before any changes in performance evaluation is possible. The Performance Alignment Work theory developed

in the present article provides one way to structure the understanding of stakeholders' experience. It could be used to deepen the understanding of performance and to drive changes in perception, addressing the concerns raised in Cedergren and Larsson's study.

Adolph et al. [2] use a grounded theory approach to examine how software developers manage the development process in software projects. Their research design shows many similarities to ours. The findings also display many similarities. Adolph et al. find their participants engaging in a four-stage process of Reconciling Perspectives, in which individuals attempt to converge their points of view or perspectives about a software project. The emphasis is on the importance of individuals' abilities to both reach out and engage in negotiations, and to protect the project from environmental noise in order to ensure a result is achieved. This is similar to our process of Performance Alignment Work in that both describe a process and both emphasise the aspect of negotiation. Our result shows that such a process is not necessarily confined to projects, but exist on all levels of the organisation. It is a more fundamental process than what may be inferred from the findings of Adolph et al. In another study by Dingsøyr and Lindsjørn [21], study participants take a "forcing" approach to code empirical data into the Salas model [61]. Our results are similar despite the lack of such an anchoring in our study design. Interestingly, our results are more similar in type to the original Salas model [61]: we describe a process rather than a list of impact factors. The question is what type of result is most relevant for a particular need, and future research should carefully choose a research design that produces the desired type of result. In any case, the similar results of the studies indicate that our results can have validity beyond the context in which it was performed.

### 5.1.2. RQ2: How do software practitioners reason about the relationships between perceived performance factors?

RQ2 concerned how practitioners reason about the relationships between perceived performance factors. Two main findings arose from our analysis. The perceived factors concern on one hand the components of social negotiation of performance (Performance Alignment Work, Sections 4.1–4.3) and on the other hand an understanding of why some teams are high-performing (Section 4.4). The components of Performance Align-

ment Work show that our participants reasoned about performance on all levels from their individual and team performance to the performance of their organisation in the marketplace and in terms of customer satisfaction. This indicates that practitioners have local or personal theories and beliefs regarding how their work influences team and organisational performance. Activities that raise these theories and beliefs into awareness and help align them could help improve actual performance in teams and organisations. For example, performance measurement programs could be conducted as participative design programs in which teams are deeply involved. A challenge then is how to balance the concerns of different stakeholders when they conflict, a question which ultimately requires consideration of ethics and values.

Based on our findings, a high-performing team is one that is exceptionally good at Performance Alignment Work (see Figures 1 and 9). Not only can it continuously organise itself internally to optimise performance, but it also engages with other teams, other parts of its organisation, and with stakeholders outside the organisation (cf. Ancona's X-Teams [4]). It both elicits their performance needs and preferences, and influences their performance expectations and alignment in a manner that is favourable to its goals. Our description of the high-performing team contributes to theory in two ways. First, it provides a proposition that is testable in specific cases: that high-performing teams are considered high-performing because they influence the criteria by which they are judged. This may appear as counter-intuitive since the conventional understanding of high performance is meeting or exceeding objectives set in advance. However, the reality experienced by the participants in this study changes so rapidly that objectives become obsolete as soon as they are set. Teams must influence stakeholder expectations and objectives if their work is to have any value at all. The objective must evolve along with the environment and the solution under development. Also, the act of setting a goal or defining criteria for success itself has already modified the environment. Being involved in that modification is what our study participants perceive as a central trait of high-performing teams. To what extent this experience stems from the Lean and Agile approach or from true changes in the market is an open question.

Second, our high-performing team description shows that there are important affective aspects, and conative aspects beyond motivation, that should be considered in

research on software team performance. The affective aspects were present in practitioners' notions of a team identity, team spirit, pride, and in their reasoning of rewards as tools for increasing performance. Identification with a team has been established as a determinant of affective outcomes such as job satisfaction [51]. We note that an interesting question is how to characterise team identity, its formation, and its relationship to company identity, corporate culture, and values. Weak indications in our data cause us to suspect that there were important differences in corporate culture between the companies in our sample. For example, participants from some companies expressed views of exclusivity – that some individuals should be explicitly excluded to create high-performing elite teams – while others expressed inclusive views – that better performance can result from diversity and a welcoming attitude. These are examples of values that form different organisational cultures.

Research results on motivation have so far been inconclusive regarding how and by what software engineers are motivated, and what the benefits of motivating them are [8, 29]. The conative aspects in the high-performing team description imply that (intrinsic) motivation on the individual level does not directly lead to performance on the team level. There are also volitional processes involved as teams go through the process of accepting a new member. Aligning affective and conative aspects within the team could be another means to reach better performance. However, due to personal characteristics, e.g., personality and values, it may be more pragmatic to strive for such alignment when constructing teams rather than trying to change its members afterwards. Another question is how to transfer the team spirit, identity, and other soft factors of the team culture to a new member. Further research on these topics is called for.

### 5.1.3. RQ3: How do performance factors experienced by software practitioners differ between different types of companies?

RQ3 concerned the similarities and differences of how practitioners experience performance in different kinds of companies. We identified a core set of ten performance-related categories which were common to all organisations under study (see Table 3). This core set represents roughly one third of the total amount of categories. When excluding the small Company A, the remainder of the companies had almost all categories

36

in common, suggesting that the complete set is relevant in medium- and larger-size companies. These findings contribute to theory by establishing the organisational context within which the results apply, and by detailing the dynamics of the Performance Alignment Work theory as an organisation changes in size. Furthermore, our study shows that many important factors, such as the effects of re-organisation, are visible in the data as differences on the organisational level. The analysis carried out in our study could thus be applied in other cases to form hypotheses regarding software development performance.

The ten-category core set shown in Table 3 concerns two of the three phases in the continuous Performance Alignment Work process: *Interpreting Performance* and *Performance Adaptation*. This suggests that the largest differences in how practitioners experience the alignment process in companies of different sizes lie in how they become aware of performance-related concerns. In smaller companies, awareness may come more directly and require less activity to solicit information to support awareness. In larger companies, it appears to be more difficult for employees to become aware of performance concerns on the individual, team, unit, organisation, and market levels. The commonalities in *Interpreting Performance* suggest that practitioners across organisations of different sizes share the experience of continuously communicating about values that underlie the definition of success. Whether this happens explicitly or is implicitly present in discussions about goals and means to achieve them, values seem to be an important factor for performance. In all companies, participants shared experiences of varying attitudes towards *Adapting to change*, and viewed adaptation as a process of experimentation – trial, error, analysis, and improved insight – with tolerance for temporary failure.

Notions of *The High-Performing Team* had a common core that concerned the creation and identity of teams, as well as the role of rewards as an influencing factor on identity, intrinsic motivation, and team spirit. In all except the small Company A, the common set of categories was larger, indicating that in medium- and larger-sized organisations, high-performing teams are based on essentially the same ingredients and can be formed, understood, and supported by similar mechanisms regardless of organisational size.

*5.2. Threats to validity*

Following Yin [70], Runeson and Höst [60] consider construct, internal, and external validity to be important criteria for case study research in software engineering. Other sets of validity criteria are commonly used in grounded theory and qualitative research, such as fit, relevance, workability, and modifiability [32]. What the researcher does when designing a study, collecting and analysing data, and reporting results, has implications on all areas of validity.

Construct validity is the extent to which the operational measures studied actually represent what the researcher intends and what is investigated according to the research questions [60]. Following Yin [70], we have strengthened construct validity by carefully establishing a chain of evidence from interview data to analysis results, and by using participant checking of intermediate results.

Internal validity concerns whether the study controls for potential confounding factors when examining causal relationships [60]. While construct validity in case studies is mostly addressed in data collection and reporting, it interacts with internal validity during data analysis. Yin [70] notes that internal validity is not a primary concern in case study research but suggests pattern matching and explanation-building as tactics to address internal validity during data analysis.

Interviews have inherent threats to validity that relate to the ability and willingness of participants to report on desired topics [70], and to the possibilities of generalising interview-based findings [48]. In planning the interviews, we chose interview questions that allow various kinds of factors to emerge: the questions do not favour one aspect of high performance over another.

Bias was alleviated in data collection by having two participating researchers, and in analysis by having four researchers, discuss and agree on the emerging results. The analysis methods ensure that an individual piece of information cannot dominate the overall result, since it must be matched with other pieces of information to form a meaningful higher-order structure. The coding methods used in this study are well known, and thoroughly documented in the literature [11, 12, 17, 32, 44, 53, 62].

External validity is the extent to which findings are possible to generalise [60]. The theoretical contributions of this paper are developed in, and describe, the local

situation in the participating companies. The theory itself has been participant to accuracy-checking procedures – triangulation on participant, company, and researcher levels – and can be judged in terms of how well it reflects the reality of the participants and the phenomena under study [3]. This is related simultaneously to construct as well as external validity. We gathered feedback on the emerged categories from one senior representative from each company who was not an interview participant. The representatives indicated that all categories were important, and they found it difficult and essentially meaningless to give the categories a forced order of prioritisation: they saw the performance phenomenon as holistic rather than in terms of individual pieces. Our interpretation is that the study has a high degree of relevance and fit, reflecting participants' reality well. Yin [70] suggests using replication logic in multiple case studies to strengthen external validity. With certain limitations, qualitative findings may be generalised to some broader theory through replication [70]. Further studies are needed for such generalisation. Our contribution has high modifiability, as it can be altered to address new relevant data – a criterion considered important for grounded theories [32].

A certain measure of convenience sampling is almost always present in practical studies [53]. There is an apparent bias in the sample towards coaches and team leaders. However, the sample represents persons who are highly regarded in their respective organisations, who have had exposure to many different teams and projects, and who have experienced both high and low performance in different situations. Results could differ with practitioners having considerably less or more experience. On the company sampling level, there is a bias towards medium- and large-size companies and a particular underrepresentation of small companies. This reduces the possibilities of generalising our findings. However, it is likely that the inclusion of more small companies would result in more variation rather than less, thus increasing the likelihood that the categories that were present in the medium- and larger-size companies would also be found among small companies. Further replication is required before this can be assessed.

Since the study was conducted in a Finnish setting, cultural bias is an important consideration. In Hofstede's cultural dimensions [39], Finland belongs to a cluster of primarily western countries, but with important unique characteristics: low power dis-

tance, high individualism and femininity (preferring quality of life over being the best), medium high uncertainty avoidance, and a short-term orientation. Cultures differently positioned on these scales could be expected to emphasise different aspects than in our material. However, the categories show a fairly even representation of perceptions across Hofstede's dimensions. Both individualistic (e.g. *Personal development* and *Intrinsic motivation to perform*) and collectivist points of view were represented (e.g. *Team identity* and *Team spirit*). It is also important to remember that corporate culture may partially override national culture at work, particularly in multinational corporations [63]. Nevertheless, culture should be considered when attempting to generalise the findings.

Finally, leadership and management styles have been shown to impact performance experiences and results in organisations (e.g. [33]). Since we did not specifically investigate this aspect in our study, we cannot make a detailed assessment of its impact on the results. Nevertheless, many of the factors are related to leadership and management style, and this factor can be considered to be embedded as a latent factor in the results. The commonalities between the case companies in our results suggest that leadership and management style did not play a large role in these cases; however, see the discussion on *Getting buy-in* at the end of Section 4.5.2.

## 6. Conclusions

In this article, we report on a study that explores how practitioners experience and reason about team performance in a changing environment. We conducted a multiple-case study in which we interviewed practitioners from five companies that use Lean and Agile approaches and operate in volatile markets. This article extends a previous study on the same subject [27].

Through the study, we showed why it is not sufficient to consider performance only as meeting predefined objectives: objectives themselves change and are subject to an interpretive dialogue in which software teams can be an influential stakeholder. Practitioners understand performance on many levels, ranging from individuals and teams to organisations, markets, and customers. They hold complex local theories and

beliefs regarding performance and the mechanisms that result in high or low performance, as interpreted by different stakeholders. Although there are variations between companies of different sizes, our findings are consistent over medium- and large-size companies, and the main findings apply to companies of all sizes in our sample. Our study illustrates how hypotheses regarding the performance of software development organisations can be formed from grounded data analysis in other cases. Our expectation is that a better understanding of the experience of performance is an important component in improving work conditions while also improving actual performance. Leading software development teams towards high performance requires sensitivity to soft factors and negotiation skills. A transparent organisation that provides access to timely performance-related information can support the creation of high-performing teams. Such teams are best positioned to offer their skills and capabilities when they are engaged in multi-directional negotiation, which may also open possibilities for using their performance advantage for strategic development in the organisation.

The present study builds on the state of the art in several important ways. First, it shows that not only managers are involved in defining the meaning of performance: all stakeholders from individuals, through teams and units, to the organisation and market, are involved in the definition. This is not unique to large organisations, and while there are some differences in the factors that are perceived in organisations of different sizes, the mechanism of experiencing performance is common to organisations regardless of size and field. Second, it shows that there are several definitions in existence at the same time, and the persons holding those definitions influence each other continuously. Performance is thus not a static but a dynamic construct. Third, our study shows a way to unpack stakeholders' perception and experience of performance: it can be viewed as a continuous process of alignment. It is thus not necessary nor advisable to begin performance improvement by trying to define precise measures of it. The starting point can rather be to understand the process by which performance gets defined at different levels of an organisation at different times. Improvement can then be focused on uncovering what the definition is at different levels, and negotiating how to bring the definitions into alignment. If necessary, measurement can then be used to monitor the mutually aligned definitions. This approach is centred more around a social and

behavioural view of software development and team performance than a technical and engineering approach. It proposes to design software development work first and foremost from a human standpoint.

We envision three future directions of research. First, we hypothesise that the experience of software development in general, and its performance in particular, arises from basic social psychology, e.g. beliefs, norms, and values. A better understanding of how to apply these in software engineering research could help explain specific understandings of performance. The interplay between individual and group factors, such as the impact of different leadership and management styles on high-performing teams, is of particular interest. Second, the link between software development paradigms and the experience of developers is open to enquiry. How, for instance, do practitioners alter their behaviour as a result of interpreting the principles of Lean and Agile software development, or other approaches to software development? Third, what are the relationships between different stakeholders' evaluations of performance? For example, what aspects of good performance do developers and customers agree or disagree on? Accounts of such agreement and disagreement could help software development organisations improve the experience of both developers and customers. Together, these directions suggest an investigation into software development as a co-experienced process. Future co-creation of software-intensive products and services could benefit from an expanded understanding of the software development process, grounded in social and behavioural science, where the aim is to manage the experience of all stakeholders during the entire life-cycle of the software product or service.

# References

[1] S.T. Acuña, Marta Gómez, and Natalia Juristo, *Towards understanding the relationship between team climate and software quality – a quasi-experimental study*, Empirical Software Engineering **13** (2008), no. 4, 401–434.

[2] S. Adolph, P. Kruchten, and W. Hall, *Reconciling perspectives: A grounded theory of how people manage the process of software development*, Journal of Systems and Software **85** (2012), no. 6, 1269–1286.

[3] David L. Altheide and John M. Johnson, *Criteria for assessing interpretive validity in qualitative research*, Handbook of qualitative research, 1994, pp. 485–499.

[4] Deborah Ancona, Henrik Bresman, and Katrin Kaeufer, *The Comparative Advantage of X-Teams*, MIT Sloan Management Review **43** (2002), no. 3, 33–39.

[5] Päivi Åstedt-Kurki and Riitta-Liisa Heikkinen, *Two approaches to the study of experiences of health and old age: the thematic interview and the narrative method*, Journal of Advanced Nursing **20** (1994), no. 3, 418–421.

[6] Nathan Baddoo, Tracy Hall, and Dorota Jagielska, *Software developer motivation in a high maturity company: A case study*, Software Process Improvement and Practice **11** (2006), no. 3, 219–228 (English).

[7] Steve Bannerman and Andrew Martin, *A multiple comparative study of test-with development product changes and their effects on team speed and product quality*, Empirical Software Engineering **16** (2011), no. 2, 177–210.

[8] Sarah Beecham, Nathan Baddoo, Tracy Hall, Hugh Robinson, and Helen Sharp, *Motivation in Software Engineering: A systematic literature review*, Information and Software Technology **50** (2008), no. 9–10, 860–878.

[9] G.R. Bergersen and J.-E. Gustafsson, *Programming Skill, Knowledge, and Working Memory Among Professional Software Developers from an Investment Theory Perspective*, Journal of Individual Differences **32** (2011), no. 4, 201–209.

[10] G.R. Bergersen, J.E. Hannay, D.I.K. Sjøberg, T. Dybå, and A. Karahasanovic, *Inferring Skill from Tests of Programming Performance: Combining Time and Quality*, International Symposium on Empirical Software Engineering and Measurement, 2011, pp. 305–314.

[11] Hugh Beyer and Karen Holtzblatt, *Contextual design*, Interactions **6** (1999), 32–42.

[12] Hugh Beyer and Karen Holztblatt, *Contextual Design*, Morgan Kaufmann Publishers, 1997.

[13] Gul Calikli and Ayse Bener, *Empirical Analyses of the Factors Affecting Confirmation Bias and the Effects of Confirmation Bias on Software Developer/Tester Performance*, Proceedings of the 6th International Conference on Predictive Models in Software Engineering, 2010, pp. 10:1–10:11.

[14] Stefan Cedergren and Stig Larsson, *Evaluating performance in the development of software-intensive products*, Information and Software Technology **56** (2014), no. 5, 516–526.

[15] European Commission, *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises (2003/361)*, 2003. Online: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32003H0361 (Retrieved 2014-08-29).

[16] Juliet Corbin and Anselm Strauss, *Basics of qualitative research: Techniques and procedures for developing grounded theory*, 3rd ed., SAGE Publications, Inc., Thousand Oaks, CA, 2007.

[17] John Creswell, *Research design: Qualitative, quantitative, and mixed methods approaces*, 3rd ed., SAGE Publications, Inc., 2009.

[18] Bill Curtis, Herb Krasner, and Neil Iscoe, *A field study of the software design process for large systems*, Communications of the ACM **31** (1988), no. 11, 1268–1287.

[19] T. Dickinson and R. McIntyre, *A conceptual framework of teamwork measurement* (Michael T Brannick, Eduardo Salas, and Carolyn W Prince, eds.), Psychology Press, 1997.

[20] T. Dingsøyr and Tore Dybå, *Team effectiveness in software development: Human and cooperative aspects in team effectiveness models and priorities for future studies*, 5th International Workshop on Cooperative and Human Aspects of Software Engineering, 2012, pp. 27–29.

[21] Torgeir Dingsøyr and Yngve Lindsjørn, *Team Performance in Agile Development Teams: Findings from 18 Focus Groups*, Agile Processes in Software Engineering and Extreme Programming, 2013, pp. 46–60.

[22] Siva Dorairaj, James Noble, and Petra Malik, *Understanding Lack of Trust in Distributed Agile Teams: A grounded theory study*, 16th International Conference on Evaluation and Assessment in Software Engineering, 2012, pp. 81 –90.

[23] Neil Douglas and Terry Wykowski, *From Belief to Knowledge: Achieving and Sustaining an Adaptive Culture in Organizations*, CRC Press, USA, 2010.

[24] Kathleen M. Eisenhardt, *Building Theories from Case Study Research*, The Academy of Management Review **14** (1989), no. 4, pp. 532–550 (English).

[25] Kathleen M. Eisenhardt and Melissa E. Graebner, *Theory Building From Cases: Opportunities And Challenges*, Academy of Management Journal **50** (2007), no. 1, 25–32.

[26] Albert Endres and Dieter Rombach, *A Handbook of Software and Systems Engineering. Empirical Observations, Laws and Theories*, The Fraunhofer IESE Series on Software Engineering, Addison Wesley, 2003.

[27] Fabian Fagerholm, Marko Ikonen, Petri Kettunen, Jürgen Münch, Virpi Roto, and Pekka Abrahamsson, *How Do Software Developers Experience Team Performance in Lean and Agile Environments?*, Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, 2014, pp. 7:1–7:10.

[28] Fabian Fagerholm and Jürgen Munch, *Developer Experience: Concept and Definition*, International Conference on Software and System Process, 2012, pp. 73–77.

[29] A. Cesar C. Franca, Tatiana B. Gouveia, Pedro C. F. Santos, Celio A. Santana, and Fabio Q. B. da Silva, *Motivation in software engineering: A systematic review update*, 15th Annual Conference on Evaluation and Assessment in Software Engineering, 2011, pp. 154–163.

[30] M. Freeman and P. Beale, *Measuring project success*, Project Management Journal **23** (1992), no. 1, 8–17.

[31] Deborah L. Gladstein, *Groups in Context: A Model of Task Group Effectiveness.*, Administrative Science Quarterly **29** (1984), no. 4, 499–517.

[32] Barney Glaser and Anselm Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine Transaction, Chicago, 1967.

[33] George B. Graen and Mary Uhl-Bien, *Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective*, The Leadership Quarterly **6** (1995), no. 2, 219–247.

[34] D. Graziotin, X. Wang, and P. Abrahamsson, *Are happy developers more productive? The correlation of affective states of software developers and their self-assessed productivity*, Lecture Notes in Computer Science **7983 LNCS** (2013), 50–64.

[35] J. Hackman, *The design of work teams*, Handbook of organizational behavior, 1987, pp. 315–324.

[36] _____ , *Why teams don't work*, Theory and research on small groups, 1998, pp. 245–267.

[37] Tracy Hall, Dorota Jagielska, and Nathan Baddoo, *Motivating developer performance to improve project outcomes in a high maturity organization*, Software Quality Journal **15** (2007), no. 4, 365–381 (English).

[38] Martin Hoegl and Hans Georg Gemuenden, *Teamwork Quality and the Success of Innovative Projects: A Theoretical Concept and Empirical Evidence.*, Organization Science **12** (2001), no. 4, 435–449.

[39] Geert Hofstede, *Culture's Consequences: International Differences in Work-Related Values*, Abridged, SAGE Publications, California, 1980.

[40] Helena Holmström Olsson, H. Alahyari, and J. Bosch, *Climbing the "Stairway to Heaven" – A Mulitiple-Case Study Exploring Barriers in the Transition from Agile Development towards Continuous Deployment of Software*, 38th EUROMICRO Conference on Software Engineering and Advanced Applications, 2012, pp. 392–399.

[41] Risa B. Hyman, *Creative chaos in high performance teams: An experience report*, Communications of the ACM **36** (1993), no. 10, 57–60 (English).

[42] Marko Ikonen and Pekka Abrahamsson, *Operationalizing the Concept of Success in Software Engineering Projects*, International Journal of Innovation in the Digital Economy **2** (2011), no. 3, 11–37.

[43] John R. Katzenbach and Douglas K. Smith, *The Wisdom of Teams: Creating the High-Performance Organization*, McKinsey & Company, Inc., New York, NY, 1993.

[44] J. Kawakita, *The Original KJ Method*, Kawakita Research Institute, Tokyo, 1982.

[45] Petri Kettunen, *Bringing Total Quality in to Software Teams: A Frame for Higher Performance*, Lean Enterprise Software and Systems, 2013, pp. 48–64.

[46] Iftikhar Ahmed Khan, Willem-Paul Brinkman, and Robert M. Hierons, *Do moods affect programmers' debug performance?*, Cognition, Technology & Work **13** (2011), no. 4, 245–258.

[47] Robert E. Kraut and Lynn A. Streeter, *Coordination in software development*, Communications of the ACM **38** (1995), 69–81.

[48] Steinar Kvale, *Doing Interviews*, SAGE Publications Ltd, Thousand Oaks, CA, 2007.

[49] Ting-Peng Liang, Chih-Chung Liu, Tse-Min Lin, and Binshan Lin, *Effect of team diversity on software project performance*, Industrial Management and Data Systems **107** (2007), no. 5, 636–653.

[50] Kurt R Linberg, *Software developer perceptions about software project failure: a case study*, Journal of Systems and Software **49** (1999), no. 2–3, 177–192.

[51] Abigail Marks and Cliff Lockyer, *Debugging the system: the impact of dispersion on the identity of software team members*, The International Journal of Human Resource Management **16** (2005), no. 2, 219–237.

[52] Joseph E McGrath, *Toward a "theory of method" for research on organizations*, New perspectives in organization research **533** (1964), 533–547.

[53] Sharan Merriam, *Qualitative research: a guide to design and implementation*, 2nd ed., Jossey-Bass, 2009.

[54] Sarma R. Nidumolu, *A comparison of the structural contingency and risk-based perspectives on coordination in software-development projects*, Journal of Management Information Systems **13** (1996), no. 2, 77–113.

[55] Ikujiro Nonaka, Ryoko Toyama, and Akiya Nagata, *A firm as a knowledge-creating entity: a new perspective on the theory of the firm*, Industrial and corporate change **9** (2000), no. 1, 1–20.

[56] C. Passos, A.P. Braun, D.S. Cruzes, and M. Mendonca, *Analyzing the Impact of Beliefs in Software Project Practices*, International Symposium on Empirical Software Engineering and Measurement, 2011, pp. 444–452.

[57] Paul Ralph, Pontus Johnson, and Howell Jordan, *Report on the First SEMAT Workshop on General Theory of Software Engineering*, SIGSOFT Software Engineering Notes **38** (2013), no. 2, 26–28.

[58] Paul Ralph and Paul Kelly, *The Dimensions of Software Engineering Success*, Proceedings of the 36th International Conference on Software Engineering, 2014, pp. 24–35.

[59] R.H. Rasch and H.L. Tosi, *Factors affecting software developers' performance: An integrated approach*, MIS Quarterly: Management Information Systems **16** (1992), no. 3, 395–409.

[60] Per Runeson and Martin Höst, *Guidelines for conducting and reporting case study research in software engineering*, Empirical Software Engineering **14** (2009), no. 2, 131–164 (English).

[61] Eduardo Salas, Dana E. Sims, and C. Shawn Burke, *Is there a "Big Five" in Teamwork?*, Small Group Research **36** (2005), no. 5, 555–599.

[62] Johnny Saldaña, *The Coding Manual for Qualitative Researchers*, SAGE Publications, Ltd., 2009.

[63] Susan C. Schneider, *National vs. Corporate Culture: Implications for Human Resource Management*, Human Resource Management **27** (1988), no. 2, 231–231.

[64] Thomas Shaw, *The Emotions of Systems Developers: An Empirical Study of Affective Events Theory*, Proceedings of the 2004 SIGMIS Conference on Computer Personnel Research: Careers, Culture, and Ethics in a Networked Environment, 2004, pp. 124–126.

[65] Goparaju Purna Sudhakar, Ayesha Farooq, and Sanghamitra Patnaik, *Soft factors affecting the performance of software development teams*, Team Performance Management **17** (2011), no. 3, 187–205.

[66] E. Whitworth and R. Biddle, *The Social Nature of Agile Teams*, Agile Conference (AGILE), 2007, pp. 26–36.

[67] M.R. Wrobel, *Emotions in the software development process*, 2013, pp. 518–523.

[68] Hee-Dong Yang, Hye-Ryun Kang, and Robert M. Mason, *An exploratory study on meta skills in software development teams: Antecedent cooperation skills and personality for shared mental models*, European Journal of Information Systems **17** (2008), no. 1, 47–61.

[69] Heng-Li Yang and Jih-Hsin Tang, *Team structure and team performance in IS development: A social network perspective*, Information and Management **41** (2004), no. 3, 335–349 (English).

[70] Robert Yin, *Case study research: design and methods*, 4th ed., SAGE Publications, Inc., 2009.

## Appendix A. Interview guide

Table A1 shows the guide used for the thematic interviews in this study. The guiding questions are examples of questions that were used to transition into the thematic areas. The application of the guide was informed by constant comparison and purposive sampling (see main text, Section 3).

Table A1: Guide for thematic interviews.

| Theme | Guiding questions | Purpose |
|---|---|---|
| Personal background | When did you start working here? In what role? What kind of role and job do you have now in this company? | Help participant recall early stages of career and state of company at that time. Elicit comparison with current role. |
| Team | Are you part of some team, or several? How long has this team existed? Have you been involved in it from the beginning? When did you become involved? How has the team composition changed along the way? | Understand participant's previous and current involvement in a team/teams. |
| Work environment / Organisation | How has this company changed during your time here? Have you encountered Agile and Lean in your work? How has this (Agile/Lean/other principle) changed the working habits of your team? | Elicit discussion about participant's views on the organisation and working environment. In particular, Agile and Lean topics inform us how the organisation functions. |
| Experience examples | Give examples of successful (unsuccessful) work experiences. On what scale did you succeed/fail? | Help participant to recall concrete and meaningful experiences of performance (high or low) and success (failure). |
| Team performance | In what way and why do you think you succeeded (failed)? What would be needed for a team to always do as well as or better than in your example? What do you think was the reason for the failures? Did the failure somehow benefit the team or the organisation? | Situate the concrete examples in a team context and elicit discussion about causes and effects, and about potential organisational learning. |
| Quality | What is a good (bad) team? What is a good (bad) organisation? What is a good (bad) result or product? What are your thoughts about good work, good ways of working, and well-being at work? | Elicit discussion about quality and meaning of teams, organisations, products, and work. |