



Do scaling agile frameworks address global software development risks? An empirical study

Sarah Beecham^{a,*}, Tony Clear^b, Ramesh Lal^b, John Noll^c

^a Lero - the Science Foundation Ireland Research Centre for Software, University of Limerick, Limerick, Ireland

^b School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

^c School of Physics, Engineering, and Computer Science, University of Hertfordshire, Hatfield, Hertfordshire, UK

ARTICLE INFO

Article history:

Received 4 October 2019

Received in revised form 21 August 2020

Accepted 11 September 2020

Available online 16 September 2020

Keywords:

Global software development (GSD)

Risks

Scaling agile frameworks

Scaled Agile Framework (SAFe)

Disciplined Agile Delivery (DAD)

Empirical study

ABSTRACT

Driven by the need to coordinate activities of multiple agile development teams cooperating to produce a large software product, software-intensive organizations are turning to scaling agile software development frameworks. Despite the growing adoption of various scaling agile frameworks, there is little empirical evidence of how effective their practices are in mitigating risk, especially in global software development (GSD), where project failure is a known problem.

In this study, we develop a GSD Risk Catalog of 63 risks to assess the degree to which two scaling agile frameworks—Disciplined Agile Delivery (DAD) and the Scaled Agile Framework (SAFe)—address software project risks in GSD. We examined data from two longitudinal case studies implementing each framework to identify the extent to which the framework practices address GSD risks.

Scaling agile frameworks appear to help companies eliminate or mitigate many traditional risks in GSD, especially relating to users and customers. However, several important risks were not eliminated or mitigated. These persistent risks in the main belonged to the Environment quadrant highlighting the inherent risk in developing software across geographic boundaries. Perhaps these frameworks (and arguably any framework), would have difficulty alleviating, issues that appear to be outside the immediate control of the organization.

© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the widespread adoption of agile methods in companies of all sizes, there is an increasing need to scale-up agile development beyond a single co-located team (Fitzgerald and Stol, 2017). This might involve multiple teams collaborating to produce different subsystems of a single product, multiple teams producing separate products that are part of a product family, or scaling agile development upward in the organizational hierarchy, from teams to program management to the executive suite (Dingsoyr and Moe, 2014). Each of these cases may also require scaling the geographic distance between teams, or team members on the same team, to accommodate Global Software Development (GSD).

Companies engage in GSD for several reasons, including access to new markets, varied skill sets, and lower labor rates (Noll et al., 2010; Ebert et al., 2008; Nurdiani et al., 2011). GSD, however, is accompanied by a *risk tariff*: software development in a co-located context is already a complex undertaking; globalization

increases this complexity, due to the “temporal, geographical and socio-cultural dispersion of stakeholders (managers, developers, clients...)” (Chadli et al., 2016), which introduce delays, misunderstandings, and mistrust (Noll et al., 2010). There are additional technical integration issues also to consider Cataldo and Herbsleb (2011), where the architecture needs to consider cross-site coupling and system inter-dependencies (Sievi-Korte et al., 2019a). It would appear therefore that when it comes to GSD, and “teams operating over the variety of boundaries, special attention has to be paid to leveraging the related risks” (Šmite, 2007).

A number of frameworks have been created to guide organizations through scaling agile development, including Scrum-of-scrums,¹ Large Scale Scrum (LESS),² the Spotify Model,³ Nexus,⁴ Scrum at Scale,⁵ Disciplined Agile Delivery (DAD),⁶ and the Scaled

¹ <http://www.agilealliance.org/glossary/scrum-of-scrums>

² <http://less.works>

³ <http://blog.crisp.se/2012/11/14/henrikkniberg/scaling-agile-at-spotify>

⁴ <https://www.scrum.org/resources/nexus-guide>

⁵ <https://www.scruminc.com/scrum-scale-case-modularity>

⁶ <http://www.disciplinedagiledelivery.com>

* Corresponding author.

E-mail addresses: sarah.beecham@lero.ie (S. Beecham),

tony.clear@aut.ac.nz (T. Clear), ramesh.lal@aut.ac.nz (R. Lal), j.noll@herts.ac.uk (J. Noll).

Agile Framework (SAFe®).⁷ Each of these has its advocates and critics, but to date, there is scant empirical evidence as to their efficacy in general, and specifically in managing risk in various settings, such as Global Software Development (GSD), where teams are distributed around the world.

This study focuses on GSD, known to have a specific set of risks to consider, that are additional to those identified by Oehmen et al. (2014) and Wallace and Keil (2004). These include task distribution, knowledge management, geographical distribution, cultural differences, and communication infrastructure (Persson et al., 2009; Verner et al., 2014). As many organizations are now adopting agile or hybrid development methodologies (Noll and Beecham, 2019; Tell et al., 2019) in globally distributed organizations (Marinho et al., 2019) that are scaling (Razzak et al., 2018), we ask,

How does the adoption of scaling agile framework practices address global software development risks?

To answer this question, we present an industrial multiple case study (Verner et al., 2009; Bass et al., 2018a) of the adoption of two scaling agile frameworks—Disciplined Agile Delivery (DAD) (Ambler and Lines, 2011), and the Scaled Agile Framework (SAFe) (Leffingwell et al., 2016).

In the next section, we provide background on project risks in Global Software Development, and give a short overview of our two scaling agile frameworks, DAD and SAFe. Following that, in Section 3, we describe our methodology, and present our results in Section 4. We discuss the implications of the results in Section 5, along with the study limitations, and conclude with a summary of our study and contribution in Section 6.

2. Background

In this section, we look to the literature to provide an outline and define types of risk in global software development projects, and then consider some of the agile frameworks that support large-scale software development. We complete this section with our research question derived from the need to have a better understanding of the impact these new frameworks might have on project success or failure.

2.1. Software project risk

The literature on software development risks extends back some years; for instance, Boehm identified a top 10 list of software project risks in 1991 (Boehm, 1991). In the early 2000s a rigorous set of studies combining a Delphi study and survey methods resulted in a comprehensive set of software project risks (Wallace and Keil, 2004; Wallace et al., 2004; Schmidt et al., 2001). Wallace and Keil (2004) structured those within a “risk categorization framework” which they derived from the Delphi study and validated through a substantial survey of over 500 project managers. That framework comprised 53 project risks mapped to the four quadrants shown in Fig. 1.

Each of the quadrants in Fig. 1 represents a focus on concerns associated with grouped aspects of risk. For example, *Customer Mandate* is perceived as having a high level of importance, and a low level of perceived control, whereas *Execution* is perceived as being of moderate importance, with a high level of perceived control. The **Customer Mandate** quadrant, for instance, concerns “risk factors relating to customers and users, including lack of top management commitment and inadequate user involvement” (Wallace et al., 2004). The quadrant **Scope and Requirements** involves “the ambiguity and uncertainties that arise

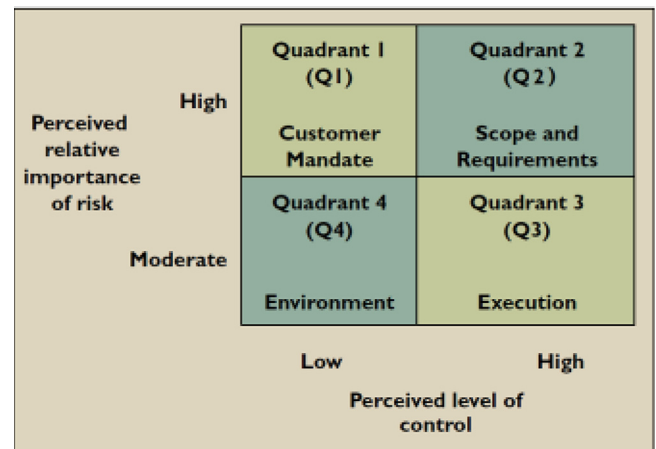


Fig. 1. Wallace and Keil's risk quadrants (Wallace and Keil, 2004).

in establishing the project's scope and requirements” (Wallace and Keil, 2004; Wallace et al., 2004). **Execution** concerns “the actual execution of the project... and many of the traditional pitfalls associated with poor project management” (Wallace et al., 2004). The **Environment** quadrant focuses on risks associated with the external or internal environment, “including changes in organizational management that might affect a project” (Wallace and Keil, 2004). The full list of the 53 risks (plus the 10 new GSD risks we add as part of this study) is shown in Fig. 2.

Apart from some work by Gotterbarn and Rogerson (2005) on ethical risk analysis, the subsequent focus of attention was on software project failures, and contributors to software project outcomes (McLeod and MacDonell, 2011; Lehtinen et al., 2014). As noted in the framework for understanding influences on software systems development and deployment project outcomes developed by McLeod and MacDonell (2011), four broad areas of focus were identified: *institutional context, people and action, development processes, and project content*. Lehtinen et al. (2014), through root cause analysis, followed on the work by MacLeod and MacDonell to identify the causes and effects of project failures, adopting a four-element framework of: *people, methods, task, and environment* to guide their analysis. While not a direct correspondence to Wallace and Keil's four quadrants (Wallace and Keil, 2004), the coverage and similarities in both of these subsequent studies are strong, suggesting that the Wallace and Keil framework still has validity and is a good candidate for evaluating how software risks are addressed in a scaling agile implementation. Wallace and Keil's risk categorization framework, with its tabulation of risks by quadrant, also supported detailed and aggregated mapping of risks addressed for each method (Section 3).

Indeed, more recent work in outsourced and global software development risks (Abdullah and Verner, 2009; Verner et al., 2014), again drew on the work of Wallace and Keil (2004). The outcome was a dual conceptual risk framework for outsourced project risks comprising the four elements from Wallace and Keil, and some additional outsourcing specific risks.⁸ The mapping of global software development risks by Verner et al. (2014), resulted in four risk components from the outsourced risk framework being identified (*project scope and requirements, project execution, project planning and control, and organizational environment*). A second mapping to the ISO 12207 software lifecycle

⁸ Wallace et al. (2004) acknowledge that their work focuses on in-house software projects, noting that outsourced projects may “exhibit a different risk profile...” and to gauge the risk in these projects may require additional risk constructs such as external relationship quality.

⁷ <http://www.scaledagileframework.com>

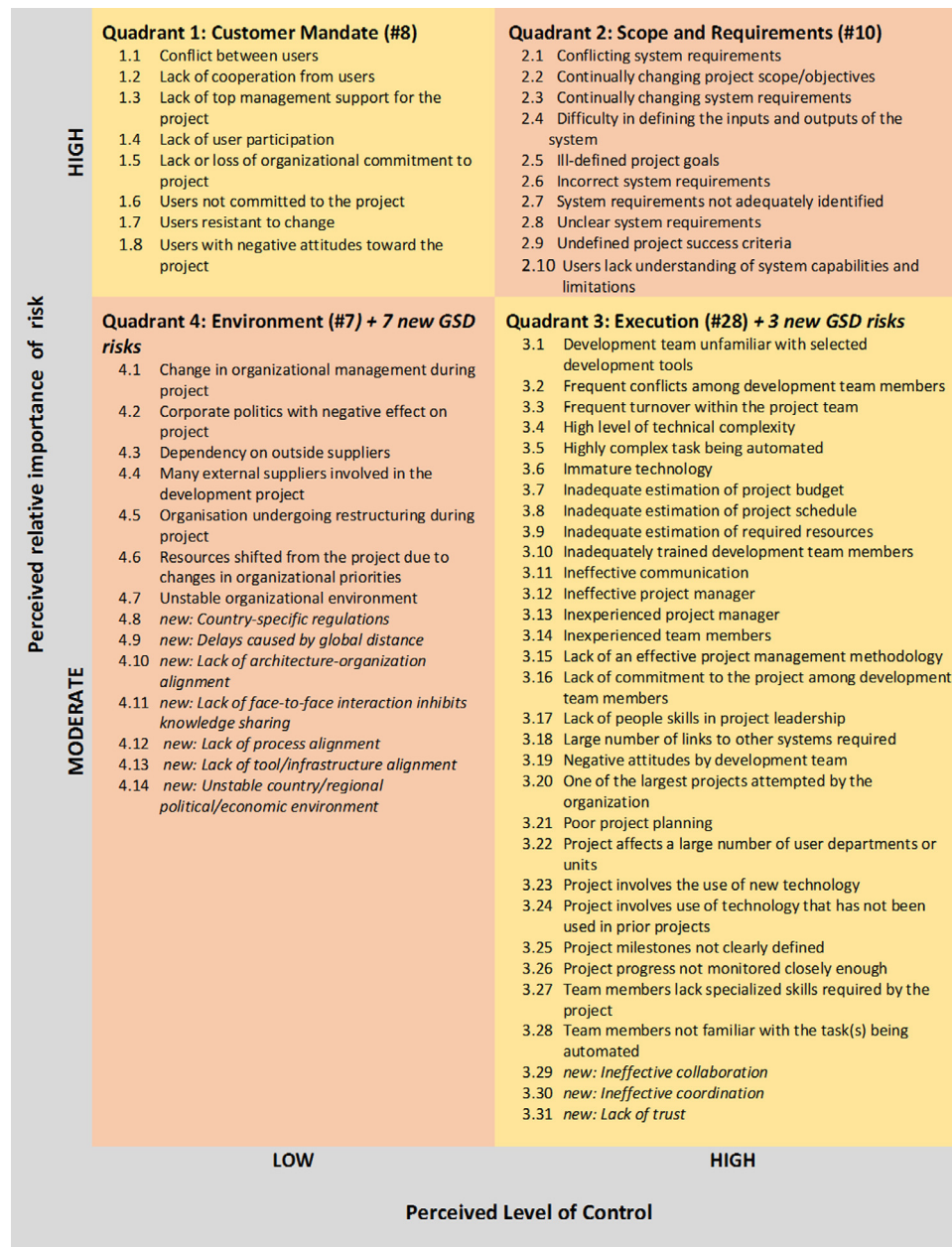


Fig. 2. GSD Risk Catalog derived from Wallace and Keil (2004) and Verner et al. (2014).

process, also by Verner et al. (2014), identified *organizational management, development process, acquisition process, and training*—but again was consistent with aspects of Wallace and Keil (2004).

In their review Verner et al. (2014) identified a broader range of GSD-specific risks and mitigations: high level and detailed level GSD vendor selection risks; requirements engineering risks; software development process risks; architectural design risks; configuration management risks; culture and social integration risks; training risks; communication and collaboration risks; planning risks; coordination risks, and control risks.

Verner et al.'s (2014) set of risks are an alternative candidate for assessing how well our scaling agile implementations (in our two case studies) addressed GSD risks. Indeed, Verner et al. (2014) suggested that with some development the Abdullah and Verner outsourcing framework (Abdullah and Verner, 2009), could be applied as a “useful framework for GSD projects”. Verner et al.'s review (Verner et al., 2014) reflects the work of 24 separate

studies of risk in GSD, including frameworks for distributed software development threats in Ågerfalk et al. (2005), and managing risks in distributed software projects in Persson et al. (2009). Since Verner et al.'s systematic literature review synthesizes the body of work in GSD risk (resulting in 85 risks), all the risks mentioned are also considered in our assessment of scaling agile framework's resistance to risk in a GSD context (see Section 3). Another candidate set of risks for GSD was enumerated by Chadli et al. (2016); however, they focus solely on project management risks, whereas we wanted to take a wider perspective to include technical risks. So, since other candidate risk frameworks were either incomplete or failed to validate their findings to the same degree as Wallace and Keil, we chose the Wallace and Keil framework of four quadrants to serve as a frame for incorporating the GSD related risks detailed in Verner et al. (2014). The combined view fully described in Section 4, and summarized in Fig. 2, shows how we augment Wallace and Keil's framework to cater for Global Software Development risks.

Key insights from this combination indicate that the risks identified in the **Customer Mandate** quadrant are not apparent in other papers on risk we have reviewed here, apart from some coverage in the “GSD vendor selection” (Abdullah and Verner, 2009) grouping, and the more recent distributed agile development risk observation of “requirements conflicts amongst multiple product owners” (Shrivastava and Rathod, 2017). In keeping with this observation, Verner et al. in their tertiary review, criticize the limited focus on developer and vendor perspectives and observe that “the client is pretty much ignored” (Verner et al., 2014) in the literature. Perhaps the vendor perspective leads to a proxy customer situation, which historically has been generally well aligned within a vendor organization; however, in a distributed agile setting this can lead to conflict between product owners (Razzak et al., 2018). In our aim to create a comprehensive risk catalog, we found no additional risks relating to quadrants capturing **Customer Mandate** and **Scope and Requirements**, that are both thoroughly covered by Wallace and Keil (2004). Unsurprisingly, the gaps were observed in the **Environment**, and **Execution** quadrants. In our GSD risk mapping exercise, we were able to categorize all of the risks (or threats) listed in Verner et al. (2014) according to the Wallace and Keil (2004) four quadrants, despite many of the GSD risks appearing at different levels of granularity, and several presented as compound risks. The new GSD risks are *Delays caused by global distance*, *Lack of architecture-organization alignment*, *Lack of face-to-face interaction inhibits knowledge sharing*, *Lack of process alignment*, *Lack of tool/infrastructure alignment*, and *Unstable country/regional political/economic environment*, all of which fall under the **Environment** quadrant. We also have three new risks in the **Execution** quadrant, namely *Ineffective collaboration*, *Ineffective coordination* and *Lack of trust*. (Beecham et al., 2020)

In summary, the overall degree of commonality of the other frameworks with the Wallace and Keil framework (Wallace and Keil, 2004) was sufficient to favor this more established framework for our comparison augmented with new GSD specific risks from Verner et al. (2014).

2.2. Scaling agile frameworks

In response to the difficulty of introducing traditional agile methods (originally designed for small co-located teams) into large-scale projects and organizations (Dybå and Dingsøyr, 2008), several scaling agile frameworks have emerged. These frameworks attempt to scale agile practices for enterprise-wide agility, to include agile for distributed teams, large projects and critical systems (Ebert and Paasivaara, 2017). Indeed, twenty such frameworks were identified by Uludağ et al. (2017), of which the Scaled Agile Framework and DAD frameworks are some of the more popular models (according to the 12th annual state of agile report Analysis.net Research, 2018).

There are debates over the precise definition of “large scale agile”, with Dikert et al. (2016) observing that “what is seen as large-scale depends very much on the context and the person defining it”. Kalenda et al. (2018), distinguish between “large scale” and “very large scale” and consider a number of further aspects of scale. However, in our study, we opt for the simplicity and comparability of a definition based on number of people, teams and locations. Therefore, building on the definition of “large scale” proposed by Dikert et al. (2016), we include the additional stipulation for *large scale agile GSD* that the definition explicitly incorporates a global focus and thus represents: “software development organizations with 50 or more people or at least six teams” (Dikert et al., 2016) and these people or teams must work across sites located in at least two different countries. Development companies, therefore, follow a geographically separated country

or company sourcing strategy according to Vizcaino et al.’s GSD ontology (Vizcaino et al., 2016, p. 74).

Our research is guided by the observation that “scaling isn’t easy; large projects often are globally distributed and have many teams that need to collaborate and coordinate” (Ebert and Paasivaara, 2017). In their column on Scaling Agile (Ebert and Paasivaara, 2017), Chris Ebert and Maria Paasivaara note that there is little support in the empirical literature on large-scale agile practice transformation.

The two scaling agile frameworks in our study both consider risk. In their introduction to Disciplined Agile Delivery, Ambler and Lines note: “The Disciplined Agile Delivery (DAD) process framework is a people-first, learning-oriented hybrid agile approach to IT solution delivery. It has a *risk-value* life cycle, is goal-driven, and is enterprise aware” (Ambler and Lines, 2011, emphasis added). In a similar vein, the Scaling Agile Framework (SAFe), places a great emphasis on risk mitigation. For example, on the first day of SAFe’s Program Increment (PI) Planning ceremony⁹ “teams identify risks and dependencies and draft their initial team PI objectives” (emphasis added). SAFe stipulates the importance of PI Planning, by saying, “if you are not doing it [PI Planning], you are not doing SAFe”.¹⁰

2.2.1. Disciplined Agile Delivery (DAD)

Disciplined Agile Delivery (DAD) is summarized in the review by Alqudah and Razzali (2017, 2016) as comprising a set of roles, practices, and phases. The full details of the method created by Scott Ambler are given in the book (Ambler and Lines, 2012) and the website for the method,¹¹ where the authors state that “the Disciplined Agile process decision framework provides light-weight guidance to help organizations streamline their processes in a context-sensitive manner, providing a solid foundation for business agility”.

We saw our case study site’s implementation of DAD unifying the four levels of the enterprise, encompassing business and software engineering functions (product management, portfolio management, program management, and project management). DAD refers to its framework as a toolkit allowing the creation of an organization-specific scaling agile approach regardless of the organization’s size. “DAD adopts practices and strategies from existing sources and provides advice for when and how to apply them together. In one sense methods such as Scrum, Extreme Programming (XP), Kanban, and Agile Modeling (AM) provide the process bricks and DAD the mortar to fit the bricks together effectively”.¹²

However, central to the DAD framework is portfolio management, which provides the ability to make the right decisions the first time, removing the bias in decision making when identifying specific business value ideas to be pursued either for development as products, features, or for further experiments, to get feedback and certainty. The governance of the DAD framework comprises a set of mandated practices spanning the entire functional setup involved in identifying, developing, and making software available for use. Hence, portfolio management is driven based on development and operations intelligence (metrics) including required guidance and suggestions from the other functional units. The other major functional process areas contributing to portfolio management are the product management, and enterprise architecture (including the IT

⁹ “Program Increment (PI) Planning a cadence-based, face-to-face event that serves as the heartbeat of the AgileReleaseTrain(ART), is a very large meeting in which all teams (remote or collocated) align on the ART to a shared mission and Vision.

¹⁰ <https://www.scaledagileframework.com/pi-planning/>

¹¹ <http://www.disciplinedagiledelivery.com/>

¹² <http://www.disciplinedagiledelivery.com/>

governance group) within the organization. Therefore, portfolio management can make informed decisions in terms of budget and resources, i.e. their development capacity to be successful in the marketplace.

With the DAD framework, a portfolio can be delivered through various IT delivery approaches (based on the size of the organization) such as program management, agile delivery, continuous delivery, or lean delivery. There is portfolio management guidance and support for IT delivery regardless of the delivery means. However, the program management approach is driven through several projects. While there are mandated practices spanning the functional setups, portfolio management emphasizes team autonomy and self-organization within a cohesive set of practices.¹³ The DAD framework identifies a set of primary and secondary roles and their responsibilities regardless of IT delivery approach, comprising leadership and more technical roles, and encouraging the development of “T-skilled engineers” (Alqudah and Razali, 2017).

DAD allows for formal people management processes,¹⁴ a missing element in many scaling agile frameworks, although as noted in an earlier study (Lal and Clear, 2017), we did not see evidence of those processes in action. While DAD is driven by agile values and principles, and therefore does not advocate “big upfront design”, in reality the process of managing work item lists from conception to readiness for development encompasses a large amount of massaging of functionality and architectural and design work prior to entering the project level phases. At project level three phases are incorporated: inception (scoping and sprint planning), construction (agile testing and coding), and transition (readiness for deployment or DevOps).

2.2.2. Scaled Agile Framework (SAFe)

SAFe was released in 2011 by Dean Leffingwell and is a living framework that is continually updated, having (at the time of writing this paper in 2020) reached version 5.0. According to Leffingwell et al. “SAFe applies the power of Agile but leverages the more extensive knowledge pools of systems thinking and Lean product development” (Leffingwell et al., 2016). According to case studies featured on the Scaled Agile Framework website¹⁵ (which admittedly may present a one-sided view), SAFe offers many business benefits, including:

- 20%–50% increase in productivity,
- 50%+ increases in quality,
- 30%–75% faster time to market,
- Measurable increases in employee engagement and job satisfaction.

However, these benefits may not be universal, or may come at a cost to work satisfaction amongst the teams (Paasivaara, 2017). In her study, Paasivaara reports that one of the teams experienced most changes as negative, where “teams felt lack of autonomy, as they could no longer decide some things on their own, such as the sprint length. With fixed increments they felt moving backward, towards the old waterfall”. Paasivaara concludes that perhaps this negative attitude is due to teams being new to SAFe and they may not have had time to witness the benefits. This is in contrast to the other case in Paasivaara’s study, where all participants described SAFe adoption as “highly successful”. The perceived lack of autonomy, and its effects on staff morale, was also found in Noll et al.’s study of SAFe, where there needs to be a balance between autonomy, the feeling of relatedness (support and trust

among colleagues), and ability and skills to take on new tasks (following a self-determination theory) (Noll et al., 2017).

SAFe offers a soft introduction to the agile world through specifying a range of structured patterns often needed when organizations transition from a more traditional environment, particularly in the context of a large project (Dybå and Dingsøyr, 2008). SAFe considers the whole enterprise, and is organized according to four levels of organization: Team, Program, Large Solution,¹⁶ and Portfolio. Each level integrates agile and lean practices, manages its own activities, and aligns with the other levels. Also, depending on the size of the operation, and the stage the company is at in terms of scaling, there are varying levels of complexity. For example, organizations can start with entry-level “Essential SAFe” (with just two organizational levels represented—Team and Program). This can build to “Large Solution SAFe” (with an additional level of the Large Solution), to “Portfolio SAFe” (in which the Portfolio level replaces the Large Solution level). Finally, there is “Full SAFe” in which all four levels of Team, Program, Large Solution, and Portfolio are represented.

The *Team* level outlines techniques similar to those used in standard Scrum, with two-week sprint cycles. As in Scrum, teams of between 5–9 members contain three roles: the Product Owner, Scrum Master, and team member. Each agile team is responsible for defining, building, and testing stories from its team backlog in a series of iterations. Teams have common iteration cadences and synchronization to align their activities with other teams so that the entire organization is iterating in unison. Teams use Scrum, XP, or Kanban to deliver prototypes every two to four weeks (Leffingwell et al., 2016). Important for scaling, all SAFe teams form part of a team of agile teams—called an Agile Release Train (ART)—that aims to deliver a continuous flow of incremental releases of value.

At the *Program* level, SAFe extends Scrum using the same ideas but on a higher level. This level defines the concept of an agile release train (ART), which is analogous to sprints at the Team level, but works at a different cadence on a larger timescale. The ART is composed of five sprint cycles. There is also a sixth “innovation planning sprint”, which allows teams to innovate, inspect, and adapt. Teams, roles, and activities are organized around the ART (Leffingwell et al., 2016).

Existing roles are stretched and new roles created to cater for the new responsibilities and practices, where a Product Manager serves as the content authority for the ART, and is accountable for identifying program backlog priorities. The Product Manager works with the Product Owners (POs) to optimize feature delivery and direct the work of POs at the team level. SAFe sees the emergence of the role of a Release Train Engineer (RTE) “who facilitates Program-level processes and execution, escalates impediments, manages risk, and helps to drive continuous improvement” (Leffingwell et al., 2016). Creating this role was a key success criterion in the case study described by Ebert and Paasivaara (Ebert and Paasivaara, 2017).

The *Large Solution* (previously *Value Stream*) level is optional, depending on the size of the organization; in larger organizations, implementation of this level calls for a value stream engineer (VSE) who plays a similar role to the RTE by facilitating and guiding the work of all ARTs and suppliers. Leffingwell et al. (2016) describe these and further roles such as Business Owner, DevOps team member, Release Manager, and Solution Manager as important.

A highest level of the SAFe hierarchy (made optional in the more recent 4.6 and 5.0 versions) is the *Portfolio level*. This set of executive-level processes completes the vertical enterprise view,

¹³ <https://www.overleaf.com/project/5e7f63a7c581ea00010f6d4f>

¹⁴ <http://www.disciplinedagiledelivery.com/%20agility-at-scale/people-management/>

¹⁵ www.scaledagileframework.com

¹⁶ The “large solution” level in SAFe v. 5.0 was termed “value stream” in version 4.0; this is the term used in our SAFe case study.

in which senior management make strategic decisions, deliver value, and prioritize ‘epics’ that are filtered down to the program level, where they are decomposed into features, which in turn are fed to the team level in the form of user stories.

While there is some early evidence in favor of SAFe and its adoption (Paasivaara, 2017; Noll et al., 2016), perhaps it is too early to judge its true merits or whether the promised benefits can be universally enjoyed. From the current literature, it is unclear how well this lean agile enterprise approach mitigates risk in global software development.

2.3. DAD and SAFe comparison

In a recent study of scaling agile strengths, weaknesses, opportunities, and threats (SWOT) in GSD, Sinha et al. (2020) identified six threats, comprising Lack of face to face communication, Improper task allocation, Cultural differences, Temporal differences, Linguistic differences, and Lack of agile coaching for scaling. Under the weaknesses quadrant, they include a Lack of knowledge sharing. These are recurrent themes in the agile and GSD risk literature, all of which we include in our GSD Risk Catalog.

When trying to decide which of the many scaling agile frameworks to adopt, Diebold and colleagues differentiate between a collection of frameworks to include DAD and SAFe (Diebold et al., 2018). They find SAFe to have a low level of flexibility (incorporating practices such as Scrum/Kanban/Lean, with specific XP practices “mandated”), whereas DAD has a medium level of flexibility (with practices including Scrum/Lean, and a mixed set of methods) (Diebold et al., 2018). They also suggest that further comparative studies are conducted to help with decision making.

Disciplined Agile Delivery (DAD) and Scaled Agile Framework (SAFe) both draw from a variety of agile and lean practices. However, according to Vaidya’s observations on three different scaling agile frameworks (Vaidya, 2014), an organization’s context is what matters most when deciding on which framework to adopt and which associated practices provide the desired results. We therefore present two case studies in two different contexts, and apply different Scaling Agile Frameworks, to provide some insight into how risk is mitigated.

Both frameworks (DAD and SAFe) place great emphasis on value and risk, so it seemed appropriate then to evaluate their efficacy at addressing global software development project risks by seeing how well they covered the software project risks identified in our GSD Risk Catalog.

With this aim in mind, we set out to address the research question, *How does the adoption of scaling agile framework practices address global software development risks?* By ‘address’ risk, we will specifically look for how the frameworks either mitigate (reduce) or eliminate the GSD risk. In the next section, we explain our approach to answering this question.

3. Method

To answer the research question noted at the end of the previous section, we take a three-phased approach. First we created a catalog of global software development risks we call the GSD Risk Catalog (Beecham et al., 2020). Second, we created a theoretical mapping of DAD and SAFe practices to the risks in our GSD Risk Catalog, which we deemed to address each risk; we also rated the degree to which the DAD and SAFe practice(s) eliminated or mitigated the given risk. Lastly, we moved from our theoretical model, to a real-world empirical setting in which we examined the extent to which each practice (in DAD and SAFe respectively) was implemented in a multiple case study comprising two GSD organizations. This served to illustrate both the extent to which each scaling agile framework addressed the risks in our catalog,

and how the two case studies implemented the risk-mitigating practices. We examined interview transcripts, survey results, and observation notes to identify instances where a risk listed in our GSD Risk Catalog was evident in either of our cases. From this evidence, we surmised the extent to which practices in each scaling agile framework addressed GSD risks.

We describe these three phases in detail in the sequel.

3.1. Phase 1: GSD risk catalog development

Given the context of our study is risk in globally distributed organizations, we first augmented the well-recognized Wallace and Keil (2004) set of risks with additional risks identified in Verner et al.’s (2014) tertiary review of risks in a GSD context.

Since there was no one clear, validated set of risks to draw on specific to GSD, we took the 85 risks detailed in Verner et al.’s review of 24 GSD risk studies (Verner et al., 2014) and mapped them onto the established Wallace and Keil risk framework (Wallace and Keil, 2004). A snippet of this extraction and mapping is shown in Table 1.

First, two researchers (author 1 and author 4) independently compared *each* risk identified by Verner and colleagues to *each* risk in Wallace and Keil’s risk catalog. If the risk from Wallace and Keil was equivalent to, or would be a consequence of, a risk from Verner et al. we created a correspondence between the two. Mapping the two sets of risks was not straightforward, since some risks in Verner and colleagues’ catalog are expressed at a high level, or as a combination of risks; in such cases, the risk from Verner et al. was mapped to multiple Wallace and Keil risks. Also, we marked those risks in Verner and colleagues’ catalog that did not correspond to a Wallace and Keil risk, or was incompletely captured by a set of Wallace and Keil risks, for later consideration.

Second, authors 1 and 4 reviewed the independent mapping of the other researcher to establish the level of agreement; disagreements were discussed between these two authors, and agreement reached. Then, author 2 moderated the final result to yield a single unified mapping. From this mapping, three broad categories emerged:

1. a mapped Verner risk, in which a Verner risk was equivalent to one or more Wallace and Keil risks,
2. an unmapped Verner risk with no Wallace and Keil equivalent, and finally
3. an unmapped Wallace and Keil risk with no Verner et al. equivalent.

In the third step, we coalesced unmapped or partially mapped Verner et al. risks into a small set of new risks. This was necessary because many risks in Verner and colleagues’ catalog were at a different level of abstraction compared to Wallace and Keil’s risks, were duplicated at the level of abstraction we needed, or appeared to be more a description of consequences of a risk rather than of the risk itself.

For example, Verner and colleagues identify two risks related to the rule of law: *Lack of protection for intellectual property rights in the vendor country* and *Problems because of differences in legal systems such as jurisdiction*. In the GSD Risk Catalog, these rather specific risks coalesced into the more general *Country-specific regulations* risk. Similarly, Verner and colleagues’ list *High organizational complexity, scheduling, task assignment and cost estimation become more problematic in distributed environments as a result of volatile requirements diversity and lack of informal communication* and *Lack of a common understanding of requirements leads to problems in system functionality*; we coalesced these (and ten others) into *Lack of face-to-face interaction inhibits knowledge sharing*.

As before, authors 1 and 4 independently reviewed the candidate GSD Risk Catalog to double-check that the mappings made

Table 1

Example of mapping risks from Verner et al. (Verner et al., 2014) to GSD Risk Catalog risks.

GSD Risk Catalog risk/Mapped risk from Verner et al.
Quadrant 1: Customer mandate
<i>Lack of user participation</i>
National, organizational, and cultural differences of participants can cause problems like rework, loss of data, confusions, etc. Lack of collaboration for RE between distributed stakeholders happens due to differences in culture, language distance and processes Application of agile practices causes problems in distributed development because of the degree of interaction between stakeholders and number of face-to-face meetings needed Requirements information not properly shared with distributed stakeholders affecting their interaction
Quadrant 2: Scope and requirements
<i>Incorrect system requirements</i>
Application of agile practices causes problems in distributed development because of the degree of interaction between stakeholders and number of face-to-face meetings needed Collaboration difficulties caused by geographic distance in agile development may cause misunderstandings and conflicts Problems caused because team members do not share equal knowledge of the domain Lack of a common understanding of requirements leads to problems in system functionality Lack of collaboration for RE between distributed stakeholders happens due to differences in culture, language distance and processes
Quadrant 3: Execution
<i>new: Lack of trust</i>
Collaboration difficulties caused by geographic distance in agile development may cause misunderstandings and conflicts Mutual trust is important but hard to obtain and lack of trust causes problems. This can be due to lack of face-to-face interaction, cultural differences, and weak social relations Fear about the future of jobs and roles, erodes trust Limited face-to-face meetings caused by geographic distance impact trust, decision quality, creativity, and general management; knowledge creation is limited within organization. This may lead to problems in creating collaboration know-how and domain knowledge Trust among stakeholders necessary to achieve innovation, flexibility, cooperation, and efficiency in distributed environment. Since often a short life span, important to achieve mutual trust rapidly, but if trust is misplaced, entire organization may suffer A vendor with poor relationship management can result in problems such as lack of trust
Quadrant 4: Environment
<i>new: Delays caused by global distance</i>
Application of agile practices causes problems in distributed development because of the degree of interaction between stakeholders and number of face-to-face meetings needed Configuration management problems cause dependency, delay and increased time is required to complete maintenance requests Temporal and physical distribution increases complexity of planning and coordination activities, makes multisite virtual meetings hard to plan, causes unproductive waits, delays feedback, and complicates simple things Inability to communicate in real time) causes collaboration problems; Stakeholders located in different time zones can lead to problems in communicating Not tailoring organizational structures to reduce delays in problem resolution causes difficulties and can result in site wars and reduce project cohesion Choosing a vendor with a lack of control over a project can result in problems such as cost and schedule overruns; and Poor schedule management

sense, and that the new risks were placed in the correct quadrant. Again, disagreements were discussed and a resolution agreed between these authors.

Finally, authors 2 and 3 reviewed the resulting GSD Risk Catalog to ensure agreement. Table 1 shows an extract of the result (see the companion technical report (Beecham et al., 2020) for the full mapping).

The mapping process is depicted in Fig. 3. As is shown, we added ten risks to Wallace and Keil's inventory, derived from Verner and colleagues' set of risks that do not map, or only partly mapped, to the Wallace and Keil risk inventory. The remaining 75 risks in Verner et al. corresponded to one or more of 41 of Wallace and Keil's 53 risks. Finally, there were 12 risks from Wallace and Keil's catalog that did not have a corresponding Verner et al. risk.

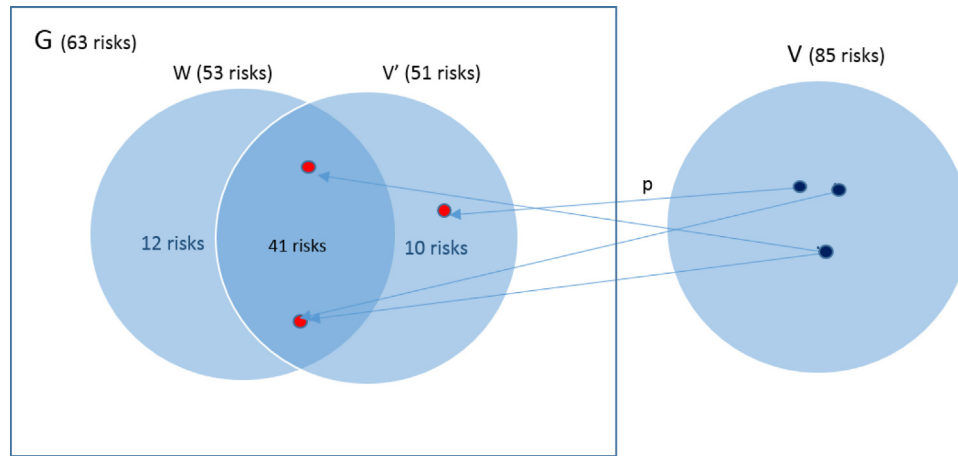
In summary, this comparison of risks in Verner et al. (2014) GSD to the well-established software project risk inventory of Wallace and Keil (2004) ensured that technical risks as well as project management risks associated with Global Software Development (GSD) are considered in our analysis. This comparison resulted in ten new risks that accommodate risks from Verner et al.'s study that were not adequately captured by a risk in the Wallace and Keil inventory. This combined set of 63 risks, which we label the "GSD Risk Catalog", is categorized according to Wallace and Keil's four quadrants: *Customer Mandate*, *Scope and Requirements*, *Execution*, and *Environment* (see Fig. 1), with the majority of new GSD risks coming under the *Environment* quadrant.

The full list of 63 risks in our derived GSD Risk Catalog is presented in Fig. 2, discussed in Section 4.2.

3.2. Phase 2: Theoretical mapping

Making use of this newly created GSD Risk Catalog, we then mapped any practice identified in our Scaling Agile Frameworks to the risk, where they appeared to mitigate risk. This mapping involved three steps:

1. Scaling Agile practice mapped to Risk factors: As part of our ongoing, longitudinal case studies, in previous work we identified sets of practices from SAFe (Noll et al., 2016), and DAD (Lal and Clear, 2017, 2018). Four researchers, working in pairs, compared each practice in the scaling agile framework, to each of the 63 risks in the GSD Risk Catalog (Fig. 2). Authors 1 and 4 mapped SAFe practices to the GSD Risk Catalog, and authors 2 and 3 mapped DAD practices to the GSD Risk Catalog. To ensure all researchers worked to the same standard, an example of how to 'map' a practice to a risk was shared amongst all researchers.
2. Strength of mitigation assessment: Once the mapping of practices to risks was completed, each risk was rated according to the degree to which the mapped practices eliminated or mitigated the risk, as to whether the practices "definitely" address the risk, address the risk "somewhat", or do "not at all" address the risk. So, if a practice or set of practices unequivocally addressed the risk, we coded the practice as "definitely"; if the practice(s) to some extent contributed to elimination or mitigation, we coded the practice "somewhat"; and when we could not identify a practice that would eliminate or mitigate the risk we coded



In the diagram above, let G be the set of risks in our GSD Risk Catalog, V be the set of risks identified by Verner and colleagues [13], and W be the set of risks catalogued by Wallace and Keil [11]. Then,

$$p : V \mapsto W$$

where p is the mapping process we employed to relate risks in Verner et al. to risks in Wallace and Keil's catalog, and V' is the subset of risks in the GSD Risk Catalog to which one or more risks in Verner et al. are mapped (i.e. the co-domain of p). The subset of G comprising ten new risks created to account for risks from Verner et al. that had no, or incomplete, mapping to risks in Wallace and Keil's catalog is the relative complement of W with respect to V' ($V' \setminus W$).

Fig. 3. GSD Risk Catalog creation based on Verner et al. (2014) and Wallace and Keil (2004).

the risk as “not at all”. Again, authors 1 and 4 assessed SAFe practices, and authors 2 and 3 assessed DAD practices.

3. Inter-rater cross-check (within frameworks): When all possibilities were exhausted, each pair of authors reviewed the mapping of his or her peer. Any disagreements were discussed within each pair until a consensus was reached.

The output from this theoretical mapping was a theory of risk mitigation according to DAD and SAFe, which is presented in Tables 11 and 12.

3.3. Phase 3: Empirical evidence

In the last stage, we examined data collected in a longitudinal multiple case study of two companies engaged in scaling agile development adoption according to the DAD (Case A) and SAFe (Case B) frameworks. A two case multiple case study has advantages over a single case, as, according to Yin, “... analytic benefits from having two (or more) cases may be substantial...” since, “analytic conclusions independently arising from two cases ... will be more powerful than those coming from a single-case ... alone” (Yin, 2018). Furthermore, Yin states that when asking “How” and “Why” types of questions, case study research is particularly relevant (Yin, 2018). Our case boundaries include time (in years), geographic locations, domain, and practice adoption. We apply the multiple case study design to address our “How” research question, and test our theoretical mapping of scaling agile framework practices (described in Phase 2 of our method) in which we hypothesize, that a given set of scaling agile practices can mitigate GSD risk. The business model for both cases is to develop, maintain and sell software to clients throughout the globe. The original aim of both of our case studies was to gain broad insight into issues and benefits of scaling agile framework adoption; as such, while not focused specifically on risk, these studies yielded a rich source of data from which we were able to identify many issues related to risk, test our theory, and compare and contrast across cases.

Table 2

New risks created to augment Wallace and Keil's catalog, and corresponding quadrant in Wallace and Keil's framework (quadrant 3 is “Execution,” quadrant 4 is “Environment”).

New risk	Quadrant
Country-specific regulations	4
Delays caused by global distance	4
Ineffective collaboration	3
Ineffective coordination	3
Lack of architecture-organization alignment	4
Lack of face-to-face interaction inhibits knowledge sharing	4
Lack of process alignment	4
Lack of tool/infrastructure alignment	4
Lack of trust	3
Unstable country/regional political/economic environment	4

3.3.1. DAD evaluation—Case A

For the empirical investigation of the DAD method, interview transcripts from Case A (described in Section 4) were examined by authors 2 and 3. In Case A, author 3 conducted interviews of participants in March of 2017 (Lal and Clear, 2017, 2018); see Section 4.3.1 for a description of the company (“Company A”) involved in Case A, and details regarding the interview participants (Table 3). The interviews were conducted at Company A's software engineering lab at Box Hill, Melbourne, Australia. A total of eight interviews (each an hour long) were conducted one week in March 2017. Participants were identified based on the various software engineering roles within this software vendor organization. These roles were: Principal Software Engineer, Senior Software Engineer, Software Engineer, Team Leader, Engineering Manager, Product Manager or Product Owner, Quality Assurance Manager, and the Director of Software Engineering. All the participants for this investigation were part of their DAD transformation. The interview instruments were based on the reasons and approach for switching to the DAD method. An agreement was made to record the interviews that were later

Table 3

List of participants interviewed in Case A.

Part. ID	Role	Team	Program	Portfolio	Country
PA1	Product owner/ Product management	1	1	1	Australia
PA2	Principal engineer/ Enterprise architect/ Program architect	1	1	1	Australia/USA/India
PA3	Software engineer	1			Australia
PA4	Quality assurance manager	1	1	1	Australia/USA/India
PA5	Senior software engineer	1	1		Australia
PA6	Team lead	1	1		Australia
PA7	Engineering manager	1	1		Australia
PA8	Director of software engineering	1	1	1	Australia
Total		8	7	4	

Table 4

List of interviewees from Case B. Locations and some roles are general, to preserve anonymity; “Sr Management” represents Architect, Product manager, Regional manager, and Development manager roles.

Part. ID	Role	Team	Program	Portfolio	Location
PB1	Scrum master, Project manager	1	1		North America
PB2	Scrum master, Project manager	1	1		UK/Ireland
PB3	Sr Management			1	Continental Europe
PB4	Scrum master, Project manager	1	1		Continental Europe
PB5	Sr Management			1	Continental Europe
PB6	Sr Management			1	UK/Ireland
PB7	Sr Management			1	UK/Ireland
PB8	Sr Management			1	North America
PB9	Developer	1			North America
PB10	Product owner	1			North America
PB11	Scrum master, Project manager	1	1		North America
PB12	Sr Developer, Tech lead	1	1		North America
PB13	Developer	1			UK/Ireland
PB14	Developer	1			UK/Ireland
PB15	Sr Developer	1	1		UK/Ireland
PB16	Product owner	1			UK/Ireland
Total		11	6	5	

transcribed. Although all interviews were conducted out of the Melbourne site only, the projects covered development across Australia, the USA, and India.

3.3.2. SAFe evaluation–Case B

Data relating to SAFe were obtained from results of an ongoing longitudinal participant-observer study (called “Case B”), with moderate researcher involvement, that began at the end of 2015 and continued through to the autumn of 2019. Similar to Case A, the main purpose of our collaboration was to observe how the company (“Company B”) transitioned from a plan-driven development process to a scaling agile development process based on SAFe practices. Company B is specifically interested in how to adopt the new agile, lean, and Kanban practices in their highly distributed setting: where teams and individual team members are globally distributed (see Fig. 5).

During our four year collaboration, authors 1 and 4, along with their colleagues, conducted 31 interviews of team members and managers in a variety of roles, at all levels of the company, chosen to be representative of all levels of the development organization.¹⁷ Interviews were conducted on-site in the company’s Dublin headquarters, and via video conference, over two years, starting November, 2015. Participants of sixteen of these interviews are directly quoted in our study, as listed in Table 4.

We also observed distributed development teams conducting Scrum ‘ceremonies,’ such as daily standups, sprint planning, and retrospective meetings; and, we observed weekly program-level “scrum of scrums” style meetings. These observations began in November 2015 and continued, focusing on different teams, until the end of 2017. Observations helped to place the interviews in context, but did not directly provide any data for this study.

¹⁷ For the questions asked in our semi-structured interviews please see our companion Technical Report (Beecham et al., 2017).

Finally, a series of three SAFe “self-assessment” surveys were administered to various teams, and program and portfolio level participants, in February 2017, July 2017, and March 2018 (Noll et al., 2016; Razzak et al., 2017, 2018); see Table 5 for details regarding the participants in the surveys. The self-assessment surveys identified the level to which the participants perceived they implemented various SAFe practices and ceremonies.

3.3.3. Within and between multiple case study evaluation

As a first step, we examined these data for evidence that the companies had experienced problems (or not) related to the risks in our GSD Risk Catalog.

Then, in the second step, we assessed the frequency at which each company performed the respective scaling agile framework practices mapped to risks in the GSD Risk Catalog. In Case A, this frequency was assessed to be “always” as the company had completed its agile adoption, at the time the interviews took place. In Case B, data from the self-assessment surveys were examined to determine the company’s self-assessed frequency of practice performance.

Finally, working in pairs (authors 2 and 3 for case A, and authors 1 and 4 for Case B), we connected the output of the previous two steps, to understand whether the scaling agile practices eliminated or mitigated the corresponding risks: 1. if the practices were implemented in the company, and no evidence of the risk was seen, the practices could have been material in *eliminating* the risk; 2. if the practices were implemented in the company, but there was evidence that the risk was a problem for the case, the practices still might have been effective at *mitigating* or reducing the risk; or, 3. this might indicate that the theoretical mapping is not effective in practice.

To determine which of these alternatives was the case, we considered three additional elements:

Table 5

List of participants in Case B “self-assessment” survey (responses aggregated to preserve anonymity).

Role	Team	Program	Portfolio	Country
Project Mgr./Scrum master	9	6	–	Canada, France, Ireland, Norway, Poland, UK
Developer	20	–	–	Canada, Ireland, Denmark
Quality assurance	8	–	–	France, Ireland, USA
Product owner	4	5	–	Canada, Ireland, USA
Director of Eng.	1	–	–	Ireland
Development manager	1	1	1	Ireland
Product manager	2	1	1	Spain
Quality assurance lead	–	1	–	Ireland
Database administrators	–	1	–	Ireland
Technical support	–	1	–	Ireland
Chief executive officer	–	–	1	Ireland
Chief technology officer	–	–	1	Ireland
Regional management/Sales	–	–	2	France, Italy
Total	45	16	6	

1. *Strength of theoretical mapping*: the degree to which the practices address the risk. Risks that are only “somewhat” addressed (by the practice), are perhaps more likely to be seen as problems.
2. *Strength of practice implementation in cases*: the frequency at which the associated practices were performed. If this was less than “always”, it is possible the practices were not effective because they were not thoroughly implemented.
3. *Level of control*: whether the risk can be *eliminated*, or only *mitigated*. Certain risks, such as *Unstable country/regional political/economic environment*, are part of the environment; they cannot be eliminated, but their impact can be reduced.

We present the results of applying this method in the sequel.

4. Results

In this section we first present a new catalog of GSD risks created by comparing and merging risks identified by Verner et al. (2014) to Wallace and Keil’s list of risks (Wallace and Keil, 2004). Then, we show the theoretical mapping of SAFe and DAD practices to these GSD risks. Finally, we present empirical evidence of the effectiveness of our theoretical mapping, that underpins the extent to which the scaling agile frameworks eliminate or mitigate risks in GSD.

4.1. GSD risk catalog development

To create a comprehensive catalog of risks faced by global software development projects, we compared 85 risks identified by Verner et al. (2014) in their tertiary study of risks in global software development, to the 53 in Wallace and Keil’s (2004) risk framework. We found many risks identified in Wallace and Keil related to GSD risks, and that many of the risks listed by Verner et al. identify more than one risk. For example, a risk listed under “Requirements engineering risks and mitigation advice” states, “A lack of suitable tools or methodologies available for requirements elicitation may lead to problems in obtaining the real requirements. (Verner et al., 2014, Table 9, p. 64)” This statement articulates two risks: a lack of suitable *tools* for requirements elicitation, and a lack of suitable *methodologies* for requirements elicitation. Over a third (32) of the risks in Verner and colleagues’ list could be classified as “compound” risks of this nature.

Other risks identified by Verner et al. are general, high-level risks that could have multiple consequences for a software development project. For example, the category “Software development process risks and mitigation advice” includes this risk: “Application of agile practices causes problems in distributed development because of the degree of interaction between stakeholders and number of face-to-face meetings needed. (Verner

et al., 2014, Table 10, p. 64)” This high-level risk leads to several risks identified by Wallace and Keil, including: Lack of user participation (quadrant 1), Incorrect system requirements (quadrant 2), and Inadequate estimation of project budget (quadrant 3).

This high-level risk also leads to several new, GSD-specific risks not found in Wallace and Keil’s catalog, including: Delays caused by global distance, Ineffective collaboration, Ineffective coordination, and Lack of face-to-face interaction inhibits knowledge sharing. Of the 85 risks identified by Verner and colleagues, 79 correspond to, imply, or result in at least one risk in Wallace and Keil’s catalog.

We also found six that had no correspondence to any of Wallace and Keil’s risks:

1. “Lack of well-defined modules causes problems with progressive integration (Verner et al., 2014, Table 11, p. 65)”
2. “Project participants have limited understanding of other project participants’ competencies (Verner et al., 2014, Table 15, p. 67)”
3. “Coordination in multi-site development becomes more difficult in terms of articulation of work as problems from communication lack of group awareness and complexity of the organization appear and influence the way the work must be structured (Verner et al., 2014, Table 17, p. 68)”
4. “Tool compatibility may prove a problem; sites are likely to prefer different programming languages, support tools, operating systems, and development tools (Verner et al., 2014, Table 18, p. 69)”
5. “No process alignment, in terms of traditions, development methods, and emphasis on user involvement, will often differentiate between sites, possibly resulting in incompatibility and conflicts (Verner et al., 2014, Table 18, p. 69)”
6. “A vendor with poor relationship management can result in problems such as lack of trust (Verner et al., 2014, Table 8, p. 63)”

And, we found 43 risks that, due to being high-level or compound in nature, not only corresponded to one or more risks in Wallace and Keil’s catalog, but also suggested a new risk, not in Wallace and Keil’s catalog. As a consequence, we formulated 10 additional risks; these are listed in Table 2.

The result of combining the 10 new risks with the 53 in Wallace and Keil’s catalog, yielded a combined GSD Risk Catalog of 63 risks, illustrated in Fig. 2. The full correspondence of risks identified by Verner and colleagues, to risks in the GSD Risk Catalog, is available in a companion technical report (Beecham et al., 2020).

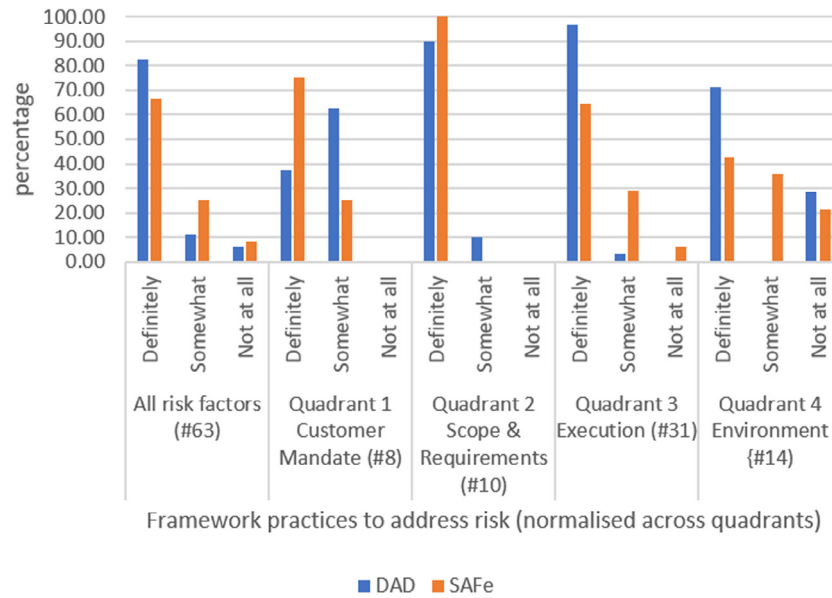


Fig. 4. Comparative strength of risk mitigation across all 63 factors.

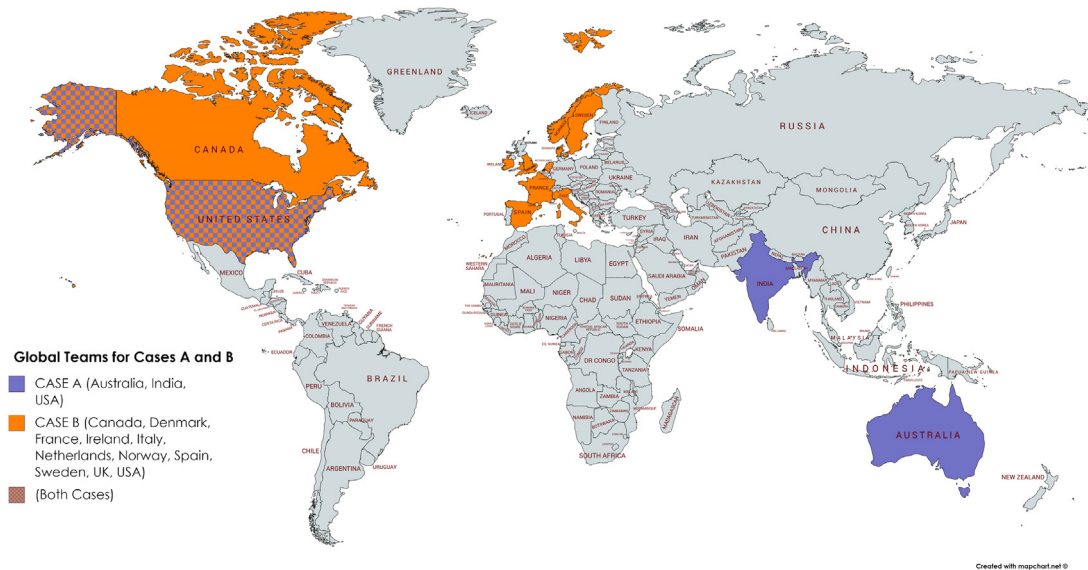


Fig. 5. Case locations.

4.2. Theoretical mapping of scaling agile practices to GSD risks

To understand the extent to which scaling agile frameworks address GSD risks, we assessed how well practices in DAD and SAFe address the risks in our GSD Risk Catalog. The result is a *theoretical mapping of scaling agile practices to GSD risks*.

Table 11 and Table 12 show our assessment of the degree to which DAD and SAFe practices (respectively) address the risks in our GSD Risk Catalog, along with examples of DAD and SAFe practices that address the given risk. Space does not allow all the associated DAD and SAFe practices to be included, so the tables present selected examples; the complete mapping is available as part of a technical report (Beecham et al., 2020).

Fig. 4 summarizes the extent to which each scaling agile framework theoretically addresses the GSD Risk Catalog risks. Looking at the total, Fig. 4 shows that both frameworks address most of the risks. The raw figures have been normalized across quadrants to allow a comparison of how each framework addresses risks in the GSD Risk Catalog quadrants outlined in Fig. 2.

Despite the frameworks addressing a similar number of risks, we see some differences when we compare risk mitigation within some quadrants.

Both frameworks address the eight risks in the **Customer Mandate** quadrant to a certain extent. SAFe appears slightly more aligned to working with the customer and user than DAD (Fig. 4).

Fig. 4 shows little difference in how the frameworks address the **Scope and Requirements** set of risks. Both DAD and SAFe address all ten risks in this quadrant completely, with the exception of DAD that has slightly weaker support for one of the factors, *Users lack understanding of system capabilities and limitations*.

The **Execution** quadrant has the most risk factors (31). As Fig. 4 shows, both DAD and SAFe address most of these 31 risk factors to some extent, although DAD appears to be the stronger framework when it comes to project execution.

The **Environment** quadrant Fig. 4 has the next highest number of risk factors (since many of the GSD specific risks are categorized in this quadrant). However, this quadrant has the most factors not addressed by the frameworks. For example, these

Table 6
Risks from GSD Risk Catalog observed in cases.

GSD Risk Catalog risk	Observed in	
	Case A	Case B
Quadrant 1: Customer mandate		
Conflict between users	Y	
Lack of cooperation from users		Y
Lack of top management support for the project		Y
Lack of user participation		
Lack or loss of organizational commitment to the project		
Users not committed to the project		
Users resistant to change		
Users with negative attitudes toward the project		
Quadrant 2: Scope and requirements		
Conflicting system requirements	Y	Y
Continually changing project scope/objectives		Y
Continually changing system requirements		Y
Difficulty in defining the inputs and outputs of the system	Y	
Ill-defined project goals		Y
Incorrect system requirements		
System requirements not adequately identified	Y	Y
Unclear system requirements		Y
Undefined project success criteria		Y
Users lack understanding of system capabilities and limitations		
Quadrant 3: Execution		
Development team unfamiliar with selected development tools	Y	
Frequent conflicts among development team members		Y
Frequent turnover within the project team		
High level of technical complexity	Y	Y
Highly complex task being automated	Y	
Immature technology	Y	
Inadequate estimation of project budget		Y
Inadequate estimation of project schedule		Y
Inadequate estimation of required resources	Y	Y
Inadequately trained development team members	Y	Y
Ineffective communication	Y	Y
Ineffective project manager		
Inexperienced project manager		
Inexperienced team members		
Lack of an effective project management methodology	Y	Y
Lack of commitment to the project among development team members		Y
Lack of people skills in project leadership		
Large number of links to other systems required		Y
Negative attitudes by development team	Y	Y
new: Ineffective collaboration	Y	Y
new: Ineffective coordination	Y	Y
new: Lack of trust	Y	Y
One of the largest projects attempted by the organization		
Poor project planning		Y
Project affects a large number of user departments or units	Y	
Project involves the use of new technology	Y	
Project involves use of technology that has not been used in prior projects		Y
Project milestones not clearly defined		Y
Project progress not monitored closely enough	Y	
Team members lack specialized skills required by the project	Y	Y
Team members not familiar with the task(s) being automated	Y	
Quadrant 4: Environment		
Change in organizational management during the project		
Corporate politics with negative effect on project		
Dependency on outside suppliers		Y
Many external suppliers involved in the development project		
new: Country-specific regulations	Y	Y
new: Delays caused by global distance	Y	Y
new: Lack of architecture-organization alignment	Y	Y
new: Lack of face-to-face interaction inhibits knowledge sharing	Y	Y
new: Lack of process alignment	Y	Y
new: Lack of tool/infrastructure alignment	Y	
new: Unstable country/regional political/economic environment		
Organization undergoing restructuring during the project		Y
Resources shifted from the project due to changes in organizational priorities		Y
Unstable organizational environment		

scaling agile methods do not appear to fully support *Unstable country/regional political/economic environment* and *Organization undergoing restructuring during the project*.

When looking across all quadrants far left in Fig. 4, the total number of risks we hypothesize each framework addresses

are very similar; DAD practices are associated with eliminating (termed “definitely”) or mitigating (termed “somewhat”) 58 risks, and SArE is associated with eliminating (termed “definitely”) or mitigating (“somewhat”) 57 risks. We now, in the next section, look to see whether two companies, who implement DAD and

SAFe practices, experience any of the associated risks, in order to test our hypotheses.

4.3. Empirical evidence

To gain some insight into the effectiveness of our theoretical mapping of scaling agile practices to GSD risks, we conducted a multiple case study involving two companies engaged in global software development. As described in our method, we examined a range of data collected from these cases for evidence of how scaling agile practices might have eliminated or mitigated GSD risks.

In this section, we introduce the multiple case study settings and summarize the types of risk we observed in the cases, as specified in the GSD Risk Catalog.

4.3.1. Case study setting

Company A, based in Melbourne, Australia, has been using DAD from some years. Company B, based in Dublin, Ireland, was undergoing a transition from a traditional, plan-driven approach to agile development using SAFe.

As illustrated in Fig. 5, both Company A and Company B have development teams around the world.

Case study A. Company A is a Melbourne-based company that produces highly intelligent enterprise asset management software for a global customer base. This software vendor has undergone a transition towards scaling agile development using Disciplined Agile Delivery (DAD) over the period from 2015 to 2017, and is continuing as of this writing in 2020. Accompanying this transition has been a move to provide the software not solely as an in-house product, but through a cloud-delivered “software as a service” (SaaS) model. The vendor has ten development teams across three different countries, with the engineering operation based in Melbourne and development teams in the USA and India (see Fig. 5). The marketing and after-sales and support teams are based in Australia and the USA.

The company has been actively implementing agile methods since 2003, so could be considered a mature agile practitioner. Nonetheless, the transition from their previous “hybrid-agile” to a scaling agile approach using DAD had been driven by pressures of scope and quality: namely, the inability to deliver the desired scope for planned releases, and inadequate quality of the software releases delivered into a SaaS environment.

With their DAD approach, programs run by the Melbourne site involve their four local project teams plus three project teams in the USA, and one in India. Case A is a global software vendor, that meets our definition of large scale agile GSD (see Section 2.2).

The DAD approach has the practice of self-contained teams. So all DAD (project) teams are local and co-located with their own work item lists (product backlogs) allocated by the program management. With the DAD approach a program is delivered through several projects. Hence, the program planning (create program portfolio which basically is the work item list) involving product manager, program manager, and enterprise architects will work at a global level—a program is at a global level whereas projects are local only. While the DAD project teams are local there is a daily tactical huddle where all the leadership roles (Architect, Tech Lead, and Product Owner) of the DAD project teams under a program meet. They (DAD project teams) also do a show and tell collectively for every sprint. Case A is conducting global software development and follows DAD recommendations for self-contained, co-located teams.

Case study B. The company we studied for Case Study B is Ocuco Ltd., a medium-sized Irish software company that develops practice and lab management software for the optical industry.

Ocuco Ltd., which we will refer to as “Company B” in the sequel, employs approximately 300 staff members in its software development organization, including support and management staff. Company B has annual sales exceeding 20 million, from customers in Britain and Ireland, continental Europe, the Nordic region, North America, and China.

Company B has ten development teams whose members are distributed across Europe and North America (see Fig. 5), involving approximately 50 developers in twelve countries; as such, Company B also meets our definition of large scale agile GSD (see Section 2.2).

As part of their transition from a plan-driven development approach, to agile software development following SAFe, Company B began introducing Scrum at the team level approximately six months before we began our study in 2015. SAFe is being rolled out to the various teams and projects in stages, with the newer projects leading the way to implementing SAFe practices such as PI Planning, Automated Testing, and Continuous Integration, whereas much of the organization is involved in SAFe recommended practices such as Communities of Practice.

While the purpose of the collaboration with both cases was similar, which was to observe how teams adopted or transitioned to scaling agile in a globally distributed setting, there are distinct differences in the study setting. Company A has been using agile methods for nearly two decades, while Company B had just begun a transition to agile development with SAFe at the time we started our collaboration. So, Company A would be considered a mature agile organization, which at the time of the study was scaling their development using DAD. Company B is more of a nascent agile company, introducing many new agile and lean practices as they attempt to scale agile development across teams and up the organizational hierarchy.

4.3.2. Empirical study results

In this section we present results of our investigation into the extent to which GSD Risk Catalog risks were observed (or not) in Cases A and B. We then establish whether the observed risks in the given case, were associated with practices implemented by Company A or B. We first provide an overview of this examination showing the risks not observed, that appear to support our mapping since the risk may have been eliminated. Then, we examine the risks that we observed in the cases.

Risks not observed in cases as issues. Risks shown in Table 9, were not observed in the case studies as having become issues; this set of practices relates to the first row in Table 8. This table provides some evidence that many risks can be eliminated through the adoption of scaling agile practices.

This category, where the scaling agile mitigation practice is implemented, and no related risks are observed, represents the ideal case, where a risk is identified and addressed before it becomes a problem. An example comes from Company A, involving spikes,¹⁸ that concerns the risk of *Team members not familiar with the task(s) being automated*: “When you got something like what we’re doing right now, rolling out a new dashboard, massive architectural spikes at the start... you got to bring that kind of stuff to the architecture upfront” (PA2). The DAD practice “*Inception phase involving entire DAD delivery team*”, along with “*Spikes*”, ensures that technical understanding is developed early in the development lifecycle.

¹⁸ A ‘spike’ is a “technical proof-of-concept” usually realized as a throwaway prototype (Ambler and Lines, 2012, p. 44).

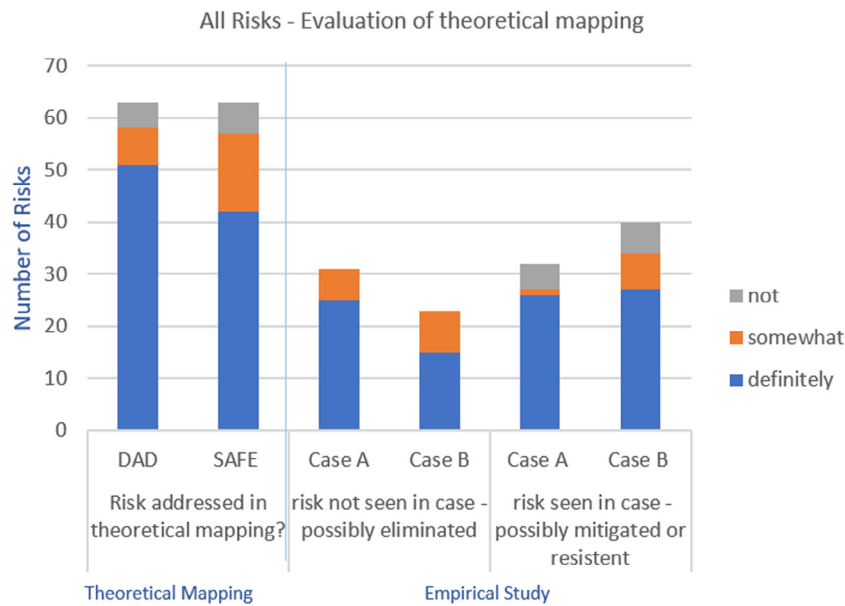


Fig. 6. Theoretical mapping evaluation – observations in two cases – All Risks. For each bar, the top (gray) segment indicates the proportion of practices *not* implemented; the middle (orange) segment indicates *partially* implemented practices; the bottom (blue) segment indicates *fully* implemented practices.

Table 7

Risks from the GSD Risk Catalog for which the associated scaling agile practices were never or rarely performed in our cases.

<i>Case A</i>	
Quadrant 3: Execution	
Large number of links to other systems required	
Quadrant 4: Environment	
Dependency on outside suppliers	
Many external suppliers involved in the development project	
new: Unstable country/regional political/economic environment	
Organization undergoing restructuring during the project	
<i>Case B</i>	
Quadrant 3: Execution	
Frequent turnover within the project team	
Ineffective project manager	
Inexperienced project manager	
Quadrant 4: Environment	
Change in organizational management during the project	
Unstable organizational environment	
Organization undergoing restructuring during the project	

Table 8

Summary of empirical evidence supporting theoretical mapping of scaling agile practices to risks in the GSD Risk Catalog.

Framework addresses?	Practice implemented?	Risk seen?	Number of risks	
			Case A	Case B
Y	Y	N	31	23
Y	N	Y	0	0
Y	N	N	0	0
Y	Y	Y	27	34
N	n/a	Y	0	1
N	n/a	N	5	5

Company B employed the SAFe practice “Develop a feature team that is organized around user-centered functionality” to address a similar instance of this risk: they employ former users of the product, who have formal qualifications in the domain, as QA staff; this ensures that implemented features are usable by actual practitioners.

Table 9

Risks from the GSD Risk Catalog not observed in either Case.

Quadrant 1: Customer mandate
Lack of user participation
Lack or loss of organizational commitment to the project
Users not committed to the project
Users resistant to change
Users with negative attitudes toward the project
Quadrant 2: Scope and requirements
Incorrect system requirements
Users lack understanding of system capabilities and limitations
Quadrant 3: Execution
Frequent turnover within the project team
Ineffective project manager
Inexperienced project manager
Inexperienced team members
Lack of people skills in project leadership
One of the largest projects attempted by the organization
Quadrant 4: Environment
Change in organizational management during the project
Corporate politics with negative effect on project
Many external suppliers involved in the development project
new: Unstable country/regional political/economic environment
Unstable organizational environment

The proportions of these risks are shown in Fig. 6 (middle bars). In Case A, practices associated with 58 risks were implemented, yet less than half (27) of these risks were observed in Case A; this suggests that DAD was effective at eliminating over half (53%) of the risks for which Company A implemented associated DAD practices. Similarly, in Case B, practices associated with 57 risks were implemented, with 23 of these risks not observed in the case, suggesting that SAFe helped eliminate up to 40% of risks for which Company B implemented associated SAFe practices.

Risks observed in cases. Table 6 shows that many GSD risks were observed to have materialized into issues in one or both cases. The evidence of risks being present is drawn from interview transcripts.

To understand whether the implementation of scaling agile practices affects the occurrence of risks, we assessed the frequency with which each company implemented each practice. Company A has been using DAD for more than five years; as

Table 10

Framework risk mitigating practice empirical evaluation. “Degree of impl.” column indicates frequency practice is performed (‘3’ is “often” performed, ‘4’ is “very often” performed, and ‘5’ is “always” performed). Table includes risks observed in one or both cases (Y) –risks in *italics* are seen in both cases. Where columns are blank indicates not addressed or seen.

Risk	DAD & Case A			SAFe & Case B		
	Framework addresses?	Degree of impl.	Risk seen?	Framework addresses?	Degree of impl.	Risk seen?
Quadrant 1: Customer mandate						
Conflict between users	somewhat	5	Y			
Lack of cooperation from users				definitely	4	Y
Lack of top management support for the project				definitely	3	Y
Quadrant 2: Scope and requirements						
<i>Conflicting system requirements</i>	definitely	5	Y	definitely	5	Y
Continually changing project scope/objectives				definitely	4	Y
Continually changing system requirements				definitely	4	Y
Difficulty in defining the inputs and outputs of the system	definitely	5	Y			
Ill-defined project goals				definitely	4	Y
<i>System requirements not adequately identified</i>	definitely	5	Y	definitely	5	Y
Unclear system requirements				definitely	5	Y
Undefined project success criteria				definitely	4	Y
Quadrant 3: Execution						
Development team unfamiliar with selected development tools	definitely	5	Y			
Frequent conflicts among development team members				somewhat	4	Y
<i>High level of technical complexity</i>	definitely	5	Y	definitely	4	Y
Highly complex task being automated	definitely	5	Y			
Immature technology	definitely	5	Y			
Inadequate estimation of project budget				definitely	4	Y
Inadequate estimation of project schedule				definitely	3	Y
<i>Inadequate estimation of required resources</i>	definitely	5	Y	definitely	4	Y
<i>Inadequately trained development team members</i>	definitely	5	Y	somewhat	4	Y
<i>Ineffective communication</i>	definitely	5	Y	definitely	5	Y
<i>Lack of an effective project management methodology</i>	definitely	5	Y	definitely	5	Y
Lack of commitment to the project among development team members				definitely	4	Y
Large number of links to other systems required				definitely	4	Y
<i>Negative attitudes by development team</i>	definitely	5	Y	somewhat	4	Y
<i>new: Ineffective collaboration</i>	definitely	5	Y	definitely	5	Y
<i>new: Ineffective coordination</i>	definitely	5	Y	definitely	5	Y
<i>new: Lack of trust</i>	definitely	5	Y	definitely	5	Y
Poor project planning				definitely	4	Y
Project affects a large number of user departments or units	definitely	5	Y			
Project involves the use of new technology	definitely	5	Y			
Project involves use of technology that has not been used in prior projects				definitely	3	Y
Project milestones not clearly defined				definitely	4	Y
Project progress not monitored closely enough	definitely	5	Y			
<i>Team members lack specialized skills required by the project</i>	definitely	5	Y	somewhat	3	Y
Team members not familiar with the task(s) being automated	definitely	5	Y			
Quadrant 4: Environment						
Dependency on outside suppliers				definitely	4	Y
<i>new: Country-specific regulations</i>	definitely	5	Y	definitely	5	Y
<i>new: Delays caused by global distance</i>	definitely	5	Y	somewhat	5	Y
<i>new: Lack of architecture-organization alignment</i>	definitely	5	Y	somewhat	5	Y
<i>new: Lack of face-to-face interaction inhibits knowledge sharing</i>	definitely	5	Y	definitely	5	Y
<i>new: Lack of process alignment</i>	definitely	5	Y	definitely	3	Y
<i>new: Lack of tool/infrastructure alignment</i>	definitely	5	Y			
Resources shifted from the project due to changes in organizational priorities				somewhat	4	Y

such, Company A always performs all except five practices in DAD that address risks in the GSD Risk Catalog. Evidence of practice implementation by Company B comes from self-assessment surveys.

Company B was in the middle of a transition to SAFe when we began our study; consequently, they perform the SAFe practices at different frequencies, ranging from “never” or “rarely”, through “occasionally” and “often”, to “very often” or “always”. We were able to establish the extent to which the SAFe practices were implemented through a series of self-assessment surveys (Razzak et al., 2017, 2018) (participants detailed in Table 5). Based on these survey responses, we determined that, in Case B six risks had no associated practices that were performed more often than “rarely”.

Table 7 lists the risks from the GSD Risk Catalog for which either company rarely or never performs any of the associated framework practices (see Tables 11 and 12 for our theoretical

mappings). These results are not surprising: in the case of Company A, they have a stable organization and operate in developed countries with a stable political and economic environment; they also do not engage outside suppliers or interface with many systems.

Company B likewise has a stable if rapidly growing organization with experienced management, and experiences very low turnover. So it is perhaps to be expected that they do not implement practices aimed at reducing the risks associated with these characteristics.

In Case B, there are instances where associated scaling agile practices were not (yet) fully implemented. There are 22 risks in Table 10 where the mode of Company B’s frequency of performance of the associated practices is “often” (‘3’ is “often” performed, ‘4’ is “very often”) indicating they don’t always perform these practices. As such, risks associated with these practices might not be fully addressed, and so could be expected

to become problems occasionally. Conversely, Case A always performs associated DAD practices. This might account for the fact that fewer risks (27 vs 34) were observed in Case A than Case B (Fig. 6, rightmost bars).

Table 8 summarizes the results in terms of the frequencies of combinations of risks addressed and seen, and practices implemented. The first column indicates a risk has been addressed by a scaling agile framework; the second column indicates whether the mapped scaling agile practices have been implemented by the company; the third column indicates a risk has been observed as an issue in the associated case; and, the fourth and fifth columns show the number of risks that are in the state indicated by the first three columns for each case.

The first row shows the number of risks from the GSD Risk Catalog that have practices mapped from the respective scaling agile framework (DAD for Case A, and SAFe for Case B), but have not been observed in the respective case study, and have mapped practices implemented by the case company. This row supports the theoretical mapping: the practices were implemented and the risks were not seen to be present in the case study organization, indicating that the practices were possibly effective in eliminating the associated risks.

The second row shows the number of risks that have practices mapped to them, were observed to have occurred in the respective cases, but the practices were not implemented. This row might indicate areas where the case companies could improve their practices to address observed risks, but no risks were found in this category.

The third row shows the number of risks that have practices mapped, but were not observed in the cases, nor were the associated practices implemented. This category is also empty. The last two rows show the number of risks that do not have associated scaling agile practices. In all but one instance these risks were not observed in either case. The one risk observed in this category—from Case B—is *Organization undergoing restructuring during the project*. This risk stemmed from the transition from a waterfall to agile development approach: the product owner of a project focused on a large customer noted, “So, we worked in waterfall fashion in the past and I think this is difficult for people to move from the waterfall way to the Scrum way” (PB10).

By contrast, the fourth row shows the number of risks that have practices mapped, were observed in the cases, and the practices were also implemented in the cases. Table 10 lists the risks involved in the fourth row of Table 8. This table shows that few (3 of 8) risks from the **Customer Mandate** quadrant were seen in either case when the associated actions were implemented. Conversely, the majority (8 of 10) of **Scope and Requirements** risks were seen. Also, nearly two-thirds (9 of 14) of the risks in the **Environment** quadrant were observed in at least one case, and nearly half (6 of 14) were seen by both cases. The majority (24 of 44) of the risks listed in Table 10 are from the **Execution** quadrant. Section 5 examines the implications of these risks.

Fig. 6 shows the proportions of these categories in rows one and four of Table 8 (risks addressed, practices implemented, and issues not seen for row one, or seen for row four). Of note is the grayed portion of risks at the top of the stacked bars (representing those risks with no associated practices); these only appear in the category of “risk seen”. Conversely, it is only those risks with associated practices that are *not seen*. The case where the risk is seen, and the practice is implemented, is harder to interpret; we discuss the possibilities in the next section (Section 5).

As Fig. 6 shows, the majority of risks that were *seen* as issues in either case are also “definitely” addressed (blue portion of the bars), while a slightly higher proportion of risks *not seen* in both cases were only “somewhat” addressed. This suggests that the theoretical strength of mitigation or elimination does not strongly affect whether the risk was observed in a case.

It seems some risks cannot be eliminated, but can be mitigated.

Fig. 7 shows the proportions of risks addressed, seen, and associated practices implemented, for the **Customer Mandate** quadrant (quadrant 1). Only three risks in this quadrant—*Conflict between users*, *Lack of cooperation from users*, and *Lack of top management support for the project*—were realized as issues in the cases.

For example, Company A experienced *Conflict between users*: “Because they’ve [product managers] been in the US and I’ve been here... That’s the sort of things that people could not appreciate and they wouldn’t even hear it because they know the product, it works like this, and they assumed the customers wanted the way it was” (PA1).

An project manager provides an example of *Lack of cooperation from users* seen in company B: “... now, in an Agile world there is no way that I could tell them when they are going to get done until the estimate is there, until we start a sprint planning... you can’t just say, ‘we are doing Agile, so, you got to wait for our next planning...’” (PB11).

Fig. 8 shows the proportions of risks addressed, seen, and associated practices implemented, for the **Scope and Requirements** quadrant (quadrant 2). The companies in both cases experienced the problems in this quadrant, including *Conflicting system requirements* and *System requirements not adequately identified*.

For example, in Case A, one participant mentioned, “... everything has to be user stories, sometimes we think that’s the underlined problem ... we get too focused on that because while it’s good to have those stories to get you going ... the magic happens every day in the team making adjustments, embracing that, rather than trying to design user stories to end”. (PA2) Similarly, in Case B, a senior manager noted, “Like I remember a few weeks ago we were with a customer from Scotland that asked for a particular improvement in a particular area, and the whole issue had gotten completely, you know, misunderstood by development and had been sitting in the backlog for a very long time, and [we] would set up a call with the customer to try and clear up what exactly what they want. Because a lot of the time there is this ‘Chinese whispers’ thing going on with... people misunderstanding things, recording the wrong request and so on” (PB4).

In Case B, the company encountered a problem where the team responsible for maintaining and enhancing the core product was devoting more effort to fixing issues raised by customers than to implementing new features; this caused the core product development to drift away from the product road-map, resulting in important features being delayed. This is another example of the *Conflicting system requirements* risk; the associated SAFe practice, “*Continuously communicate emerging requirements and opportunities back into the program vision through product owner*,” mitigates this risk by ensuring the product vision takes into account issues raised by customers; also, the product owner would be fully aware of the product vision and associated road-map, and therefore would be able to make correct decisions about the relative priorities of fixes and new features. As a result of recognizing the link between this risk and the corresponding SAFe practice, Company B moved from part-time product owners who doubled as technical support staff, to dedicated personnel who focus solely on the product owner role. This does not *eliminate* the risk of conflicting system requirements, but it does reduce the impact by ensuring the product owner prioritizes conflicting requirements properly.

Another example from Case A, relating to conflicting requirements, shows the power of spikes to mitigate (rather than eliminate) this risk: an interviewee noted, “absolutely, it [requirements conflict] mainly happens in inception, but throughout the

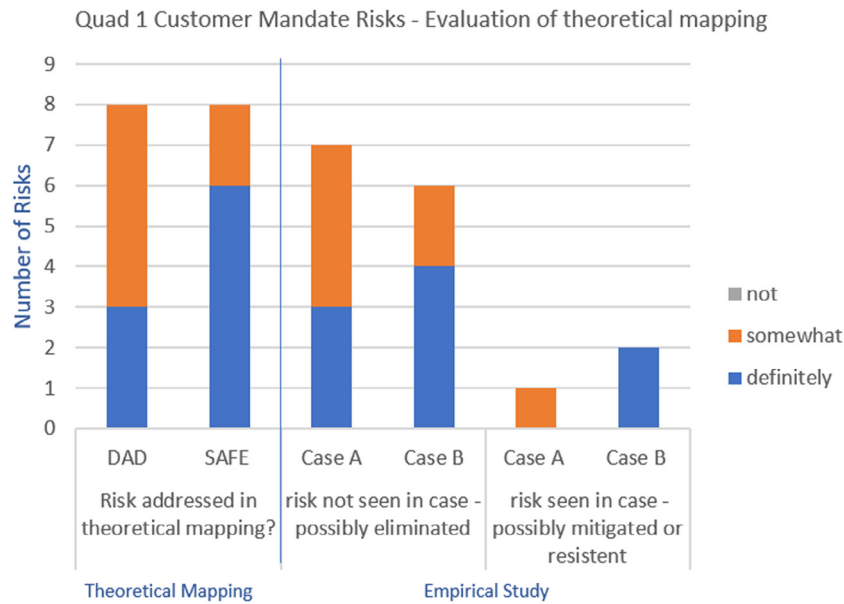


Fig. 7. Theoretical mapping evaluation – observations in two cases – Quad 1.

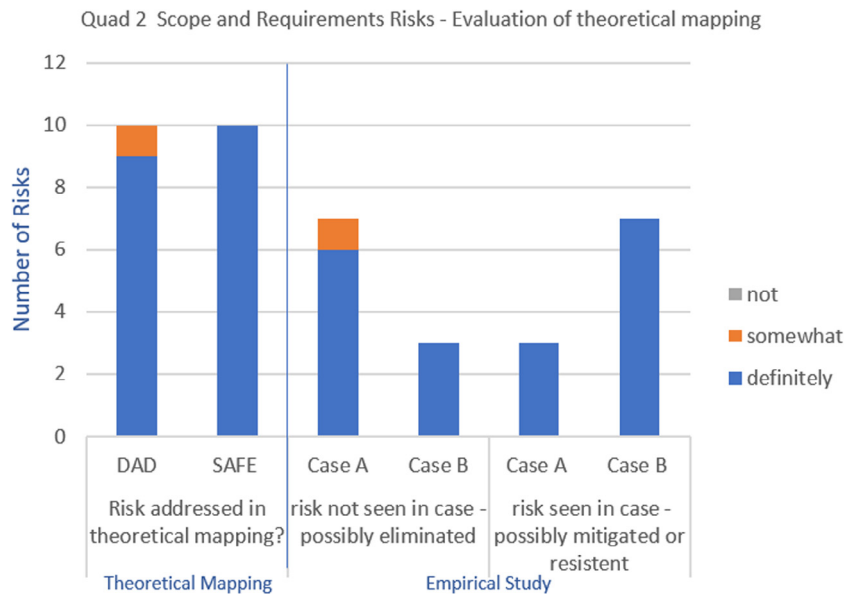


Fig. 8. Evaluation of theoretical mapping – observations in two cases – Quad 2.

project as we get closer to needing to execute in a specific story, we analyze risks and potentially do spikes on things, hopefully before those stories are done" (PA7). The DAD practice, "*Engineers in the inception phase minimize risk through spikes,*" helps to resolve conflicts when they arise.

Some risks, all from the **Execution** quadrant (quadrant 3, Fig. 9), suggest problems related to agile development in a GSD context: 1. *Lack of an effective project management methodology*, 2. *Ineffective coordination*, 3. *Ineffective collaboration*, 4. *Ineffective communication*, 5. *Lack of trust*, 6. *Negative attitudes by development team*, and 7. *Inadequate estimation of required resources*.

For Case A, *Lack of an effective project management methodology* manifests itself in technical issues. For example, one interviewee described how automation affects their ability to deliver: "Automation has played a bigger role for us, we reduced the release cycle from 6 weeks to 2 weeks, without automation to give us that nightly check ability, it's very difficult, and our product is

very complex, just setting up the environment and do manual testing is weeks of effort". But they are making progress: "And also, in the criteria being done, it's as close as possible to 100% automation... it's something we didn't have before. We used to concentrate on a lot of manual work, now we say to the team, it's just your responsibility as QA to automate" (PA6).

Both companies experienced *High level of technical complexity*, and *Team members lack specialized skills required by the project*, risks that are related to technical complexity of development. For example, a participant from Case A reflected on the long-term effects of legacy code: "The thing that we have to deal with is that we have a lot of legacy code that was already built in those 3 tiers, and we brought a lot of that across... we were sort of forced to stay with it because the need to reuse our code, we just didn't have time to build everything... lot of room for improvement I think, if we started from scratch I think things will be a bit different" (PA2). A scrum master in Case B observed, "Even

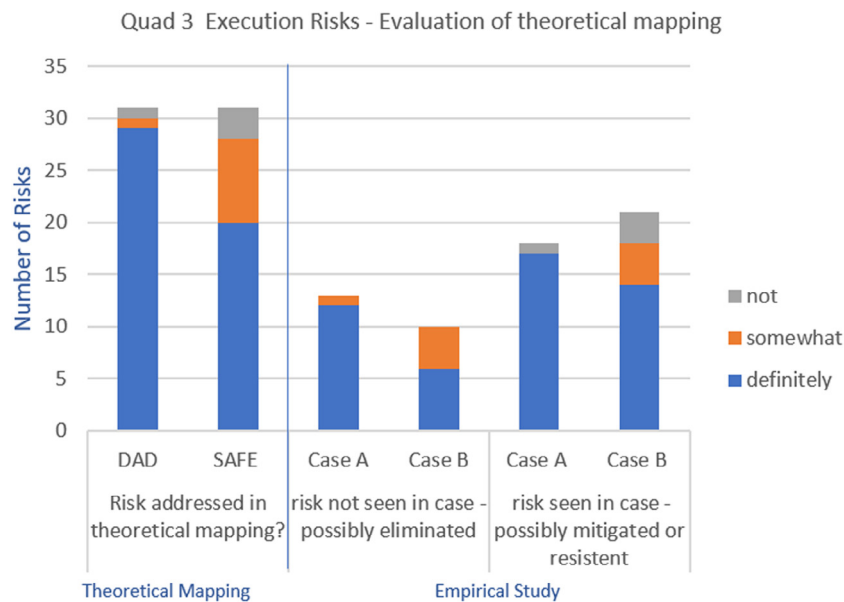


Fig. 9. Theoretical mapping evaluation – observations in two cases – Quad 3.

though I have [Sr. Developer] here in Portland [to assist] but he is very heavily involved with other things. So, he is quite busy” (PB10). Similar to the **Environment** quadrant risks, these two risks are inherent aspects of the project and so can be mitigated (for example, by enlisting more experienced developers) but not eliminated.

Fig. 10 shows the proportions for the **Environment** quadrant (quadrant 4). Both companies experienced the *Lack of architecture-organization alignment* risk in this quadrant. In Case A, this risk manifests itself in their move to the cloud: “Yes, we have on the cloud for example, we have about 5 sort of databases, each hosting a number of tenants, at least 2 different geographies, US and Europe. I think our biggest database has got like 50 tenants in it, and the other are starting to fill up as well, and we’ve reached the limit before, and we just simply start up another server. It is not a simple matter of one server, you have to, it’s all clustered machines, you have to through 3–4 different servers you have to bring up to every new database” (PA2). In Case B, where growth by acquisition in different countries has changed the organizational structure, a developer commented, “... it happens regularly in every week—that someone forgets to unlock the unit. So, when that happens—previously all developers were in Dublin and that shouldn’t be a problem—now we have a problem” (PB13). This introduces *Delays caused by global distance* also in quadrant 4: “The difficulty is that, when I [PO] am here on-site I only get couple of hours in the morning to deal with Dublin stuff. If I do not get the things that I need from them even though these two hours I am pretty much isolated for the day” (PB10).

We identified several instances where the risks observed related to the very process of transition to the new framework (Case B), or in the process of ongoing adaptation and adopting new practices as the process matured (Case A). For instance, among the risks observed in each case we noted that *Inadequate estimation of project budget* was an issue, and, for Case A, that DAD was hoped to offer some solutions “... now the DAD process is coming” (PA4). *Inadequately trained development team members* appeared to be an issue, again with Case A indicating that new DAD practices and roles involved “a completely different way of thinking for us” (PA2). The risk that *Team members lack specialized skills required by the project* was evidenced differently in Case A as team members “lacking the expertise of integrating all the

different products and offerings under one company product” (PA2), and in Case B “don’t fully understand [the] job” (PB10), so both new technical demands and new roles posed challenges.

In the next section, we discuss the implications of these results.

5. Discussion

This study was motivated by the lack of empirical evidence as to the efficacy of scaled agile frameworks in general, and specifically in managing risk in Global Software Development (GSD) settings, where teams are distributed around the world. Our results in Section 4 provides a promising set of responses to address this gap, as captured by our research question:

How does the adoption of scaling agile framework practices address global software development risks?

The comprehensive GSD Risk Catalog we derived complements the earlier software risk framework of Wallace and Keil (2004), by incorporating a further ten new GSD specific risks, making a total of 63 risks. The new GSD specific risks are situated within the **Environment** and **Execution** quadrants of Wallace and Keil (2004).

The theoretical mapping suggests that the two scaling agile frameworks investigated, DAD and SAFe, could contribute strongly to eliminating or mitigating risks in our GSD Risk Catalog. However, the empirical assessment of issues related to those risks provides a more nuanced picture of the frameworks and their strengths and limitations. On the one hand, nearly half (31 of 58 or 53% for Case A, and 23 of 57 or 40% for case B) of risks were *not* seen as issues for the companies when they implemented the associated practices. This suggests that the respective scaling agile frameworks are effective at totally *eliminating* a subset of GSD risks.

Fig. 6 shows that a majority of risks are addressed by both DAD and SAFe, both from a theoretical and empirical point of view. The vast majority of risks seen in Case A are addressed by DAD, while a slightly lower proportion of risks seen in Case B are addressed by SAFe. Looking at the figure, in both cases it appears that the *strength* of mitigation has little influence on whether the company will experience the risk.

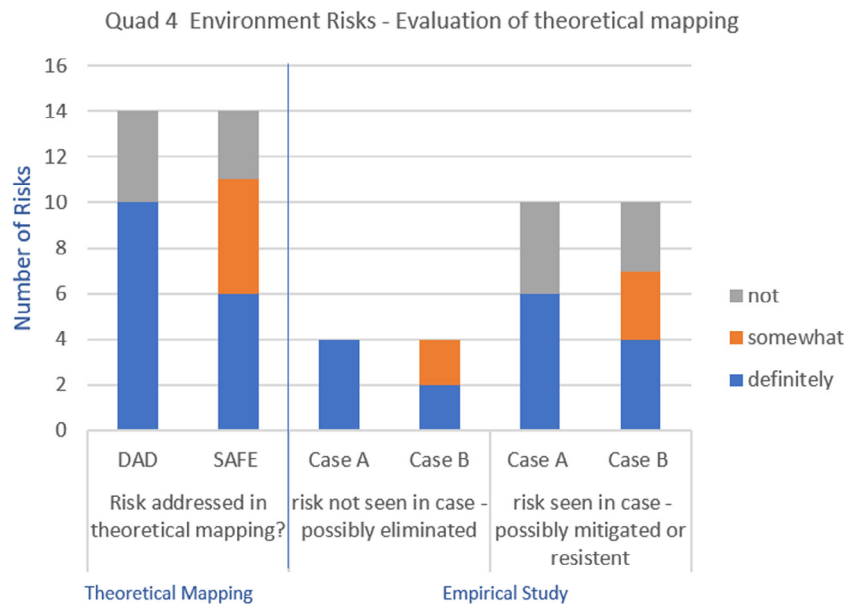


Fig. 10. Theoretical mapping evaluation in two cases — Quad 4.

At the time of our study, Company B was in the middle of their transition from plan-driven to agile development with SAFe. As such, for the risks that were seen as issues in Case B, a majority (22 of 34, or 65%) of associated SAFe practices were performed less than “always” (Table 10, column five). So this could account for the somewhat higher number of issues seen in Case B than in Case A.

The remaining risks (that were observed in the cases) are possibly *mitigated* rather than eliminated. That is, the risk became a problem, but its *impact* was reduced by the associated scaling agile practices. Wallace and Keil note that companies have low control over the risks in the **Customer Mandate** (quadrant 1) and **Environment** (quadrant 4) quadrants (see Fig. 1); we would therefore not expect risks in these quadrants to be eliminated, because they result from external forces out of control of the companies.

Delays caused by global distance is an example of an **Environment** risk: if teams are located in Vancouver and Dublin, or Melbourne and New York, the only way to eliminate delays caused by lack of timezone overlap is to shift working hours, or close one location; neither of these is likely to be practical. Verner and colleagues also recognize that global distance (Noll and Beecham, 2016) introduces difficulties for agile development, noting that “Lack of synchronous communication in agile development causes problems” and “Collaboration difficulties caused by geographic distance in agile development may cause misunderstandings and conflicts” (Verner et al., 2014, p. 64). While Verner and colleagues don’t offer any mitigation advice for these risks (unlike other risks associated with agile methods in GSD), we assert that SAFe practices such as “Calculate the Cost of Delay” and “Manage and optimize the flow of value through the program using various tools, such as the Program and Value stream Kanbans and information radiators” (Beecham et al., 2020) can reduce the impact of these risks by highlighting time-critical information.

Risks in the **Customer Mandate** quadrant (quadrant 1) are determined by the users, who are also largely out of the control of the development organization. This would be the case for the *Conflict between users and Lack of cooperation from users* risks. DAD practices such as “Product manager and product owner roles, business case, feature funnel, Mandated DOD [definition of done], TDD [test-driven development] practices, and including manual systems testing before deployment into production environment” and

“DevOps practices—collaboration between Operations and SE (program and project) teams” reduce the impact of this risk (Beecham et al., 2020), but do not eliminate it. Verner and colleagues also recognize continuous integration and test-driven development as ways to mitigate risks associated with agile development in GSD (Verner et al., 2014, Table 10, item 2, p. 64).

Some risks in the **Scope and Requirements** quadrant (quadrant 2) also depend on the customer. *Continually changing system requirements* is an example: if the customer decides to change some of the requirements, for example in response to changing market conditions, the development teams need to react accordingly or the customer will not be satisfied. SAFe practices such as “Continuously communicate emerging requirements and opportunities back into the program vision through product owner,” “Work with stakeholders to understand the specific business targets behind the user-system interaction,” and “Perform system demo as near as possible to the end of the iteration” help ensure that requirements changes are detected and accounted for as soon as possible. DAD has similar practices, including having a “product manager who does market investigation and feedback on potential features and functionalities from potential and current customers” and a “product owner with UX responsibility, for storyboarding, do user research in the field”.

We note that overall, few risks in the **Customer Mandate** quadrant were seen by either company (Fig. 7), despite this quadrant being classified as having a “low” perceived level of control (Wallace and Keil, 2004) (Fig. 1). Conversely, Company B experienced nearly two-thirds (18 of 28) of the risks in the **Scope and Requirements** quadrant (see Fig. 9), despite this quadrant having a “high” perceived level of control (Wallace and Keil, 2004).

Agile methods emphasize the need to *accept* rather than prevent or control requirements change through requirements freezes and control change boards in order to deliver the most value to customers and end-users (Beck and Gamma, 2000); possibly, this shift in attitude toward requirements and users means that Wallace and Keil’s original classification of the perceived level of control of the **Customer Mandate** and **Scope and Requirements** quadrants appears to be reversed in the context of agile software development.

Agile methods also promote intense interactions (both formal and informal) among stakeholders, including the development organization, customers, and users. This means that there is a closer

relationship between developers and other stakeholders than in traditional plan-driven approaches. Yet, according to Pikkarainen et al. agile methods applied in larger development situations involving multiple external stakeholders “can sometimes even hinder the communication” (Pikkarainen et al., 2008).

In our theoretical mapping, we contended that many scaling agile practices, such as “all hands” PI ceremony, recognize the importance of all distributed team members meeting face-to-face, and that coupled with enhanced tools for video-conferencing and information sharing, and daily remote stand-up meetings, would have contributed to a highly collaborative environment. Yet the major issues experienced by both cases appear in the **Environment** and **Execution** quadrants. This may be due to the situation where the “application of agile practices causes problems in distributed development because of the degree of interaction between stakeholders and number of face-to-face meetings needed” (Verner et al., 2014). It would appear that Global Software Development impedes this kind of interaction (Noll et al., 2010): of the 17 risks observed in both cases, eight are the new GSD risks added to the GSD Risk Catalog to augment Wallace and Keil’s inventory:

1. *Ineffective collaboration,*
2. *Ineffective coordination,*
3. *Lack of trust,*
4. *Country-specific regulations,*
5. *Delays caused by global distance,*
6. *Lack of architecture-organization alignment,*
7. *Lack of face-to-face interaction inhibits knowledge sharing, and*
8. *Lack of process alignment.*

Lack of architecture-organization alignment is a recurring issue in GSD (Sievi-Korte et al., 2019a) that appears to be a risk that can at best be mitigated; recent studies into architectural design in GSD indicate that the architecture does not always reflect the structure of the organization, where architects interviewed stated that working across geographic boundaries required new strategies (Sievi-Korte et al., 2019b).

Also, both cases experienced *Inadequate estimation of required resources*. Estimation is a persistent problem in any software development context; in their study of 145 software projects, Kitchenham and colleagues found that less than two-thirds of the projects produced estimates within 25% of the actual time required (Kitchenham et al., 2002). So, inadequate estimation may be a fact of life in software development, that cannot be eliminated, especially in a global setting.

Agile methods accept that initial estimates are not accurate, but that they will improve as development teams gain more experience with the requirements and their own capabilities. As such, this risk is one that is likely to be a problem initially, but will be mitigated over time; a scrum master in Case B confirmed this, observing, “For the first few sprints we completely over-committed to a lot of stuff which we just couldn’t deliver. So, we are trying to fit in with the velocity that is based on the size of the team” (PB10).

It is surprising that *Lack of trust* is an issue in both cases, since the Agile Manifesto values “individuals and interactions over processes and tools”, and among its twelve principles are “Build projects around motivated individuals. Give them the environment and support they need, and trust them to get the job done” (Beck et al., 2001). As noted in Section 4.3.2, in Case A, this is related to *Ineffective communication* that stems from GSD, while in Case B, some of the problem derives from the fact that the company is still in the process of moving from a plan-driven to an agile development approach.

Verner and colleagues also recognized that GSD presents difficulties for agile software development. Six risks from Verner et al.’s list (Verner et al., 2014, Table 10, p. 64) specifically concern agile development; five of these explicitly cite the impact of GSD on communication, collaboration, or coordination: 1. “Application of agile practices causes problems in distributed development because of the degree of interaction between stakeholders and the number of face-to-face meetings needed”, 2. “Lack of synchronous communication in agile development causes problems”, 3. “Collaboration difficulties caused by geographic distance in agile development may cause misunderstandings and conflicts”, 4. “Poor communication bandwidth for agile development causes problems with communication and knowledge management”, 5. “Lack of tool support for agile development causes problems with agile practices”, and 6. “Large teams involved with agile development can cause problems related to communication and coordination”.

The fact that both cases experienced these risks suggests that certain risks are endemic to GSD, and, due to their impact on communication among developers, customers, and users, may be beyond the capabilities of scaling agile frameworks to eliminate.

However, we know from observation that both companies can exploit communication and coordination technologies such as video conferencing, real-time “chat”, and issue management software, to effectively implement scaling agile practices related to communication, coordination, and collaboration. And we know from associated studies with Case B, that practitioners are highly motivated (with a few exceptions) (Noll et al., 2017), a status further supported by very low staff turnover (Bass et al., 2018b; Beecham and Noll, 2015; Hall et al., 2008). So even though we still observed instances of risks related to communication and coordination becoming issues, this does not mean that scaling agile frameworks are ineffective at addressing GSD risks; rather, we hypothesize that scaling agile practices can reduce the probability of GSD risks becoming problems, and potentially reduce the impact of such problems when they do occur, but that these risks cannot be eliminated entirely.

In summary, the scaling agile philosophy of collaboration, both horizontally and vertically throughout the enterprise (Leffingwell et al., 2016; Razzak et al., 2018), appears to eliminate or mitigate software development risks through better sharing of information, joint decision-making, progressive refinement, and adaptation of goals. As noted in Ambler and Lines (Ambler and Lines, 2012, p. 8), “high collaboration is a hallmark of agility”. A strong governance structure, and set of supporting roles and practices within each framework, enable such behavior, and contribute to reducing risk and thereby to the improved outcomes advocated by each method (Noll et al., 2016; Lal and Clear, 2017).

5.1. Threats to validity

5.1.1. Construct validity

We have used qualitative data from interviews and observations to determine whether risks from our GSD Risk Catalog were experienced by either of the companies in our case studies. This data collection did not specifically set out to examine risks that the organizations were experiencing. However, the broader questions asked as part of these longitudinal studies were to understand the challenges and issues that the practitioners experienced in their development, whilst transitioning to their target scaling agile frameworks, in a GSD setting.

The measures we used to test whether a scaling agile practice mitigated a given risk, was to look for the absence of qualitative evidence that a company experienced a risk, combined with evidence that the company had implemented practices deemed to address the risk, as an indication that the practices did address

the risk by eliminating it. But this relationship may be coincidental; the risk might not have materialized anyway, or might have been eliminated by some other means.

Also, we speculated that the *presence* of risks combined with practice implementation may indicate that the practices do not eliminate risks. Again, this could be coincidental.

Future work could involve following-up with the participants to verify that the scaling agile practices helped, or did not help, to eliminate risks, and why.

5.1.2. Internal validity

Our approach to mapping agile practices to risks, which we deemed to mitigate them, was careful, involved inter-rater cross-checks, and was supported by evidence from each framework's documentation and case data. But, we may have misinterpreted some of the risks and corresponding practices. Moreover, the complexity and interlocking nature of many roles and practices were such that making a direct mapping was challenging in some cases. To what extent the identified practices were mutually self-reinforcing across cases was hard to determine especially since the cases represent different application domains. The observed differences and similarities therefore, across frameworks, may be due to situational differences rather than framework differences.

There are trade-offs between types of case study (Bass et al., 2018a) where for example a multiple case study has several validity gains over a single case study in having higher transferability (broader relevance to other cases, and observations are made in a naturalistic real-world setting) and higher confirmability (with more support for findings across studies) (Lincoln and Guba, 1985). Yet, the level of credibility is considered lower, where, for example, confusion can exist in cause and effect across cases due to the varied settings, that are only visited at one point in time (Leonard-Barton, 1990).

As noted above, the presence or absence of risks might not be related to scaling agile framework practices, but rather could be the result of confounding variables. For example, we did not observe the *Unstable country/regional political/economic environment* risk in either case. But this is almost certainly because both companies are located in, and sell to, stable economies and countries. Similarly, Case A experienced the *Project involves the use of new technology* risk; since, at the time of our study, they were moving to SaaS, this risk may simply have been new enough that it was not yet mitigated.

There is also a possibility that researcher bias influenced the mappings that form the core of this study. To reduce this possibility, we first did independent mappings as individuals, then compared the results. We resolved all disagreements by discussion within each pair of researchers, thus achieving complete agreement in the end.

The data collected from Case A might be limited by the fact that interviews and observations were done in one location. Although the Melbourne site has interactions with teams and customers across the globe, and so the participants would have a good understanding of Company A's processes, they might not have a complete view of the issues faced by Company A, nor the real frequency with which DAD practices are performed at other sites. As this would tend to underestimate the issues, and overestimate the frequency of performance, we feel our interpretations stemming from Case A are conservative.

5.1.3. External validity

Our findings suggest that these conclusions may hold for other companies implementing scaling agile frameworks, but in particular, the DAD and SAFe frameworks, as we observed that two companies with different characteristics and product domains nevertheless experienced similar risks, especially as related

to Global Software Development. That said, each organization's implementation will inevitably have its own characteristics, as DAD and SAFe are large, complex, and adaptable frameworks, frequently supported by experienced agile coaches (Ambler and Lines, 2012, p. 73), who are needed to help tailor the frameworks to the circumstances of each adopting organization. Therefore, the contributing role of context would need further investigation through additional studies.

6. Conclusion

In this paper, through a three-phase process, we have illustrated how two scaling agile frameworks—DAD and SAFe—largely address the 63 software development risks in the GSD Risk Catalog.

The first phase involved identifying Global Software Development risks faced by software development organizations, by examining the literature on risks in both conventional and GSD contexts. The result was a GSD Risk Catalog of 63 risks, divided into four quadrants following Wallace and Keil (2004): 1. **Customer Mandate**, 2. **Scope and Requirements**, 3. **Execution**, and 4. **Environment**.

In the next phase, we identified potential risk mitigation and elimination practices in the two scaling agile frameworks. We did this by constructing a theoretical mapping of scaling agile practices to risks in the GSD Risk Catalog. We compared practices from DAD and SAFe to risks in the GSD Risk Catalog, creating a mapping of practices to risks that shows how scaling agile practices eliminate or mitigate those risks.

To assess the strength of the scaling agile frameworks to mitigate or eliminate risk, and avoid criticism that we “speculated that the strategy would have helped observed problems” (Verner et al., 2014), we performed an empirical assessment of the resulting theoretical mappings. This empirical assessment determined the frequency with which practices in each framework were performed in two companies, and the risks encountered by those companies. Through examination of observation and interview notes and transcripts, and self-assessment survey results, from two case studies of global software companies, we were able to support much of our theoretical mapping, and provide evidence that both DAD and SAFe appear to eliminate or mitigate the majority of risks in the GSD Risk Catalog.

Thus, this study adds to the limited empirical evidence of the efficacy of scaling agile frameworks. It suggests that the claims that these frameworks are risk and value-driven approaches have some validity.

Of the four quadrants in the GSD Risk Catalog, the **Customer Mandate** quadrant appears to be better addressed through the SAFe framework than DAD. **Scope and Requirements** risks are addressed well by both methods. **Execution** risks are better mitigated by DAD than SAFe, and **Environment** risks are less well addressed by either approach. This suggests that the **Environment** set of risks are less amenable to being addressed by a process framework.

A further outcome of creating the GSD Risk Catalog is the addition of ten new risks related to Global Software Development, which were not identified in the Wallace and Keil inventory (Wallace and Keil, 2004) (see Table 2). These new risks appear to be endemic and suggest a *risk tariff* in GSD: all of these except *Lack of tool/infrastructure alignment* and *Unstable country/regional political/economic environment* were experienced by both companies.

The result of these three phases is a scaling agile risk theoretical mapping that shows how two scaling agile frameworks—Disciplined Agile Delivery and the Scaled Agile Framework—can potentially eliminate or mitigate software project risks in global software development.

Table 11

DAD practices mapped to GSD Risk Catalog risks.

<i>GSD Risk Catalog risk/Practices</i>	<i>Level</i>
Quadrant 1: Customer mandate	
<i>Conflict between users</i>	Somewhat
Product manager responsible for features and functionalities, product owner with UX responsibility. Business case, feature funnel, storyboarding, user stories, DOD including acceptance tests, iteration show and tell, beta testing, production environment test, DevOp practices, real time monitoring of the product environment.	
<i>Lack of cooperation from users</i>	Somewhat
Product manager, product owner. Business case, feature funnel, storyboarding, DevOp practices, real time monitoring of the product environment by end-users, feedback from end-users.	
<i>Lack of top management support for the project</i>	Definitely
Portfolio manager, portfolio planning team. Portfolio planning including feature funnel practices, business case including six monthly presentation to the executives for approval.	
<i>Lack of user participation</i>	Somewhat
Product management, vision planning, program management, story boarding. Product manager, proxy user product owner, on-site customer with UX responsibility.	
<i>Lack or loss of organizational commitment to the project</i>	Definitely
Product management, portfolio management, program management teams. Product planning, portfolio planning, and program planning. Work item list.	
<i>Users not committed to the project</i>	Somewhat
Product manager, product owner, domain expert. Business case, feature funnel, storyboarding, DevOp practices, real time monitoring of the product environment by end-users, feedback from end-users.	
<i>Users resistant to change</i>	Definitely
Product management, vision planning, program management, story boarding. Product manager, proxy user product owner, on-site customer with UX responsibility. Both roles require individuals with solid local technical product knowledge.	
<i>Users with negative attitudes toward the project</i>	Somewhat
Product manager of features and functionalities, product owner with UX responsibility). Business case, feature funnel, storyboarding, user stories, DOD including acceptance tests, iteration show and tell, beta testing, production environment test, DevOp practices, real time monitoring of the product environment.	
Quadrant 2: Scope and requirements	
<i>Conflicting system requirements</i>	Definitely
Product manager who does market investigation and feedback on potential features and functionalities from potential and current customers. Vision planning, product owner with UX responsibility, for storyboarding, do user research in the field. Engineers in the inception phase minimize risk through spikes. Vision planning, business case, feature funnel.	
<i>Continually changing project scope/objectives</i>	Definitely
Program management creates work item list for the entire program. Delivery teams choose user stories for their work item list (project). Program daily huddles. DOD, UAT. Programs are time-boxed, defined, and relatively short duration.	
<i>Continually changing system requirements</i>	Definitely
Product manager who does market investigation and feedback on potential features and functionalities from potential and current customers. Vision planning, product owner with UX responsibility, for storyboarding, do user research in the field. Engineers in the inception phase minimize risk through spikes. Vision planning, business case, feature funnel.	
<i>Difficulty in defining the inputs and outputs of the system</i>	Definitely
Program management including Enterprise architects and Product owners. Program work item list, architecture design, story boarding, user stories. Project-level product owners, team architects, software, and QA engineers. Inception phase including spikes to understand users stories. UAT during construction phase.	
<i>Ill-defined project goals</i>	Definitely
Product owner, Team lead, Architecture owner. Project work item list, DOD, inception phase, construction phase, transition phase.	
<i>Incorrect system requirements</i>	Definitely
Product management. Vision planning to identify high level/epic stories. Product owner, story boarding.	
<i>System requirements not adequately identified</i>	Definitely
Program management including Product owners, enterprise architects, program manager. Program work item list, story boarding, user stories. Project level product owners, team architects, software and QA engineers. Inception phase project item list, spikes, user stories estimated and prioritized using user story points,	
<i>Unclear system requirements</i>	Definitely
Program management including enterprise architects. Program work item list. Architect owner. Spikes.	
<i>Undefined project success criteria</i>	Definitely
Portfolio, program management, project management. Portfolio manager, program manager, product owners and enterprise architects. Release plan (roadmap) at portfolio level; program level work item. Iteration show and tell, program daily hurdles, project level DOD and story points.	
<i>Users lack understanding of system capabilities and limitations</i>	Somewhat

(continued on next page)

Table 11 (continued).

GSD Risk Catalog risk/Practices	Level
Product managers, customer success team. Vision planning. Program management does storyboarding for UX.	
Quadrant 3: Execution	
Development team unfamiliar with selected development tools	Definitely
DAD delivery teams are empowered to make tool decision	
Frequent conflicts among development team members	Definitely
Program management, including program manager, product owners, Enterprise architect, team leads, Architect owners. Daily tactical huddle. DAD delivery team. Primary and secondary roles. Daily stand-up meeting, self-organizing teams, empowerment.	
Frequent turnover within the project team	Definitely
Self-contained DAD delivery team. Primary and secondary roles, including three leadership roles and separate HR manager. Self-organizing teams, empowerment, task sharing, T-skilling.	
High level of technical complexity	Definitely
Entire DAD delivery team involved in inception phase. Spikes. Provide story points and estimates. Sprint planning.	
Highly complex task being automated	Definitely
Entire DAD delivery team involved in inception phase. Spikes. Provide story points and estimates. Sprint planning.	
Immature technology	Definitely
DAD delivery teams are empowered to make tool decision. Temporary roles in DAD delivery teams. Coaching and training for upskilling on unfamiliar technology.	
Inadequate estimation of project budget	Somewhat
Features are delivered through program.	
Inadequate estimation of project schedule	Definitely
User stories and estimates. Priority setting. Spikes. User stories, story points. Re-estimating.	
Inadequate estimation of required resources	Definitely
Self contained DAD delivery team with up to 13 individuals. Primary and secondary roles three leadership roles and separate HR manager. Empowered self-organizing teams. Task sharing. T-skilling.	
Inadequately trained development team members	Definitely
Project management including Product owner, Architecture owner. Self-organizing teams, T-skilled, coaching by Architecture owner and Product owner, pair programming. Bring in outside expertise to develop and upskill team members.	
Ineffective communication	Definitely
Project management: empowerment, self contained delivery team (PO, AO, TL, SE, QA, T. Writer etc.), self-organizing, pair programming.	
Ineffective project manager	Definitely
Shared leadership roles (Product owner, Architecture owner, Technical lead) for DAD delivery teams. Project management tasks are shared by the entire DAD delivery team (individuals in primary roles).	
Inexperienced project manager	Definitely
Shared leadership roles (Product owner, Architecture owner, Technical lead) for DAD delivery teams. Project management tasks are shared by the entire DAD delivery team (individuals in primary roles).	
Inexperienced team members	Definitely
Self-contained DAD delivery team with up 13 individuals. Primary and secondary roles. Three leadership roles. Task sharing, pair programming.	
Lack of an effective project management methodology	Definitely
Clearly defined program management roles and practices. Clearly defined project management practices and roles, including DAD method phases and activities.	
Lack of commitment to the project among development team members	Definitely
Project management including Team lead. Self-organizing teams, inception phase engineers empowered to run spikes, provide story points and re-estimate each user story. Sprint planning and commitment to sprint objectives.	
Lack of people skills in project leadership	Definitely
Clearly defined leadership roles and practices with program and project management including based on self-organizing DAD delivery teams.	
Large number of links to other systems required	Not at all
Negative attitudes by development team	Definitely
Self-organizing teams. Inception phase engineers empowered to run spikes, provide story points, and re-estimate each user story. Sprint planning and commitment to sprint objectives.	
new: Ineffective collaboration	Definitely
All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles – every one are treated as equal, empowered, and make collective decisions. Three leadership roles with a DAD project team – team lead, AO and PO. All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles (software engineers). Three leadership roles with a DAD project team – team lead, AO and PO. Co-location, shared workspace, daily stand-up meetings, pair programming. All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles. Three leadership roles with a DAD project team – team lead, AO and PO. Co-location, shared workspace, daily stand-up meetings, pair programming.	

(continued on next page)

Table 11 (continued).

GSD Risk Catalog risk/Practices	Level
<i>new: Ineffective coordination</i>	Definitely
All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles (software engineers). Three leadership roles with a DAD project team – team lead, AO and PO. Co-location, shared workspace, daily stand-up meetings, iteration planning, iteration reviews. All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles (software engineers). Three leadership roles with a DAD project team – team lead, AO and PO. Co-location, shared workspace, daily stand-up meetings, pair programming.	
<i>new: Lack of trust</i>	Definitely
All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles – every one are treated as equal, empowered, and make collective decisions. Any training need for the team is allowed through the secondary role- a coach, consultant or skilled individual member from another team can join for a period of time to upskill those who need upskilling. All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles – every one are treated as equal, empowered, and make collective decisions. T-skilled means every team must learn on the fly all the broad skills required to deliver projects. All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles only-every one are treated as equal, empowered, and make collective decisions. However, AO, PO and Team lead can make decisions so that teams can make progress if stuck with a problem.	
<i>One of the largest projects attempted by the organization</i>	Definitely
Program management, including Enterprise architect and product owners. Work item list and DAD delivery teams. Iteration show and tell.	
<i>Poor project planning</i>	Definitely
Program Manager, Enterprise architects, Product owners. Work item list. Story boarding. Architecture planning. Product owners, Architect owners, Team lead, DAD delivery team. Inception phase, iteration planning, DOD, UAT. Re-estimation, story points.	
<i>Project affects a large number of user departments or units</i>	Definitely
Product management, portfolio management, project management, IT operations including cloud deployment.	
<i>Project involves the use of new technology</i>	Definitely
DAD delivery teams are empowered to make tool decision. Technical experts as temporary roles in DAD delivery teams. Coaching and training for upskilling on unfamiliar technology.	
<i>Project involves use of technology that has not been used in prior projects</i>	Definitely
DAD delivery teams are empowered to make tool decisions. Temporary roles in DAD delivery teams. Coaching and training for upskilling on unfamiliar technology.	
<i>Project milestones not clearly defined</i>	Definitely
Project management – each phase can be a milestone including creating a work item list, delivering sprint/iteration based on DOD.	
<i>Project progress not monitored closely enough</i>	Definitely
Program manager responsible to deliver a program. Program daily huddles. Iteration show and tell. Product owner, Architecture owner, team leader. User story points. Time-boxed construction phase has several iteration cycles, all time boxed. Daily stand-up meeting.	
<i>Team members lack specialized skills required by the project</i>	Definitely
Product owner, Architecture owner. Self-organizing teams, T-skilling, coaching by Architecture owner and Product owner, pair programming. Bring in outside expertise to develop and upskill team members.	
<i>Team members not familiar with the task(s) being automated</i>	Definitely
Inception phase involving entire DAD delivery team. Spikes. Story points and estimates, sprint planning.	
Quadrant 4: Environment	
<i>Change in organizational management during the project</i>	Definitely
IT governances drives DAD organization structures. Product management, portfolio management, program management, and project management.	
<i>Corporate politics with negative effect on project</i>	Definitely
Product management, portfolio management, program management and project management. Collective decision making at various levels with DAD framework.	
<i>Dependency on outside suppliers</i>	Not at all
<i>Many external suppliers involved in the development project</i>	Not at all
<i>new: Country-specific regulations</i>	Definitely
DevOp practices – collaboration between Operations and SE (program and project)teams. Portfolio planning at enterprise level, product management, program planning at engineering level, PO, enterprise architects.	
<i>new: Delays caused by global distance</i>	Definitely
DevOp practices- collaboration between Operations and SE (program and project)teams. Mandated DOD, TDD practices and including manual systems testing before deployment into production environment. Program management- program manager, enterprise architect, PO.	
<i>new: Lack of architecture-organization alignment</i>	Definitely
DevOp practices at program and project levels. Performance monitoring practice of product environment for real time data for program portfolio planning and project planning. Enterprise architectures provide technology roadmap and guidance for project teams. Teams must run spikes. Enterprise architecture team (EAT) [VP architect + 8 enterprise architects (EA) from different geographic locations] plan and provide architecture strategy and guidance for the entire organization through weekly virtual EAT meetings. provide are Program architects owners form the enterprise architecture team”	

(continued on next page)

Table 11 (continued).

GSD Risk Catalog risk/Practices	Level
<i>new: Lack of face-to-face interaction inhibits knowledge sharing</i>	Definitely
All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles (software engineers). Three leadership roles with a DAD project team – team lead, AO and PO. Co-location, shared workspace, daily stand-up meetings, iteration planning, iteration reviews. DevOp practices- collaboration between Operations and SE (program and project) teams. Project level (DAD project teams)- inception phase-spikes, story-points and re-estimate, construction phase- look ahead planning, iteration planning- re-estimate	
<i>new: Lack of process alignment</i>	Definitely
All DAD project teams are feature team, self-contained and self-organizing teams, T-skilled and all team members are in primary roles (software engineers). Three leadership roles with a DAD project team – team lead, AO and PO. Co-location, shared workspace, daily stand-up meetings, iteration planning, iteration reviews. Flexibility and empowerment for project teams to also adopt their team specific practices. DAD project teams- a coach or consultant in secondary role Product management and program management- mandated practices for projects.	
<i>new: Lack of tool/infrastructure alignment</i>	Definitely
Booking shared resources. Mandated DevOp practice, collaboration between operations and engineering teams. Mandated practice- DevOp practice, collaboration between operations and engineering teams.	
<i>new: Unstable country/regional political/economic environment</i>	Not at all
<i>Organization undergoing restructuring during the project</i>	Not at all
<i>Resources shifted from the project due to changes in organizational priorities</i>	Definitely
Program level defines projects and resources. DAD delivery teams remain fixed based on small, medium and large teams (each has a defined number of individuals).	
<i>Unstable organizational environment</i>	Definitely
IT governance drives DAD organization structures. Product management, portfolio management, program management, and project management.	

Table 12

SAFe practices mapped to GSD Risk Catalog risks.

GSD Risk Catalog risk/Practices	Level
Quadrant 1: Customer mandate	
<i>Conflict between users</i>	Somewhat
Help to derive collaboration among the key stakeholders in the business. Use capabilities as end-to-end solution services that support the achievement of user goals. Complete user stories from Team Backlog and assure that each story meets its (local) Definition of Done.	
<i>Lack of cooperation from users</i>	Definitely
Encourage and facilitate incremental development and fast customer feedback. Facilitate continuous improvement by quantitative metrics, customer feedback, and the Inspect and Adapt retrospective cycle. Engage stakeholders and provide feedback–time-boxed to an hour or less.	
<i>Lack of top management support for the project</i>	Definitely
Allocate budget for development of value stream. Define Strategic Themes. Allocate budget authority to the decision makers.	
<i>Lack of user participation</i>	Definitely
Facilitate continuous improvement by quantitative metrics, customer feedback, and the Inspect and Adapt retrospective cycle. Develop a feature team that is organized around user-centered functionality. Each team is capable of delivering end-to-end value. Feature teams operate primarily with user stories, refactors, and spikes. Fast and intimate feedback from customers.	
<i>Lack or loss of organizational commitment to the project</i>	Definitely
Allocate budget for development of value stream. Define Strategic Themes. Allocate budget authority to the decision makers.	
<i>Users not committed to the project</i>	Definitely
Encourage and facilitate incremental development and fast customer feedback. Facilitate continuous improvement by quantitative metrics, customer feedback, and the Inspect and Adapt retrospective cycle. Engage stakeholders and provide feedback–time-boxed to an hour or less.	
<i>Users resistant to change</i>	Somewhat
Provide the opportunity for 'just the right amount' of Architecture and User Experience guidance. Identify user-business value. Continuously validate user experience via user experience testing.	
<i>Users with negative attitudes toward the project</i>	Definitely
Encourage and facilitate incremental development and fast customer feedback. Facilitate continuous improvement by quantitative metrics, customer feedback, and the Inspect and Adapt retrospective cycle. Engage stakeholders and provide feedback–time-boxed to an hour or less.	
Quadrant 2: Scope and requirements	
<i>Conflicting system requirements</i>	Definitely
Apply enabler for exploration that provides a way for development teams to flesh out the details of requirements and design. Continuously communicate emerging requirements and opportunities back into the program vision through product owner. Work with stakeholders to understand the specific business targets behind the user–system interaction.	
<i>Continually changing project scope/objectives</i>	Definitely
Perform system demo as near as possible to the end of the iteration. Integrate every other iteration. Actively participate in ongoing agreements to maintain business and development alignment as priorities and scope are inevitably changed.	
<i>Continually changing system requirements</i>	Definitely
Perform system demo as near as possible to the end of the iteration. Integrate every other iteration. Continuously communicate emerging requirements and opportunities back into the program vision through product owner.	

(continued on next page)

Table 12 (continued).

GSD Risk Catalog risk/Practices	Level
<i>Difficulty in defining the inputs and outputs of the system</i>	Definitely
Define PI planning's primary outputs. Demonstrate each new feature in an end-to-end use case. Develop a feature team that is organized around user-centered functionality. Each team is capable of delivering end-to-end value. Feature teams operate primarily with user stories, refactors, and spikes.	
<i>Ill-defined project goals</i>	Definitely
Implement value stream coordination to ensure that the enterprise moves forward with each value stream in lockstep with the enterprise objectives. Align development to business via business context, vision, and Team and Program PI Objectives. Create a set of 'SMART' team PI objectives for each individual team with business value assigned.	
<i>Incorrect system requirements</i>	Definitely
Epics and lightweight business cases. Ensure that Epics and Enablers are reasoned and analyzed prior to reaching a Program Increment boundary, are prioritized appropriately, and have established acceptance criteria to guide a high-fidelity implementation. Define Roadmap.	
<i>System requirements not adequately identified</i>	Definitely
Epics and lightweight business cases. Ensure that Epics and Enablers are reasoned and analyzed prior to reaching a Program Increment boundary, are prioritized appropriately, and have established acceptance criteria to guide a high-fidelity implementation. Define Roadmap.	
<i>Unclear system requirements</i>	Definitely
Epics and lightweight business cases. Ensure that Epics and Enablers are reasoned and analyzed prior to reaching a Program Increment boundary, are prioritized appropriately, and have established acceptance criteria to guide a high-fidelity implementation. Define Roadmap.	
<i>Undefined project success criteria</i>	Definitely
Define Success criteria to validate the implementation. Impacts the identification, success criteria, and prioritization of epics in the funnel and backlog states. Success criteria provide a mechanism to understand progress towards the intent.	
<i>Users lack understanding of system capabilities and limitations</i>	Definitely
Perform system demos. Integrate to illustrate a particular feature, capability, or nonfunctional requirement. Demonstrate each new feature in an end-to-end use case.	
Quadrant 3: Execution	
<i>Development team unfamiliar with selected development tools</i>	Somewhat
Understand requirements for working on "technical infrastructure, tooling, and other systemic impediments". It may be more efficient to perform an upgrade or migration at a time when there isn't a critical demo just a few days away. Perform Daily Stand-up to understand team's status, "escalate problems, and get help from other team members"	
<i>Frequent conflicts among development team members</i>	Somewhat
Vote of confidence/commitment from the entire program to these objectives. Develop PI commitment. Establish an agreement to determine how the work is performed for each activity type.	
<i>Frequent turnover within the project team</i>	Not at all
<i>High level of technical complexity</i>	Definitely
Break down the architectural enabler into small enabler stories that can fit in iterations. Perform initial exploration of epics and rank them roughly by using Weighted Shortest Job First (WSJF) to determine which epics should move to the next step for deeper exploration. Apply Weighted Shortest Job First prioritization method for job sequencing.	
<i>Highly complex task being automated</i>	Definitely
Perform initial exploration of epics and rank them roughly by using Weighted Shortest Job First (WSJF) to determine which epics should move to the next step for deeper exploration. Apply Weighted Shortest Job First prioritization method for job sequencing. Define Weighted Shortest Job First (WSJF).	
<i>Immature technology</i>	Definitely
Work with Agile Teams that perform research spikes, create proof of concepts, mock-ups, etc. Support technology/engineering aspects of program and value stream kanbans. Local stories representing new functionality, refactors, defects, research spikes, and other technical debt are identified, written as enabler stories, estimated, and sequenced.	
<i>Inadequate estimation of project budget</i>	Definitely
Develop Incremental implementation by keeping the epics in the portfolio backlog until there is implementation capacity available. Avoid overhead and enables the train to make fast and local decisions within the constraints of the allocated budget. Provide WIP limits to ensure that the teams responsible for analysis undertake into responsibly and do not create expectation for implementation or time frames that far exceed capacity and reality.	
<i>Inadequate estimation of project schedule</i>	Definitely
Agile Estimating and Planning. Adopt Agile estimating and planning by using the currency of story points. Epic Progress Measure.	
<i>Inadequate estimation of required resources</i>	Definitely
Develop Incremental implementation by keeping the epics in the portfolio backlog until there is implementation capacity available. Provide WIP limits to ensure that the teams responsible for analysis undertake into responsibly and do not create expectation for implementation or time frames that far exceed capacity and reality. Use capacity allocation to estimate portfolio epic based on the given knowledge of program velocities.	
<i>Inadequately trained development team members</i>	Somewhat
Conduct specialized training to keep up with advancements in their respective fields.	
<i>Ineffective communication</i>	Definitely
Facilitate continuous improvement by quantitative metrics, customer feedback, and the Inspect and Adapt retrospective cycle. Establish high-bandwidth communication across all team members and stakeholders. Perform system demo as near as possible to the end of the iteration.	

(continued on next page)

Table 12 (continued).

GSD Risk Catalog risk/Practices	Level
Ineffective project manager	Not at all
Inexperienced project manager	Not at all
Inexperienced team members	Somewhat
Involve a subject matter expert in basic exploration and sizing. Interact with analyst and subject matter experts during specification workshops. Work with stakeholders and subject matter experts to define the epic and its potential benefits.	
Lack of an effective project management methodology	Definitely
Coach leaders, teams, and Scrum masters in lean-Agile practices and mindsets. Use "Agile Project Management Tools to capture stories and status, defects, test cases, estimates, actuals, assignments, burn-down chart".	
Lack of commitment to the project among development team members	Definitely
Vote of confidence/commitment from the entire program to these objectives. Develop PI commitment. Limit the commitments to longer-term work, because some other item may come along that is more important than a prior commitment.	
Lack of people skills in project leadership	Somewhat
Define Scrum Master role. Exhibits Lean-Agile leadership. Protects and communicates.	
Large number of links to other systems required	Definitely
Perform initial exploration of epics and rank them roughly by using Weighted Shortest Job First (WSJF) to determine which epics should move to the next step for deeper exploration. Apply Weighted Shortest Job First prioritization method for job sequencing. Define Weighted Shortest Job First (WSJF).	
Negative attitudes by development team	Somewhat
Perform Iteration Retrospective as the check step for the overall iteration. Use "iteration retrospective to drive program level changes to process either immediately or in the Inspect and Adapt workshop". Perform Iteration Retrospective to identify way to improve.	
new: Ineffective collaboration	Definitely
Manage dependencies by applying extensive degree of cooperation; a common value stream backlog; implementation of new, crosscutting capabilities; additional system integration; additional roles and responsibilities; special considerations for pre-, post-, and PI planning activities; different degree and types of DevOps support. Product manager create features in collaboration with product owner and other key stakeholders. Features are also created as a result of decomposition of epics. Encourage the collaboration between teams and System and Solution Architects, Engineering, and User Experience designers.	
new: Ineffective coordination	Definitely
Manage dependencies by applying extensive degree of cooperation; a common value stream backlog; implementation of new, crosscutting capabilities; additional system integration; additional roles and responsibilities; special considerations for pre-, post-, and PI planning activities; different degree and types of DevOps support. Implement value stream coordination to ensure that the enterprise moves forward with each value stream in lockstep with the enterprise objectives. Align development to business via business context, vision, and Team and Program PI Objectives.	
new: Lack of trust	Definitely
Manage dependencies by applying extensive degree of cooperation; a common value stream backlog; implementation of new, crosscutting capabilities; additional system integration; additional roles and responsibilities; special considerations for pre-, post-, and PI planning activities; different degree and types of DevOps support. Adopt SAFe's set of seven transformational patterns that can be used to move the organization to Lean-Agile Program Portfolio Management ([Decentralized decision-making], [Demand management; continuous value flow], [Lightweight; epic-only business cases], [Decentralized, rolling-wave planning], [Agile estimating and planning], [Lean-Agile budgeting and self-managing Agile Release Trains], [Objective, fact-based measure and milestones]). Self-managing Agile Release Trains.	
One of the largest projects attempted by the organization	Definitely
Perform initial exploration of epics and rank them roughly by using Weighted Shortest Job First (WSJF) to determine which epics should move to the next step for deeper exploration. Apply Weighted Shortest Job First prioritization method for job sequencing. Define Weighted Shortest Job First (WSJF).	
Poor project planning	Definitely
Implement value stream coordination to ensure that the enterprise moves forward with each value stream in lockstep with the enterprise objectives. Align development to business via business context, vision, and Team and Program PI Objectives. Create a set of 'SMART' team PI objectives for each individual team with business value assigned.	
Project affects a large number of user departments or units	Somewhat
Demonstrate each new feature in an end-to-end use case. Develop a feature team that is organized around user-centered functionality. Each team is capable of delivering end-to-end value. Feature teams operate primarily with user stories, refactors, and spikes. Use capabilities as end-to-end solution services that support the achievement of user goals.	
Project involves the use of new technology	Definitely
Work with Agile Teams that perform research spikes, create proof of concepts, mock-ups, etc. PI Features are broken into stories and placed on team backlog. Local stories representing new functionality, refactors, defects, research spikes, and other technical debt are identified, written as enabler stories, estimated, and sequenced.	
Project involves use of technology that has not been used in prior projects	Definitely
Work with Agile Teams that perform research spikes, create proof of concepts, mock-ups, etc. Ensure that the demo environments are adequate to the challenge of reliably demonstrating new solution functionality. Local stories representing new functionality, refactors, defects, research spikes, and other technical debt are identified, written as enabler stories, estimated, and sequenced.	
Project milestones not clearly defined	Definitely
Implement value stream coordination to ensure that the enterprise moves forward with each value stream in lockstep with the enterprise objectives. Align development to business via business context, vision, and Team and Program PI Objectives. Create a set of 'SMART' team PI objectives for each individual team with business value assigned.	

(continued on next page)

Table 12 (continued).

GSD Risk Catalog risk/Practices	Level
<i>Project progress not monitored closely enough</i>	Definitely
Perform system demos. Perform system demo as near as possible to the end of the iteration. Demonstrate each new feature in an end-to-end use case.	
<i>Team members lack specialized skills required by the project</i>	Somewhat
Involve a subject matter expert in basic exploration and sizing. Interact with analyst and subject matter experts during specification workshops. Work with stakeholders and subject matter experts to define the epic and its potential benefits.	
<i>Team members not familiar with the task(s) being automated</i>	Definitely
Develop a feature team that is organized around user-centered functionality. Each team is capable of delivering end-to-end value. Feature teams operate primarily with user stories, refactors, and spikes. Refine the backlog. Involve with Agile team for short period of time.	
Quadrant 4: Environment	
<i>Change in organizational management during the project</i>	Not at all
<i>Corporate politics with negative effect on project</i>	Somewhat
Provide decision-making filters in the portfolio kanban, thereby influence the portfolio backlog. Define Strategic Themes. Formulate Strategic themes.	
<i>Dependency on outside suppliers</i>	Definitely
Business Owner prepares to communicate the business context, including milestones and significant external dependencies, such as those of suppliers. Work with Customers, stakeholders, and Suppliers to establish high-level Solution Intent; help establish the solution intent information models and documentation requirements. Work with Suppliers, making sure the requirements for supplier-delivered capabilities are understood, and assist with the conceptual integration of these concerns.	
<i>Many external suppliers involved in the development project</i>	Definitely
Business Owner prepares to communicate the business context, including milestones and significant external dependencies, such as those of suppliers. Work with Customers, stakeholders, and Suppliers to establish high-level Solution Intent; help establish the solution intent information models and documentation requirements.	
<i>new: Country-specific regulations</i>	Definitely
File patent, certify the system, audit certain regulatory requirements for economic success of product development. Defines how the system builder's system intent must be organized, packaged, and integrated for use by the Customer to meet any compliance, certification, and other objectives. Supports compliance and contractual obligations.	
<i>new: Delays caused by global distance</i>	Somewhat
Establish high-bandwidth communication across all team members and stakeholders. Calculate the Cost of Delay. Manage and optimize the flow of value through the program using various tools, such as the Program and Value stream Kanbans and information radiators.	
<i>new: Lack of architecture-organization alignment</i>	Somewhat
Identify dependencies and forecast cross-team and cross-ART coordination. Create enabler features by architects or engineers to pave the architectural runway which is maintained in the program backlog alongside business features. Create architectural enablers to pave the runway.	
<i>new: Lack of face-to-face interaction inhibits knowledge sharing</i>	Definitely
Provide guidance for cross-team design and implementation synchronization. Manage and optimize the flow of value through the program using various tools, such as the Program and Value stream Kanbans and information radiators. RTE facilitate periodic synchronization meetings, including the ART sync at the Program level and the Value Stream (VS) sync at the value stream level.	
<i>new: Lack of process alignment</i>	Definitely
Adopt decentralized, program, and team-based rolling-wave planning via routine and cadence-based PI Planning activity. Program Portfolio Management (PPM) team continuously assesses and improves their processes using a structured, periodic Self-Assessment. Perform a structured, root cause analysis-based, problem-solving workshop for large, systematic program-level problems.	
<i>new: Lack of tool/infrastructure alignment</i>	Definitely
Manage dependencies by applying extensive degree of cooperation; a common value stream backlog; implementation of new, crosscutting capabilities; additional system integration; additional roles and responsibilities; special considerations for pre-, post-, and PI planning activities; different degree and types of DevOps support. Create infrastructure enablers to be ready to develop, test, and integrate the initiatives. Develop infrastructure in different levels of SAFe to support frequent or continuous integration and testing.	
<i>new: Unstable country/regional political/economic environment</i>	Somewhat
Build an economically viable solution. File patent, certify the system, audit certain regulatory requirements for economic success of product development. Require evolution of the Customer's deployment environment for new solutions.	
<i>Organization undergoing restructuring during the project</i>	Not at all
<i>Resources shifted from the project due to changes in organizational priorities</i>	Somewhat
Apply WSJF to prioritize the items in the Value stream and Program backlogs. PI cadence offers time box experimenting approach which later help to produce the desired and long term results. Apply WSJF prioritization to compare business and enabler epics against each other.	
<i>Unstable organizational environment</i>	Not at all

These findings in a global software development scaling agile context echo Oehmen and colleagues' (Oehmen et al., 2014) assertion that risks cannot be avoided; at best organizations manage risk by applying practices that lead to a structured reduction of uncertainty.

CRediT authorship contribution statement

Sarah Beecham: Conceptualization, Writing, Methodology, Investigation, Original draft preparation, Visualization, Project Administration. **Tony Clear:** Writing, Conceptualization, Review &

edit. **Ramesh Lal:** Investigation, Writing, Data curation. **John Noll:** Investigation, Methodology, Writing, Conceptualization, Data curation, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are indebted to our Case Study organizations—Company A and Ocuco Ltd (Company B)—and the many members of those organizations we observed and interviewed over a four-year period. Extra special thanks to Clodagh Nic Canna, of Ocuco, who was the main driver behind our collaboration. We also offer our thanks to Abdur Razzak, who helped with data collection and some early analysis of issues. This work was supported, in part, by Science Foundation Ireland grant 13/RC/2094.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jss.2020.110823>.

References

- Abdullah, L., Verner, J., 2009. Outsourced strategic IT systems development risk. In: Third International Conference on Research Challenges in Information Science. RCIS 2009, IEEE, pp. 275–286.
- Ågerfalk, P.J., Fitzgerald, B., Holmström, H., Lings, B., Lundell, B., Conchúir, E.Ó., 2005. A framework for considering opportunities and threats in distributed software development. In: International Workshop on Distributed Software Development. DiSD 2005, Paris, August 29, 2005, Austrian Computer Society, pp. 47–61.
- Alqudah, M., Razali, R., 2016. A review of scaling agile methods in large software development. *Int. J. Adv. Sci. Eng. Inf. Technol.* 6 (6), 828–837.
- Alqudah, M., Razali, R., 2017. Key factors for selecting an agile method: A systematic literature review. *Int. J. Adv. Sci. Eng. Inf. Technol.* 7 (2), 526–537. <http://dx.doi.org/10.18517/ijaseit.7.2.1830>.
- Ambler, S., Lines, M., 2011. *Disciplined Agile Delivery: An Introduction*. IBM Software, Somers, NY.
- Ambler, S.W., Lines, M., 2012. *Disciplined Agile Delivery: A Practitioner's Guide to Agile Software Delivery in the Enterprise*. IBM Press.
- Analysis.net Research, 2018. 12th Annual State of Agile Report. Tech. Rep., CollabNet VersionOne.
- Bass, J.M., Beecham, S., Noll, J., 2018. Experience of industry case studies: A comparison of multi-case and embedded case study methods. In: Workshop on Conducting Empirical Studies in Industry. Gothenburg, Sweden.
- Bass, J.M., Beecham, S., Razzak, M.A., Noll, J., 2018. Employee retention and turnover in global software development: Comparing in-house offshoring and offshore outsourcing. In: International Conference on Global Software Engineering. Gothenburg, Sweden.
- Beck, K., Gamma, E., 2000. *Extreme Programming Explained: Embrace Change*. Addison Wesley, Boston, USA.
- Beck, K., Grenning, J., Martin, R.C., Beedle, M., Highsmith, J., Mellor, S., Bennekum, A.v., Hunt, A., Schwaber, K., 2001. *Manifesto for agile software development*. <http://agilemanifesto.org/>.
- Beecham, S., Clear, T., Lal, R., Noll, J., 2020. Scaling Agile Frameworks and Global Software Development Risk: Supplementary Data. Tech. Rep. 2020_TR_03, Lero, The Irish Software Research Centre, URL <https://www.lero.ie/publications/technicalreports>.
- Beecham, S., Noll, J., 2015. What motivates software engineers working in global software development? In: Abrahamsson, P., Corral, L., Oivo, M., Russo, B. (Eds.), *Product-Focused Software Process Improvement. PROFES 2015*, In: Lecture Notes in Computer Science, vol. 9459, Springer, pp. 193–209. http://dx.doi.org/10.1007/978-3-319-26844-6_14.
- Beecham, S., Noll, J., Razzak, M.A., 2017. Lean Global Project Interview Protocol. Tech. Rep. TR_2017_02, Lero, the Irish Software Research Centre, URL http://www.lero.ie/sites/default/files/Lero_TR_2017_02_Beecham_Noll_Razzak-Lean%20Global%20Project%20Interview%20Protocol.pdf.
- Boehm, B., 1991. Software risk management: principles and practices. *IEEE Softw.* 8 (1), 32–41.
- Cataldo, M., Herbsleb, J.D., 2011. Factors leading to integration failures in global feature-oriented development: an empirical analysis. In: Proceedings of the 33rd International Conference on Software Engineering, pp. 161–170.
- Chadli, S.Y., Idri, A., Fernández-Alemán, J.L., Ros, J.N., Toval, A., 2016. Identifying risks of software project management in global software development: An integrative framework. In: 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications. AICCSA, pp. 1–7.
- Diebold, P., Schmitt, A., Theobald, S., 2018. Scaling agile: how to select the most appropriate framework. In: Proceedings of the 19th International Conference on Agile Software Development: Companion, pp. 1–4.
- Dikert, K., Paasivaara, M., Lassenius, C., 2016. Challenges and success factors for large-scale agile transformations: A systematic literature review. *J. Syst. Softw.* 119, 87–108.
- Dingsøyr, T., Moe, N.B., 2014. Towards principles of large-scale agile development: A summary of the workshop at XP2014 and a revised research agenda. In: 15th International Conference on Agile Software Development (XP), May 26–30, 2014, Rome, ITALY, pp. 1–8.
- Dybå, T., Dingsøyr, T., 2008. Empirical studies of agile software development: A systematic review. *Inf. Softw. Technol.* 50 (9–10), 833–859.
- Ebert, C., Murthy, B.K., Jha, N.N., 2008. Managing risks in global software engineering: Principles and practices. In: 2008 IEEE International Conference on Global Software Engineering, pp. 131–140.
- Ebert, C., Paasivaara, M., 2017. Scaling agile. *IEEE Softw.* 34 (6), 98–103. <http://dx.doi.org/10.1109/MS.2017.4121226>.
- Fitzgerald, B., Stol, K.-J., 2017. Continuous software engineering: A roadmap and agenda. *J. Syst. Softw.* 123, 176–189.
- Gotterbarn, D., Rogerson, S., 2005. Responsible risk analysis for software development: Creating the software development impact statement. *Commun. AIS* 15, 730–750.
- Hall, T., Beecham, S., Verner, J., Wilson, D., 2008. The impact of staff turnover on software projects: the importance of understanding what makes software practitioners tick. In: Proceedings of the 2008 ACM SIGMIS CPR Conference on Computer Personnel Doctoral Consortium and Research. ACM, pp. 30–39.
- Kalenda, M., Hyna, P., Rossi, B., 2018. Scaling agile in large organizations: Practices, challenges, and success factors. *J. Softw.: Evol. Process* 30 (10).
- Kitchenham, B., Pfleeger, S.L., McColl, B., Eagan, S., 2002. An empirical study of maintenance and development estimation accuracy. *J. Syst. Softw.* 64 (1), 57–77.
- Lal, R., Clear, T., 2017. Scaling agile at the program level in an Australian software vendor environment: A case study. In: Australasian Conference on Information Systems, pp. 1–12.
- Lal, R., Clear, T., 2018. Enhancing product and service capability through scaling agility in a global software vendor environment. In: Espinoza, R., Šmite, D. (Eds.), *Proceedings 2018 IEEE 13th International Conference on Global Software Engineering*. IEEE, Los Alamitos, California, pp. 59–68. <http://dx.doi.org/10.1145/3196369/3196378>.
- Leffingwell, D., Yakyma, A., Knaster, R., Jemilo, D., Oren, I., 2016. *SAFe 4.0 Reference Guide: Scaled Agile Framework for Lean Software and Systems Engineering*. Addison-Wesley Professional.
- Lehtinen, T., Mäntylä, M., Vanhanen, J., Ikonen, J., Lassenius, C., 2014. Perceived causes of software project failures—an analysis of their relationships. *Inf. Softw. Technol.* 56 (6), 623–643.
- Leonard-Barton, D., 1990. A dual methodology for case studies: Synergistic use of a longitudinal single site with replicated multiple sites. *Organ. Sci.* 1 (3), 248–266.
- Lincoln, Y.S., Guba, E.G., 1985. *Naturalistic Inquiry*, first ed. SAGE Publications, Inc., Beverly Hills, Calif.
- Marinho, M., Noll, J., Richardson, I., Beecham, S., 2019. Plan-driven approaches are alive and kicking in agile Global Software Development. In: *Empirical Software Engineering and Measurement. ESEM, IEEE*.
- McLeod, L., MacDonell, S., 2011. Factors that affect software systems development project outcomes: A survey of research. *ACM Comput. Surv.* 43 (4), 24.
- Noll, J., Beecham, S., 2016. Measuring global distance: A survey of distance factors and interventions. In: International SPICE Conference on Process Improvement and Capability Determination in Software, Systems Engineering and Service Management. Dublin, Ireland, URL http://lero.ie/sites/default/files/2015_TR_02_Sarah%20Beecham_John%20Noll_Measuring%20Global%20Distance%20A%20Survey%20of%20Distance%20Factors%20and%20Interventions.pdf.
- Noll, J., Beecham, S., 2019. How agile is hybrid agile? An analysis of the HELENA data. In: *Product-Focused Software Process Improvement. PROFES 2019*, In: *Product-Focused Software Process Improvement*, Springer International Publishing, pp. 341–349.
- Noll, J., Beecham, S., Richardson, I., 2010. Global software development and collaboration: Barriers and solutions. *ACM Inroads* 1 (3).
- Noll, J., Beecham, S., Richardson, I., NicCanna, C., 2016. A global teaming model for global software development governance: A case study. In: *International Conference on Global Software Engineering*. IEEE, Irvine, California, USA, pp. 179–188.

- Noll, J., Razzak, M.A., Beecham, S., 2017. Motivation and autonomy in global software development: An empirical study. In: Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering, pp. 394–399.
- Nurdiani, I., Jabangwe, R., Šmite, D., Damian, D., 2011. Risk identification and risk mitigation instruments for global software development: Systematic review and survey results. In: 2011 IEEE Sixth International Conference on Global Software Engineering Workshop, pp. 36–41.
- Oehmen, J., Olechowski, A., Kenley, C.R., Ben-Daya, M., 2014. Analysis of the effect of risk management practices on the performance of new product development programs. *Technovation* 34 (6), 441–453.
- Paasivaara, M., 2017. Adopting SAFe to scale agile in a globally distributed organization. In: 2017 IEEE 12th International Conference on Global Software Engineering. ICGSE, IEEE, pp. 36–40.
- Persson, J.S., Mathiassen, L., Boeg, J., Madsen, T.S., Steinson, F., 2009. Managing risks in distributed software projects: an integrative framework. *IEEE Trans. Eng. Manage.* 56 (3), 508–532.
- Pikkariainen, M., Haikara, J., Salo, O., Abrahamsson, P., Still, J., 2008. The impact of agile practices on communication in software development. *Empir. Softw. Eng.* 13 (3), 303–337.
- Razzak, M.A., Noll, J., Richardson, I., Nic Canna, C., Beecham, S., 2017. Transition from plan driven to SAFe®: Periodic team self-assessment. In: International Conference on Product-Focused Software Process Improvement. PROFES '17, Springer, pp. 573–585.
- Razzak, M.A., Richardson, I., Noll, J., Nic Canna, C., Beecham, S., 2018. Scaling agile across the global organization: an early stage industrial SAFe self-assessment. In: 13th International Conference on Global Software Engineering. ICGSE 2018, IEEE, pp. 116–125.
- Schmidt, R., Lyytinen, K., Keil, P., Cule, M., 2001. Identifying software project risks: An international Delphi study. *J. Manag. Inf. Syst.* 17 (4), 5–36.
- Shrivastava, S.V., Rathod, U., 2017. A risk management framework for distributed agile projects. *Inf. Softw. Technol.* 85, 1–15.
- Sievi-Korte, O., Beecham, S., Richardson, I., 2019a. Challenges and recommended practices for software architecting in global software development. *Inf. Softw. Technol.* 106, 234–253.
- Sievi-Korte, O., Richardson, I., Beecham, S., 2019b. Software architecture design in global software development: An empirical study. *J. Syst. Softw.* 158, 110400. <http://dx.doi.org/10.1016/j.jss.2019.110400>.
- Sinha, R., Shameem, M., Kumar, C., 2020. SWOT: Strength, weaknesses, opportunities, and threats for scaling agile methods in global software development. In: Proceedings of the 13th Innovations in Software Engineering Conference. ISEC 2020, ACM, New York, NY, USA, <http://dx.doi.org/10.1145/3385032.3385037>.
- Šmite, D., 2007. Project outcome predictions: Risk barometer based on historical data. In: International Conference on Global Software Engineering. ICGSE 2007, pp. 103–112.
- Tell, P., Klünder, J., Küpper, S., Raffo, D., MacDonell, S.G., Münch, J., Pfahl, D., Linssen, O., Kuhrmann, M., 2019. What are hybrid development methods made of?: an evidence-based characterization. In: Proceedings of the International Conference on Software and System Processes. IEEE Press, pp. 105–114.
- Uludağ, Ö., Kleehaus, M., Xu, X., Matthes, F., 2017. Investigating the role of architects in scaling agile frameworks. In: 21st International Enterprise Distributed Object Computing Conference. EDOC, pp. 123–132. <http://dx.doi.org/10.1109/EDOC.2017.25>.
- Vaidya, A., 2014. Does DAD know best, is it better to do LeSS or just be SAFe? Adapting scaling agile practices into the enterprise. In: Proceedings of the 32nd Pacific Northwest Software Quality Conference. PNSQ, pp. 1–18.
- Verner, J., Brereton, O., Kitchenham, B., Turner, M., Niazi, M., 2014. Risks and risk mitigation in global software development: A tertiary study. *Inf. Softw. Technol.* 56 (1), 54–78.
- Verner, J.M., Sampson, J., Tasic, V., Bakar, N.A.A., Kitchenham, B.A., 2009. Guidelines for industrially-based multiple case studies in software engineering. In: 2009 Third International Conference on Research Challenges in Information Science. pp. 313–324. <http://dx.doi.org/10.1109/RCIS.2009.5089295>.
- Vizcaíno, A., García, F., Piattini, M., Beecham, S., 2016. A validated ontology for global software development. *Comput. Stand. Interfaces* 46, 66–78.
- Wallace, L., Keil, M., 2004. Software project risks and their effect on outcomes. *Commun. ACM* 47 (4), 68–73.
- Wallace, L., Keil, M., Rai, A., 2004. Understanding software project risk: a cluster analysis. *Inf. Manag.* 42 (1), 115–125.
- Yin, R.K., 2018. Case Study Research and Applications: Design and Methods, sixth ed. Sage Publications, Inc., Thousand Oaks, CA.

Sarah Beecham is a Senior Research Fellow in Lero, the Science Foundation Ireland Research Centre for Software (<https://www.lero.ie>). Sarah joined Lero in 2009, to conduct research into software quality and process improvement. Her wider interests are in socio-technical aspects of software engineering to include Software Engineer motivation, scaling agile methods, distributed software development, and how technology influences the lives of the older adult. She works closely with industry, where her research is problem driven. She is also interested in the education of the next generation of software engineers. In her empirical research, she employs both qualitative and quantitative methods. She has supervised (and examined) several national and international Ph.D. students. Sarah sits on many Program Committees, is an Associate Editor of the Journal of Systems and Software, was General Chair for the Evidence and Assessment in Software Engineering conference (EASE), and Co-chair for The Education and Training Track at ICSE in 2019, and is a Co-chair of 2021 ICSE SCORE contest.

Tony Clear is an Associate Professor within the Software Engineering Research Lab - <https://serl.aut.ac.nz/> - at Auckland University of Technology. He holds positions as an Associate Editor for ACM Transactions on Computing Education (TOCE), for the journal Computer Science Education, and ACM Inroads for which he is also a regular columnist and Editorial Board member. He is active in research within the software engineering and computer science education communities. With Professor Daniela Damian of University of Victoria, Canada he is working on a Royal Society of NZ International Leaders Fellowship Grant titled - "Leading the Way in Software Ecosystems for NZ". Tony has chaired or served on the program committee for conferences such as ICGSE, EASE, ITICSE, ICER, ACE, FIE, CITRENS, APRES, ECIS, SIESC, and reviewed for journals such as TSE, IST, JSS, JSEP, IJEE, CLEIJ. He supervises and has examined doctoral students in Global Software Engineering, CS Education and interdisciplinary topics, and has chaired or participated in several doctoral consortia.

Ramesh Lal is a senior lecturer in IT and Software Engineering at Auckland University of Technology(<http://www.aut.ac.nz>). Ramesh currently is teaching papers at undergraduate and postgraduate levels, including supervising several Masters and Ph.D. students doing software engineering and project management related research. Ramesh is also an academic leader and supervisor for the final year undergraduate students undertaking industry-based R & D projects in software development. Ramesh has a strong industry connection with software development teams and vendors. He has been researching agile approach for software development since 2005.

John Noll is a Senior Lecturer at the University of Hertfordshire, where his research focuses software development processes, especially agile methods, global software development, and open source software development. He has published numerous conference and journal papers, and has held a variety of academic and industrial research positions at institutions and companies including Lero, the Irish Software Research Centre, Santa Clara University, the University of Southern California, Network Appliance, Inc., Perceptronics, Inc., and Hewlett-Packard Laboratories.