A Risk Management Framework for Distributed Agile Development Projects

SYNOPSIS

Submitted in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY

By
SUPRIKA VASUDEVA SHRIVASTAVA

Under the Guidance of

Dr. URVASHI RATHOD
Professor, Symbiosis Centre for Information Technology (SCIT)
Symbiosis International University (SIU)

Faculty of Computer Studies
SYMBIOSIS INTERNATIONAL UNIVERSITY
Pune - 412115

Research Project Funded By

PROJECT MANAGEMENT INSTITUTE OF EDUCATIONAL FOUNDATION (PMIEF)
Pennsylvania, USA

2014
1. Introduction

Software organizations are working under tight time constraints and development of software occurs in highly volatile environment. The unpredictability in the system requirements and changing business needs have transformed the development approach from traditional heavyweight processes to lightweight incremental and iterative methods. Hence, many organizations are adopting agile methodology for software development, which helps them to accelerate delivery schedules, adapt to the changing business needs, align business and technology goals and generate competitive advantage (Holler, 2010). Along with this, there has been a steady, irreversible trend toward the globalization of business, and of software-intensive high-technology businesses in particular. This type of development of software which allows the team members to be located in various remote sites during the software lifecycle and thus making up a network of distant sub-teams is called Distributed Software Development (DSD) (Jiménez, Piattini, Vizca´ıno 2009). DSD is gaining recognition because it helps in saving cost and reduces time to market. In order to reap the benefits of both the approaches, software organizations are blending the distributed development and agile approach. A survey conducted by VersionOne, states that organizations are constantly scaling agile beyond single team and single project (VersionOne, 2013).

Distributed agile development (DAD), helps the organization to build low cost solutions to cater to the changing needs of the business. Although, distributed agile development is able to deliver products close to customer requirements and faster than the traditional method (Kahkonen & Abrahamsson, 2003), there are many risks and challenges involved. These risks primarily occur due to the fact that distributed development and agile methods differ significantly in their key tenets. Agile focuses on close and frequent interaction between the teams and the customer for software development. On the contrary, DSD is characterized by spatial distance, time-zone differences and cultural distance which hinders effective communication and collaboration between the project stakeholders. Hence, there is need to explore important risks which are involved in DAD, and suggest appropriate risk management methods which can be used to reduce their impact.

Recent survey show that as the level of geographic distribution increases, the risks of communication also increases, resulting in lower success rates (Ambler, 2012a). Another survey result shows that 60% of co-located agile project are successful, while roughly 25% can be considered as failed projects. On the other hand, although, more than 50% distributed agile projects have been successful, but 50% of them have failed too (Ambler, 2012b). Higher failure rate of software projects indicate that risk management is critical for the success of DAD projects.

Research literature reveals the scarcity of work in the area of risk management for distributed agile projects (Jalali and Wohlin 2010). Although, there are research studies which discuss about the issues in DAD project, but they are either based on systematic review of literature (Hossain et al. 2009; Jalali and Wohlin 2010), or are case studies, which suffer with the limitation of situation based perspective (Simon, 2002; Fowler 2006; Mattsson et al., 2010). There are other studies, which deal with risks in distributed development (DSD) and risks in agile development separately. Few studies provide an overall view of the risks in DAD projects by including inputs from the existing research and current trends in the industry (Mattsson, Azizyan ad Magarian, 2010). The paucity of work in this area has motivated us to take up this research
A Risk Management Framework for Distributed Agile Development Projects.
Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)

The study and unveil the significant risks which plague the software development process in distributed agile environment.

In this study, initial exploration involving in-depth interviews and project work document analysis has been done for identification and categorization of risk factors in DAD projects. The derived risk categories could be mapped to the Leavitt’s model of organizational change in order to put the findings in organizational context. Important risk factors and most suitable risk management methods were identified by conducting quantitative descriptive study. The findings were then consolidated to create a risk management framework. Along with this, it was observed that the conflict in the properties of DSD and agile principles are primarily responsible for the risks occurring in DAD. The effect of risk factors on project constraints of time, cost and quality was also studied as a part of the descriptive study. The risk management framework was validated by implementing partially in DAD projects of three different companies and was found to be beneficial for them.

The risk management framework identifies the important risk factors which may impact the success of DAD project. The identified risks are related to all the core areas of software development including, communication and collaboration amongst the project stakeholders, software engineering activities, project and team management, and technology setup. Further, the research work shows that, how the contradiction which exists between the distributed development properties and agile principles and practices become the primary reason for those risks and hence, suggests the best practices which can be adopted by DAD teams to control those risks.

2. Review of Literature

Due to the increase in the complexity of the modern software systems and the changing business needs, the traditional ‘waterfall’ (Royce, 1970) approach is not considered to be suitable for development (Hneif & Hockow, 2009) (Guntamukalla, Wen & Tarn, 2006). Organizations started looking at ‘iterative and incremental models’ which enabled the developers to create more complete versions of software (Pressman, 2005). As the problems like changing requirements, lack of customer involvement, low budgets, tight deadlines and miscommunication are becoming more critical, a new set of lightweight methodologies known as ‘Agile Methods’ are being used by the industry for developing software.

Agile approach strips away much of the heaviness associated with traditional methods in order to promote quick response to the changing environment and changing user needs (Erickson, Lytinen & Siau, 2005). These methods use multiple, small iterations, where each iteration involves all the software development activities like requirement analysis, design, programming and testing. The goal is to create a stable, integrated and tested, partially complete system at the end of the iteration. Various agile methods being used include Scrum (Schwaber & Beedle, 2002), Extreme Programming (XP) (Beck 2005), Feature-Driven Development (FDD) (Palmer & Felsing, 2002), Agile Modeling (AM) (Ambler, 2002), The Crystal Family (Cockburn, 1998) and Dynamic Systems Development Method (DSDM) (Stapleton, 1997). Although numerous methods have been identified, two concepts lie at the heart of each of these methods: working code and effective people.

A parallel trend of the software industry is towards global sourcing and distributed development, which helps them to create low cost solutions in lesser time while getting the advantage of global pool of talented professionals. Distributed Software Development (DSD) is
defined by Prickladnicki, as a software development process where at least one involved actors (project team, or customer or user), is physically distant from the others (Prickladnicki, Audy & Evaristo 2003). The term, Global Software Development (GSD) also represents development occurring at inter-continental scale. Besides GSD, there are other related terms which are used in the literature including Virtual Teams, Multisite Development, Dispersed Development, Offshoring, Outsourcing (does not always equate to distribution). In this study we are using ‘Distributed Software Development (DSD)’ term since, it is a general term to represent software development with geographically dispersed teams and does not have connotations of great geographical distances found in other designations (Layman, Williams, Damian & Bures, 2006). DSD is characterized by the properties like spatial distance (geographic distance), time-zone distance and socio-culture distance.

Organizations are combining agile with distributed development in order to get the benefits of both the approaches. They are able to create low cost solutions, which fulfill the changing requirements of businesses and solution can be delivered in shorter time. DSD and agile work on different principles which makes the distributed agile projects difficult to manage. DSD requires formal communication amongst the distributed teams to communicate effectively across the globe. Agile, on the other hand is based on informal communication with co-located teams working in close collaboration. Moreover, several agile best practices, including collaboration, face to face communication, self organizing teams, retrospectives, showcases, etc become more challenging in the distributed model (Thoughtworks, 2008). As a result, there are significant challenges, which make risk management in DAD crucial for successful project delivery.

Risk is defined as the net impact of the exercise of vulnerability, considering both the probability and the impact of occurrence (Boehm 1991). Further, a risk factor can be defined as a condition that can present a serious threat to the successful completion of a software development project (Schmidt et al. 2001). Risk management is the process of identifying risk, assessing risk and taking steps to reduce the risks to an acceptable level (Stoneburner, Goguen & Feringa, 2002). Four basic methods can be used for controlling the risks, namely, risk mitigation, risk avoidance, risk transfer and risk acceptance. This study involves identification of the risk factors which may become a threat to the execution of DAD projects and the most suitable risk management approaches which are frequently used by the practitioners for controlling them.

It is also required to assimilate the risks in the organizational context for which an existing organizational model must be related to the findings of the study. We chose Leavitt’s model for this purpose, as the basic elements of our research findings corresponded to the components of Leavitt’s model of organizational change. According to Leavitt’s model, an organization is composed of four interacting components, namely, task, structure, actor and technology (Leavitt, 1964). These components can easily be translated into well known elements of software development. ‘Task’ signifies expected outcomes in terms of goals and deliverables. ‘Actors’ cover all stakeholders including users, managers and designers. ‘Structure’ denotes project organization and other institutional arrangements. ‘Technology’ means development tools, methods and hardware and software platforms. Leavitt’s model of organizational change was found to be suitable for this study due to its intuitive appeal, simplicity and appropriateness for the problem situation. Moreover, Leavitt’s Model has been used in the information system literature extensively as it covers most of the important aspects of software development from the organizational perspective (Lyytinen et al., 1998)
2.1 Consolidation of Literature Review
We did an extensive literature survey to find out the challenges faced by teams executing DAD projects and the suggested approaches to manage those problems. Along with this, we also reviewed the literature which considers the issues in distributed development, agile development and traditional software development. We found that there is scarcity of research work in this area and there are certain issues which have not been explored at all. The following studies were considered for review:

1. Systematic review of literature on Distributed Agile Development: (Jalali and Wohlin 2010), (Hossain et al. 2009)
2. Case Studies in Distributed Agile Development: (Simon 2002), (Fowler 2006), (Mattson et al. 2010), (Miller, 2008), (Lee, Banerjee, Lim, Kumar, Hillegersberg, and Wei 2006), (Balasubramanion et al. 2006), (Therrain 2008).
3. Risk Management Frameworks or Risk Classification for Distributed Agile Development: (Hossain et al. 2009), (Mudumba and Lee 2010), (Mattsson et. 2010)
4. Risk Management studies in Agile Development: (Nyfjord and Mattson 2008), (Boehm and Turner, 2005), (Misra, Kumar and Kumar 2009), (Chow and Cao, 2008)
5. Risks Management studies in Distributed Agile Development: Prickladniki et al. (2006), Prickladniki et al. (2003), (Perrson 2009), (Jim´enez , Piattini , Vizca´ino 2009), (Carmel and Agrawal, 2001), (Herbsleb, 2007), (Smite, Wohlin, Gorschek & Feldt, 2009)

Examination of literature shows that there is dearth of research work, which provides a comprehensive view of risk management in DAD. There are few systematic review studies which highlights the risks in DAD, but they unable to provide the industry perspective of the area under study. Several case studies are available, but the challenges and the corresponding solutions presented in them are situation based. Various studies are discussing risks in agile development and risks in DSD separately, but few are considering the challenges when the two approaches are combined. Hence, this study has been undertaken, in order to unveil the most important risks faced by DAD teams and suggests the frequently used risk management methods which can be used by them.

3. Scope and Statement of the Problem

Study of the existing studies reveals that the software development organizations are embracing distributed agile development to create low cost solutions which can deliver high business value in lesser time. This study aims at investigating the most significant risk factors which impact distributed agile development (DAD) projects and the most suitable methods to control them.

There exists a relationship between the DAD and distributed development (DSD) and between DAD and agile development since DAD is actually a combination of DSD and Agile approach of software development. Hence, it is apparent that while we study the risks in DAD, problems caused due to distributed development and those caused due to agile development will also emerge and hence will be required to include them in the research findings.
Software project risk points to an aspect of development task, process or environment, which if ignored tends to adversely affect the outcome of a software project (Lyytinen et al. 1993). We planned to unearth the risk factors impacting DAD projects through qualitative exploratory study and then identify the most important ones through a quantitative descriptive study. Hence, we define our research problem in the following research questions:

**Research Question 1:** What are the most significant risk factors in DAD projects as perceived by the practitioners?  
**Research Question 2:** What are the risk management techniques frequently used in practice for managing the identified risk factors?  
**Research Question 3:** How are the risk factors categorized to facilitate real world adoption?  
**Research Question 4:** Which of the project constraints of Time, Cost and Quality are impacted most significantly by specific risk factors identified in the exploratory study from the perspective of the practitioners?

In order, to study the risks in DAD, we plan to approach the organizations which execute DAD projects. Since, the base location of the researchers is India, majority of the respondents in the sample will be from Indian software organizations. Some part of our sample may consist of respondents who belong to organizations outside India too, but such cases will be very few. In our view, the location attributes of the population may not have great effect on responses as the subjects in sample are expected to have an experience of working in distributed environment for sufficiently long time and so, would have much enlarged and common vision of the risks faced in DAD projects. This would eventually subdue the effect of specific location on their experience and thus, views.

### 4. Research Objectives

The research questions lead to the following research objectives for this study:

1. To identify the conditions (risk factors) which impact the success of DAD projects.
2. To identify the dimensions of those risk factors which include, description, cause of the risk, software development methodology which is the source of the risk and the corresponding risk management methods which can be used to reduce their impact.
3. To create a categorization of risk factors impacting DAD project by grouping related risk factors together.
4. To segregate the most important risk factors from the identified set of factors which impact DAD projects.
5. To provide the frequently used risk management approach for the risk factors so that the practitioners can control the risk more effectively.
6. To identify the risk factors which may lead to time overrun, those risk factors which may lead to cost overrun and those risk factors which may lead to low quality solutions.
7. To develop a risk management framework that can easily be adopted by the organizations and signifies the important risk factors, corresponding risk management techniques in practice and the impact of risk factors on the project constraints of time, cost and quality.
5. Hypothesis

Based on our research problem, we propose that
1. The identified risk factors for distributed agile development projects are perceived by practitioners to have an impact on the success of the projects.
2. The proposed risk management approaches for the risk factors are frequently used to reduce the impact of the risk in distributed agile development projects.
3. The identified risks occurring in DAD projects also have impact on the project constraints, namely, time, cost and quality.

Our proposition is operationalized by the following set of hypothesis:

Hypothesis for Risk Factors in DAD Projects
1. H₀RC: The identified Risk Categories do not have an impact on the success of DAD project.
   H₁RC: The identified Risk Categories have an impact on the success of DAD project

2. H₀RA: The identified Risk Areas do not have an impact on the success of DAD project.
   H₁RA: The identified Risk Areas have an impact on the success of DAD project.

3. H₀RF: The identified Risk Factors do not have an impact on the success of DAD project.
   H₁RF: The identified Risk Factors have an impact on the success of DAD project.

Hypothesis for Risk Management Methods for DAD Projects
1. H₀RM: The suggested Risk Management Approaches are not used frequently for reducing the impact of the risk factors in DAD projects.
   H₁RM: The suggested Risk Management Approaches are used frequently for reducing the impact of the risk factors in DAD projects.

Hypothesis for impact of Risk Factors on Project Constraints in DAD Projects
1. H₀T: The identified Risk Factors do not have an impact on the Time of the DAD Project
   H₁T: The identified Risk Factors have an impact on the Time of the DAD Project

2. H₀C: The identified Risk Factors do not have an impact on the Cost of the DAD Project
   H₁C: The identified Risk Factors have an impact on the Cost of the DAD Project

3. H₀Q: The identified Risk Factors do not have an impact on the Quality of the DAD Project
   H₁Q: The identified Risk Factors have an impact on the Quality of the DAD Project

We plan to test the above listed hypothesis by conducting the quantitative descriptive study.

6. Research Study Methodology

The existing literature gave us insights about the risks encountered by teams executing DAD projects and the approaches used to reduce their impact. We planned to perform an extensive exploration to reveal the risk factors impacting DAD projects and their corresponding methods to
control them. We decided to conduct descriptive quantitative study in order to identify the most important risk factors and the most suitable risk management methods as found by practitioners. The qualitative and quantitative study will be consolidated to create a risk management framework for DAD projects. It was decided to evaluate the framework for its benefits and usage by getting it validated from the software organizations which execute DAD projects.

6.1 Methodology for Qualitative Exploratory Study

Since, there was a dearth of literature addressing the challenges in DAD projects; we decided to perform an exploratory study. It was considered appropriate to conduct ‘In-depth interviews’ of practitioners in order to identify the risks in DAD projects. It was planned to conduct a pilot study in order to improve the questionnaire before collecting data. The in-depth interview would provide us the data pertaining to risks in DAD projects, cause of the risk, the software development approach which is the originating source and the corresponding risk management approaches. The risks which are caused due to blending of distributed development and agile methods will be referred to as ‘DAD Risk Factors’, while other risk factors will be termed as ‘Non DAD Risk Factors’. The ‘Non DAD Risk Factors’ will refer to those risks which are caused due to the use of other software development approaches like agile software development (ASD), distributed development (DSD) and traditional software development (SWD). While, we will be including both DAD risk factors and Non DAD risk factors in our framework, the focus of the discussions of the study will be on DAD risk factors only.

In order to strengthen the findings of our initial exploration, we decided to analyze the project work documents of DAD projects to reveal the risks and issues. This would reduce the possible bias in our findings (Denzin, 1990). Document Analysis is a systematic procedure for reviewing or evaluating documents—both printed and electronic material (Bowen 2009).

We used ‘Constant Comparative Method’ for analyzing the qualitative data and create an initial risk categorization for DAD risks (Glaser and Strauss, 1967). Constant comparison method is concerned with generating and plausibly suggesting (not provisionally testing) many properties and hypotheses about a general phenomenon (Glaser, 1965), which in our case, is the Risk Management in Distributed Agile Development (DAD) projects. It is iterative such that although we proceed from one stage to another, the previous stages remain in operation and continuously contribute to the evolution of data analysis results (Seaman, 1999).

This method involves four stages including, comparing incidents applicable to each category, integrating categories and their properties, delimiting theory and writing theory (Glaser, 1965). In our study, the information on risk factors, which we will obtain from the respondents will correspond to the ‘incidents’ given in the explanation of constant comparison method (Glaser, 1965). These incidents will be compared, the properties of the risk factors will emerge which will lead to the formation of the risk categorization. The categorization will be composed of risk factors at one level, risk areas at the second level and then risk categories at a higher level. This method will be applied iteratively to consolidate the data obtained by conducting in depth-interviews and by project work document analysis which will lead to the evolution of a risk categorization for DAD projects. The process of updating and strengthening of the risk categorization matrix will lead to repetition of certain patterns. At this stage the categories will be saturated and it will correspond to ‘delimitation of theory’. Further, the risk categories will be mapped to the existing theoretical themes. This step can be referred to as ‘writing theory’ and will help us to take into account all the aspects of software development from the organization’s perspective.
6.2 Methodology for Quantitative Descriptive Study

Exploratory study will lead to the creation of a ‘risk factor categorization for DAD projects’ containing risk factors impacting DAD projects grouped in relevant risk areas and then broad risk categories. In order to identify those risk factors which are most significant and the most appropriate risk management approaches in DAD projects from the view of the practitioners, we decided to conduct a quantitative descriptive study of risk factors. Besides this, we also decided to gather data corresponding to the project constraints which are being impacted by those risks also. It was decided to use the triple constraints, time, cost and quality for this analysis.

6.2.1 Analysis Method for Rank Order Data for Risk Factors

In order to find out the most important risk factors for DAD projects, we decided to obtain the rank order data corresponding to risk category, risk area and risk factor at three different levels. A self-administered questionnaire will be used, which will be pilot tested by sending it to experts and obtaining their views on the content and quality of the questionnaire.

We will obtain the rank order data corresponding to risk factors, risk areas and risk categories based on the ‘severity of impact they have on the project success’ on a scale of 5 to 1. The scale to be used will be defined as: 5: very strong impact, 4: strong impact, 3: some impact, 2: negligible impact, 1: no impact (0: don’t know). At the first level, the risk categories will be ranked depending upon the extent to which they impact the project success. At second level, the relative ranks of risk areas under each category will be obtained. At third level relative ranks of risk factors under each risk area, based on the severity of impact they have on the project will be obtained.

In order to measure the level of agreement of the respondents on the ranking of the risk factors, we decided to use Kendall’s test of Concordance on the rank order data (Kvam and Vidakovic, 2007). The same applies to rank order data of risk areas and risk categories. We planned to use the significance level of 0.05 for analyzing the data obtained after Kendall’s test at all the three levels.

6.2.2 Analysis Method for Rating Data for Risk Management Methods

Corresponding to the risk factors in DAD projects, the suitable risk management methods to control them will also be collected during exploration. We intend to find out the most frequently used risk management approaches for controlling each of the risk factor for DAD project. We decided to obtain the rating of risk management methods corresponding to each risk factor based upon the ‘frequency of use’ in the practice with help of a self-administered questionnaire. The scale to be used will be defined as follows: 5: Always Used, 4: Often Used, 3: Sometimes Used, 2: Seldom Used, 1: Never Used (0: Don’t Know)

In order to find out the statistically significant ‘frequency of usage’ of each risk management method to control a particular risk, we used Chi-square test of Independence test. This test is used to analyze the frequencies of two variables with multiple categories to determine whether the two variables are independent (Black, 2009). The significance level for performing analysis of data after applying Chi-Square test used will be 0.05. By using chi-square test, we will be able to find out the risk management methods identified during exploratory study are being used more frequently in the practice and those methods which are used relatively with lesser frequency.
6.2.3 Analysis Method for Project Constraint Data
Along with the rank order data, we decided to obtain the project constraint (time/cost/quality) which is being impacted most significantly when a risk occurs in DAD projects. We planned to use Chi-Square test of Independence for identifying the risk factors which are statistically significantly impacting one of the project constraints. We decided to use the significance level of 0.05 for analyzing the data. This will help us to segregate all the risk factors which are impacting the ‘time of the project’, those which impact the ‘cost of the project’ and those which impact the ‘quality of the project’. This will in turn help the managers executing DAD projects, to concentrate only those risk factors which are impacting the project constraint, which is important for a particular project.

6.3 Methodology for Validation
The findings of the qualitative descriptive study and quantitative descriptive study will be consolidated to create the risk management framework for DAD projects. We plan to evaluate its applicability and performance of the risk management framework in the real world by getting it validated from the companies. Companies will be provided with the partial framework due to the paucity of time. According to the relevance of risk category to the present status of the project, one or two of the risk categories will be implemented. The risk management framework will be suggesting the risk factors in the order of their importance and the most appropriate approaches which can be used to control the risk in DAD projects. The risk exposure (magnitude of risk) will be measured based on the perception of the project team before and after implementing the framework. The benefit of implementing the framework will be obtained by the difference between the risk magnitude before and after the implementation of the framework. The percentage of risk factors whose risk magnitude has reduced and the percentage or risk factors which were altogether eliminated will tell us about the benefit of the framework.

We will also ask the respondents to specify the risk management methods which they used for controlling the risk. This will help us to evaluate the usage and benefit of the framework for DAD projects.

7. Findings
This study involved the exploration of risk factors impacting distributed agile development projects, through a qualitative exploratory study, which was followed by the quantitative descriptive study. The results obtained after consolidating the findings from these two studies led to the formation of the risk management framework for DAD projects.

7.1 Qualitative Exploratory Study: Data Analysis and Results
Exploratory study aimed at unveiling the risk factors which may become a threat to the success of DAD projects and the suitable risk management methods to control them. We conducted in-depth interviews of the practitioners and then analyzed project work documents of DAD projects. Purposive sampling was used for data collection, wherein, the sample conforms to the criterion that the subjects are industry practitioners, who have managed DAD projects. This sample included program directors, project managers, delivery managers, test managers, developers, agile coaches, scrum masters and team members having experience of maximum 10-20 years in IT development and minimum 3 years of experience in handling DAD projects. The size of the Information Technology (IT) companies, we visited for interviewing the practitioners and project
work document analyses varied from 1000 employees to above 100,000 employees. We conducted 13 in-depth interviews, which were predominantly face-to-face meetings (12) and few were telephonic (1) lasting for 45 minutes to 1 hour. We recorded data by making verbatim notes (Seaman, 1999) and further contacted the respondents to obtain the missing information.

For each risk factor, we obtained the cause of the risk, the source data in terms of the software development methodology and the risk management approaches suitable to manage those risk factors. We used constant comparison method to analyze the qualitative data pertaining to risks in DAD projects. This led to the creation of an initial risk factor categorization. We also examined the existing review of literature to support the identified risk factors and the management techniques.

We analyzed work documents of 28 DAD projects of specific projects in 13 different IT companies located in cities within and outside India. Work documents like risk registers, retrospectives, and sprint backlogs, documents in tools, e-mails, status reports and statement of work, were analyzed at the company location and was supported by interaction with the project. The findings from the in-depth interview and the work document analysis were further consolidated to form a risk categorization for DAD projects. The risk categorization comprised of eighty-one risk factors grouped under relevant risk areas and five broad risk categories. Table 1 lists the identified risk factors, the risk areas that the factors belong to and the corresponding risk categories.

The risk categories which emerged out of this process were Software Development Life Cycle, Project Management, Group Awareness, External Stakeholder Collaboration and Technology Setup which had relevant risk areas and risk factors listed under them. It was required to relate our findings to an organizational model for facilitating its adoption in industry. We considered various existing organizational models, namely, Leavitt’s model, Weisbord’s six-box model, Nadler-Tushman Congruence model and Tichy’s framework and assessed for their suitability for our research findings (Burke 2011). We chose ‘Leavitt’s model of Organizational Change’ since, the relationship between the components of Leavitt’s model and the major risk categories was easily apparent. The risk category ‘Software Development Life Cycle’ was mapped to the ‘Task’; ‘Project Management’ to ‘Task-Structure Interdependencies’, ‘Group Awareness’ to ‘Structure’, ‘External Stakeholder Collaboration’ to ‘Actor’ and ‘Technology Setup’ to ‘Technology’ component of Leavitt’s model. The core inside of the model in figure 1 depicts the mapping of risk categories to the dimensions of Leavitt’s model.

We also categorized the risk factors based on the amount of literature support available, in order to identify those factors which are encountered by the DAD teams but not addressed in research studies and need further exploration. This helped us to highlight the risk factors which are having feeble visibility in the existing literature and hence need to be considered by future studies.

7.2 Quantitative Descriptive Study: Data Analysis and Result
The aim of the quantitative descriptive study was to identify the most important risk categories, risk areas and risk factors and the most frequently used risk management methods to reduce their effect. It was required to collect objective data related to ranking of risk factors on the basis of their impact on the project success. Further, it was decided to collect the rating data for risk management methods corresponding to each risk factor based on the ‘frequency of use’ in the industry. Along with that, we also aimed to collect data on the project constraint (time, cost, quality) which is being impacted most significantly when a risk factor in DAD project occurs.
We used a self-administered questionnaire, which was improved by obtaining feedback on its content from 11 experts. The questionnaire was split into two parts due to its length and hence the data collection was done in two separate phases. Part I of the questionnaire aimed at collecting data pertaining to ranking of risk factors and Part II of the questionnaire was corresponding to the rating data of risk management methods.

7.2.1 Risk Categories, Risk Areas and Risk Factors: Data Analysis and Results

Before collecting the data, Part I of the questionnaire was pilot tested by 15 respondents, which led to the creation of an additional risk category, namely, ‘Business Objectives and Goals’. The risk category deals with the problems which arise when the development team in the software organization is not aligned with the business objectives and goals of the customer’s organization and vision of the project.

We used purposive sampling for this study since the sample must fulfill the criteria of being a practitioner who has minimum three years of experience in managing DAD projects. Our sample included CEO’s of companies, program directors, program managers, senior project managers, project managers, chief architect, consultants, agile coaches, scrum masters, business analyst, and senior software engineers with the required experience of executing DAD projects. They belonged to various countries like UK, USA, Japan and India. We contacted 76 practitioners, out of which 11 respondents did not respond. Hence, we sent our questionnaire to 65 respondents, out of which 39 sent us their response back. For some of the responses which had missing data, we had to contact the respondent again through mail or telephone to obtain the required data. Part I of the self-administered questionnaire was used for the collection of rank order data for risk categories, risk areas and risk factors at three different levels.

We applied Kendall's test of Concordance on the risk factors rank order data in order to evaluate the level of agreement amongst the respondents. The Kendall’s test of Concordance provided us with the mean rank of risk factors within each risk area, Kendall’s coefficient of Concordance (W), observed p-value. The mean rank was used to arrange the risk factors in their respective risk areas in the descending order of mean rank obtained. As a result the risk factor with ‘strongest impact’ to the factor came on the top of the list, while the factor with least impact was at the bottom, which helped in highlighting the important risk factor. The value of ‘W’ was used to assess the level of agreement on the rank order of the risk factors amongst the respondents. The ‘observed p-value’ was used to identify the risk factors which are statistically significant. The same method was applied to risk area under each risk category and to the risk category also.

At the risk category level, the ranking was statistically significant (p-value 0.0000), although there was a low level of agreement amongst the respondents (W: 0.221). The risk category ‘Group Awareness’ was ranked highest which was followed by ‘External Stakeholder Collaboration’ and then ‘Software Development Life Cycle’. Next in the order was the risk category ‘Project Management’ and then ‘Technology Setup’. Similarly we obtained Kendall’s coefficient of concordance, p-value and mean rank for each of the risk area and risk factor.

A consolidated view of the risk factors, risk areas and risk categories, arranged as per the mean rank along with the level of significance can be seen in the table below (Table 1). The statistically significant risk areas and risk factors are differentiated from the insignificant ones by using various symbols in the Table 1. The DAD risk factors are marked with superscript D and the Non DAD risk factors are marked with the superscript ND. This distinction is made based on the inputs we obtained from the respondents during our exploratory study. Overall, it was
observed that the level of agreement for the ranking by the respondents was low. The reason for low ‘W’ could be the heterogeneity in the designation of the respondents, as certain respondents belong to the executive category while others were in the middle management level. Secondly, due to comprehensiveness of the exploratory findings, possibly, every respondent has not encountered certain risks and hence their views differ.

We also analyzed the ‘Don’t Know Responses’ corresponding to ranking data of risk factors in DAD projects. We applied the chi-square test (with level of significance as 0.05) on the data and obtained insignificant results. Hence, we did not include the results of the ‘Don’t know response’ in our risk management framework. This analysis helped us to segregate the risk factors which have lesser visibility amongst the practitioners, although they impact DAD projects, and hence must be considered for further exploration by future studies.

Table 1: Risk factors in DAD Projects listed ordered as per the mean rank and segregated as per the significance value.

<table>
<thead>
<tr>
<th>Risk Categories ordered based on the Mean Rank</th>
<th>Risk Area ordered based on the Mean Rank</th>
<th>Statistically Significant Risk Factors ordered based on Mean Ranks</th>
<th>Statistically Insignificant Risk Factors ordered based on Mean rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Awareness **</td>
<td>Communication (p-value: 0.467)</td>
<td>Lack of Communication between the Team and the Client **</td>
<td>ógement of Risk Factors between the Team and the Client **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of Communication between Team Members **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uncommon Language **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maintaining Communication in Large Distributed Teams **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underinvestment on Travel by the Management **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsuitability of Flat Communication Structure to Large Organization</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of Documentation **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>Coordination &amp; Collaboration (p-value: 0.05)</td>
<td>Poor Collaboration between Different Sites **</td>
<td>**</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td>Poor Collaboration between Developers and Testers **</td>
<td>**</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td>Poor Coordination between Different sites **</td>
<td>**</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>Trust</td>
<td>**</td>
<td>Trust (p-value : 0.527)</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of Trust between the Client and the Offshore team **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of Trust between the Onshore and Offshore Agile Teams **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>External Stakeholder Collaboration **</td>
<td>Customer Collaboration ** (p-value: 0.019)</td>
<td>**</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Product owner is not Empowered **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unavailability of Product Owner **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Product Owner’s Attitude towards the Team **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Customer not aligned to Agile Practices **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>Multiple Vendor Involvement ** (p-value: 0.014)</td>
<td>**</td>
<td>Poor Coordination between Multiple Vendors **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Challenges in Code Integration with Multiple Vendors **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inappropriate User Story Estimates with Multiple Vendors **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dependency on Third Party **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>Software Development Life Cycle **</td>
<td>Requirement Elicitation ** (p-value: 0.000)</td>
<td>**</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requirements Unclear to the Product Owner **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requirements Unclear to the Team **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requirement conflicts amongst multiple Product Owners **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-functional Requirements are not fulfilled **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inadequate communication about End User Requirements **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unclear Project Objectives **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Loss of sight of Big Picture to Agile Teams **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>Release Management **</td>
<td>**</td>
<td>(p-value: 0.827)</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trouble in making Release Plan **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Difficulty in System Release Management and **</td>
<td>ógement of Risk Factors between Team Members **</td>
</tr>
<tr>
<td>Risk Categories ordered based on the Mean Rank</td>
<td>Risk Area ordered based on the Mean Rank</td>
<td>Statistically Significant Risk Factors ordered based on Mean Ranks</td>
<td>Statistically Insignificant Risk Factors ordered based on Mean rank</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
</tr>
<tr>
<td>Testing and Integration **</td>
<td>(p-value : 0.031) Automated Testing ND</td>
<td>Missing Technical Debts D</td>
<td>Deployment **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Un-maintainable Code ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Code Integration across Multiple Sites D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross Functional Teams Insufficient for testing of Large Projects D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deferring End-of-Sprint code improvement to Code Hardening ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test Data Management ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deployment **</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(p-value : 0.000)</td>
<td>Inadequate Prioritization of Requirements D</td>
<td>(p-value : 0.334) Rework Caused by Architectural Changes D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Product owners keep changing the Requirements and their Priorities Frequently D</td>
<td>Unidentified User Interface Design ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inappropriate User Story Estimates in a Team ND</td>
<td>Inconsistency in Design Standards of Distributed Teams D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Difficulty in Maintenance of Product Backlog ND</td>
<td>Inadequate Focus on Design ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Difficulty in Story Point Technique for Estimation ND</td>
<td>Unavailability of User Interface (UI) Designers ND</td>
</tr>
<tr>
<td></td>
<td>(p-value : 0.0000)</td>
<td>No common Definition of Done between the Distributed Teams D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ineffective Standup Meetings D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Different Software Development Practices and Standards followed by Multiple Teams D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(p-value 0.077)</td>
<td>Project Organization *</td>
<td>(p-value : 0.216) Higher Interdependency between the teams D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Project Planning and Control *</td>
<td>Using Component team instead of Feature team D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Project Initiation *</td>
<td>Power Struggle between the Teams ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Acquiring and development of</td>
<td>Team Reorganizing in Every Sprint D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Growth in Team Size or Development Sites D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(p-value : 0.361) Lacks of Domain Knowledge ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unavailability of Business Analyst D</td>
</tr>
</tbody>
</table>
Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)

<table>
<thead>
<tr>
<th>Risk Categories ordered based on the Mean Rank</th>
<th>Risk Area ordered based on the Mean Rank</th>
<th>Statistically Significant Risk Factors ordered based on Mean Ranks</th>
<th>Statistically Insignificant Risk Factors ordered based on Mean rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>project team *</td>
<td>(p-value: 0.002) Attraction due to Work Pressure and developer’s burnout **</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Managing the Project Team *</td>
<td>Lack of Special Skills for Agile Development ND</td>
<td></td>
</tr>
<tr>
<td>Technology Setup **</td>
<td>(p-value 0.162) Infrastructure and Resources</td>
<td>Lack of Experience in Managing Distributed Project ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tool Selection</td>
<td>Lack of Uniformity in Multisite Team’s Capabilities D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Factors in italics means same rank)</td>
<td></td>
</tr>
</tbody>
</table>

Legends used in the Table (above)
Significance value
1. **: (Significance Value: p-value <= 0.05): obtained after applying Kendall’s test of Concordance on all the ranks given to the risk areas under the marked category
2. * : (Significance Value: p-value > 0.05 and < 0.10): obtained after applying Kendall’s test of Concordance on all the ranks given to the risk areas under the marked category
3. Those risk areas which do not possess any asterisk are the ones which have p-value > 0.10
4. Superscript D: DAD Risk Factors and Superscript ND: Non DAD Risk Factors

7.2.2 Risk Management Methods: Data Analysis and Results
Similarly, we obtained the rating data corresponding to the risk management methods of DAD risks by using Part II of the questionnaire for quantitative data collection. We contacted only those respondents who had provided us with the response for the first part of the questionnaire, since they were aware of the context of the study and also had the minimum experience required for responding our questionnaire. We could obtain the consent from 28 respondents for filling the second part of the questionnaire out of which 19 respondents provided us with their response.

Chi-square test was applied on this data which provided us with a chi-square coefficient, p-value and the rating of each risk management method as given by maximum number of respondents corresponding to each of the risk factor. The risk management methods for each of the risk were ordered based on the rating data. It was observed that the results were statistically insignificant, due to the less number of respondents (19) who provided us the rating data for risk management approaches. Along with quantitative data, we also obtained the qualitative reviews on the risk management approaches listed in the questionnaire. We would like to assert that, since the rating of risk management methods is about the frequency of use in practice and should not be mistaken with the rigorously analyzed and reasoned perfection.

We tried to reduce the effect of this limitation by obtaining qualitative views from the respondents on the suggested methods. We identified certain risk management approaches which are ideal methods from the agile perspective, but due to the immaturity in the organizations processes, these are not being used frequently in the industry. Secondly, there were certain methods, which were poor techniques, but are used frequently and hence may have received
higher ranks. Some respondents even suggested additional risk management approaches for the risks, besides the methods suggested, which were included in the results.

7.3 Impact of Project Constraints: Data Analysis and Results
As we collected the ranking data for risk factors, we also obtained the project constraint which is most significantly being impacted when a particular risk occurs. Three project constraints, namely, time, cost and quality were considered. We applied the Chi-square test on this data and obtained highly significant results (p-value 0.000) for all the risk factors. Hence, we assert that we could clearly identify the risk factors which impact ‘time’, those which impact ‘cost’ and those which impact ‘quality’ in DAD projects.

7.4 Business Goals and Objectives: Data Analysis and Results
We identified the following risks which in the risk category ‘Business Objectives and Goals’ and they are listed in the order of the ranks obtained based on the impact they have on software development projects.

2. No Idea of the Damage that could be caused to the Customer’s Business due to the poor quality of the software, delay in development or spending more than the project budget.
3. Lack of Clarity in the Business Value that the Developed Software will create for the Customer.
4. Unrealistic Expectations on Customer and Development sides.
5. The Information System (IS) Objectives and Plans of the solution provider are not aligned with the Business Objectives and Plans of the customer.

Although, we obtained the mean rank after applying the Kendall’s test of concordance, but the results were insignificant (p-value 0.223) and had a low value of W (0.037). We suggest some of the risk management approaches, which can be adopted to reduce the impact of these risks in the section below. Due to time and resource constraints, we did not do an in-depth analysis of these factors, and hence, it can be considered for future research. We also performed the analysis of the project constraint data pertaining to the ‘Business Objectives and Goals’ risks and we obtained highly significant results (p-value: 0.000). We also obtained the rating of the risk management approaches corresponding to each of the risk in this category. We obtained highly significant results for risk management methods rating in case of risk number 2 and 3 (p-value: 0.002 and 0.000 respectively), less significant results in case of risk factor number 1 (p-value: 0.069) and insignificant results in case of risk factors number 4 and 5 (p-value: 0.480 and p-value: 0.229). These findings would be useful for the practitioners for reducing the risks caused due to the lack of clarity of business vision of the project amongst the distributed team members.

7.5 Result Validation: Data Analysis and Results
The DAD risk management framework was validated for evaluating its applicability and performance in the real world. Partial framework, consisting of a particular risk category depending upon the company’s requirements was provided for implementation in real world DAD project. We contacted 20 experts who are practicing distributed agile in 10 different companies. Three multi-national IT companies agreed to perform validation of the risk management framework (partial). We provided the partial framework corresponding to the ‘Software Development Life Cycle’ risk category to Company 1 and the ‘Project Management’
Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)

risk category to Company 2 and Company 3, based on their requirements. Objective data corresponding to the risk exposure for each of the considered risk factor was obtained before and after implementation of the framework. The analysis of the observed difference in the risk exposure (risk magnitude) of the considered risks before and after implementation of the framework was used to measure the benefit and performance of the framework.

7.5.1 Validation Results and Analysis: Company 1
The partial framework corresponding to ‘Software Development Life Cycle’ was implemented in the seventh sprint of a medium sized project. The project was an ‘Application development for Banking and Financial Services’. The project had three teams which were located within India as well as outside India. Each team had 15 members including business analyst (1), developers (9), testers (3), scrum master (1) and product owner (1). Consolidation of the validation results as provided by the three teams is given as in Table 2.

<table>
<thead>
<tr>
<th>Team No.</th>
<th>Total no. of Risk factors analyzed (Total 34 risks)</th>
<th>%age of risk factors which did not occur</th>
<th>%age of risk factors with reduced impact</th>
<th>%age of risk factors with unchanged impact</th>
<th>%of risk factors with inflated impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>34</td>
<td>27/34 (79.41%)</td>
<td>3/34 (8.8%)</td>
<td>0</td>
<td>3/34 (8.8%)</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3/6 (50%)</td>
<td>3/6 (50 %)</td>
<td>0</td>
<td>2/6 (33%)</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>8/15 (53.33%)</td>
<td>1/15 (6.6% )</td>
<td>5/15 (33.3 %)</td>
<td>1/15 (6.6%)</td>
</tr>
</tbody>
</table>

7.5.2 Validation Results and Analysis: Company 2
In the second company, we provided the ‘Project Management’ part of the risk management framework. The project was a ‘Product Development’ with teams located at 5 different locations all across the globe. The team was using a hybrid of agile methods and traditional approach for development. The framework was implemented in the third sprint of the project, at one location in India which had the team size of 20 members. There were eight sprints and length of each sprint was 2 months. The project manager considered 18 (out of total 23) risk factors from the project management risk category. We noticed that for 9 out of 18 risk factors (50%), the risk magnitude reduced. There were also 2 factors (11%) for which the risk magnitude remain the same while 7 risk factors (43%) for which the risk magnitude increased.

7.5.3 Validation Results and Analysis: Company 3
We again provided the ‘Project Management’ part of the risk management framework to Company 3. We noticed that the total number of risk factors considered was 23 (all risk factors in the Project Management risk category) and the number of risks whose magnitude reduced was 12 factors (53%). There were 6 risk factors for which the risk magnitude remained the same (26%) and none of the risk factors for which the risk magnitude inflated.

The validation results shows that in all the three cases, majority of the risk factors considered before the framework implementation either did not occur, or their risk magnitude reduced after the framework was implemented. This indicates that the framework was beneficial for project and helped in effective risk management.
8. Interpretation of Results

We indentified the risk factors which are more important from the practitioner’s views and statistically significant also. Based on the results, we observed that risks in distributed agile projects occur primarily because distributed development properties and agile practices differ from each other in their key tenets (Balasubramaniam et al. 2006).

8.1 Interpretations: Qualitative Descriptive Study (DAD Risk Factors)
We identified five distinct properties of distributed development namely, Spatial distance, Temporal Distance, Work/development culture, Language Barrier and Large Project Scope. These properties which characterize distributed development are in conflict with the principles and practices of agile development and hence are causing the DAD risk factors to occur.

As we analyzed the results obtained after qualitative study, we identified the DAD risk factors which area occurring primarily due to one of the DSD property and also identified agile principles and practices which is conflicting with the DSD property identified. Here we provide and ‘overall view of the risk factors’ which are caused due to a particular DSD property. Specific risk factors corresponding to each of the risk category, segregated based on the DSD property to which they corresponds to, is presented in the figure 2.

1. Spatial Distance: Geographic separation minimizes face-to-face interactions in distributed teams causing miscommunication and mistrust amongst the stakeholders. Agile methods revolve around frequent communication and collaboration between the team members and the customer. Risk factors corresponding to lack of communication/collaboration amongst the teams and the client become a major hurdle. Underinvestment on travel by the management and lack of documentation increases the severity of trouble causes to ineffective communication. Agile practices like removing technical debts, pair programming, continuous testing and integration cannot be done effectively when teams are non collocated. Risk factors related to in project planning and team organization, structuring and management also can be attributed to spatial distance.

2. Temporal Distance: When teams are distributed and there is long time zone difference, it becomes difficult to arrange for daily meetings and other planning events like release planning, sprint planning etc., which are important ceremonies in agile development environment. Time zone difference lead to difficulty in coordination related problems. Team members need to work outside office hours which lead to frustration and burnouts (Mattsson et al. 2010).

3. Language Barrier: Differences in languages and working styles of teams of different countries cause ineffective communication and directly affect the productivity (Cusumano 2008). This hinders agile philosophy of open, direct communication, trust, and close cooperation between business people and developers (Mattsson et al. 2010). Good language skills are imperative for agile teams because they need to interact more frequently with the client and showcase their work regularly.

4. Work/Development Culture: Continuous evolution of software needs development of high quality source code which is easier to understand and maintain. This is achieved when teams adopt rigorous engineering practices like refactoring, test driven development, continuous integration (Cusumano 2008). Difference in the work cultures of distributed teams means significant differences in the engineering processes and practices and the tools being used that further reduces productivity (Bavani, 2012).
5. Large Project Scope: Distributed environments involve large network of heterogeneous, autonomous and distributed models (Jiménez et al. 2009). Agile methods are designed for projects involving small and medium sized collocated teams to address the need for fast informal communication (Misra, Kumar and Kumar 2009; Turk et al. 2002). With larger teams, the need for large number of communication lines can reduce the effectiveness of agile practices such as various review and project planning meetings (Turk et al. 2002), (Woodward et al. 2010).

8.2 Interpretations: Quantitative Descriptive Study (DAD Risk Factors)
It was observed that the risk factors, risk areas and the risk categories which were ranked above in their respective list were primarily caused due to spatial distance between the project stakeholders. Next in the order of importance was temporal distance. Then there some factors which occur due to work/development culture, language barrier and few due to large project scope.

8.2.1 Interpretation for Risk Category, Risk Area and Risk Factors Ranking

8.2.1.1 Interpretations for Risk Category Ranking
On applying Kendall’s test of Concordance on the rank order data of risk categories it was observed that the results were statistically significant. ‘Group Awareness’ risk category was ranked as highest, since the category deals with problems arising due to lack of communication and collaboration amongst the teams and the client which goes in contrast with agile practices of frequent and close communication. ‘External Stakeholder Collaboration’ comes after Group Awareness as it encompasses risks due to lack of coordination between team and the customers and other vendors /third parties. This can also be attributed to spatial and temporal distance. After this, we have ‘Software Development Life Cycle’ which consists of risks due to inability to effectively do various engineering activities, involving frequent interaction between project stakeholders, due to their physical separation. To some extent, temporal distance and work /development culture difference also causes these types of problems. The next risk category ‘Project Management’ is dealing with project planning and team management. ‘Technology Setup’ considers problems due to communication infrastructure and tool selection. These two categories are ranked low due to maturity in the development processes and infrastructure facilities required to support DAD in most of the organizations, but the risks are again caused predominantly due to spatial distance.

8.2.1.2 Interpretations for Risk Area Ranking
We observe in the risk category, Group Awareness, the risk areas are ordered as Communication, Coordination and Collaboration, Trust. Lack of communication and ineffective coordination amongst the DAD teams happens primarily due to spatial and temporal distance. The risk category ‘External Stakeholder Collaboration’ contains risk areas which deal with risk arising due to lack of coordination between the development team and other stakeholders like customer, vendors, and third parties. The risk area ‘Customer Collaboration’ is ranked higher, since it deals with unavailability of the customer or misalignment of customers with agile practices due to spatial distance. The other risk area, Multiple Vendor Involvement, is ranked comparatively low. It has issues related to coordination problems between team and the vendors which attributed to difference in work/development culture and to some extent spatial distance also. Software
development activities like requirement engineering, release management, testing and integration are placed above other areas like design, coding and integration and standards of agile ceremonies. All the higher ranked risk areas in the ‘Software Development Life Cycle’ risk category are the ones which need frequent and effective communication. Further, we observe that risk areas in ‘Project Management’, namely, ‘Project Organization’, ‘Project Planning and Control’, ‘Project Initiation’ and ‘Acquisition and Development of Project Team’ have almost similar ranks. All of these are related to spatial distance and few risks due to difference in work development culture. Risk areas in Technology setup, which are related to unavailability of communication infrastructure and tool selection are also related directly to spatial distance, although have received low ranks.

8.2.1.3 Interpretation for Risk Factor Ranking
We analyzed the DAD risk factors and observed that most of the important risk factors are caused due to spatial distance between the stakeholders. In DAD projects, lack of communication, collaboration and coordination leads to major risks, which are attributed to spatial and temporal distance. The team finds it difficult to collaborate with the customer due to physical separation which causes poor understanding of the project vision. Majority of risk factors occurring due to spatial distance relate to poor requirement elicitation, difficulty in testing, difficulty in pair programming, unresolved technical debts and ineffective agile ceremonies which are core practices in DAD projects.

Temporal distance is also leading to important risks in DAD projects. Large difference in time zone between the team members lead to lack of collaboration and coordination. Moreover, it also causes obstacles in testing and code integration which further causes delayed releases.

Some risks are caused due to differences in work/development culture of multiple teams. These risks factors are related to use of different engineering practices, different agile practices, and different tools by multiple teams. Few risks are also caused due to uncommon language of distributed teams which also lead to ineffective communication which can cause serious trouble for DAD projects. Large project size leads to few risk factors which are related to difficulty in product backlog management, inadequate prioritization and difficulty in testing and integration of software projects.

The DAD risk factors in each risk category (and in relevant risk areas) have been segregated based on the DSD property which is causing them and are shown in Figure 2. In the figure, the significant risk factors are bold-faced and the insignificant ones have been marked with the label (ISF). Further, the highly significant risk areas (observed p-value <= 0.05) are marked with ‘***’ and the less significant risk areas are marked with ‘*’ (observed p-value >0.05 and < 0.1). The insignificant risk areas are not marked with any symbol. All the risk categories were highly significant, hence all have been marked with ‘***’.
8.3 Interpretation: Risk Management Methods for DAD risks
For all the DAD risk factors, we identified the risk management methods which are used more frequently and those which are used less frequently. We obtained important insights about the best practices which can be used by the DAD teams to reduce the risks.

The practices suggested by this study, focuses on reducing the communication gap between the teams and increased interaction between the teams and the customer for product development. We suggest that enhancing collaboration between the teams and even other stakeholders like, vendors, third parties will lead to better results. In case of multiple customers or vendors, having a single person to take the responsibility and interacting with the team would
Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)

help. Following good engineering practices, like continuous testing and integration, pair programming, backlog management, sharing a common definition of done becomes absolutely necessary for DAD teams. Regular team meetings like scrum-of-scums, daily stand-ups and periodic collocation of teams help in removing the impediments in project success. Distribute teams should be preferably long-lived feature teams with minimum inter-dependencies, which helps in having increases trust, team-spirit and productivity. DAD teams must be supported with dedicated business analyst and they should be provided with appropriate communication infrastructure and tools. The recommendations for risk management are based on the views of the practitioners and hence may not represent the ideal set of practices for managing the identified risk.

8.4 Interpretation: Validation
We had provided the partial framework to the three companies for implementation in real time DAD projects. The difference in the risk magnitude was found before and after implementation of the framework to assess the benefit of the framework. The results and the reviews obtained from the validation shows that, the use of partial framework helped the DAD teams to foresee the risks which may impact the success of their respective projects and hence perform a better risk management. Significant number of risk factors contained in the risk management framework, which were provided to them, was perceived as important by the companies. Although, all the companies have their own risk management methods, the suggested practices were also useful for controlling the risks. Hence, we claim that if the complete implementation of the framework is done for a significant time during the execution of a DAD project, it would help the team to do effective risk management.

9. Conclusion and Recommendations
Software organization are using distributed agile development (DAD) to develop quality solutions in limited time, while making best use of globally available talent. However, these benefits come with various risks like reduced likelihood of project success, increased delivery time, reduced team performance and increased dysfunction (Miller 2008).

This study explores the risk factors impacting DAD project success and the most suitable risk management approaches to control them. The related risk factors were grouped to form risk areas and further broad risk categories, namely, Group Awareness, External Stakeholder Collaboration, Software development Life Cycle, Project Management and Technology Setup. These risk categories were then related to the Leavitt’s model of organizational change, in order to give an organizational perspective to our findings. The study further identified the risk factors which are most significant for DAD projects and the risk management approaches which are most frequently used in practice.

This study discusses about the conflict in the inherent properties of DSD, namely, spatial distance (geographic separation), temporal distance (time-zone differences), language barrier, work/development culture and large project scope and the principles of agile like face-to-face communication, short iteration, continuous integration, customer involvement and others. Most of the important risks in DAD projects are caused due to the spatial distance between the stakeholders. There are certain important risks occurring due to temporal distance and difference in work/development culture. Very few risks occur due to language difference between the teams and large scope of the DAD project. Based on the research findings obtained from the
Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)

practitioner’s perspective, we suggest the best practices which can be adopted by DAD teams to manage the risks and challenges. Further, the study also suggests those risk factors which impact the ‘Time’ of project delivery, those which impact ‘Cost’ and those which impact the ‘Quality’ of the project separately.

We contend that the findings of this study have rendered some very important insights on DAD projects. These insights will help the practitioners to deal with the complexity which is introduced in DAD projects due to the inherent properties of software development, distributed development, agile methods and the interplay of these properties. The findings of the study will help DAD teams to anticipate the risks which may become a threat to their project and take appropriate measure to control those risks by referring the suggested risk management methods for those risks. Traditional view of managing risks will be inapt to deal with this complexity and so, software engineering community has a mammoth work to understand, organize and formulate the mechanisms that can rule over this new world of software development.

10. Limitations of the study

Although, we have taken measures to ensure that the results of the study are reliable and dependable, but certain conditions while carrying the study have posed threats to the validity of the results.

Extensive exploratory study was conducted which involved carrying in-depth interviews and analysis of the project work documents. It was difficult to take verbatim notes, while having simultaneous discussions with the respondents. This problem was overcome by contacting the respondents again to fill up the missing or unclear information. Project work document analysis was difficult to conduct, since the respondents were unwilling to provide us the documents of projects due to confidentiality constraints. Explicit efforts were taken by us to convince the respondents to provide us with the relevant documents. Except for few projects (2), the documents corresponding to all the other projects were viewed and analyzed at the company’s location. Collection and analysis of large amount of qualitative data in limited amount of time has its inherent complexity.

In case of quantitative study, the length of the questionnaire posed major difficulty for data collection. We had to split the data collection in two phases, due to which more time and efforts was spent in collecting quantitative data. The respondents, who had helped us multiple times for this study, were reluctant to fill the questionnaire in two phases. This caused less response rate for the second part of the quantitative study, which further became the reason for not getting significant results for rating for risk management methods. Collecting the data pertaining to project constraint was also difficult because, the respondents were unable to spot one project constraint for a particular risk. We asked the respondents to give us multiple project constraint for a particular risk factor if they are unable to select on particular constraint from the given ones. In such cases, we asked the respondent to select one of the constraints, which is relatively more important. Overall, due to the length of the questionnaire, there were many cases when the data was missing, and we had to contact the respondents multiple times for collecting the missing data.

The partial risk management framework was given to the companies for validation. It took significant efforts to get the framework implemented on real projects as the companies were bounded by contractual restriction with the client. We took substantial efforts to convince the company officials for getting the validation done in limited time.
11. Future Scope

Research on risks in DAD projects is in its nascent stage. Present study contributed by providing significant insights on the risks which could impact the success of DAD project and approaches which could be used to control them. There were certain risk factors which were identified during exploration, but we could not find significant amount of support from the literature. These factors have been separated in our study and the researchers can consider such factors for further exploration. We have explored the risks which occur due to the lack of understanding of the vision of the project with the development teams. These risks were grouped under the risk category ‘Business Goals and Objectives’. The findings of this study were statistically insignificant; hence it can be considered for more in-depth analysis by the future researchers. The risk management methods suggested in this study for controlling the identified risks in DAD projects, is based on their frequency of use in practice, and may not represent the ideal solutions. A detailed study of risk management methods can be done to provide solutions, which are ideal with respect to agile practices and are suitable for implementation, such that they can help in effective risk management for DAD projects. We have studied the impact of risk factors on three project constraints, namely, time, cost and quality. Besides, the constraints considered in the study, other constraints like ‘Business Value’, ‘Organizational Culture’ are also impacted with risks in DAD projects occur. Future studies can explore the impact of the risks in DAD projects on such constraints as well.

12. Research paper Publication

Following are research publications in the area of research

1. **Title: Distributed Agile Software Development: A Review**
   Authors: Suprika V Shrivastava and Hema Date
   Journal of Computer Science and Engineering, Volume 1, Issue 1, May 2010
   Cited by: 12

2. **Title: A Framework for Risk Management in Globally Distributed Agile Software Development (Agile GSD)**
   Authors: Suprika V Shrivastava and Hema Date
   Interscience Management Review (IMR), Volume 2, issue 1, 2010
   http://interscience.in/IMR_Vo2,No1/Paper_6.pdf

3. **Title: Categorization of Risks Factors for Distributed Agile Projects**
   Sent for review to the Journal of Information and Organization (Elsevier) on December 30, 2013 and are awaiting their response.
   The same paper was sent to the Empirical Software Engineering Journal (Springer) on March 29, 2013 and had undergone three major revisions till finally rejected on December 15, 2013 for being lengthy. The editor and the reviewer appreciated the rigor of the work and the research methodology. According to their views, the work is significant but it needs to be presented in more concise way.
Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)

References

Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)


Submitted by: Mrs. Suprika Vasudeva Shrivastava (PRN: 10039001006)