10-1-2010

The Effects of IT, Task, Workgroup, and Knowledge Factors on Workgroup Outcomes: A Longitudinal Investigation

Srikanth Mudigonda
University of Missouri-St. Louis, spmn9b@umsl.edu

The Effects of IT, Task, Workgroup, and Knowledge Factors on Workgroup Outcomes: A Longitudinal Investigation

Srikanth Mudigonda

M.S. Information System, Dakota State University, 2003.

M.Sc.(Hons.) Biological Sciences and Master of Management Studies, Birla Institute of Technology and Science, Pilani, India, 1999.

A Thesis Submitted to The Graduate School at the University of Missouri – St. Louis in partial fulfillment of the requirements for the degree

Ph.D. in Business Administration with an Emphasis in Information Systems

September, 2010.

Advisory Committee:

Dr. Rajiv Sabherwal
(Advisor)

Dr. Mary Lacity

Dr. Haim Mano

Dr. Ashok Subramanian

Dr. Vicki Sauter

Acknowledgments

I am greatly indebted to my dissertation advisor, Dr. Rajiv Sabherwal, for his patient guidance throughout my doctoral studies. His advice and words of encouragement have helped me overcome the hurdles I faced, particularly during my dissertation stage. I am also grateful to members of my dissertation committee, Dr. Mary Lacity, Dr. Haim Mano, Dr. Ashok Subramanian and Dr. Vicki Sauter, for their generosity with their time, advice and encouragement. I am also grateful to members of the IS faculty: Dr. Marius Janson, Dr. Kailash Joshi, Dr. Dinesh Mirchandani and Dr. Joseph Rottman, and members of the Marketing faculty: Dr. Michael Elliott and Dr. Paul Speck for their help and advice during my tenure at UMSL. Through their words and actions, my professors have molded my approach to research and teaching. For that, I shall remain eternally grateful.

I am thankful to all my doctoral colleagues, past and present: Anand Jeyaraj, Charles Chowa, Danai Tsotra, Michael Griggs, Jay Wang, Jian Wang, Chris Kang, Dennis Shi, Hui Zhou, Denise Franke, Vidya Iyer, Aihua Yan, Kyootai Lee, Shaji Khan, R.S. Prasad, Daniel Hollander, Mike Cornish, Robert Talbott, Rooji Sugathan, Stanislaus Simon-Solomon, Kathy Ntlaja and Svetlana Krasteva. Their help during my coursework and camaraderie have helped me grow as a person and a scholar. In particular, I am grateful to Chris Kang for all his help and advice over the years, that went way beyond what is expected of a friend.

I am grateful to Dr. Judith Walker de Felix, Ms. Elizabeth Sampson, Ms. Mary Ellen Heckel, Ms. Clara Jackson and Ms. Nykia Watts, whose help and cooperation made my job as a graduate assistant enjoyable and a great learning experience.

I am grateful to Dr. Keith Womer, Dr. Thomas Eysell, Mr. Karl Kottemann and their staff for all the facilities and help they have been providing to doctoral students, making our work and study at UMSL a pleasure and an enjoyable and worthy learning experience.

My special thanks to Arvind Mani and Ranjith Boggarapu for their friendship and encouragement over the years.

Finally, I would like to express my deepest gratitude to my family.

Abstract

Organizations work towards achieving their goals by integrating and utilizing the knowledge available within their boundaries. In order to successfully manage the knowledge-related processes occurring in their workgroups, organizations need to understand how different contingency factors affect the knowledge-related processes of a workgroup, ultimately affecting the workgroup's knowledge outcomes and performance. Knowledge processes, by their nature, are dynamic, time-dependent. A review of extant literature revealed a gap: few studies exist that studied the research question, using a longitudinal methodology. Hence to obtain a deeper understanding of the longitudinal effects of different contingency factors on knowledge outcomes and performance of workgroups and, consequently, to contribute to the literature in this area, this dissertation was conducted. Specifically, this dissertation investigated the longitudinal effects of contingency factors that were grouped into five categories, on three outcome variables, via workgroup processes that were grouped into three categories. The research question, which combines the above aspects of the investigation and guided this dissertation is:

Which factors, from the five categories of factors (a) characteristics of the workgroup; (b) characteristics of the tasks assigned to the workgroup; (c) the interface between the workgroup and the tasks; (d) characteristics of the knowledge required to complete the tasks; and (e) characteristics of the information technologies, affect workgroup outcomes, including (i) average consensus among a workgroup's members about each other's areas of knowledge; (ii) average accuracy of knowledge; and (iii) performance of the workgroup, over time, and in what way?

These workgroup processes included in the study can be categorized into three groups: *processes related to scheduling of tasks*, *processes related to completion of tasks* and *processes accompanying those related to completion of tasks*.

An agent-based model, that was derived from findings and theory drawn from extant literature was used in this investigation. Key aspects of the model were validated using data obtained from a series of four qualitative, semi-structured interviews. The results of simulations of the agent-based model were analyzed using the methodology of panel data analysis.

The results indicate that only a subset of contingency factors from each category affect each of the workgroup outcomes. Specifically, *average task priority*, *average knowledge-intensity of subtasks*, *average propensity to share*, *time in training phase*, *probability of non-specific exchange*, *number of agents*, *number of locations* and *average project intensity* were found to have a positive effect on average consensus, while *average task intensity*, *average self-knowledge* and *average number of tasks per agent* had negative effect on *average consensus*. In the case of *average accuracy of knowledge*, *average knowledge level* and *number of agents* were found to have a positive significant effect. Finally, in the case of *percentage of project completed*, *average propensity to share*, *average knowledge level*, *average self-knowledge*, and *time in training phase* were found to have a positive significant effect, while *average knowledge intensity of subtasks*, *richness of email*, and *average direction time* were found to have a negative significant effect. *Average number of tasks per agent* was found to have a significant negative effect between workgroups and positive significant effect within workgroups.

The dissertation contributes to literature by describing the simultaneous, longitudinal effects of a large set of contingency factors on the outcome variables and, of those, identifying those that are have a significant longitudinal effect on the workgroup outcomes. By doing so, it provides a shortlist of contingent factors that could be used in future empirical, confirmatory studies. Additionally, the specifications of the agent-based model and the accompanying source code provide a basis for future work that can explore workgroup-related phenomena in greater depth. For practitioners, the dissertation offers recommendations regarding the factors on which they should focus to increase the likelihood of favorable workgroup outcomes. It also helps them identify those contingency factors whose negative effects on the workgroup outcomes can be mitigated through appropriate policies and procedures.

Table of Contents

Table of Tables and Appendices

Table of Figures

1. Introduction

Organizations work towards achieving their goals by integrating and utilizing the knowledge available within their boundaries (Grant, 1996a). Organizational knowledge may be dispersed across various knowledge artifacts, and, more importantly, among their employees who work together as members of workgroups. In the present day organizations, three major phenomena are at work, affecting the ability of organizations to reach their goals. The first phenomenon is the increased transience of workforce, particularly, the relatively shorter durations of tenures of younger members of the workforce (Bureau of Labor Statistics, 2008). This implies a corresponding increase in the transience, or, stated differently, a decrease in the continued relevance, of the knowledge available to organizations. Additionally, shorter tenures of employees imply that there would be fewer instances of employees passing on their hard-earned knowledge to their less-knowledgeable and less-experienced colleagues. The second phenomenon is the ever-increasing amount of data being produced within and outside organizations. In order to survive and exploit the opportunities that arise, organizations require a commensurate amount of knowledge to process the data into information, make sense of it and use it. The third phenomenon is the sharing of knowledge between individuals, verbally, through socialization, without a corresponding capture of the knowledge in an explicit form. Given these phenomena, it is essential to understand how different contingency factors affect the sharing of knowledge knowledge and, consequently, performance, among a group of employees in an organization. Such an understanding would allow for a better management of the contingency factors that affect the processes involved in the sharing of knowledge and it can be ensured that those who need knowledge are able to obtain it from their colleagues in a timely manner.

Expressed in the form of a question, the motivation for this dissertation is, “Which contingency factors affect a workgroup's knowledge-related outcomes over time, and in what way?” To the answer the question, it is useful to view organizations as “social communities that specialize in the creation and internal transfer of knowledge” (Kogut and Zander, 2003). While employees of an organization may work as part of workgroups, the overall structure of an organization results from decisions that can be a combination of top-down, and organic, emergent decision making (Robey, 1991). The employees, who are spread throughout the organization, working in different departments, functional areas or workgroups, can be seen as forming the basic cognitive components of an organization. Therefore, organizations can be seen as being involved in distributed processing of information, and in the creation and exchange of knowledge (Tsoukas, 1996; Ackoff, 1981). An organization's structure affects the flow of information and exchange of knowledge among its employees, thereby affecting the overall performance of the organization (Dalton et al. 1980). Hence, while investigating the effects of various contingency factors that affect a workgroup's outcomes over time, attention must be paid to the structure of the workgroup and its processes pertaining to the exchange of knowledge that would help it reach its assigned goals.

Within a workgroup, employees who lack the knowledge that is required to complete an assigned task would contact other employees from whom they can obtain the required knowledge. Their choice of colleagues who can act as knowledge sources is based on what they know about their colleagues' areas of knowledge. This notion of a group-level memory of who knows what is captured by the construct of transactive memory (Brandon and Hollingshead, 2004). The success of an employee in completing a task assigned to him/her depends on whether the employee possesses the required knowledge or is able to locate and obtain the required knowledge from a colleague who would act as a knowledge source. Consequently, the success of a workgroup, across time, is predicated on the successful completion of the tasks assigned to each member of the workgroup. Thus, it is essential to study how different processes of a workgroup lead to outcomes of knowledge accuracy, transactive memory and workgroup performance as a function of different contingency factors over time.

Some of the contingency factors affecting the likelihood of a workgroup member completing her/his assigned tasks include the characteristics of the tasks themselves, their inter-dependencies (Steiner, 1972) and the fit between the tasks and the employees' knowledge (von Hippel, 1994). Additionally, the use of appropriate media for communications that lead to exchange of knowledge is also an important contingency factor. When face-to-face communication is not possible, employees working together on a project use appropriate information technologies for communication to ensure that their exchanges related to knowledge have minimal equivocality (Daft and Lengel, 1986). Thus, a workgroup's choice of information technologies that facilitate communication, and its policies associated with employees' training, and the usage of the information technologies, tend to affect the effectiveness of communications and, ultimately, the success of a project assigned to the workgroup. In addition to the characteristics of a workgroup of employees working together on a project, its processes related to exchange of knowledge and information about knowledge, and the use of information technologies, successful completion of a project depends on the characteristics of the knowledge that affects the ease with which it can be transferred (Szulanski, 1996).

Using the contingency factors that were described in the previous paragraphs as a basis, the research question can be restated, in more specific terms, to define the scope of the current dissertation. The identification of specific categories of contingency factors, processes and outcomes, and their incorporation into a more specific and detailed research question is presented next.

1.1. Research purpose

In the context of the three phenomena described earlier, it is important to understand how, over time, different contingency factors affect the evolution of knowledge, transactive memory at the workgroup-level, and in addition to affecting the performance of a workgroup. Such an understanding would help the decision-makers and members of a workgroup to manage their workgroup processes better, thereby leading to a greater possibility of positive outcomes. A survey of extant literature revealed that few studies exist that studied the research question, using a longitudinal methodology.[1] Thus, the motivation of this dissertation was to provide a deeper understanding of how the three outcomes of interest evolve over time, as a function of contingency factors and workgroup processes. One consequence of such an investigation is the contribution to the literature in this area of research. The approach chosen was exploratory in nature, since the goal was not to test hypotheses that could be developed from prior literature, but to obtain findings that would form a basis for one or more empirical studies that would use a confirmatory research

1 Lewis (2004) was a longitudinal study that investigated how transactive memory evolves over time and affects performance of the workgroup across time. However, the set of contingency factors included in that study is a subset of the set of contingency factors (identified and described in chapter 2) considered here. Additionally, the longitudinal nature of the study was restricted to one project, whereas this dissertation looked at the outcomes across a series of projects.

approach. An additional goal of this dissertation is to assist practitioners by helping them identify those contingency factors that are within their control and can be modulated such that the outcomes of the workgroups that they supervise, or of which they are a member, can be enhanced.

The contingency factors that were considered relevant were based on a review of extant literature and are grouped into the following five categories: (a) characteristics of a workgroup; (b) characteristics of the tasks assigned to the workgroup; (c) the interface between a workgroup and its tasks; (d) the characteristics of knowledge; and, (e) characteristics of information technologies. The organizational processes are categorized as: (i) processes related to the completion of tasks, (ii) processes accompanying those processes that are related to the completion of tasks, and (iii) processes related to scheduling of tasks The outcome measures of interest are: (1) average consensus among a workgroup's members about each other's areas of knowledge, which indicates the transactive memory of the workgroup; (2) aggregate knowledge level of the workgroup, indicated by average accuracy of knowledge; and, (3) performance of the workgroup, measured in terms of the percentage of project completed. Detailed descriptions of the contingency factors, the processes and the outcomes are provided in chapter 2. Their operationalizations are described in chapter 3.

By using the specific categories of contingency factors and the three outcomes of interest, the research question can be restated in more specific terms as follows:

Research question: Which factors, from the five categories of factors (a) characteristics of the workgroup; (b) characteristics of the tasks assigned to the workgroup; (c) the interface between the workgroup and the tasks; (d) characteristics of the knowledge required to complete the tasks; and (e) characteristics of the information technologies, affect workgroup outcomes, including (i) average consensus among a workgroup's members about each other's areas of knowledge; (ii) average accuracy of knowledge; and (iii) performance of the workgroup, over time, and in what way?

1.2. Outline of the dissertation

The remainder of this dissertation explains how the categories of contingency factors and the outcomes, which were mentioned in the research question, were operationalized, how the research question was answered and what the results imply. The contributions of this dissertation, its limitations, its implications for practice and potential directions for future research are also described.

Chapter 2, *Model development*, summarizes the literature that informs the inclusion of the contingency factors, workgroup processes and outcome measures used in this study. Chapter 3, *Methodology*, describes details of the methodology used, which involved the implementation of an agent-based model whose specifications were derived from extant literature and partially validated through support that was obtained through a series of semi-structured interviews. The details of the panel data analysis method, which was used to analyze the data obtained through the simulation of the agent-based model, is also described in the chapter. Chapter 4, *Results*, describes the results and presents a preliminary interpretation of the the results that were obtained via panel data analysis regressions of the data that are outputs of the simulation. Finally, in chapter 5 titled *Discussion*, presents detailed analyses of the results and how they relate to various empirical findings reported in current literature. The chapter also describes the contributions made by this dissertation to the literature and its implications to practitioners. Limitations that define the context in which the results should be understood and interpreted are also identified. The chapter then identifies potential directions for future research and concludes by summarizing the study.

1.3. Summary

In this chapter, three main issues, viz., an increasingly-transient workforce, ever-increasing levels of data, and the associated increase in the knowledge required to process the data, and the large proportion of knowledge staying in individual's minds, rather than being converted into knowledge artifacts, were identified and how these three phenomena are affecting an organization's ability to successfully operate in its environment was described. It was explained that these issues provide the context for this dissertation. Then, it was explained that current literature does not fully address the question of how, over a period of time, different contingency factors affect knowledge outcomes and performance of a workgroup. The need to understand the longitudinal effects of the contingency factors on the outcomes of interest and the gap in extant literature were explained as the motivations for this dissertation, which was then formally described in the form of a research question. It was also explained that this dissertation is exploratory in its approach, given a paucity of literature that is based on longitudinal investigations of the effects of contingency factors. Finally, the outline of the organization of this dissertation was presented.

In the next chapter, Development of the model, first, detailed descriptions of the concepts associated with knowledge and transactive memory are presented. This is followed by summaries of extant literature that informed the choice of the contingency factors, workgroup processes and outcomes that were included in this dissertation. The chapter concludes with a summary of its contents and a brief overview of the subsequent chapter.

2. Development of the model

As described in chapter 1, *Introduction*, there are three phenomena affecting the effectiveness of an organization's knowledge over time: (1) an increasingly-transient workforce; (2) ever-increasing levels of data, and the associated increase in the knowledge required to process the data; and, (3) the large proportion of knowledge staying in individual's minds, rather than being converted into knowledge artifacts. These three phenomena imply that each individual is less likely to possess all the knowledge he/she needs to complete the tasks assigned to him, and, consequently, has to rely on his/her colleagues for help. Hence, it is important to understand how different contingency factors affect the processes of (a) how information about who knows what in a workgroup is shared among members of the workgroup; (b) how knowledge is shared among members of the workgroup; and, (c) how these contingency factors also affect the workgroup's performance as a whole. The motivation of this dissertation was to provide the above-described understanding. Specifically, this dissertation investigated the longitudinal effects of contingency factors that were grouped into five categories (1) *characteristics of workgroup*, (2) *characteristics of task*, (3) *interface between the workgroup and the tasks assigned to it*, (4) *characteristics of knowledge,* and, (5) *characteristics of information technologies* on the three outcomes (i) *average accuracy of knowledge*, (ii) *average consensus*, and (iii) *percentage of project completed*. These effects were investigated via a set of workgroup processes that were grouped into three categories: (1) *processes related to scheduling of tasks*, (2) *processes related to completion of tasks* and (3) *processes accompanying those related to completion of tasks*.

This chapter summarizes the literature related to the three outcomes of interest and the contingency factors that affect them. A summary of the studies reviewed while developing support for this dissertation is presented in appendix A1. It also explains how the findings from literature associated with knowledge, transactional memory and workgroup performance, have informed the development of an agent-based model that was used to answer the research question. In the remainder of this chapter, first, the concepts associated with knowledge and processes related to knowledge are presented in sections 2.1 – 2.3. Then, sections 2.4 and 2.5 explain the concepts associated with transactive memory and transactive memory systems. Subsequently, sections 2.6 – 2.8

summarize studies that describe the relationships between contingency factors and outcome variables and the associated workgroup processes that were considered in this dissertation. Finally, section 2.9 presents a summary of this chapter and a brief overview of the next chapter.

2.1. Multiple views of knowledge

The multiplicity of units and levels of analysis at which knowledge has been studied appears to arise from a need to link individual knowledge and organizational knowledge, on the one hand, to the performance of individuals within the organization, and the overall performance of the organization, on the other hand. Knowledge in organizations can be seen as existing at multiple levels, based on how it is aggregated. In their review of extant literature pertaining to knowledge management that encompasses individual and organizational knowledge, Alavi and Leidner (2001) observe that knowledge is viewed by various researchers as being personalized information that is more than data and information; as a state of mind, which has both individual and collective (that is, organization-level) components; as an object that can be created, stored and manipulated; as a process of applying of expertise that various individuals in an organization possess; as being access to information; and, as a capability that has a potential for action. Other views of knowledge include the distinction made between declarative and procedural knowledge at the individual level. This distinction may be described as the distinction between know-what and know-how (Singley and Anderson, 1989; Kogut and Zander, 1992). Despite this diversity in the definitions of knowledge and perspectives on it, in the literature, the consensus is that knowledge is a valuable resource that can provide competitive advantage to organizations that obtain, create and apply it.

Next, descriptions of knowledge, as it is considered at different levels of analysis, are presented.

2.2. Knowledge at multiple levels

Depending on the level at which it is operationalized and applied in theorizing in the literature, knowledge is used as a construct at various levels of analysis. At the individual level, knowledge has been defined as a set of justified true beliefs (Nonaka, 1994; Becerra-Fernandez, et al., 2004). Nonaka explains that knowledge is dynamic: it is evaluated continually in the light of new experiences (Nonaka, 1994, pg 15). When new information, which is received from the environment, supports a person's beliefs about something, those beliefs about that something are reinforced. In contrast, if the new information received is contradictory, then the existing knowledge is brought into question, causing the person to re-evaluate it to decide whether the knowledge continues to be justifiable and whether it needs revision. Additionally, at the individual level, scholars examining knowledge have discussed knowledge as being tacit or explicit (e.g., Polanyi, 1962; Nonaka et al., 2000; Nonaka and Toyama, 2003); teachable or codifiable (Zander and Kogut, 1995); general or specific (Hayek, 1945; Jensen and Meckling, 1996); situation-specific (O'Reilly and Pondy, 1979; Hayek, 1945), technology-specific (Choudhary and Sampler, 1997) or both (Sabherwal and Becerra-Fernandez, 2005). The aspects of knowledge that are mentioned above may be seen as attributes of knowledge, that is, characteristics of knowledge that would affect the ease with which individuals can understand it, explain it to others, or draw upon it to provide instructions to others.

Individual knowledge can be aggregated at various levels in an organization to indicate the potential of the specific collective, such as a workgroup, a department, or the organization as a whole, to achieve its goals through an application of the knowledge. Scholars studying organizations explain that, as an aggregate, organizational knowledge is a valuable resource that confers dynamic capabilities to firms (Kogut and Zander, 1992). At an aggregate level, knowledge may be encapsulated in collections of artifacts such as documents, memos, software, policies and procedures, in an explicit form and remain in the minds of organizational employees in a tacit form (Alavi and Leidner, 2001). However, it is the individuals that learn; organizations (or sub-units such as workgroups or departments) "learn only through the learning of its members" (Simon, 1991, pg 125).

Given the importance that has been assigned to organizational knowledge in the literature[2], it is essential to understand the processes that are associated with its creation, storage, dissemination and application.

Based on whether they describe knowledge held by individuals who can readily

2 As seen in the 37,500 citations that resulted from a search using "organizational knowledge" as the key phrase on Google Scholar on July 5, 2010. http://scholar.google.com/scholar?hl=en&q="organizational+knowledge"

apply or share with others, or whether they describe knowledge (beliefs or perceptions) that individuals have of others' areas of knowledge, two key aspects of memory of individuals need to be recognized. They are: (a) knowledge possessed by an individual, which allows the individual to complete the tasks assigned to the individual, and (b) knowledge that the individual has about others' areas of knowledge, which would lead the individual to seek knowledge from others in case the individual lacks knowledge in a required area. This knowledge of others' areas of knowledge, when aggregated at the workgroup level, is described as transactive memory (Brandon and Hollingshead, 2004).

There are four types of knowledge processes – knowledge discovery, knowledge capture, knowledge sharing and knowledge application – that correspond to the aspect (a) described above. These are described in section 2.3. Then, sections 2.4 and 2.5 describe aspect (b), transactive memory, and the processes related to it.

2.3. Knowledge-related processes

The knowledge-related processes in an organization can be categorized as (a) discovery-related processes, (b) capture-related processes, (c) sharing-related processes, and (d) application-related processes (Grant 1996a, 1996b; Nahapiet and Ghoshal, 1998; Becerra-Fernandez et al., 2004, pp. 32-35). This section describes each of these categories.

The processes categorized as *knowledge discovery* consist of *combination* and *socialization* activities. *Combination* activities are those where an individual synthesizes multiple bodies of explicit knowledge to create new, more complex sets of explicit knowledge (Nonaka, 1994). Such activities could result in an incremental or a radical change to the knowledge-base of an individual (Nahapiet and Ghoshal, 1998). *Socialization* is the synthesis of tacit knowledge by individuals while working together on common activities (Becerra-Fernandez et al., 2004). Examples of *socialization* include apprenticeships, and training where new-comers to an organization or a workgroup learn about how others think and generate new ideas and knowledge.

The processes categorized as *knowledge capture* consist of *internalization* and *externalization* activities. *Internalization* is the conversion of explicitly coded knowledge into tacit knowledge by an individual. An example of internalization is the situation where an individual learns about a specific new area of knowledge in his field of interest and assimilates it; this process of assimilation is facilitated by what the individual already knows in that field (Nonaka and Takeuchi, 1995). *Externalization* is the conversion of an individual's tacit knowledge into explicit form, such as verbalizing or documenting one's knowledge about a particular topic (Nonaka and Takeuchi, 1995).

Knowledge sharing processes consist of *exchange* activities, which involve the sharing of knowledge in explicit form by two or more individuals (Grant, 1996b). In addition to being part of the category *knowledge creation*, *socialization* is described as being a sharing process as well. During socialization, individuals learn knowledge that is in tacit form from one another.

Knowledge application processes consist of *direction* and *routine* activities. *Direction* is the process by which an individual who possesses knowledge in a certain area directs the actions of another individual, without transferring the underlying knowledge to the directed individual (Grant, 1996a). *Routines* are the procedures, rules and norms that guide individuals' behavior in an organization. Routines develop over time, and require constant repetition. They economize communication more than direction-related activities because the routines needed to complete a task are embedded in the procedures and the technologies (Grant, 1996a).

2.4. Transactive memory

In the previous sections, concepts related to knowledge, how it is viewed in the literature, in terms of various attributes and aggregation at different levels was presented. Additionally, the processes associated with knowledge were also presented. Complementary to the knowledge that helps them complete their work, individuals in organizations, also need knowledge about others' areas of knowledge. This complementary knowledge, when aggregated among a group of individuals, is known as transactive memory. This section explains the concept of transactive memory and its associated concept of transactive memory system.

A workgroup's transactive memory is a distributed knowledge-base that is internal to each member. It consists of beliefs that members of a workgroup develop about each other's areas of knowledge. Re-evaluations of existing transactive memory within each member of a workgroup happen as a result of information that is obtained through communication and observations of how other members perform their tasks (Brandon and Hollingshead, 2004; Wegner, 1995). The knowledge-base also consists of each individual's unique knowledge that is distinct from the knowledge held by other members of the workgroup. Transactive memory is so called because it develops via *transactions* among the group[3] members (Wegner, 1986). Just as knowledge in a project results from workgroup processes associated with its creation, encoding, storage, retrieval, sharing, etc., (as described in sections 2.2 and 2.3), so too does transactive memory – it results as a product of certain other workgroup processes. As mentioned in chapter 1 and earlier in this chapter, organizations are finding themselves operating in environments where the amount of data being produced and needs to be monitored is increasing rapidly. Their success in such environments is predicated on how well their employees can obtain new knowledge and apply their existing knowledge to solving problems. Complementary to the knowledge that individual employees have within themselves, is the knowledge that they have about their colleagues' areas of knowledge. The latter knowledge helps them identify the sources from whom they can seek help, when needed, or to whom they can direct new information, based on their perceptions of the sources' areas of knowledge. These two types of knowledge develop over time, via different collective-level processes. The latter type of knowledge is transactive memory. A workgroup's transactive memory, considered in conjunction with the processes associated with its creation, modification and retrieval is called a transactive memory system. The processes associated with transactive memory are described next.

2.5. Processes resulting in a transactive memory system

Associated with transactive memory, which is a product (a memory of colleagues' areas of knowledge) are processes that help in the development, modification and use of the product. This section discusses those processes. Wegner (1995) classified the group processes that result in the emergence of a transactive memory into the following

3 Wegner originally studied groups in general, and did not focus on groups within an organizational setting. Hence the more general term "group" is used here and later in the document instead of "workgroup", whenever the original studies did not involve a business setting.

categories: (a) directory updating, whereby people learn what others are likely to know; (b) information allocation, where new information is communicated to the person whose expertise will facilitate its storage; and (c) retrieval coordination, which is a plan for retrieving needed information on any topic based on knowledge of the relative expertise of the individuals in the memory system, i.e., a workgroup. Taken together, the three processes, along with the memory store, which is the knowledge-base described in section 2.4, constitute a functional transactive memory system (Hollingshead, 1998). Brandon and Hollingshead (2004, pg. 633) describe a transactive memory system as , “the basic idea is that people in relationships develop an implicit structure for assigning responsibility for information based on their shared conception of one another’s expertise. As a result, the cognitive burden on each individual member is reduced, yet a larger pool of information is available to each member than could be managed by any one person alone”.

The responsibility of managing new information that is obtained, and processing it based on prior knowledge, is either implicitly or explicitly assigned to the member of a workgroup who is believed to have the most amount of knowledge related to the new information (Anand et al., 1998; Hollingshead, 1998a; Hollinghead 2000; Moreland and Myaskovsky, 2000). Thus, knowing who knows what allows a workgroup's members to share the cognitive burden associated with processing new information and learning new knowledge, based on how well the new information and knowledge relate to the knowledge possessed by a given member. While doing so, members use the directory updating and information allocation processes described above. In situations where a member requires knowledge in area that he/she currently lacks knowledge, he/she is going to use the retrieval coordination process to determine, based on his/her beliefs about others' areas of knowledge, who is the best source of knowledge and seek the needed knowledge from that member.

In order to study the emergence of the workgroup's outcomes over time, both the knowledge processes described in section 2.3, and the processes that lead to the formation and development of transactive memory describe above, must be considered. In chapter 3, detailed descriptions of the subsets of processes belonging to the knowledge exchange and the transactive memory groups of processes, which are implemented in the current dissertation's agent-based model, are presented. Next, the contingency factors belonging to five categories, workgroup processes belonging to three categories, and three outcome variables are described. Justifications, drawn from prior literature, for their inclusion in the agent-based model, which forms the core of this dissertation, are presented as well.

2.6. Contingency factors affecting the knowledge outcomes and performance

A review of prior literature suggested several variables that were considered as contingency factors that affect the outcomes of interest, via processes followed by a workgroup. In the subsections of this section, the contingency factors[4] that were included in the agent-based model used in this dissertation are described in conjunction with justifications for their inclusion, which are based on findings reported in the literature. Then, in sections 2.7 and 2.8, workgroup processes that were implemented in the agent-based model, and the operationalizations of the outcome variables are presented. As mentioned earlier, in this dissertation, five categories of contingency factors were studied in terms of their effects on the outcome variables. Next, each of the five categories of contingency factors are discussed, beginning with the category, *characteristics of the workgroup.*

2.6.1. Characteristics of the workgroup

This category of contingency factors describes the various attributes of the members of a workgroup. The factors were identified based on findings reported in prior literature and are expected to have an effect on the workgroup's processes.

Cognitive-interdependence among the members of a workgroup results when the workgroup members lack knowledge in all the areas that are associated with the work assigned to the members (cf. Wegner, 1986). The following example illustrates the notion of cognitive-interdependence: consider a project with two tasks, assigned to a workgroup which consists of two agents[5].

- Task A, which was assigned to agent1 requires knowledge in areas KA1, KA2 and KA3

---

4 The specific simulation parameters that operationalize these factors are identified in table 3.5.

5 Here, and in the following sections, the term *agent* is used as a synonym of *a workgroup member*. As described in chapter 3, an agent is a computational representation of a workgroup member. Given that agent-based simulation is one of the methods used in this dissertation, the term agent is used in lieu of workgroup member throughout the document.

- Task 2, which was assigned to agent7, requires knowledge in areas KA2, KA4, and KA5
- Agent1 has knowledge in areas KA1, KA3 and KA4
- Agent7 has knowledge in areas KA2, KA5 and KA6

In the above-described situation, task1 can be completed only if agent1 acquires knowledge, or instructions that are based on the knowledge in area KA2 and would help in completing the task. Similarly, task2 can be completed only if agent7 acquires knowledge, or instructions that are based on the knowledge in area KA4, and would help in completing the task. Thus, the two agents have a cognitive-interdependence, because each agent has knowledge in an area required by the other. Hence, it is essential to investigate the effects of some of the attributes of a workgroup's members on the processes that lead to the development of transactive memory. Wegner et al. (1985) and Hollingshead (2001) describe that cognitive interdependence is a key requirement in the development of transactive memory systems in a group. Hence, it is essential to include its operationalization in the study. As can be inferred from the above example, the greater the number of areas in which an agent has knowledge, the lesser is its cognitive-dependence on other agents in a workgroup. At the workgroup level the contingency factor, *average knowledge level*, at the beginning of a simulation run, indicates the degree of cognitive-interdependence among the members of a workgroup.

The performance of a workgroup depends on the degree of skills and knowledge possessed by the members of the workgroup (Larson and Lafasto, 1989). The variation in the aggregate level of competence is captured in the contingency factor *average knowledge level*. The operationalization of this contingency factor is described in subsection 3.2.1.7.

Hollingshead (2000) reported that in workgroups, the members who have a more accurate perception of their own areas of knowledge are less likely to interact with other members for seeking knowledge that is needed for completing their own tasks. The finding implies that in such workgroups, members are less likely to interact with others than the members of workgroups whose members have less accurate perceptions of their own areas of knowledge. Hence, the level of self-awareness of agents in a workgroup is treated as a contingency factor and is operationalized via the contingency factor (aggregated at the workgroup-level) average self-knowledge. Its operationalization is described in subsection 3.2.1.6.

Wasko and Faraj (2005) report that members of an organization would contribute knowledge to others for several reasons. The cited reasons include enhancement of professional reputation, being part of a social network at their organization, and having an inherent motivation to share their knowledge and experiences with others. Prior work (Gray, 2001; Davenport and Prusak, 1998) also indicates that members of a workgroup may choose not to fully share one's areas of knowledge due to their fear of losing power or value in the workgroup. Hence, the current study included a contingency factor, *average propensity to share* (aggregated at the workgroup-level), which captures the likelihood that a workgroup's agent is going to share (a) its knowledge, when requested by an agent in the context of completing a task assigned to it that requires the requested knowledge, or, (b) information about the presence of knowledge with other members of the workgroup. Details of its operationalization are presented in subsection 3.2.1.8.

Hollingshead and Brandon (2003) describe the key role played by communication in the development of transactive memory in a group. Building on this finding, it can be argued that future communications would lead to the discovery of new information about others' areas of knowledge, and consequently to revision of one's perception of others' areas of knowledge. Furthermore, it is reasonable to assume that this individual-level tendency to modify one's perceptions of others' areas of knowledge would vary across individual agents in a workgroup. Additionally, in the case of workgroups, where agents obtain knowledge from other agents, modification of perceptions about others' propensity to share the requested knowledge or information that indicates presence of knowledge, can also vary. Hence, it is essential to operationalize an agent's openness to modify its perceptions of (1) other agents' propensity to share and, (2) the accuracy of other agents' knowledge in various areas. For the sake of simplicity, a single contingency factor, *average openness to change*, is assumed to operationalize both aspects described above. The operationalization of *average openness to change* is described in subsection 3.2.1.8.

Turnover in workgroups is reported to lead to both positive and negative outcomes. On the positive side, workgroups with new members replacing existing ones have an influx of new knowledge and ideas, thereby increasing the repertoire of knowledge and skills available, thereby increasing the workgroup's ability to handle a variety of tasks (Carley, 1992; March, 1991). Additionally, turnover involving the removal of ineffective members, or members whose presence in the workgroup may be detrimental to the workgroup's morale is also expected to have a positive effect on the workgroup. On the negative side, according to human capital theory (Strober, 1990), voluntary turnover would lead to a loss of skills that might prove to be critical to the workgroup's performance and may also affect the morale of the workgroup negatively.

In addition to the above effects, the entry of a newcomer into a workgroup implies that the remaining members of the workgroup would only be able to learn about the newcomer's areas of knowledge through interactions. Since there is no prior history of interactions, the newcomer would not be part of the workgroup's existing transactive memory system. For the above mentioned reasons, it is essential to study the effect of turnover on the processes and therefore on the outcomes. In the agent-based model, turnover is represented by two contingency factors: (1) *probability of turnover*, which determines the likelihood of a randomly-chosen member in a workgroup being replaced during a given time period, and, (2) *average proportion of knowledge areas common with the replaced agent*, which indicates the proportion of areas of knowledge that are common to the replaced and replacing agents (averaged across all turnover episodes). The operationalization of the above contingency factors is described in subsection 3.2.1.9.

According to Wittenbaum et al. (1998), the size of a workgroup can influence transactive memory by affecting communications within the group: members in larger groups need to maintain relatively more information in order to be accurate in their assessment of other members' areas of knowledge. Hence, workgroup size was operationalized via the contingency factor, *number of agents.*

Kanawattachai and Yoo (2007) and Kotlarsky and Oshri (2005) report that geographic separation of members of workgroups affects the workgroup's transactive memory systems and performance, by affecting the frequency and richness of communications negatively, during the initial time periods in the history of a workgroup. In later stages of the history of the workgroup, transactive memory positively affects the exchange of knowledge and information about knowledge among a workgroup's members. Hence, an operationalization of the geographic location of workgroup members was included in the study in the form of the contingency factor, *number of locations*, so that its effects on the workgroup processes that are determine the choice of a communication medium, could be studied.

Davenport and Prusak (1998, pg. 90) describe how members of a workgroup, while engaged in non-task-related activities, may exchange knowledge and/or information about different areas in which they have knowledge. Hence, an operationalization of this phenomenon and its effect on the emergence of transactive memory systems and performance of a workgroup was considered pertinent. The contingency factor, *probability of exchange of information about a non-task-specific-knowledge area*, which indicates the likelihood of an exchange of information in a non-task-related area during a given time period, operationalizes the phenomenon.

Prior studies (Liang et al., 1995; Lewis et al., 2005; Ren et al., 2006) indicate that the positive effect of training on the development of transactive memory and the performance of a workgroup. Training of the members of a workgroup is a phase in the lifetime of the workgroup that is scheduled before the commencement of the project phase. During this phase, members of the workgroup interact with each other and learn about each other's areas of knowledge. Hence, the contingency factor, *time in training phase*, was included in the study to operationalize the pre-project training phase. Specifically, it indicates the time periods spent by agents of a workgroup interacting with each other, and learning about each other's areas of knowledge, before the commencement of the workgroup's projects. Its operationalization is described in subsection 3.2.1.18.

Members of a workgroup might not always succeed in completing all the tasks assigned to them. In the context of information systems projects, the key reasons for the abandonment of projects were found to be (a) the nature of the work, (b) the structural attributes of the members involved in the projects, and (c) level of coping with uncertainty (Ewusi-Mensah and Przasniki, 1991; Pan and Pan, 2006). Lesca and Caron-Fasan (2008) identified several other factors that contribute to the abandonment of projects: poor project impetus, uninvolved management, unqualified people, inaccurate expectations, project mismanagement, strategy misalignment, poor participation, hostile culture, insufficient budget, conflating technical and managerial problems, previous project trauma and underestimated complexity. Since the successful completion of a project is predicated on the successful completion of each component of the project that is assigned to agents of the workgroup, it is essential to include in the study contingency factors that capture the amount of failure an agent can manage before abandoning the work-unit assigned to it. As indicated by prior literature, the causes for abandonment of tasks in a project are numerous. However, including all, or even a reasonably large subset of the factors identified in literature would add significantly to the complexity of the agent-based model simulated in the dissertation. Hence, a single contingency factor, *maximum number of failed tries*, is used to serve as a proxy for the combined effects of the contextual factors and member attributes on the likelihood of a subtask (that is part of a project) being abandoned by the agent to whom it is assigned.[6] Its operationalization is described in subsection 3.2.1.4.

2.6.2. Characteristics of the tasks

This category of contingency factors describes the various attributes of the tasks performed by members of a workgroup that are expected to have an effect on the workgroup's processes.

A workgroup consisting of several members, each working on a small part of a larger problem, at times cooperating with each other by obtaining/providing the required knowledge, is an example of distributed problem-solving (cf. Smith and Davis, 1981). In addition to breaking down a larger problem, a project (as described in subsection 3.2.1.2), into smaller problems assigned to individual members of the workgroup, called tasks, a greater flexibility, in terms of choosing to work on a specific component of a task, would be possible for the members if the tasks themselves are decomposed into subtasks (individual components of a task). This decomposition allows a member to choose to work on a different subtask, if he/she could not locate a knowledge source for a subtask, in a given time period. Once the current subtask is complete, the member can resume his/her search for a source for the knowledge associated with the subtask that he/she tried to complete earlier. The likelihood of a member finding a source for the knowledge tends to increase with time, because, over a period of time, other members, through their own knowledge-

6 As described in the methodology chapter, subsection 3.2.1.4, a subtask is considered to be abandoned after a certain number of tries. This rule encapsulates the role of two factors: (1) complexity of the project, as indicated by the number of subtasks and their associated knowledge areas, and (2) lack of qualified members - members in a workgroup are modeled as having incomplete knowledge and partially correct perceptions of the accuracy their own and others' areas of knowledge. These two factors are a subset of the entire set of factors identified by Lesca and Caron-Fasan (2008).

seeking interactions, might have acquired the knowledge that they are seeking, and may be potential sources of the knowledge being sought by the member. If there were no decomposition of a task into subtasks, then the member searching for the knowledge required to complete his/her assigned task, is not free to work on a different task (the reasons for this would be explained in section 2.6.3, where the notion of network of task-interdependencies is explained). Hence, the decomposition of a task into subtasks is essential. To study how a variance in the number of subtasks associated with each task affects the workgroup's processes, a contingency factor, *task intensity*, is included in the study.

In addition to varying in terms of the number of subtasks that constitute them, tasks can also vary in terms of the number of knowledge areas associated with them. Stated differently, tasks, and more specifically, subtasks, can vary in difficulty as measured by the number of unique areas of knowledge required for their completion. The notion of subtask difficulty, as indicated by the number of areas of knowledge associated with each subtask, on an average, is captured by the contingency factor, *knowledge intensity of subtasks.* Its implementation is described in subsection 3.2.1.4.

Espinosa and Pickering (2006) indicate that members of a workgroup that are separated geographically, across timezones, are able to coordinate their activities better in instances where coordination difficulties that arise due to task-priority conflicts are lower. This finding implies that, in addition to the temporal sequencing of tasks (described in 2.6.3), the notion of task priority, and its role in the choice of a task by the member of a workgroup, must be investigated. The contingency factor, *task priority*, captures the notion. Its implementation is described in subsections 3.2.1.11 and 3.2.1.16.

2.6.3. The interface between the workgroup and the tasks

This category consists of those contingency factors that describe the context created by the assignment of tasks to various workgroup members during their lifetime[7]. In their review of various factors that affect the effectiveness of workgroups, Campion et al. (1993) identified the workload assigned to each member of a workgroup as a factor that affects the

---

7 Here and elsewhere in this document, the lifetime of a workgroup is a series of projects in which the workgroup participates. It is possible for the membership of the workgroup to change, due to turnover.

performance of the workgroup. Urban et al. (1995) found that with increased workload, the performance a group deteriorates. Littlepage et al. (2008) have observed that the development of a workgroup's transactive memory systems is affected by the relative workloads borne by members of a workgroup. The effect is positive, when the increase in relative workload per member results in the assignment of tasks to members who are more knowledgeable in areas associated with the tasks assigned to them. However, when the match between members and the tasks assigned to them was lower, greater workloads resulted in lowered overall performance of the workgroup. Hence, the notions of workload of the workgroup as a whole and the workload of each workgroup member are considered as two pertinent contingency factors. They are operationalized as *average project intensity*, which indicates the total number of tasks to be completed by a workgroup as a whole, and *average number of tasks per agent*, which indicates the workload, on an average, of each member of the workgroup. Their implementation is described in sections 3.2.1.2 and 3.2.1.6.

To answer the research question that was stated in chapter 1, the effects of contingency factors on workgroup outcomes must be studied across multiple time periods. Lewis (2004) studied the emergence of a transactive memory system and its effect on workgroup performance over time. But his study was limited to the duration of a single project. Other studies (e.g., Austin, 2003; Ren et al., 2006) have studied the emergence of transactive memory over time by studying workgroups across multiple projects. Building on these ideas, this dissertation studied a workgroup across several projects, each of which spanned multiple time periods. The contingency factor[8] *number of projects per workgroup* (number of projects per simulation run, since, as explained in subsection 3.2.1.18, each simulation run represents a single workgroup's lifetime) operationalizes the longitudinal nature of the investigation.

In addition to studying a workgroup across multiple projects, the possibilities that (1) any two successive projects in the lifetime of a workgroup need not be entirely similar, and, (2) they may be similar to an extent, whose value lies in the range of entirely dissimilar to entirely similar[9], must also be investigated. The contingency factor *similarity of projects* addresses the above issues. Specifically, it determines the degree to which two successive

8 While the importance of this contingency factor is recognized, given the exploratory nature of the study, this factor is kept constant across all simulation runs.

9 It must be noted that for the sake of simplification, this attribute is held constant across all the workgroups; this assumption can be relaxed, but doing so was deemed to beyond the scope of the current study.

projects are similar to each other in terms the proportion of subtasks (which are the most basic units of work) that are common to each of the two successive projects. Subsection 3.2.1.2 describes the implementation of this contingency factor.

Steiner (1972) and Campion et al. (1993) indicate that task interdependence is one of the determining factors that can positively affect the effectiveness of a workgroup. Therefore, task interdependence was included as a contingency factor in this study, operationalized by *connectedness of network of task-interdependencies*. Its implementation is described in subsection 3.2.1.3.

2.6.4. Characteristics of knowledge

This category consists of those attributes of knowledge that affect the degree of difficulty encountered by a member in two types of circumstances. First, while receiving the knowledge in a given area from another member who has the knowledge, and second, while receiving instructions that are based on the knowledge in the given area from another member. Instructions, rather than knowledge, are provided if it is deemed that, in the context of applying the knowledge in completing a subtask, through an agreement between the source and the recipient members, receiving the instructions is quicker (in terms of time spent) than receiving and comprehending the knowledge.

Using Szulanski's (1995) description of the characteristics of knowledge that make its transfer difficult because of the 'stickiness' of the knowledge, the contingency factor *stickiness time* was included in the study. It indicates the time taken to transfer knowledge from a source to a recipient such that, at the end of the transfer, the recipient comprehends the received knowledge and can apply it to complete the subtask with which the specific area of knowledge is associated. Complementary to the notion of *transfer of knowledge* is the notion of *direction* (Grant, 1996a) which describes the phenomenon where a knowledge source does not provide the needed knowledge. Instead, the source provides instructions that are derived from the knowledge in the requested area, to the recipient. The recipient can follow the instructions and complete the associated subtask, without having to receive and comprehend the required knowledge. The notion of *direction* is captured in the contingency factor *direction time*, which indicates the amount of time taken by a recipient to receive instructions in the required area of knowledge from a source. The implementations of *stickiness time* and *direction time* are described in sections 3.2.1.7 and 3.2.1.15.

2.6.5. Characteristics of information technologies

In the current study, two types of information technologies were considered: (1) those that facilitate communication, e.g., telephone, email software, instant messaging or chat software, and (2) those that facilitate the workgroup members' ability to locate a knowledge source, e.g., an expert-seeker type of software (Becerra-Fernandez, 2000). These two types of information technology are discussed next.

As described by Media Richness Theory (Daft and Lengel, 1986), different types of media, such as face-to-face, electronic media such as email, etc., facilitate communication with different degrees of richness. Kane and Alavi (2007) studied how the use of different information technologies associated with knowledge-exchange in a workgroup affect the knowledge level of individuals and the workgroup as a whole. Hence, the use of four types of communication media is included as four contingency factors. Apart from face-to-face, which involves direct member-member communications, three information technologies were included: email, text-based chat and telephone. The use of a particular electronic medium for communication by a member is predicated on contextual factors such as whether it allows synchronous communications, communications between members who are geographically separated, the value of richness as perceived by the member who is the recipient of knowledge or instructions, as a function the member's own attributes, etc. These implementation details as described in sections 3.2.1.11– 3.2.1.14 and 3.2.1.16.

In addition to facilitating communications, information technology can aid a member in locating source for knowledge in a needed area, by providing a white-pages type of service, described as 'expert-seeker' (Becerra-Fernandez, 2000). Hence, a workgroup's use of expert-seeker-type functionality is included as a contingency factor, *use expert-seeker.* Its implementation is described in subsection 3.2.1.10.

2.7. Categories of processes

The processes followed by a workgroup can be classified into three categories: (1) *processes related to the completion of tasks*, (2) *processes accompanying those related to completion of tasks*, and (3) *processes related to scheduling of tasks*. The specific processes in each of the above three categories are identified below; details of their implementation are presented in appropriate locations in section 3.2.

The category *processes related to the completion of tasks* consists of the following actions on the part of agents of a workgroup: *(P1) search for potential sources of knowledge in the required area*; *(P2) choice of a knowledge source for the set of potential sources*; *(P3) choice of a medium for the transmission of the requested knowledge*; *(P4) choice of the knowledge transmission mechanism (direction or transfer)*; *(P5) actual transmission of knowledge*; *(P6) verification of the accuracy of the obtained knowledge.*

The category *processes accompanying those related to task-completion* consists of the following processes: *(P7) exchange of information pertaining to non-task-related area of knowledge*;

(P8) modification of perceptions of workgroup member's propensity to share; and, *(P9) learning new knowledge*. These processes are initiated after the processes in the category *processes related to the completion of tasks* are initiated.

The category *processes related to scheduling of tasks* consists of the following processes: *(P10) identification of task to complete*, and *(P11) identification of subtask to complete*. The processes in this category precede those presented in the category *processes related to the completion of tasks.*

Details of each of the above processes, along with the likelihood and sequential order of occurrence, are provided in subsection 3.2 in chapter 3, *Methodology*. Figure 2.1 presents a high-level overview of the relationships between the categories of contingency factors, and the workgroup processes on the one hand, and the workgroup processes and the outcome variables on the other hand. Table 2.1 provides a detailed listing of the contingency factors, workgroup processes and outcomes.

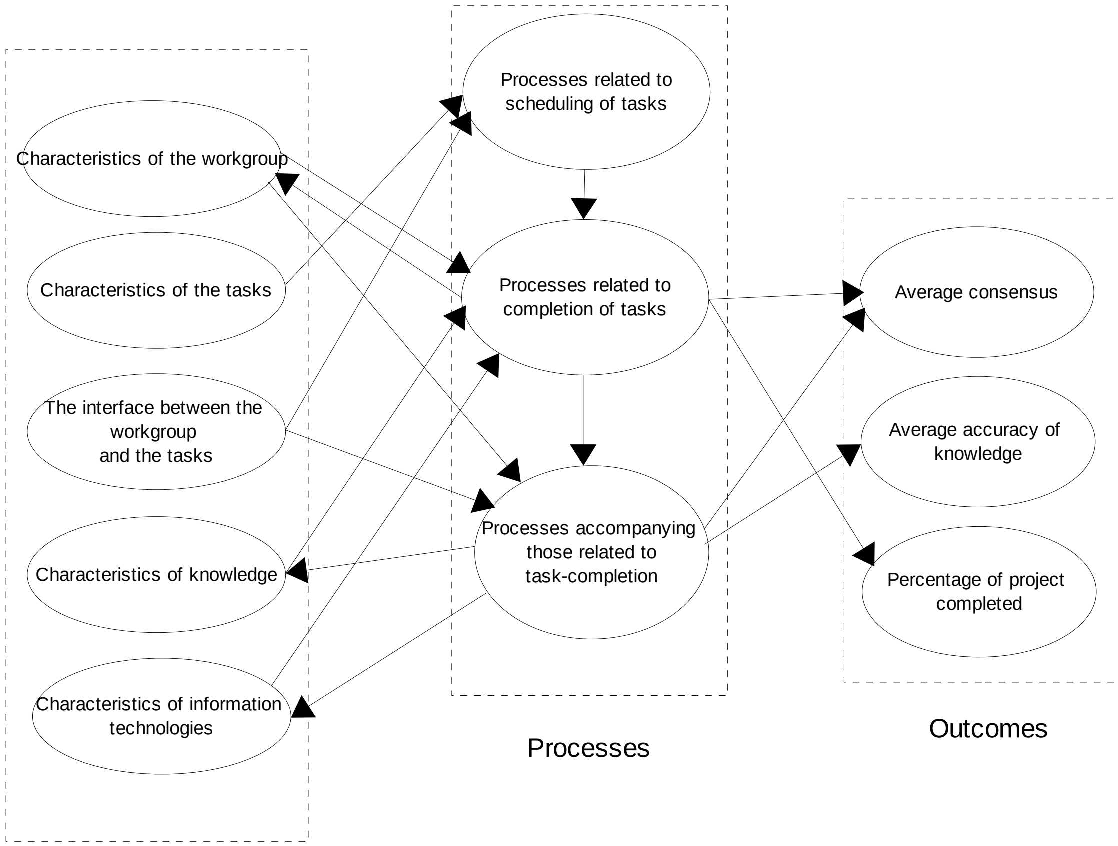


Figure 2.1. Overview of the relationships between contingency factors, processes and outcome variables[10]

2.8. Outcome variables

As mentioned earlier, in chapter 1, there are three outcomes of interest: (a) a measure of a workgroup's transactive memory, *average consensus of the workgroup*; (b) a measure of the workgroup's overall knowledge level, *average accuracy of knowledge of the workgroup;* and, (c) a measure of the workgroup's performance, *percentage of project completed*. This section describes the outcome variables.

The inclusion of the two outcome measures, *average consensus of the workgroup* and *average accuracy of knowledge of the workgroup* is derived directly from Austin's (2003) study, where he described consensus and accuracy of a workgroup's knowledge as

10 The backward pointing arrows between the group of processes and the contingency factors represent the fact the parameters representing the contingency factors are initialized at the beginning of a simulation run and are modified over the course of the simulation run as a result of the workgroup processes that act on them.

two measures that indicate the workgroup's transaction memory[11]. *Average consensus of the workgroup* indicates the overall degree of agreement about the knowledge areas of each member of the workgroup, averaged across all members of the workgroup. The consensus measure does not take into account whether the beliefs about the members' areas of knowledge are correct, that is, whether the members who are perceived to have knowledge in certain areas, do indeed possess accurate knowledge in those areas.

The purpose of the second outcome measure *average accuracy of knowledge of the workgroup* is to complement the first measure, by indicating the proportion of areas in which the agents of the workgroup have accurate knowledge. Accurate knowledge is knowledge, which, when applied, helps the member possessing it complete the subtask with which the given knowledge is associated.

The third performance measure, *percentage of completion of work*, is consistent with Austin's (2003) operationalization of workgroup performance as the attainment of workgroup goals. The purpose of the third outcome measure is to determine the efficacy of a workgroup. In this dissertation, each workgroup follows the processes described in section 2.7, under a given set of contingency factors, described in section 2.6, in completing the work (the tasks and their constituent subtasks) assigned to it.

2.9. Summary

This chapter described the key concepts used to answer the research question: knowledge, and transactive memory, and discussed the important facets of these concepts as described in the literature. Then, drawing on prior literature, it identified the pertinent contingency factors, which were categorized into five groups, and found to affect a workgroup's processes. The processes themselves are categorized into those that pertain to the scheduling of tasks, those that pertain to the completion of tasks and those that are associated task-completion processes. Next, the three outcome variables were described.

The next chapter, *Methodology*, presents details of the semi-structured interviews, which provided support for certain key aspects of the agent-based model, and the details of the agent-based model and its simulation. Then, the statistical method of panel data

11 Total knowledge stock and knowledge specialization are the other two aspects of transactive memory that Austin used in his study.

analysis, which was used to analyze the data collected through an agent-based simulation, is described.

Table 2.1. Details of Contingency Factors, Processes and Outcomes

Contingency Factors	*Processes*[12]	*Outcomes*
Characteristics of the workgroup (F1) average knowledge level (F2) average openness to change (F3) average self-knowledge (F4) probability of turnover (F5) average proportion of knowledge areas common with the replaced agent (F6) number of agents (F7) number of locations (F8) probability of exchange of information about a non-task-specific-knowledge area (F9) time in training phase (F10) average propensity to share (F11) maximum number of failed tries *Characteristics of the tasks* (F12) average task intensity (F13) average task priority (F14) average knowledge intensity of subtasks *The interface between the workgroup and the tasks* (F15) average project intensity (F16) average number of tasks per agent (F17) number of projects per workgroup (F18) similarity of projects (F19) connectedness of network of task-interdependencies *Characteristics of knowledge* (F20) average direction time (F21) average stickiness time	*Processes involved in the completion of tasks* (P1) search for potential sources of knowledge in the required area (P2) choice of a knowledge source for the set of potential sources (P3) choice of a medium for the transmission of the requested knowledge (P4) choice of the knowledge transmission mechanism (direction or transfer) (P5) actual transmission of knowledge (P6) verification of the accuracy of the obtained knowledge *Processes auxiliary to the completion of tasks* (P7) exchange of information pertaining to non-task-related area of knowledge (P8) modification of perceptions of workgroup member's propensity to share (P9) learning new knowledge *Processes involving the scheduling of tasks* (P10) identification of task to complete (P11) identification of subtask to complete	(O1) Average consensus of the workgroup (O2) Average accuracy of knowledge of the workgroup (O3) Percentage of completion of work

12 These processes occur at the individual workgroup-member-level. There are additional processes such as (a) creation of knowledge repertoire; (b) creation of subtask repertoire; (c) assignment of subtasks to tasks; (d) assignment of tasks to projects; (e) assignment of knowledge areas to subtasks; (f) allocation of knowledge areas to agents; (g) assignment of tasks to agents; (h) creation of a random network represents task-interdependencies; (i) creation of a set of agents; (j) assignment of agents to workgroups; (k) implementing the training-phase; and (l) execution of turnover. These processes are not listed here because they happen before the project-phase of a workgroup, and are handled by a 'supervisor' who does not participate in any of the task-completion activities.

Contingency Factors	*Processes*	*Outcomes*
Characteristics of information technologies (F22) use of expert-seeker (F23) characteristics of telephone (F24) characteristics of email (F25) characteristics of text-based chat (F27) characteristics of face-to-face		

3. Methodology

To determine how different contingency factors affect the outcome variables over time, via a set of workgroup processes, an agent-based model was developed. Its goal was to simulate the processes that a workgroup whose characteristics are defined by a set of parameters. In addition to the workgroup's characteristics, the parameters of the agent-based agent-based model also operationalized all the contingency factors that were described in chapter 2. Data recorded during the simulation of workgroups, via computational experiments, helped understand how the outcome variables changed over time as a function of the parameters and the workgroup processes that were simulated. As described in chapter 2, prior literature informed the choice of contingency factors and workgroup processes that were included in the agent-based model. Additional support for key aspects of the agent-based model was obtained via a series of four qualitative, semi-structured interviews.

In this chapter, first, summaries of the interviews, and the changes that were made to the specifications of the agent-based model based on the support obtained from each interview, is presented in section 3.1. Next, in section 3.2, detailed specifications of each aspect of the agent-based model are presented, along with explanations of the choice of the simulation development and execution environment and procedures followed for validating the model's implementation. Then, section 3.3 explains the panel data analysis methodology that was used to analyze the data obtained from the simulation. Finally, a summary of this chapter is presented, along with a brief overview of the subsequent chapter (chapter 4, *Results*).

Next, summaries of the interviews, and how the data obtained from the interviews were used to validate and inform the agent-based model, are presented.

3.1. Qualitative Interviews

In order to examine the appropriateness of the various aspects of the simulation specification described in the proposal document, and to determine if any additions are needed, a series of four semi-structured interviews were conducted. These interviews were preceded by a pilot interview, whose purpose was to determine the suitability of the questions included in the interview-guide that was used in the main interviews. Additionally, feedback was sought from the interviewee to determine whether any revisions to the interview-guide were required. Additionally, the pilot interview helped in the estimation the approximate time required to complete an interview based on the then-current interview-guide. This estimate helped in the scheduling of the main interviews.

The main interviews were semi-structured. They sought to validate various aspects of the agent-based model. Where needed, answers were obtained to questions that probed certain topics in greater depth compared to the depth of responses elicited by the original questions. The interviews were conducted sequentially. After analyzing the data obtained from each interview, necessary additions and changes were made to the specifications of the agent-based model. Prior to each subsequent interview, the interview-guide was modified to obtain support for the revised set of simulation specifications. This iterative process of revising the simulation specifications in the light of new evidence, and seeking support for the updated version of the specifications, continued until no additional changes were needed. The above approach was informed by the principle of "constant comparative method" (Glaser and Strauss, 2009, pg. 104), where evidence is collected in an iterative manner to support or refute previously-formed conceptions about the phenomenon being investigated. The iterative process terminates when no unique interpretations may be drawn from the most recent data-collection efforts.

Next, a summary of the pilot interview is presented. Table 3.1 provides summaries of the characteristics of the interviewees and the contexts to which they related while responding to the interview questions.

3.1.1. Summary of the pilot interview

Data from the summary interview was grouped into the following categories: (a) *choice and use of information technologies*; (b) *characteristics of knowledge and their affect on knowledge-exchange*; (c) *search for source of knowledge in a particular area*; (d) *determining the accuracy of acquired knowledge*; and, (e) *formation and change of perceptions of a knowledge-source's areas of knowledge and propensity to provide accurate knowledge*. Details of the findings from the first interview are presented next.

Choice and use of information technologies

- The type of communication medium chosen depends on the richness and immediacy of communication it provides.
- Synchronous voice communication (Google talk® and Skype®) used for audio chat
- Collaborative data/application sharing (Microsoft Netmeeting®)
- E-mail was used for asynchronous communication.
- Audio chat is used primarily; video did not appear to provide any additional information; audio chat served as a proxy for face-to-face communications.

Use of electronic media is preferable after there have been prior face-to-face interactions. These prior interactions provide information cues regarding the patterns of speech, etc., that the other person uses in communication.

Characteristics of knowledge and their affect on knowledge-exchange

- The description of tacit knowledge should be more elaborate: tacit knowledge is knowledge that is by nature difficult to articulate, therefore is more difficult to share.
- A lack of prior face-to-face experience makes it more difficult to share tacit knowledge via an electronic medium

Search for source of knowledge in a particular area

- Publicly available information is sought in order to determine a potential source's areas of expertise before contacting that person for knowledge (e.g., prior published work, posts on various online forums, specific role in an organization that creates and uses knowledge in the area that is being sought by the knowledge-seeker, e.g., developers of software which is being used by the knowledge seeker and about which the knowledge seeker needs clarification/new knowledge.
- Potential sources (individuals) include those that are recommended by the individuals that a knowledge-seeker trusts.
- A knowledge-seeker is more likely to seek knowledge from those who made at least some of their knowledge areas public, e.g., posting in a blog.
- A person willing to share in one area is assumed to be equally willing to share in other areas.

Determining the accuracy of acquired knowledge

- Perceived accuracy of knowledge is determined by its usefulness to the the knowledge seeker.

Formation and change of perceptions of a knowledge-source's areas of knowledge and propensity to provide accurate knowledge

- If repeated requests from a seeker of a source were met with inaccurate knowledge being shared in all or most of the requests, then that source is judged as having less accurate knowledge.
- Evaluation of a source's knowledge is based on one's own judgment of the accuracy of the source and on the evaluation of that source by others whom the seeker trusts as being good judges of the source. A greater weight is given to others' judgment of the source than to one's own judgment of the source.
- If the knowledge given by the source was unclear/not understandable then the source is probably not willing to share knowledge. Inaccurate knowledge sharing tends to indicate lesser level of honesty. However, this judgment is not always made. Length of association with a source affects the tendency of a recipient to associate inaccuracy in the knowledge received to dishonesty. If inaccurate knowledge was shared by a source with whom the recipient has had a long relationship, then the source is seen as being honest but having incorrect knowledge. If the duration of relationship with the source has been short, then the source's inaccurate knowledge is seen with a 50% chance as resulting from a tendency of the source to be dishonest and with a 50% chance of the source having inaccurate knowledge.
- Initially, the source might not have had the time to provide knowledge in a more comprehensive and comprehensible form. So, if the knowledge obtained in the initial interaction is inaccurate, then the source is not necessarily seen as being dishonest. Dishonesty might be inferred if repeated interactions result in inaccurate knowledge being shared every time. However, if during repeated interactions, the knowledge that is shared is perceived as being more comprehensible and accurate, then the perception of the source's level of honesty associated with sharing knowledge goes up.
- A source might not be willing to share knowledge or be dishonest if sharing

knowledge would affect the person's well-being, e.g., affecting the source's standing in the organization. A tendency to share knowledge by a source is described to be a result of positive experiences that the source had in the past, when the source was a recipient of knowledge. A person who has had negative experiences when seeking knowledge from others, for example, being given inaccurate knowledge, might choose to either behave negatively towards others who seek knowledge from him/her by providing inaccurate knowledge, or might choose to act differently from his sources, by sharing accurate knowledge with those who seek it from him/her. The specific choice a person makes depends on his/her past experiences and other aspects of his/her personality.

Based on the above information, in the main interviews, the following additional questions were considered for inclusion.

1. If you (as a recipient) received inaccurate knowledge, would you deliberately provide inaccurate knowledge to others when they seek knowledge from you?
2. Would you share inaccurate knowledge with those sources who shared inaccurate knowledge with you?
3. Would you share accurate knowledge with sources who, in the past, shared inaccurate knowledge with you?

Table 3.1. Profiles of interviewees and their contexts

Interview	*Role of the interviewee in the context*	*The context*
Pilot	Researcher and collaborator in an academic research project in the field of Information Systems	Academic research, with a goal of publication in top journals in the field of Information Systems. The interviewee worked with his co-authors (two, in addition to the interview) who were geographically separated, but were from similar time zones. The research team employed multiple media for communication: face-to-face, e-mail, video and text chat, and collaborative-document-editing tools
1.	Systems analyst and programmer at a large US financial institution with operations in multiple US locations and with systems-development-related activities situated in multiple US locations and one off-shore location.	Design and implementation of a new software-based system for loan-origination. The team employed multiple media for communication: face-to-face, e-mail, voice and text chat, and collaborative-document-editing tools

Interview	*Role of the interviewee in the context*	*The context*
2.	Storage systems engineer working for a systems integrator company located in a large metropolitan area in the US.	A project to centralize storage and set up remote replication processes for a client of the company. The team utilized multiple media for communication: face-to-face, video conferencing, telephone, text-messaging and e-mail.
3.	Team-leader, in charge of a software development team with members located at one US-based (located in a large metropolitan area) and one off-shore location.	The team worked on the design and development of a time-entry system. The team utilized multiple media for communication: face-to-face, e-mail, text-chat and telephone.
4.	Project manager at a property-development firm in a large metropolitan area in the US.	A property development project that involved the rehabilitation of apartment buildings in the region. The team utilized multiple media for communication: face-to-face, e-mail, text-chat and telephone.

4. On what basis would you decide on whether to share knowledge accurately or inaccurately?

The above line of questioning (1 – 4) was extended to a person's tendency to divulge the different areas in which the person has knowledge, to result in the following questions:

5. Under what circumstances would you choose to divulge the areas of in which you have knowledge?
6. When would you choose not to reveal all the areas in which you have knowledge?

Additionally, questions were included that determine the likelihood of a person choosing to seek knowledge from the same source even after receiving inaccurate knowledge from that source. These questions were:

7. If, following your initial request for knowledge, you were provided with knowledge that was not entirely accurate, under what circumstances would you continue to use that person as a knowledge source?
8. Under what circumstances would you stop seeking knowledge from that source?

However, including questions 1 – 6 and incorporating their responses into the specifications of the agent-based model was deemed to lead to additional complexity associated with the implementation of the agent-based model. Hence, to keep this study manageable,the processes related to decisions made by members of a workgroup that are related to knowledge-sharing, as indicated by questions 1 – 6 were considered beyond the scope of the study.

Next, summaries of the main set of interviews are presented. Complete transcripts of the interviews, in question-answer format, along with the support each question was seeking and any associated comments, are provided in appendix A2.

3.1.2. Summary of interview 1

Based on a review of the original specifications of the agent-based model and the interview transcript, the two variables (described in the simulation specification present in the proposal document) *propensity to divulge* (*ptd*) and *propensity for honesty* (*pfh*) were seen to be not entirely orthogonal. Specifically, the following definitions of *ptd* and *pfh*:

> " Each agent, based on its *propensity to divulge,* that is, how open the agent is in terms of revealing information about its areas of knowledge – let this parameter be called *ptd* – updates the expert seeker directory at the end of each project. Higher values of *ptd* imply that the agent provides information about more of its areas of knowledge." (proposal document, page 44)

> "During each such interaction, each agent in the pair of interacting agents learns about the existence (or absence) of knowledge in one randomly chosen area in the other agent. This exchange of information about the presence/absence of knowledge is constrained by the probability that an agent would honestly reveal information to the outside world (to the expert-seeker directory and to the agents with which it interacts), that is, by the parameter pfh." (ibid., page 43)

along with the following description

> "The constructs propensity to divulge and propensity for honesty described in point 35 capture the dimensions of organizational pertaining to being honest and participatory in knowledge sharing and job-related activities. Propensity to divulge is considered to capture the notion of people contributing to the knowledge sharing activities in an organization because of the joy they derive in helping others (Kankanhalli, et al., 2005 ; Wasko and Faraj, 2005). Propensity for honesty captures the notion that employees in an organization do not fully reveal the various areas in which they have expertise. Reasons for not being completely honest about one's areas of expertise include the fear of losing one's power or value (Gray, 2001; Davenport and Prusak, 1998)" (ibid., page 44)

imply that a workgroup member has a tendency to share (or not share) and that this tendency varies. The variation in this tendency can be construed as having two components: (1) a generalized tendency to share and, (2) a tendency to share with a specific person in a dyadic relationship. A combination of these two components determines the likelihood of a person sharing his/her knowledge with a specific recipient, whether the recipient is a corporate database such as an expert-seeker, a person whom the source trusts, or a stranger. The decomposition of the tendency to share into two components, while interesting from a theoretical and explanatory perspective, would add considerably to the complexity of the simulation model, if the two components are implemented separately. The workgroup-level parameter, *generalized propensity to share* represents an average (at the workgroup-level) measure of an agent's general tendency to share knowledge, or information about the presence of knowledge in a given area, with other agents. This parameter is used to set individual-level parameter values of *propensity to share*, for each agent of the workgroup.

At the level of an individual agent, *propensity to share* captures the notion that when an agent is in a situation, where it has to share its knowledge with a knowledge seeking entity (another agent, or the 'expert-seeker'), it shares only a certain proportion of its knowledge. Specifically, when another agent contacts the focal agent for knowledge in a particular area, *propensity to share* determines the probability that a source agent provides the knowledge-seeking agent the knowledge requested. That is, *propensity to share* determines the probability of sending a 1 instead of a -1 to the knowledge-seeking agent, or a 1 instead of a 0, to an information seeking agent. In the case of a knowledge-seeking agent, a 1 represents the transmission of correct knowledge in the request area and a -1 represents the transmission of incorrect knowledge in the requested area. In the case of an information-seeking agent, a 1 represents information that knowledge is present in a randomly chosen area and a 0 represents absence of knowledge in a randomly-chosen area.

In addition to having its own propensity to share encapsulated as an attribute *propensity to share*, an agent would also have perceptions of other agents' *propensity to share* values. Specifically, each agent will have an attribute *perceived propensity to share*, for every other agent in the workgroup. This implies that in a workgroup of size *n*, each agent would have one attribute *propensity to share*, representing its likelihood of sharing knowledge, or information about knowledge, with other agents, and *n - 1 perceived propensity to share* attributes, one for each of the other agents in the workgroup. Initially, when no tasks have been performed, and no task-specific knowledge exchanges have occurred, a knowledge-seeking agent will have *perceived propensity to share* values that are equal to its own *propensity to share* value. When a knowledge-seeking agent contacts a source agent for knowledge and receives it, it modifies its *perceived propensity to share* attribute value, which is associated with the source agent, as a function of the accuracy of knowledge that it received during the completion of its latest task. The following set of equations describe this process.

$$perceived\_pts_{task\ 0} = generalized\_propensity\_to\_share \qquad -(3.1)$$

$$perceived\_pts_{task\ n} = perceived\_pts_{task\ n-1} * 1.01 * chg \qquad -(3.2)$$

where *chg* = 1 if the knowledge provided was accurate (helped complete the subtask)
chg = -1 if the knowledge provided was inaccurate (did not help complete the subtask)

Change occurs in *perceived propensity to share* at the completion of each subtask, not at the completion of each time period. This is because only when all the knowledge that is needed to complete a subtask has been received can the accuracy of the knowledge be judged in terms of whether it can be used to complete the subtask for which it was sought. If the knowledge is accurate, then the *perceived propensity to share* value associated with the source agent is increased by 10%; if the knowledge is inaccurate, then the *perceived propensity to share* value associated with the source agent is decreased by 10%.[13]

3.1.2.1. Support

The first interview provided support for the following aspects of the simulation specification:

- The use of different types of IT that are used for communication in synchronous and asynchronous modes.
- Synchronous modes of communication include face-to-face, group chat, instant messaging and phone calls.
- Asynchronous modes of communication include email and document-sharing via document repositories.
- The inclusion of IT that provides 'expert-seeker' or 'corporate white-pages' type of functionality in the agent-based mode is also supported.

13 It is possible to have this value be included in the agent-based model as a parameter, whose value is modified within and across workgroups. However, to keep the implementation of the agent-based model simple, this value is assumed to be constant within each workgroup and across all workgroups.

- Accuracy of knowledge is defined in terms of its applicability in completing a task

3.1.2.2. Changes suggested

The responses provided during the interview indicated the following changes and additions to the specification of the agent-based model:

- Agents can be geographically separated. Hence, each agent should have a *location-id* attribute
- The location of the source and the recipient determines, along with the stickiness of knowledge to be exchanged, richness of the communication medium and whether the medium supports synchronous communication, the probability of choosing a particular medium. So, among the various communication media available, the medium which provides the best fit with respect to synchronicity of communication, richness of communication and the stickiness of the knowledge to be exchanged would be chosen. The above rule regarding the choice of a communication medium replaces the rules detailing the use of various media as determined by the richness parameters *richness of face-to-face*, *richness of phone*, *richness of email* and *richness of chat.*
- Friends (defined as agents that are directly connected to a focal agent via prior interactions either during the pre-project phase or during the project-phase) of friends can act as sources of knowledge. Each agent's friends will let the agent know, while exchanging knowledge that is not related to the task that the knowledge-seeking agent is trying to complete, about their friends' areas of expertise. This type of knowledge exchange may be governed by a parameter *probability of utilizing friend's expertise*, which captures the probability of a source agent providing a knowledge-seeking agent information about an area of knowledge of one of the source agent's friends. Additionally, it is assumed that when an agent needs knowledge that none of the agents with whom it has direct connections possess, and there is no indication in the 'expert-seeker' database about who among the unconnected agents has the knowledge, then the knowledge-seeking agent will choose, iteratively, in decreasing order of *perceived propensity to share,* each agent to whom it is connected and tries to obtain information about who among the agent's

(that the knowledge-seeking agent has contacted) friends has the knowledge that it needs. If, after iterating through all its friends, it fails to identify a potential source for the knowledge it needs, it is going to abandon the current task because it lacks the knowledge necessary for completing the task. However, if it does find, via its friends, one of more agents that are believed (by the friend reporting the information to the knowledge-seeker) to have the required knowledge, it is going to contact that agent for knowledge.

- A new parameter, *propensity to share*, to replace the two parameters, *propensity to divulge* and *propensity for honesty*.
- *Propensity to share* determines an agent's tendency, as a source of knowledge, to share knowledge with a knowledge-seeking agent or 'expert-seeker'. Correspondingly, *perceived propensity to share* is a value that a knowledge-seeking agent has associated with each source agent. Among a set of potential source agents that are believed, by the knowledge-seeking agent, to have the required knowledge, the agent with whom the value of *perceived propensity to share* associated is the highest, is chosen.
- An agent that is contacted for knowledge in a certain area by a knowledge-seeking agent can provide the knowledge-seeking agent with inaccurate knowledge.

3.1.3. Summary of interview 2

In this subsection, a summary of the findings from interview 2 are presented.

3.1.3.1. Support

The second interview provided support for the following aspects of the simulation specification

- Geographical separation of a workgroup's members
- The use of different types of IT based on the location, stickiness of knowledge, synchronicity of communication, and the richness of the communication medium
- The network of relations between a knowledge-seeking agents and other members of the workgroup can help a knowledge-seeking agent in determining who among the “friends-of-friends” can be a potential source for knowledge.

- The pre-project phase (described in the “training phase” portion of the simulation specification) where workgroup members interact with each other and find out about each others areas of knowledge.
- 'Expert-seeker' type IT
- The use of *generalized propensity to share* to indicate a base-level tendency to share knowledge or information about the presence of knowledge in a particular area, averaged across all agents in a workgroup
- Accuracy of knowledge defined as whether it can used in completing a task.
- *Perceived propensity to share* values, which are updated for each agent with respect to the specific agent that it contacted for knowledge, are used in determining whom to contact in the future. This makes it possible for an agent, which provided inaccurate knowledge during one of the previous instances when it was contacted for knowledge by a knowledge-seeking agent, can be chosen as source of knowledge in one of the future tasks, if it turns to be the agent with the required knowledge and has highest *perceived propensity to share* value associated with it
- The inclusion of the provision of inaccurate knowledge by a source agent to a knowledge-seeking agent is neither contradicted nor confirmed.

3.1.3.2. Changes suggested

The following changes were made to specifications of the agent-based model:

- The probability of choice of IT depends, in addition to the stickiness of knowledge, richness of the communication medium and synchronicity of communication, on the priority of the task that is assigned to the knowledge-seeking agent. Specifically, in the case of tasks with higher priority values, agents are expected to choose a medium that provides the maximum richness, so that the required knowledge, or instructions based on the associated knowledge, may be transmitted from the source to the recipient in the shortest duration of time possible.

3.1.4. Summary of interview 3

In this subsection, a summary of the findings from interview 3 are presented.

3.1.4.1. Support

The third interview provided support for the following aspects of the simulation specification:

- Geographical separation of workgroup's members
- The use of different types of IT based on the location, stickiness of knowledge, synchronicity of communication, the richness of the communication medium and priority of the task assigned to the knowledge-seeker
- The network of relations between a knowledge-seeking agents and other members of the workgroup can help a knowledge-seeking agent in determining who among the “friends-of-friends” can be a potential source for knowledge.
- The pre-project phase (described in the “training phase” portion of the simulation specification) where workgroup members interact with each other and find out about each other's areas of knowledge.
- 'Expert-seeker' type IT.
- Generalized propensity to share
- Accuracy of knowledge in terms of its applicability in completing a task.
- The possibility of a knowledge-seeking agent being provided inaccurate knowledge by a source
- Modification of a knowledge-seeker's *perceived propensity to share* associated with a source agent, as a result of a knowledge-exchange-related interaction with the source
- Possibility of seeking knowledge from a source, despite receiving inaccurate knowledge in one of the previous instances

3.1.4.2. Changes suggested

There is no support for the phenomenon of abandonment of a task, and consequently a project. However, a simulation run would terminate only after all the projects, and their constituent tasks and sub-tasks, are complete. In an instance, where (a) an agent that is trying to complete a sub-task does not have the required knowledge, and, either (b) there is no other agent in the workgroup that has the required knowledge or, (c) there is at least one agent with the required knowledge but that agent's areas of knowledge remain hidden from the knowledge-seeking agent, then the knowledge-seeking agent will continue searching for knowledge indefinitely, resulting in the simulation not terminating. To avoid this situation, a terminating condition needs to implemented. Hence, the 'abandonment' described in the summary of interview 1 was considered to be reasonable and required in the simulation. The parameter *maximum number of failed tries*, which is assumed to be constant across all tasks and all projects, determines the number of time periods an agent should spend before abandoning a task and, consequently, the project. The parameter, *maximum number of failed tries*, ensures that the simulation does not run indefinitely, which might occur if an agent keeps does not abandon its search for a source agent.

3.1.5. Summary of interview 4

In this subsection, a summary of the findings from interview 4 are presented.

3.1.5.1. Support

The fourth interview provided support for the following aspects of the simulation specification

- Geographical separation of workgroup's members
- The use of different types of IT based on the location, stickiness of knowledge, synchronicity of communication, the richness of the communication medium and priority of the task assigned to the knowledge-seeker. Two new types of IT were mentioned: AutoCAD and Powerpoint. These tools were used in face-to-face meeting scenarios, hence the simplifying assumption that they are part of the face-to-face medium was made
- The pre-project phase (described in the “training phase” portion of the simulation specification) where workgroup members interact with each other and find out about each other's areas of knowledge.
- Inclusion of *propensity to share* and *perceived propensity to share*
- Inclusion of the modeling of inaccuracy in the exchange of knowledge
- Definition of accuracy of knowledge in terms of its use in completing a task. The interviewee's role was that of a coordinator of tasks among his workgroup's members (supervisory). Hence, the interpretation of completion of a task in this case implies whether the interviewee was able to successfully manage the

completion of a task that he delegated to a workgroup member. The accuracy of knowledge in this case can be seen as whether his belief that the task he delegated to another member is true, that is, whether the delegated task was successfully completed and whether a knowledge of that tasks successful completion could be used by the interviewee in working on his next task.

3.1.5.2. Changes

No changes were deemed necessary since the evidence from the fourth interview did not indicate any new aspects of the phenomenon that were not considered in the prior interviews and the current version of simulation specifications.

3.1.6. Additional remarks and follow-up interviews

Evidence from interviews 2,3 and 4 supports the inclusion of rules in the simulation specification to operationalize the notion of a change in the attribute *propensity to share* of an agent, as a function of its previous knowledge-seeking experiences. However, using the criterion of keeping the implementation of the agent-based model relatively simple, while ensuring that the phenomenon of interest is still modeled with sufficient richness, it is assumed that an agent's *propensity_to_change* remains constant throughout the duration of a simulation run.

In addition to the set of main interviews, there were follow-up interviews of the four interviewees. The responses from the follow-up interviews are presented, in a collated form, in appendix A3, which is structured as follows:

- the responses of the four interviewees are presented in a collated form, via a table. Each table is preceded by a question that describes the context in which the interviewee is going to communicate with another member of his/her workgroup via a communication medium. Each interviewee's response is indicated by an *x* in the cell corresponding to a communication medium and its likelihood of being chosen.
- the interviewee's rationale for the particular choice is provided in a table – one response per row; this table follows the above-described table.

The goals of the follow-up interviews were

(a) obtain information about the interviewee's tendency to use a communication medium under various situations provided in the interview-guide, and

(b) understand the rationale the interviewees used while making a choice of a communication medium under the conditions provided in the interview-guide.

In all four follow-up interviews, the interviewees were asked to provide their preferences for choosing each of the four media for communication: face-to-face, phone, email and chat/videoconferencing. However, at one of the interviewee's organization, two different types of software were used for instant messaging, which was text-only, and chatroom/video conferencing. Hence, the interviewee added a fifth item to his interview-guide and provided his preference for it, which is distinct from his preference for videoconferencing/chatrooms. Based on this information, and the responses provided by other interviewees in the second and third set of interviews, *text chat* is used to describe a medium that allows for exchange of messages that are primarily text-based and the exchanges can be synchronous or asynchronous (e.g., the offline-messages features incorporated by several instant-messaging programs).

As mentioned earlier, the purpose of the interviews was to seek validation of key aspects of the agent-based model. Next, in section 3.2, the agent-based model is described in detail. The section begins with a description of the rationale that informed the choice of agent-based simulation as the methodology for investigating the research question, which is then followed by detailed specifications of the agent-based model.

3.2. Simulation using an agent-based model

The simulation portion of the study involved the specification, implementation and execution of an agent-based model. Agent-based modeling of social phenomena has been gaining acceptance among researchers in the area of social sciences (Epstein and Axtel, 1996). Agents are entities in a computer simulation that operationalize decision-making units in organization of a workgroup, e.g., members of a workgroup. The specification also includes precise descriptions of various relevant aspects of a workgroup to which the agents belong. Examples of the pertinent aspects of a workgroup that were represented in the simulation include the size of the group, the prevalence of specific norms, rules of interaction among the group's members, etc. The micro-level specifications of agents' behavior guide the interactions within and among themselves and with their environment. The outcome of these interactions is the emergence of macro-level phenomena, which are not pre-specified and cannot be deduced in a deterministic manner.

In addition to replication of prior findings via computation-based experiments, as a methodology, simulation can be used for building theory (Davis et al., 2007). The key to the latter use of simulation is the implementation of computation-based experiments. This technique involves the development of a set of specifications for a model that represents the phenomenon being investigated in a parsimonious manner, and is based on the researcher's observations and/or results from prior studies. Then, the specification is iteratively made more elaborate (and thus more complex), with two goals: (1) to make the model have as high a fidelity as feasible to the phenomenon being investigated, such that all key aspects of the phenomenon are captured, and (2) to ensure that the model can be implemented and executed within the time constraints that bound the study. This progressive increase in the complexity of the specification allows for the testing of a greater variety of experimental conditions, and more importantly, interactions among the conditions, a process that might prove to be relatively infeasible were it implemented via a lab-experiment, a field-experiment or a field study. Also, the agent-based modeling approach makes it possible to delineate the specific effects of individual experimental conditions and the effects of their interactions. Since the data are generated via computations, the problem of collecting a sufficiently large sample, which might occur in real-life settings (including lab-experiments), is alleviated to the extent that the generation of a data set of required size is feasible in terms of the resources required for running the required number of simulations. Additionally, novel and extreme experimental conditions can be introduced into previously-specified models, making it possible to observe the effects of conditions that might occur infrequently or might be too difficult to implement and/or replicate in a lab-experiment or a field-experiment setting.

The idea of using agent-based simulation for building and testing theory pertaining to organizations has received researchers' attention (cf. Carley, 2002). Reviewing the extant literature pertaining to organizations, Carley (2002) describes the acceptance of agent-based modeling as a methodology for both theory-testing and theory-building. She categorized existing literature as belonging to the computation-based social and organization science perspective. She described this perspective as being based on the arguments of distributed cognition, transactive memory, and social construction of knowledge. She states that at the heart of this perspective is the argument that "... organizations are complex, computational and adaptive synthetic information processing agents. This new perspective urges a formalization of the roles of networks, learning and agency in affecting social and organizational change. Organizations are composed of intelligent adaptive agents who are constrained and enabled by their positions in networks linking agents, knowledge, resources and tasks" (Carley, 2002, pg. 257). In terms of the conception of organizations as distributed knowledge systems (Tsoukas, 1996), agents metaphorically represent individual cognitive units or knowledge sub-systems that create, store, process, disseminate and apply knowledge.

Epstein (2007) explains that theories, which aim to explain specific, empirically-observed phenomena, can be tested by developing an agent-based model based on those theories and comparing the outcomes of the agent-based model with the empirically-observed phenomena. He explains that the theories are validated to the extent that the results of the simulations, which are based on the theories, are consistent with empirical observations of the phenomena in question. Davis et al. (2007, pg. 483) state that "simulation is especially useful for theory development when the focal phenomena involve multiple and interacting processes, time delays, or other nonlinear effects such as feedback loops and thresholds. In these situations, simulation is likely to reveal non-intuitive elaborations of simple theory that are difficult to uncover using other methods, including armchair thought processes."

Given the above support for agent-based modeling as methodology for studying organizational processes, this study used it for answering the research question presented in chapter 1. Next, the specifications of the agent-based model are described.

3.2.1. Specifications of the simulation

As explained in section 3.2, the agent-based simulation component of the study was informed by information obtained from prior studies in the area of TMS and from the evidence gathered via the interviews (described in section 3.1). In this section, details specifications of the simulation are presented. These specifications reflect additions and changes to the original set of specifications, which were part of the dissertation proposal, in the light of support obtained during the interviews. Table 3.5 lists the simulation parameters their ranges and support for their inclusion, drawn from literature and interviews.

3.2.1.1. Agents

The key component of a workgroup is its members. Agents are computational representations of workgroup members. Each agent object is a data-structure that captures the essential attributes of a workgroup member:

(1) an identifier
(2) location: an agent can belong to one and only one of several possible geographic locations
(3) potential to share: determines the likelihood of the agent sharing (a) the knowledge requested by another agent; (b) sharing information with another agent about the knowledge it possesses in a knowledge area
(4) perceived richness of different communication media: (a) face-to-face; (b) telephone; (c) e-mail; (d) chat (instant messaging via desktop or hand-held/mobile communication devices)
(5) knowledge in different areas: an agent can have accurate, inaccurate or no knowledge in different areas of knowledge
(6) perceptions of its own knowledge in different areas: an agent might have correct or incorrect perceptions of the knowledge it possesses in a given area
(7) perceptions related the areas of knowledge and propensity to share of other agents in the workgroup

3.2.1.2. Projects and their completion

A project consists of a series of tasks. At the initialization of an experimental run, which represents the lifetime of a workgroup, all the agents are initialized with their attribute values, and the projects are assigned their various tasks (and their subtasks). Two successive projects in a simulation run can be similar to the extent defined by the proportion of subtasks that are common to both projects. The parameter *similarity of projects* defines the proportion of common subtasks. Similarity of projects is not transitive: given projects A, B and C, if A is similar to B, and B is similar to C, then it does not necessarily follow that A and C are similar to each other. Similarity between two projects varies in the continuous interval [0, 1].

A simulation run is complete when all the projects (fixed, determined by a parameter, *number of projects per workgroup*) are complete. Simulation runs are mutually-independent:

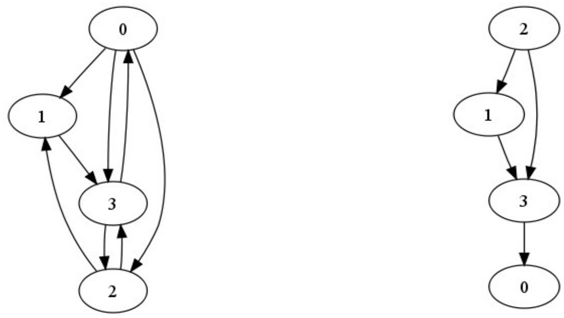
the parameters defining each simulation run are assigned values that are independent of the values assigned to the parameters of every other simulation run.

3.2.1.3. Tasks and their completion

Each task in a project is associated with a group of unique subtasks. A task is considered complete when each subtask associated with it is either completed or abandoned. In a project, each task is assigned to one and only one agent. Each agent, however, may be assigned more than one task. At the beginning of each project, tasks are assigned to agents randomly. Additionally, the tasks have sequential-interdependence, that is, some tasks might precede other tasks, forming a network of tasks. An agent can work on its assigned task only if (a) the task has no predecessors, or (b) all the preceding tasks are completed.

Additionally, it is possible for an agent to be assigned multiple tasks. After accounting for the precedence constraint, an agent will start working on the task with the highest priority among the tasks that have met the precedence constraint. At the beginning of each simulation run, a network with the required number of tasks is creating by assigning randomly inter-dependencies between two randomly-chosen pair of tasks. It is required that the graph of task-interdependencies not have cycles in it, and is thus a directed-acyclic-graph. An absence of this condition may lead to situations where, say for three tasks A, B and C, A depends on B, B on C and C on A, thereby leading to the situation that the project can never be completed, because each task is dependent on the completion of the other tasks. Figure 3.1 illustrates this.

In figure 3.1, network (a) has cyclic dependencies, for example, 0 → 1 → 3 → 0, 1 → 3 → 2 → 1. Network (b) illustrates the same network of tasks, with the cycles removed. Hence, the sequence in which the tasks can be performed is: 2 → 1; 2 and 1 → 3; 3 → 0. In the simulation, the network of inter-task dependencies should be generated randomly. For example, if the number of tasks, as determined by the parameter, *average project intensity*, is 4, then the initial randomly-generated network of inter-task dependencies may resemble 3.1.a. Once the cycles are removed from the network, then a score, indicated by the variable *connectedness of network of task-interdependencies*, which captures the notion of connectedness of a graph (cf. Butts, 2010), is computed. This variable is required in order to distinguish between the configurations of two networks of tasks that have the same number of tasks and inter-task dependencies but differ in terms of specific configurations of inter-task dependencies. Figure 3.2 illustrates this notion. For task-networks where no inter-task dependencies exist, connectedness of the task-network is zero.



(a) Cyclical dependencies present **(b) Cyclical dependencies absent**

Figure 3.1. Illustrating mutual dependency among tasks

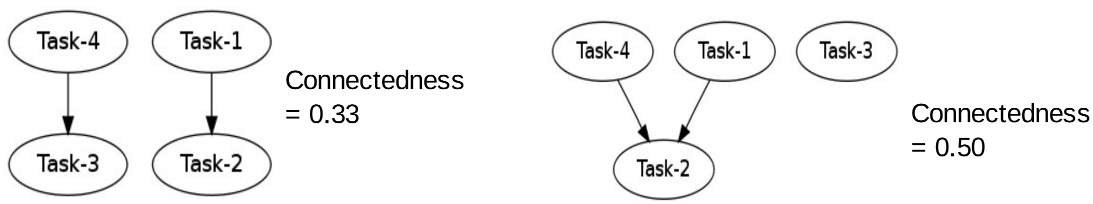


Figure 3.2. Illustrating the difference in connectedness values of two task graphs having the same number of edges and nodes, but different configurations.

3.2.1.4. Subtasks and their completion

Subtasks are the most elementary units of work in a project. The number of subtasks per task is determined by a parameter, *task intensity*. Each subtask is associated with a given number of knowledge areas. This value is determined by a parameter, *knowledge intensity of subtasks*. In a project, each subtask is associated with only one task. Each knowledge area may be associated with more than one subtask. The following rules define the completion of a subtask:

(a) there is no specific order of completion for the subtasks associated with a given task

(b) a subtask is considered complete once an agent has accurate knowledge in all the knowledge areas associated with the subtask[14]

14 In reality, time is spent completing a subtask. However, the time spent in completing a task is not of interest in this study,

(c) if an agent receives inaccurate knowledge from its source, then the knowledge in that area fails the accuracy test

(d) one or more areas of inaccurate knowledge result in an abandoned subtask, if the assigned agent fails to obtain accurate knowledge associated with those subtasks within a pre-specified (indicated by the parameter *maximum number of failed tries*) number of times

(e) once a subtask is completed or abandoned, the agent works on completing the next subtask associated with the task, by obtaining the required areas of knowledge, if needed

3.2.1.5. A workgroup's repertoire of subtasks and knowledge

Given the similarity constraint associated with two consecutive projects in a simulation run (described in the subsection *projects and their completion*), at the beginning of a simulation run, a set of subtasks is assembled such that the constraint is met. The minimum size of this set of subtasks is given in equation 3.8. Based on the size of this repertoire and the value of *knowledge intensity of subtasks*, the set of knowledge areas is determined. The knowledge areas in this set represent the total knowledge required to complete the theoretical set of all the subtasks. During a simulation run, a subset of tasks from the workgroup's repertoire of subtasks is chosen. Similarly, the knowledge areas assigned to various agents form a subset of all the knowledge areas associated with the workgroup's repertoire of subtasks.

3.2.1.6. Assignment of knowledge areas and tasks to agents

At the beginning of each simulation run, each agent is assigned knowledge in every area belonging to the knowledge repertoire described in the previous subsection. The knowledge in each area can be in one and only one of three possible states: *accurate*, *inaccurate*, and *absent*. The initial proportion of areas (at the beginning of a simulation run) in which an agent has accurate knowledge is determined by a parameter, *average knowledge level.* Additionally, the proportion of areas in which the agent has inaccurate knowledge is given by *average inaccurate knowledge level.* Note that the sum of the proportions of accurate and inaccurate knowledge levels need not be 1; the difference

---

so is not considered.

between the sum and 1 indicates the proportion of areas in which the agent has no knowledge. In addition to the actual knowledge (which can be either *accurate*, *inaccurate* or *absent*), an agent has beliefs about the knowledge it posses, that is perceptions of self-knowledge. The parameter, *average self-knowledge*, is used to determine the initial number of areas about which an agent has correct perceptions, that is, an agent believes that it has knowledge when it does have accurate knowledge, or that it lacks knowledge when it does not have knowledge.

It should be noted that the conceptions of each task as a set of subtasks, each subtask being associated with a set of knowledge areas and each agent being assigned a task (and possessing knowledge in a few areas) are based on the Task-Expertise-Person unit proposed by Brandon and Hollingshead (2004).

Tasks, too, are assigned to agents in a random fashion. If the number of tasks is less than the number of agents in a workgroup, then there is at least one agent in the group that will remain idle in the project. This condition represents an “under-worked” workgroup of agents. If the number of agents matches the number of tasks, then an “evenly-matched” condition is said to exist in the workgroup. In projects where the number of tasks exceeds the number of agents, the workgroup is said to be “overworked”. As indicated in table 3.5, the range of the parameters representing number of agents and number of tasks are set such that all three conditions occur at least once.

Assignment of tasks to agents happens at the beginning of each project in the project phase. The assignment is carried out in two phases. First, each randomly-chosen agent is assigned one randomly-chosen task. After this step, if unassigned tasks remain, then a randomly-chosen number of tasks is selected and assigned to a randomly chosen agent. The second step described above is repeated until all the tasks have been assigned to agents. The first step in the two-step process described above ensures that in the “under-worked” and “evenly-matched” conditions, each agent is assigned a task, to the extent that the range of tasks allows such an assignment. This way, a uniformity in the workload of agents is maximized.[15]

3.2.1.7. Knowledge and its exchange

15 Consider a situation where only the second step is implemented. As determined by trial runs, this approach would result in a greater number of simulation runs where a larger proportion of agents would be assigned less than one task, while a smaller number of agents would be assigned more than one task. To reduce the skewness in the distribution of workload, the two-step process was implemented.

In order to complete a subtask, an agent should have knowledge in all the areas associated with the subtask. Each knowledge area is associated with two attributes that describe the efforts needed in articulating it and transmitting it. The attributes *transfer time* (synonymous with "direction", Grant, 1996a) and *stickiness time*, which is based on "stickiness" of knowledge proposed by Szulanksi (1995) capture the notion of difficulty of communicating knowledge by an agent to the agent that requested the knowledge. The following rules describe the completion of subtasks and the associated exchange of knowledge:

1. if an agent possesses the knowledge in all the areas that are required to complete a given subtask, then the subtask is considered to be complete
2. if an agent lacks the required knowledge, then it tries to obtain it from another agent in the workgroup
3. the search for a source for the required knowledge is a two-step process:
   (a) the agent tries to identify, based on the information it has about the areas of knowledge of other agents, the set of agents that are potential sources. From this set, it chooses the agent that it has the highest value of *perceived propensity to share*
   (b) if the agent fails to identify a source, as described in step (a), it utilizes the information available via the expert-seeker, to identify the set of potential source agents. From this set, it chooses an agent randomly, since it perceives all the agents in this set as having equal values of *perceived propensity to share*
4. once a source agent is identified, the recipient determines, based on the priority of the task, location of the source agent and the perceived richness values of various communication media, the medium to use
5. for a task (and subtasks) with a given priority, and the attributes of a communication medium that facilitate (or do not facilitate) synchronous communication, and the location of the source agent and itself, the recipient agent chooses a communication medium with the best richness value (the details of this computation are provided in section 3.2.1.16)
6. next, using the amount of time associated with *transfer* (where the source agent ensures that the recipient receives and understands the knowledge being given)

and *direction* (where the source agent only provides the instructions needed to complete the subtask, without expending the effort needed to ensure the understanding of the received knowledge by the recipient) of the knowledge being provided, *transfer* or *direction* is chosen as the mode of knowledge transmission

7. the result of transmission of knowledge via *transfer* is that the recipient, if it needs to complete another subtask that requires the knowledge that it just received via *transfer*, can complete it without requiring the knowledge from another agent; if the knowledge was transmitted by *direction*, the recipient will complete the current subtask, but if it encounters another subtask in the future requiring the same knowledge, then it needs to obtain the knowledge once again
8. in addition to the task-related knowledge transfer, agents can also exchange information about non-task-related knowledge – this exchange is analogous to the "water-cooler exchanges" described by Davenport and Prusak (2000)
9. it is assumed that the likelihood of exchange of the information is independent of the medium selected
10. non-task-related-knowledge-exchanges are implemented as follows:
    (a) in the training phase, pairs of agents are randomly-chosen
    (b) for each pair, the value of *probability of exchange of information about a non-task-specific-knowledge area* parameter determines whether exchange of information occurs from one agent to another, pertaining to the existence of knowledge in the first agent, in a randomly-chosen area
    (c) next, a randomly-chosen value is compared with the value of *perceived propensity to share* of the first agent to determine whether the agent provides correct or incorrect information about the presence of knowledge in the randomly-chosen area of knowledge
    (d) step (c) is repeated with the original recipient as the source and the original source as the recipient
    (e) steps (a) – (d) above are repeated for all pairs of agents (note: if the number of agents is an odd number, then there will be an agent that does not participate in information exchange during the current time period
    (f) in the project-phase, for every pair of agents that are involved in knowledge-

transmission the following steps are performed

(1) the value of *probability of exchange of information about a non-task-specific-knowledge* is used to determine whether a non-task-related-knowledge-exchange will occur

(2) if (f)(1) indicates that an information exchange will occur, then the source agent's *propensity to share* is used to determine whether it will provide correct information to the recipient about the existence of knowledge in a randomly-chosen area (this area will be different from the area in which knowledge is currently being transmitted)

(3) if step (f)(2) occurs, then it is repeated with the roles of source and recipient reversed with respect to exchange of information regarding the presence of knowledge in a randomly-chosen area in the knowledge-receiving agent

11. at any given instance, during the project phase, an agent can be in of three mutually-exclusive states:
    (a) the agent is a source of knowledge
    (b) the agent is a recipient of knowledge
    (c) the agent is neither a source nor a recipient of knowledge
12. at the beginning of a knowledge transmission transaction, the likelihood that an agent whose knowledge has been sought, does indeed provide the required knowledge to the requesting agent is determined by the value its attribute *propensity to share*; once an agent agrees to provide knowledge, then, in subsequent time periods, it will continue to transmit knowledge (either by *direction* or *transfer*) until all the knowledge has been transmitted
13. the likelihood of an agent reporting information to the expert-seeker about the presence or absence of knowledge in a randomly-chosen area is also governed by the value of its *propensity to share*

Initially, at the beginning of a simulation run, each agent is provided knowledge in a certain number of areas, the proportion of which, averaged across all agents, is defined at the workgroup level by the value of the parameter *average knowledge level.* The proportion of areas in which each agent is assigned knowledge is determined by drawing from an N(mean, mean/3) distribution, where mean = *average knowledge level*. The values of other agent attributes (*self knowledge*, *propensity to share*, *maximum number of failed tries*, *openness to change*) are computed analogously, with the appropriate workgroup-level parameter as the mean.

3.2.1.8. Revision of mutual perceptions by agents

Each agent, through its interactions, develops an understanding of the areas in which the other agents have knowledge. These perceptions of other agents' areas of knowledge and their propensity to share accurate information undergo changes as follows:

1. as discussed in the summary of interview 1, an agent's *propensity to share* consists of two components: (a) a *generalized propensity to share*, which indicates a base-level tendency on part of the agent to share knowledge and information with other agents and expert-seeker*,* and (b) an agent-agent-level *perceived propensity to share* – this indicates agent-agent specific tendency to share
2. given the complexity involved in maintaining source's *propensity to share* values for other agent in the workgroup, it is assumed an agent's *propensity to share* will not change, that is, stated differently, an agent, while acting as a source, will have the same tendency to share with other agents and the expert-seeker knowledge (in the case of agents) and information about the presence of knowledge (in the case of both other agents and expert-seeker), and this tendency remains constant for the duration of a simulation run
3. while the tendency to share is assumed to be constant, a knowledge-seeking agent's perception of other agents' (who could be potential or actual sources) propensity to share, *perceived propensity to share*, is assumed to vary
4. initially, a knowledge-seeking agent's *perceived propensity to share* value with respect to a source agent is same as its own *propensity to share*
5. the *perceived propensity to share* value that a recipient has with respect to a source agent varies based on the accuracy of the knowledge received
6. if the recipient agent receives accurate knowledge, which it uses in completing a subtask, the *perceived propensity to share* of the source agent is incremented by a constant value, which is assumed to be 0.01
7. if the received knowledge is inaccurate, the value of *perceived propensity to share* is

decremented by a constant value, which is assumed to be 0.01

8. the likelihood that an agent revises its value of the other agent's *perceived propensity to share* in the light of the accuracy of the knowledge it receives from a source agent is determined by the parameter *openness_to_change*: higher values of *openness_to_change* imply a greater likelihood of the recipient agent to increment or decrement the *perceived propensity to share* value associated with the source by 0.01

3.2.1.9. Turnover

Turnover is defined as the replacement of an agent by another agent, during the course of a project. The likelihood that, at a given time instance, an agent can be replaced by another agent is determined by a parameter*, turnover*. When an agent replaces another agent, the following constraints should be satisfied:

1. the replaced agent and the replacing agent share a certain amount of common knowledge, given by a parameter, *average proportion of knowledge areas common with the replaced agent*
2. if the replaced agent is in the process of providing knowledge to another agent, then the knowledge transmission is terminated, the recipient increments the number of failures associated with obtaining knowledge in the particular area by one, and, in the next time period, conducts a search for a new knowledge-source agent

3.2.1.10. Expert-seeker

In addition to relying on its own beliefs regarding who knows what[16], an agent may consult an expert-seeker database to determine potential sources of required knowledge. This consultation occurs when the knowledge-seeking agent is unable to determine a source on its own. The expert-seeker simulates the concept of an organizational-skills database (or corporate white-pages) and contains self-reported (that is, agents report to the expert seeker the areas in which they believe that have knowledge) information about all agents' areas of knowledge.

16 Note that these beliefs are based on past interactions with other agents during the training phase, during task-related knowledge-exchanges and non-task-related information-exchanges during the project phase

The likelihood that an agent reveals information to the expert-seeker is determined by the agent-level parameter *propensity to share*. The likelihood that the expert-seeker functionality is available during an experimental-run is 50%, represented by the parameter *use_expert_seeker*. The expert-seeker updates its database by contacting all the agents, each time all of the agents have completed one transaction involving transmission of knowledge.[17]

3.2.1.11. Key attributes of the context and the medium

Based on an interpretation of the second and third sets of interviews, the likelihood of choosing a particular medium for communication is described as being determined by two sets of variables: (a) the attributes of the communication medium, and (b) the relevant aspects of the situation. The relevant attributes associated with a medium are:

1. *synchronousness*: the ability of the medium to facilitate synchronous communication
2. *actual richness of a medium*: which captures the notion of bandwidth for communication of information cues that is physically provided by the medium[18]
3. *locality*: whether the use of the medium is predicated upon both the knowledge-seeker and the knowledge-source being present in physical proximity

The relevant attributes associated with the context in which communication aimed at knowledge exchange occurs are:

1. *stickiness of knowledge*: captures the notion of difficulty associated with transferring knowledge from a knowledge-source to a knowledge-seeker (based on Szulanski, 1995)
2. *task priority*: indicates the level of importance of the task that is assigned to a knowledge-seeker
3. *shared knowledge*: the amount of project-relevant knowledge that is shared by the knowledge-seeker and the knowledge-source

A knowledge-seeker's choice of a particular communication medium is determined by a match between the two sets of attributes that are described above. This is explained next.

3.2.1.12. Relationships between the attributes of the context and the attributes of the

17 It is possible for an agent to remain idle during a time period; the expert-seeker will still contact it.

18 The richness of a medium can be broken down into absolute (discussed here) and perception-based (based on the subjective perceptions and shared knowledge of those using it) components. This decomposition is presented on page 5.

medium

A knowledge-seeker's choice of a communication medium is determined by how well the attributes of a communication medium address the needs as determined by the attributes of the context. Specifically,

1. *stickiness time* of a knowledge area determines the minimum level of *perceived richness* a medium should provide in order to transfer the knowledge. Knowledge of greater *stickiness time* requires a greater amount of effort (as described in the specification document, takes a greater number of time-steps) to be transferred. Hence, a communication medium that facilitates the transmission of a greater number of information cues facilitates the transfer of a greater amount of knowledge per unit of time. Conversely, fewer time-steps will be needed to transfer the same amount of knowledge by a *richer* medium than by a *leaner* medium.
2. *task priority* determines whether the medium that will be selected should facilitate synchronous communication medium, which is captured in the attribute *synchronousness*
3. *locality* is a requirement that determines whether the use of a medium is predicated on both parties involved in communication are located at the same physical location. For instance, *face-to-face*, as a medium of communication, is possible only when both the knowledge-source and the knowledge-seeker are at the same physical location. The use of the remaining three media is not constrained by *locality*
4. *shared knowledge* determines the *perceived richness* of a medium. Ngwenwayama and Lee (1997, pg. 157) explain that organizational context provides a reference schema within which the employees interpret each other's actions and that this shared context would enable the use of a *leaner* medium, that is, one that can convey a fewer number of signals or cues, for conveying a greater amount of information. As explained by Becerra-Fernandez et al. (2004, pg. 45), who cite Argote and Ingram (2000), “common knowledge supports knowledge transfer within the organization”[19]*. By relying on the context-specific knowledge that he/she perceives to share with other members of his/her workgroup, a member of a workgroup can choose an appropriate encoding for his/her messages, e.g., by using specific jargon, metaphors, etc., which he/she believes would be interpreted*

19 Here, *shared knowledge* is treated as being synonymous with *common knowledge*.

appropriately by the recipient of the message, for conveying knowledge that is requested by the recipient of the message. Thus, the sender of a message can convey a greater amount of information via a fewer number of message cues, since the recipient is expected to glean the required information from the message by interpreting the message appropriately, through an application the relevant, context-specific knowledge. Hence, a *leaner* medium can be used to convey *richer* (in terms of information content) messages via a fewer number of cues. In contrast, if the sender believes that the recipient of a message does not have the required context-specific knowledge (that is, the perceived amount of shared knowledge is low), then the sender needs to use more-elaborately-encoded messages for conveying the same amount of information. This implies that the sender of the message will have to use a *richer*, that is, one that provides a greater bandwidth for the transmission of message cues, medium. Thus, when the perceived amount of shared knowledge is high, then a medium, which has a low amount of physical bandwidth, may be perceived to be rich because it is perceived to facilitate the transmission of a greater amount of knowledge via appropriately encoded messages.

Based on the above explanation, *perceived richness* of a medium can modeled as a function of *actual richness* of the medium and *shared knowledge*, i.e., knowledge that is shared between the knowledge-seeker and the knowledge-source.

The richness of a communication medium can be described as consisting of three components:

(a) *actual richness of a medium*, as discussed in point 4 above

(b) *perceived richness inherent to a medium,* which captures the notion that each knowledge-seeking member of a workgroup has a generalized belief regarding the richness of communication that is facilitated by a communication medium

(c) *perceived richness of a medium in relation to a knowledge-source*, which captures the notion that each member of a workgroup, based on his/her belief regarding *perceived richness inherent to a medium,* and the amount of knowledge shared with another member of his/her workgroup who is believed to be the source of knowledge being sought, assigns a certain level of richness to a communication medium that facilitates an exchange of knowledge with the source

3.2.1.13. Computation of the perceived richness of a communication medium in relation to a knowledge-source

Over the course of a single project, and across multiple projects that span the lifetime of a workgroup, workgroup members exchange knowledge in order to complete the tasks that are assigned to them. This exchange can lead to a change in the amount of shared knowledge for any given pair of members. Consequently, for any given pair of knowledge-seeking member and another member who is seen as a knowledge-source by the knowledge-seeking member, the perceived amount of shared knowledge, from the perspective of the knowledge-seeking member, may change. Thus, the *perceived richness of a medium in relation to a knowledge-source* can be modeled as follows.

perceived richness inherent to a medium = f(actual richness of the medium) – (3.3)

perceived richness of a medium in relation to a knowledge-source = f(perceived richness inherent to medium, perceived amount of shared knowledge) – (3.4)

The above-described components of richness of a medium will be implemented in the simulation as follows:

1. the values of *actual richness of the medium* for each communication medium are constant across all simulation runs; they are initialized as described in table 3.2
2. for each agent, the values of *perceived richness inherent to a medium* are initialized for each medium at the beginning of each simulation run, by drawing from an *N(x, x/3)* distribution (values that are negative or lie outside the range of 3 standard-deviations are not assigned), where *x* stands for *actual richness of a medium*
3. the *perceived richness of a medium in relation to a knowledge-source* can be modeled as a function of two variables a) a baseline value that is time-invariant and is represented by *perceived richness inherent to a medium*, and b) a differential, time-variant value that is determined by *perceived amount of shared knowledge*

Mathematically, the description provided in point 3 above can be represented as

$$\textit{perceived richness of a medium in relation to a knowledge-source}_t = \textit{perceived richness inherent to a medium} + c * \log_{n+1}(k_t + 1) \quad \text{– (3.5)}$$

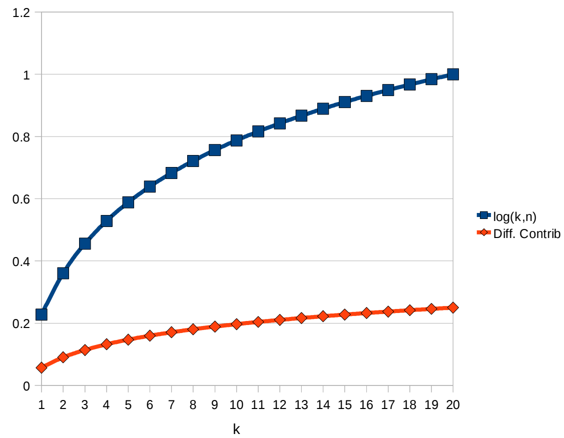
In equation 3.5, k_t represents *perceived amount of shared knowledge* at time *t,* measured as a non-negative integer, *n* represents the total amount of knowledge, measured as a positive integer, and *c* is a proportionality-constant, whose value is chosen to be 0.1. Its purpose is to scale the value of the differential contribution made by *perceived amount of shared knowledge*, such that the value of *perceived richness of a medium in relation to a knowledge-source*$_t$ does not exceed 1.0.[20] The time-based variation in the contribution of *perceived amount of shared knowledge* is a consequence of exchange of knowledge among various pairs of members of the work group.

The expression $log_n k_t$ can be interpreted as the contribution, in terms of information cues, made by k_t number of perceived areas of shared knowledge with respect to *n* (total) number of knowledge areas, to *perceived richness of a medium in relation to a knowledge-source*$_t$. The use of $log_{n+1}(k_t +1)$ in equation 3.5, instead of $log_n k_t$ is based on the observations that, a) if $k = 0$, then the value of the expression $log_n k_t$ is undefined; b) if $k = 1$, then the expression $log_n k_t$ evaluates to 0, which implies that the contribution made by a single perceived area of shared knowledge is zero and is therefore indistinguishable from the contribution made by 0 perceived areas of shared knowledge. Hence, in order to ensure that a) there is distinction made between the contribution made by 0 areas of shared knowledge and 1 area of shared knowledge, as perceived by the knowledge-seeking member, and b) the contribution made by 1 area of shared knowledge (as perceived by the knowledge-seeking member) is not zero, as would be the result of using $log_n k_t$, a value of 1 is added to both *n* and *k* in the expression $log_n k_t$. An additional benefit of doing so is that the value of $log_{n+1}(k_t +1)$ would be 1, when $n = k$, which would be the case even when the expression $log_n k_t$ is used. Hence, the use of $log_{n+1}(k_t +1)$, rather than $log_n k_t$ is appropriate in equation 3.5.

Figure 3.3a depicts the differential contribution made by perceived amount of shared knowledge, which is given by the value of the expression $c * log_{n+1}(k_t +1)$. In the figure, $c = 0.1$ and $n = 20$. Figure 3.3b presents the values of perceived richness of a medium in relation to a knowledge-source for different values of perceived amount of shared knowledge, for three values of perceived richness inherent to a medium: average, maximum possible (average+3 standard-deviations), minimum possible (average-3 standard-

20 The effect of *c* was estimated by assigning it values from the range 0.1 – 0.5. Values of $c > 0.1$ resulted in the values of *perceived richness of a medium in relation to a knowledge-source*$_t$ that exceeded 1.0. Hence, 0.1 is the appropriate value of *c*.

deviations).



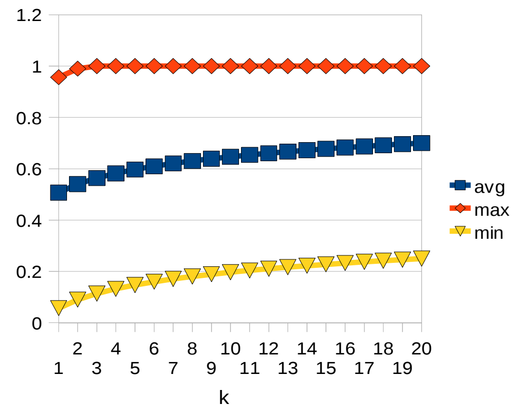


Figure 3.3a. The differential contribution of *perceived amount of shared knowledge*

Figure 3.3b. The value of *perceived richness of a medium* in relation to a knowledge-source in three scenarios

3.2.1.14. Synchronousness and locality

Different communication media offer different possibilities of facilitating synchronous communication and impose the constraint of whether both the source and recipient of knowledge need to be present at the same physical location. The interviewees indicated that in a given instance, in addition to the urgency (seen as synonymous with task priority), their choice of a communication medium depends on whether the communication had to be synchronous and also whether the individuals with whom they were communicating were in close physical proximity.[21] This information is summarized in table 3.3.

Table 3.2. Baseline values of *perceived richness inherent to a medium* of various communication media

Medium	***Actual Richness***	***Rationale***
Face-to-face	1.0	Face-to-face communication facilitates the transmission of all three types of cues: aural, visual and textual, e.g., through the use of notes, whiteboard, etc. The use of a whiteboard in a face-to-face scenario has been mentioned in two of the follow-up interviews (first and second), and in the first main interview of the second set of interviews. *Face-to-face*, as a medium, is identified in literature as providing the maximal richness of cues. Hence, it is assigned a value of 1.0, of which 0.6 is attributed to visual cues (including body language, diagramming on paper/whiteboard, etc.), 0.3 to aural cues (including tone of the voice) and 0.1 to textual cues (notes, equations, etc.)

21 Specifically responses to questions *d* and *i* from interview 1; responses to questions *c* and *e* from interview 2 and response to question *f* from interview 3.

Medium	Actual Richness	Rationale
Phone	0.45	While only one (voice) out of the above-mentioned three types of information cues can be transmitted via this medium, it is possible to convey additional information, for instance, by both source and recipient using a whiteboard or a piece of paper to describe a diagram to each other and discuss the diagram over the phone. However, the transmission of any type of visual cues, such as body-language, images or video is not possible. Hence, the value is computed as 0.3 + 0.1 + 0.05 = 0.45 (0.3 for aural cues, 0.1 for information that is transmitted via aural cues, but pertains to visual information, e.g., description of a picture, a diagram, a scene, etc., and 0.05 for the transmission of aural cues pertaining to textual information, e.g., description of a mathematical equation). Admittedly, these assignments are subjective. However, they capture the notion that a phone call can facilitate the transmission of more than just aural information.
Text chat	0.25	0.1 + 0.0 + 0.15 (using the rationale described above for textual, aural and visual (transmission of images/diagrams, etc.) information, respectively)
E-mail	0.25	0.1 + 0.0 + 0.15 (using the rationale described above for textual, aural and visual (transmission of images/diagrams, etc.) information, respectively)

Table 3.3. *Synchronousness* and *Locality* of various media

Medium	Synchronousness	Locality
Face-to-face	Yes	Yes
Phone	Yes	May be
Chat	Yes	May be
E-mail	No	May be

3.2.1.15. Stickiness

From an analysis of the interviews in sets 2 and 3, it is evident that *stickiness time*, which indicates the difficulty associated with transferring knowledge from a knowledge-source to a knowledge-seeker, affects the choice of the medium: the interviewees preferred choosing a medium that provided higher level of perceived richness when the knowledge to be exchanged had a higher value of *stickiness time*. Each agent has a value of *perceived stickiness time* associated with each knowledge area in its knowledge vector. The above definition of *stickiness time* (and therefore, of *perceived stickiness time*) needs to be extended in order to simplify the modeling of the effect of *perceived stickiness time* on the choice of a communication medium. Let *stickiness time* of knowledge in a particular area be defined as the number of time periods that are needed for the knowledge in that area to be transferred from a source to a recipient when *face-to-face* is the medium used for communication.[22] This definition implies that, when a medium with a lesser value of

22 The definition of *perceived stickiness* should be analogously extended, as described in point 45 of the simulation

perceived richness of a medium is used to communicate knowledge in a specific area of knowledge, the number of time-periods needed would be greater than those needed when *face-to-face* is the medium of communication. Thus, the actual time[23] taken when a medium *m* other than *face-to-face* is used for transferring knowledge in an area *ka* would be:

$$actual\ time_{m,ka} = face\text{-}to\text{-}face\text{-}specific\ stickiness_{ka} / (\ perceived\ richness\ of\ m\ in\ relation\ to\ a\ knowledge\text{-}source) \qquad -(3.6)$$

Given that *stickiness time* values are measured in positive integers, and *perceived richness* values[24] in positive real numbers, the value of the expression on the right-hand-side of equation 3.6 is rounded up to the next highest integer.

In contrast to the actual time that is taken for transferring knowledge of a particular stickiness, via a medium with a specific value of perceived richness with respect to a knowledge-source, i.e., the value computed in equation 3.6, an analogous value that is described as the *expected time needed to transfer knowledge via a medium of a specific value of absolute richness* can be computed (equation 3.7). This value, denoted by *expected time*$_{m,ka}$, is based on the observations that (a) each medium provides a specific (limited) amount of physical bandwidth, and, (b) for each pair of knowledge-seeker and knowledge-source, the actual amount of shared knowledge can be determined.[25] The term *expected time*$_{m,ka}$ represents a theoretical value that indicates the time that should be taken, ideally, to transfer knowledge via a medium, if the knowledge-seeker has perfect information about all aspects of his/her environment, that is, the knowledge-seeker knows the exact amount of bandwidth that is provided by a medium and knows the exact amount of knowledge that the knowledge-source has, and, therefore, can correctly interpret the cues that are sent via the medium and understand them in the exact way that was intended by the knowledge-source. Described differently, the term *expected time*$_{m,ka}$ represents the time required to transfer knowledge from a knowledge-source to a knowledge-seeker when (1) every member of the workgroup has perfect information about the physical bandwidth

specification.

23 Recall that the unit for measuring of stickiness of a knowledge area is number of time periods

24 The labels *stickiness* and *perceived richness* are used here for the sake of brevity. They imply the appropriate, fully-named parameter names.

25 This value differs from the *perceived amount of shared knowledge*, because it is computed by the observer (that is by the simulation) and not by the *knowledge-seeker*; also, this value would not be used by the *knowledge-seeker*

provided by each medium, (2) will correctly encode and decode messages that are framed to completely utilize the available bandwidth (that is, there is no noise associated with the signal) and (3) there is no ambiguity associated with the messages. Hence, no time is spent by the knowledge-seeker/sender in dealing with ambiguity associated with incorrect interpretation of the messages.

The term *expected time*$_{m,ka}$ represents a theoretical value against which the value of *actual time*$_{m,ka}$ can be compared. Such a comparison would indicate the efficiency of communications, which occur in the workgroup, as a result of the choices that the workgroup members make regarding the communication media.[26] The term *expected time*$_{m,ka}$ will be computed as:

$$\text{expected time}_{m,ka} = \text{face-to-face-specific stickiness}_{ka} \,/\, (c * \log_{n+1}(K_t+1)) \qquad -(3.7)$$

where K_t indicates the actual amount of shared knowledge between the knowledge-seeker and the knowledge-source. The terms *face-to-face-specific stickiness*$_{ka}$ and *c* have the same meaning as they do in equations 3.5 and 3.6, respectively.

3.2.1.16. Computation of the choice of a communication medium

Analyses of the interviews 2 and 3 indicate that, in addition to considering the factors discussed in the previous sections, interviewees choose a medium based on the priority of the tasks that they are trying to complete. Specifically, it was found that tasks of higher priority led the interviewees to use media that facilitated synchronous communication. Table 3.4 enumerates the relationship between task priority and synchronousness.

Table 3.4. Linking *task priority* to *synchronousness*

Task priority	*Synchronousness*
High	Yes
Medium	Sometimes/may be
Low	No

26 It should be noted that the assignment of *perceived value of the medium* values to each agent, captures the fact that, in reality, a person's perception of the richness of a medium can differ from the actual richness (physical bandwidth) of the medium. Hence, the agents in the simulation make biased (partially determined by the initial random assignment of value to the time-invariant component of the perceived richness of the medium) choices of communication media. The efficiency of their knowledge-exchange activities, which indicates how quickly they can exchange knowledge with each other to complete the group's tasks, is also determined, to some extent, by this inherent (and randomly-introduced) bias.

Based on the discussions presented so far, the assignment of a communication medium to an agent, at each time-step in the simulation, will based on the information presented in tables 3.2 – 3.4. A knowledge-seeking agent follows the following steps, once a source-agent is identified by the knowledge-seeking agent:

i. the *locality* of the source agent is determined; this determines whether *face-to-face* can be considered as a possible medium (table 3.3)
ii. the *perceived richness* of each medium is computed for the given pair of agents (table 3.2 and equation 3.5)
iii. based on the priority of the task, the *synchronousness* need is determined (table 3.3)
iv. the media, which have the required values of *locality* and *synchronousness* are selected
v. from the above set of media, the medium with the highest value of *perceived richness of a medium in relation to a knowledge-source* is chosen

3.2.1.17. Computation of the size of the workgroup's repertoire of subtasks

The minimum number of unique tasks required to ensure that the similarity constraint (similarity between two successive projects) is determined as follows.
Let x_1 be the number of subtasks in $project_1$ and x_2 be the number of subtasks in $project_2$ and p_{sim} be the value of project similarity, that is, the proportion of subtasks common to $project_1$ and $project_2$, such that $project_2$ immediately follows $project_1$. Then the number of unique subtasks in $project_1$ is given by $x_1 – x_1*p_{sim}$. Similarly, the number of unique subtasks in $project_2$ is given by $x_2 – p_{sim}*x_2$. Therefore, the total number of unique subtasks in $project_1$ and $project_2$ is given by

$$Number\ of\ unique\ subtasks = (1-p_{sim})*x_1 + (1-p_{sim})*x_2 + p_{sim}*x_1 \qquad – (3.8.i)$$

Since project similarity is not transitive, $project_3$ that immediately follows $project_2$ may be entirely (or partly) similar to $project_1$, so we only need to ensure that the uniqueness constraint is satisfied only for any two consecutive projects. To ensure that enough unique subtasks area available across all the projects in a workgroup's lifetime (assuming that the number of projects is fixed, that is, every workgroup has the same number of projects), we need

$$minimum\ number\ of\ subtasks = (1-p_{sim}) * x + (1-p_{sim}) * x + p_{sim} * x \qquad -(3.8.ii)$$

where *x* refers to the maximum of the two values x_1 and x_2. Rearranging the above equation, after substituting

x = *maximum value of number of tasks per project * maximum number of subtasks per task.*

we get

*minimum value of workgroup's repertoire of subtasks = (2 – value of similarity) * maximum value of number of tasks per project * maximum number of subtasks per task – (3.8)*

By making the simplifying assumption that the value of number of tasks per project (described by the parameter *project_intensity*) is fixed across all simulation runs (to ensure that the dataset that results from the simulation is a balanced panel (Baum, 2007), the values of minimum and maximum in the above equation may be substituted by similarity by value of workgroup's repertoire and value of number of tasks per project, respectively.

3.2.1.18. Experiment

Each experiment, or a simulation run, consists of two phases:

1. a training phase, where:
   (a) agents interact with each other and exchange information about the areas in which they have knowledge
   (b) exchange of information occurs a fixed number of times
   (c) during each exchange, a randomly-chosen pair of agents exchange information about a randomly-chosen chosen area of knowledge
   (d) each agent indicates to the other agent whether it possesses knowledge in the randomly-chosen area
   (e) during an exchange, each agent takes turns acting as a source and recipient of information
   (f) the likelihood of providing information is governed by a parameter, *probability*

of exchange of information about a non-task-specific-knowledge area

2. a project phase, where:
   (a) agents work on completing a series of projects
   (b) the project phase is complete once all the projects associated with the simulation run are complete
   (c) upon completion, a project will have a proportion of completion ranging from zero (all tasks were abandoned) to 1 (all tasks were completed)

3.2.1.19. Outcome variables

At the end of each project, the following metrics are computed:

1. *average consensus*: *average consensus* about each other's areas of knowledge; consensus indicates the workgroup's transactive memory (Austin, 2003)
2. *average accuracy of knowledge* (averaged across all agents) in which an agent has knowledge
3. *percentage of project completed*: number of subtasks completed/total number of subtasks

The following equations describe the above-mentioned computations[27]:

$$\textit{Percentage of project completed} = \frac{\textit{no. of tasks completed}}{\textit{no. of tasks completed} + \textit{no. of tasks abandoned}} \qquad \textit{– (3.9)}$$

$$\textit{Accuracy of knowledge in one agent}^{28} = \frac{\textit{no. of areas where an agent has accurate knowledge}}{\textit{no. of knowledge areas}} \qquad \textit{– (3.10)}$$

$$\textit{Average accuracy of knowledge across all agents} = \frac{\sum \textit{accuracy of knowledge of an agent}}{\textit{no. of agents}} \quad \textit{for all agents}$$

– *(3.11)*

–

27 Where appropriate, the standard deviation and the skewness values are computed using procedures that are analogous to the computation of averages (means)

28 While knowledge can be seen from objective and subjective perspectives, in this study it is assumed that the accuracy of knowledge all agents of a workgroup is evaluated by the same objective standard: can the knowledge be applied to the completion of the specific subtask for the completion of which it was obtained by the knowledge-seeking agent to whom the subtask is assigned.

$$Agreement\ between\ two\ agents\ about\ another\ agent's\ areas\ of\ expertise =$$

$$\frac{\sum no.of\ areas \text{ in which } the\ pair\ believes\ the third\ agent has\ accurate knowledge + \sum no.of\ areas \text{ in which } the\ pair\ believes the\ third\ agent has\ inaccurate knowledge}{total\ no.of\ areas \text{ in which } at\ least\ one\ agent \text{ in } the workgroup believes that\ atleast\ one other agent\ has knowledge} \quad -(3.12)$$

$$Consensus\ of\ the\ workgroup = \frac{\sum agreement\ values\ for\ every\ agent \in the\ workgroup}{total.no.\ of\ agents \in the\ workgroup} \quad -(3.13)$$

The following table (table 3.5) describes the various parameters that describe various aspects of the simulation.

Table 3.5. Simulation Parameters

Parameter	***Description***	***Range***	***Basis for inclusion***
Average proportion of knowledge areas common with the replaced agent	Average proportion of knowledge that is common between a replaced and a replacing agent during the execution of turnover ┼	[0, 0.9]	Builds on the model assumptions presented in March (1991) and Kane and Alavi (2008), where *turnover* is operationalized as an agent replacing a existing agent. The extension here refers to making explicit, and controlling, the degree of similarity of knowledge areas between the replacing and the replaced agents.
Average direction time	Average time for transmitting knowledge via direction ┼	[1, 3] (integers)	Grant (1996a) describes the notion of transmission of knowledge via direction.
Average propensity to share	Average value of generalized potential to share ‖	[0, 0.9]	Interviews (section 3.1 and appendices A2 and A3)
Average inaccurate knowledge level	Average proportion of areas of knowledge in which an agent has inaccurate knowledge ┼	[0, (1 – average knowledg e level]	March (1991) operationalized the notions of correct and incorrect knowledge in the case of organizational exploration and exploitation of knowledge.
Average knowledge level	Average proportion of areas of knowledge in which an agent has accurate knowledge┼ at the beginning of a simulation run	[0, 0.9]	Kane and Alavi (2007) operationalized the notion of knowledge level of organizational agents
Average openness to change	Average openness to change – determines the likelihood of an agent modifying its perception of the *propensity to share* of its source agent based on the accuracy of the knowledge it received ‖	[0, 0.9]	Interviews (particularly #2) indicate that a recipient of knowledge is open to changing his/her opinion of the source of the knowledge, based on the accuracy of the knowledge provided.
Average self knowledge	Average proportion of areas in which an agent has correct perceptions about the accuracy of its knowledge ‖	[0, 0.9]	Assumption, derived from Wegner 1987, Wegner et al., 1991, Nickerson, 1999, Brandon and Hollingshead, 2004, according to whom, in order for transactive memory to develop, each individual in a group should have an awareness

Parameter	***Description***	***Range***	***Basis for inclusion***
			of one's own knowledge which would then inform one's decision of seeking knowledge from others in the group in areas where one lacks knowledge.
Average knowledge intensity of subtasks	Average number of knowledge areas per subtask [§]	[1, 3]	
Average stickiness time	Average time for transmitting knowledge via transfer [+]	[1, 3] (integers)	Szulanski's (1995) notion of stickiness of knowledge.
Average task intensity	Average number of subtasks per task [a]	[1, 3]	Ren et al., (2006, pg 676)
Average task priority	Average priority of a task [a]	[1, 3]	Espinosa and Pickering (2006) indicate the importance of task priority on coordination of team efforts and outcome (success/failure) of projects.
Average number of tasks per agent	Average number of tasks assigned to an agent		Urban et. al. (1995) indicate the importance of workload (interpreted here as the average number of tasks assigned to an agent) on team performance.
Turnover	Likelihood of an agent being replaced by another agent during a given time period [¤]	[0, 1.0]	0, 0.1, 0.3 (Kane and Alavi, 2007, pg 804) - the values used here are from a wider domain of values
Maximum number of failed tries	Maximum number of tries an agent tries to obtain knowledge in an area associated with a subtask before abandoning the subtask [¤]	[6, 10]	Trial runs indicated that this range resulted in a greater variance in the value of project completion percentage
Media richness of chat	Average richness of chat (text-based instant messaging/SMS messaging) as the communication medium [¤] used to compute the perceived richness of a medium, for each agent by drawing from N(0.25, 0.25/3)	0.25	
Media richness of email	Average richness of e-mail as the communication medium [¤], used to compute the perceived richness of a medium, for each agent by drawing from N(0.25, 0.25/3)	0.25	
Media richness of face-to-face	Average richness of face-to-face as the communication medium [¤]	1.0 (fixed)	Interviews (section 3.1and appendix A3)
Media richness of telephone	Average richness of telephone as the communication medium	0.45	Interviews (section 3.1and appendix A3)

Parameter	***Description***	***Range***	***Basis for inclusion***
	▫used to compute the perceived richness of a medium, for each agent by drawing from N(0.45, 0.45/3)		
Number of agents	Number of agents per workgroup ▫	[2, 15] (integers)	3 agents, 60 groups (Ren et al., 2006, pg 676). 3-person groups (Lewis et al., 2005, pg 588).
Number of locations	Number of geographic locations ▫	[1, 5]	Faraj and Sproull's (2000) study indicates the effect of locational diversity on the development of transactive memory system in a geographically distributed team
Number of projects per workgroup	Number of projects per simulation run ▫	10[fixed] ¢	Austin (2003), Lewis et. al. (2005) studied teams working on multiple projects, over a period of time. These studies inform the inclusion of multiple projects. The size is fixed at 10 to (a) ensure that panel data is balanced (Baum, 2007) and (b) initial trial runs indicated that projects in the range 7-15 produced a greater variance in the outcome variables.
Probability of exchange of information about a non-task-specific-knowledge area	Likelihood of a non-task-specific-knowledge-related-exchange occurring ▫	N(0.5, 0.5/3)	Captures the notion of non-task-specific-knowledge exchanges described as “water cooler conversations” by Davenport and Prusak (1998)
Similarity of projects	Similarity between two consecutive projects in a simulation run ▫	[0, 1]	Lewis et al. (2005, pg. 596) referred to functional similarity across tasks as contributing to the formation of TMS in a group; however, they described this as a direction for future research
Average project intensity	Number of tasks associated with a given project.	[1, 12] integer values	
Time in training phase	Time allocated to the training phase of a simulation run	10[fixed]¢	Ren et al., 2006, pg 676, used 50 training periods
Use expert-seeker	Availability of the expert-seeker database by an agent in the workgroup ▫	Yes/No	Becerra-Fernandez (2000) describes the “expert-seeker” people-finder software used at NASA; evidence from interview 1.
Connectedness of network of task-interdependencies	Indicates the degree of to which the network graph representing the tasks and their precedences is connected. Greater the connectedness value, greater is		

┼ averaged across all knowledge areas in a simulation run

‖ averaged across all agents in a simulation run

§ averaged across all subtasks

▫ constant across simulation run

ª averaged across all tasks in a simulation run

¢ fixed, to ensure that the data represents a balanced panel (Frees, 2004; Baum, 2007)

Several steps were taken to ensure the fidelity of the implementation of the model with the specifications of the model. An overview of the procedure is presented in subsection 3.2.3, with details describing the verification of the implementation in appendix A4. A flowchart depicting the flow of the implementation of the model is presented in appendix A5. The source of the implementation is provided in appendix A6.

3.2.2. Implementation of the model

The choice of the programming language and environment for implementing the simulation model was based on the following criteria, derived from my personal experience with various computer languages:

1. Support for multiple programming paradigms, specifically, object-oriented and functional programming. This criterion was necessary because (a) object-orientation facilitates the development of the implementation in a modular fashion, with classes and their objects representing various components of the model, and (b) the syntax associated with functional programming makes it easy to write code that operates on collections of entities in a concise and elegant form.
2. Availability of libraries necessary for performing the necessary statistical and network-related computations. This criterion is essential because the stochastic nature of various components of the model and the network of relationships among the agents require the use of statistical and social-network related computations.
3. Dynamic, strong typing in the programming language: strong typing ensures that objects once created are treated uniformly and correctly across their lifetime; dynamic typing ensures that a variable can be used without having to declare and fix its type ahead of time. While this leads to a compromise in the execution time, it also leads to a decrease in the coding time and this compromise was deemed acceptable.
4. Availability of tools to run in a GNU/Linux environment: this is based on my personal preference of using free/open source software over proprietary software, including proprietary operating systems.

Based on the above criteria, the environment chosen for implementing the model consisted of the Python programming language[29], along with the Numpy numerical algorithms

29 http://www.python.org

library[30], and the R statistical environment[31] for doing network-related computations. The RPy library[32] was used as a bridge through which the code written in Python could create and manipulate and read output of code in R that dealt with the network-related computations. The simulation was implemented and executed on hardware running Ubuntu desktop and server operating systems[33].

3.2.3. Verification of the implementation of the simulation

To verify the implementation of the simulation the following procedures were followed:

1. The source code of the simulation's implementation was carefully reviewed to ensure that it addresses all aspects of the model that were described in chapter 3. The use of object-oriented paradigm for creating the program ensured that the mapping of the key concepts between the model's description and their implementation, in terms of various classes, their instantiations, their attributes and accompanying methods, along with appropriate modules and helper functions ensured that the complexity of the implementation was manageable and also facilitated the mapping of the model's specifications to source code. Comments are included at appropriate locations in the source code to help a reader understand which aspects of the model are being implemented, along with any low-level implementation details, where necessary.
2. The implemented program was run in a “debug” mode, where messages are logged from each individual method associated with each object in the code, and from each key decision point within each method of each object. The log files were then reviewed carefully to ensure that the following conditions were met:
   (a) the code implemented all aspects of the model
   (b) the changes in state of each object of the simulation and the simulation's state, as a whole, were valid and correct
   (c) the values of each parameter that was modified and each outcome variable were within expected ranges
3. Additionally, the results of the simulation output were analyzed using the following procedure:

30 http://numpy.scipy.org/
31 http://cran.r-project.org/
32 http://rpy.sourceforge.net/
33 http://www.ubuntu.com/

(a) For each contingency factor, the highest and lowest values were identified.
(b) For each contingency factor and for each of the maximum and minimum values, the set of values corresponding to each of the three outcome variables were identified. The result is that for each contingency variable, there were two sets for values for each of the three outcome variables.
(c) Next independent-sample (Welch's) t-tests were run, to determine whether workgroups corresponding to the "minimum" condition of a contingency factor are different from workgroups corresponding to the "maximum" condition of a contingency factor, for a given outcome variables.
(d) Steps (a)-(c) were repeated for each contingency factor and outcome variables.

The results of the t-tests are presented in table A4.1. The implications of the results in terms of whether they validate the implementation of the simulation are presented in the next. The results of the t-tests indicate that, overall, the results are consistent with expectations that were based on intuition and reasoning. Thus, it was considered that the implementation of the simulation is valid. This approach is consistent with the “predictive validity” criterion used for validating the implementation of agent-based models (Tesfatsion, 2010)

3.3. Overview of the analysis method

The data obtained via the simulation were analyzed using panel data regressions (Frees, 2004; Baum, 2007). Panel data refers to data that are organized such that each unit of measurement – a panel, which can be an individual, a workgroup, a nation, etc., - is surveyed repeatedly over a fixed number of time periods, to record the values of different predictor (regressor) variables on different outcome (regressed) variables. When the number of time periods for which the data are collected is the same across all panels, the dataset is said to be 'balanced'. If data are not available for each panel for all the time periods, then there are said to be missing data and the panel dataset is said to be 'unbalanced' (Baum, 2007).

The predictor variables in a panel dataset may vary within a panel across time periods (within a workgroup, in this dissertation, since each workgroup is treated as the unit of analysis), they may vary between workgroups, or they may vary both within and between panels. Based on whether a group of predictors varies within or between panels, the regression coefficients of the predictors are estimated by assuming fixed effects of the regressors, if the regressors vary within a workgroup or by assuming between effects, if the regressors vary between panels. Fixed-effects-based regressions "permit each cross-sectional unit to have its own constant term while the slope estimates are constrained across units, as is the variance" (Baum, 2007; pg 221). Consider the following regression equation:

$$y_{it} = x_{it}\beta_k + z_i\delta + u_i + \varepsilon_{it} \qquad -(3.14)$$

In equation 3.14, x_{it} *is a* 1 x *k* vector of variables that vary over panels and time, *β* is the *k* x 1 vector of coefficients on *x*, z_i is a 1 x *p* vector of time-invariant variables that vary only over panels and *δ* is the *p* x 1 vector of coefficients on *z*, u_i is the individual level effect, and ε_{it} is the disturbance term.[34] In both fixed-effects and random-effects regressions, the u_i values are always assumed to be uncorrelated with ε_{it}. However, the u_i values may or may not be correlated with the regressions in x_{it} and z_i. If the u_i values are uncorrelated with the regressors, the regression model is called *random-effects* regression, whereas if they are correlated, the regression model is called a fixed-effects model. To determine which model is most appropriate, Hausman test and test of over-identification via the Sargan-Hansen statistic (Schaffer and Stillman, 2006) of the models need to be performed. These tests indicate whether the assumption that the individual-level effects are uncorrelated with the regressors is warranted, and therefore, whether a random-effects regression is appropriate.

In between-effects regressions, "the group means of *y* [the dependent variable] are regressed on the group means of *x* [the vector representing predictor variables] in a regression of *N* observations. This estimator ignores all the individual-specific variation in *y* that is considered by the within estimator, replacing each observation for an individual with his/her mean behavior" (Baum, 2007; pg 226).

Based on whether they varied within a workgroup or across workgroups, the contingency factors were categorized into "within-group" contingency factors and "between-groups" contingency-factors and used as predictors in panel regressions involving the corresponding estimation methods, that is, within-group contingency factors were used as

34 Equation 3.14 and the associated descriptions of variables were obtained from Baum, 2007; pg 220

predictors in panel regressions with within-group estimation and between-group contingency-factors were used as predictors in panel regressions with between-group estimations. Details of these regressions, along with the results that identify which contingency factors have significant within-group effects and which contingency factors have significant between-group effects, are presented in chapter 4, *Results*.

3.4. Summary

In this chapter, details of the two components of the methodology used in the dissertation are presented. Section 3.1 described the purpose of the semi-structured interviews, and explained the support obtained via the interviews for some key aspects of the agent-based model. In section 3.2, the specifications of the agent-based model were described in detail. Finally, section 3.3 provided an outline of the statistical method of data analysis called panel data analysis, which was used to analyze the data obtained via simulation runs of the agent-based model.

In the next chapter, *Results,* details of the different panel data regressions are presented, along with their interpretations in the light of extant literature and the specifications of the agent-based model. The subsequent chapter, *Discussion*, explains the results further and explains the contributions made by this dissertation to research in the area of transactive memory and workgroup process outcomes. It identifies the limitations of this dissertation and its implications to practitioners and for future research that can be based on the current work.

4. Results

This chapter presents the results from the regression analyses, followed by brief interpretations of the results. As described in section 3.3, the methodology of panel data analysis was used to investigate the effects of the five categories of contingency factors on the three outcome variables. These effects were studied using two sets of regressions: the first set involved those factors which were shown to have only time-invariant components of standard deviations The second set of regressions included those contingency factors that were observed to have both time-invariant and time-variant components of standard deviations. The details of the regressions and their results are presented in this chapter.

This chapter is organized as follows. First, the classification of the contingency factors into “within-group” and “between-group” categories is described. Then, the results of regressions of the contingency factors on each of the three outcome variables are presented and interpreted. Finally, a summary of the current chapter is presented, along with an brief introduction to the next chapter, *Discussion*.

4.1. Summaries of independent variables

The design of the simulation experiments helped categorize the contingency factors into within-group predictors and between-group predictors. Specifically, the following contingency factors, which are initialized once, at the beginning of a simulation run, are expected to be predictors that are included in between-group effects regression[35]: *average knowledge level*, *average openness to change*, *average self-knowledge*, *probability of turnover*, *average proportion of knowledge areas common with the replaced agent*, *number of agents, number of locations*, *probability of exchange of information about a non-task-specific-knowledge area*, *time in training phase*, *average propensity to share*, *maximum number of failed tries*, *average direction time*, *average stickiness time*, *average project intensity*, *use expert-seeker*, *richness*[36] *of telephone*, *richness of email*, and *richness of text-based chat*. The contingency factors *time in training phase*, *project intensity*, *task intensity*, and *subtask intensity of knowledge* were considered to be the predictors in the between-

35 In later sections of this chapter, this set of contingency factors is cited as the time-invariant set of contingency factors; those contingency factors that are known to have (as described in table 4.1) both within-group and between-group components of standard-deviations, are cited as the time-variant set of contingency factors.

36 It should be noted that even though other attributes of electronic communication media, *synchronousness* and *locality* are incorporated into the simulation model, only the richness attributes of the communication media varied across groups; the former two attributes of the communication media are held constant across all simulation runs (workgroups) and are therefore not included in the analyses.

effects regressions. Support for the above categorization was sought via a summarization of the variables used in the regressions using the *xtsum* function of Stata, which provides a decomposition of the standard deviation of each predictor into within-group standard-deviation and between-group standard-deviation.

It must be noted that based on the values of *average direction time*, *average stickiness time*, *average project intensity*, *average task intensity*, *average subtask intensity of knowledge*, and *average task priority* at the beginning of a simulation run, the value of the specific number of tasks for a given project, and the specific subtasks assigned to the specified number of tasks, are selected at the beginning of each project in the series of projects that constitute a simulation run (a workgroup's lifetime). Hence, once the project-level assignments of tasks, subtasks to tasks and assignment of tasks to agents are made, the values of each of the above contingency factors, when computed, are going to be different for each project. This implies that the above contingency factors vary *within* each project in a workgroup's lifetime, that is, each of the above contingency factors varies within each panel, across time.

In table 4.1, the values in bold indicate the specific, significant standard-deviation components of a variable. The results include all the contingency factors and outcome variables. As described earlier in this section the contingency factors *average direction time*, *average stickiness time*, *average project intensity*, *average task intensity*, *average subtask intensity of knowledge*, and *average task priority*, all have significant within-group and between-group components of standard-deviations.

Table 4.1. Summaries of variables

Variable		Mean	Std. Dev.	Min	Max	Observations
Use expert-seeker	overall	0.517	0.500	0.000	1.000	N = 25000
	between		**0.500**	0.000	1.000	n = 2500
	within		0.000	0.517	0.517	T = 10
Average propensity to share	overall	0.000	1.000	-1.722	1.732	N = 25000
	between		**1.000**	-1.722	1.732	n = 2500
	within		0.000	0.000	0.000	T = 10
Average incorrect knowledge-level	overall	-0.000	1.000	-1.207	3.027	N = 25000
	between		**1.000**	-1.207	3.027	n = 2500
	within		0.000	-0.000	0.000	T = 10

Average knowledge-level	overall	0.000	1.000	-1.543	1.546	N = 25000
	between		**1.000**	-1.543	1.546	n = 2500
	within		0.000	0.000	0.000	T = 10
Average openness to change	overall	-0.000	1.000	-1.728	1.733	N = 25000
	between		**1.000**	-1.728	1.733	n = 2500
	within		0.000	-0.000	0.000	T = 10
Average self-knowledge	overall	-0.000	1.000	-1.541	1.589	N = 25000
	between		**1.000**	-1.541	1.589	n = 2500
	within		0.000	0.000	0.000	T = 10
Average maximum number of failed tried	overall	-0.000	1.000	-1.426	1.414	N = 25000
	between		**1.000**	-1.426	1.414	n = 2500
	within		0.000	0.000	0.000	T = 10
Time in training phase	overall	0.000	1.000	-1.563	1.565	N = 25000
	between		**1.000**	-1.563	1.565	n = 2500
	within		0.000	0.000	0.000	T = 10
Probability of exchange of non-task-related knowledge	overall	-0.000	1.000	-2.960	3.089	N = 25000
	between		**1.000**	-2.960	3.089	n = 2500
	within		0.000	0.000	0.000	T = 10
Average direction-time	overall	0.000	1.000	-8.282	8.285	N = 25000
	between		**0.409**	-1.426	1.277	n = 2500
	within		**0.912**	-7.963	7.456	T = 10
Average stickiness-time	overall	0.000	1.000	-8.296	8.281	N = 25000
	between		**0.405**	-1.422	1.603	n = 2500
	within		**0.914**	-8.049	8.160	T = 10
Average richness of chat	overall	0.000	1.000	-0.183	16.983	N = 25000
	between		**1.000**	-0.183	16.983	n = 2500
	within		0.000	0.000	0.000	T = 10
Average richness of email	overall	0.000	1.000	-0.225	12.577	N = 25000
	between		**1.000**	-0.225	12.577	n = 2500
	within		0.000	0.000	0.000	T = 10
Average richness of phone	overall	0.000	1.000	-0.190	10.574	N = 25000
	between		**1.000**	-0.190	10.574	n = 2500
	within		0.000	0.000	0.000	T = 10
Average Consensus	overall	0.000	1.000	-1.735	1.432	N = 25000
	between		**0.996**	-1.735	1.426	n = 2500

	within		0.095	-1.897	1.537	T = 10
Average accuracy of knowledge	overall	0.000	1.000	-1.803	2.322	N = 25000
	between		**1.000**	-1.803	2.322	n = 2500
	within		0.002	-0.061	0.070	T = 10
Percentage of project completed	overall	0.000	1.000	-1.040	4.712	N = 25000
	between		**0.807**	-1.040	3.766	n = 2500
	within		**0.591**	-3.552	5.119	T = 10
Average project intensity	overall	0.000	1.000	-1.589	1.602	N = 25000
	between		**0.308**	-1.009	1.051	n = 2500
	within		**0.951**	-2.408	2.495	T = 10
Average knowledge-intensity of subtasks	overall	0.000	1.000	-4.538	4.534	N = 25000
	between		**0.413**	-1.529	1.458	n = 2500
	within		**0.911**	-5.176	5.179	T = 10
Average task intensity	overall	0.000	1.000	-2.749	1.895	N = 25000
	between		**0.523**	-1.247	1.601	n = 2500
	within		**0.852**	-3.939	2.662	T = 10
Average task priority	overall	0.000	1.000	-2.782	2.766	N = 25000
	between		**0.315**	-1.051	1.235	n = 2500
	within		**0.949**	-3.498	3.380	T = 10
Average number of tasks per agent	overall	0.000	1.000	-0.596	4.342	N = 25000
	between		**0.726**	-0.596	3.009	n = 2500
	within		**0.688**	-3.358	3.753	T = 10
Number of Tasks	overall	0.000	1.000	-1.589	1.602	N = 25000
	between		**0.308**	-1.009	1.051	n = 2500
	within		**0.951**	-2.408	2.495	T = 10
Task connectedness	overall	0.000	1.000	-1.543	0.812	N = 25000
	between		**0.313**	-1.045	0.812	n = 2500
	within		**0.950**	-2.119	1.856	T = 10
Average proportion of common knowledge areas	overall	0.000	1.000	-1.575	1.549	N = 25000
	between		**1.000**	-1.575	1.549	n = 2500
	within		0.000	0.000	0.000	T = 10
Average turnover	overall	0.000	1.000	-1.738	1.746	N = 25000
	between		**1.000**	-1.738	1.746	n = 2500
	within		0.000	0.000	0.000	T = 10

Average number of agents	overall	0.000	1.000	-1.540	1.569	N = 25000
	between		**1.000**	-1.540	1.569	n = 2500
	within		0.000	0.000	0.000	T = 10
Average number of locations	overall	0.000	1.000	-2.300	1.469	N = 25000
	between		**1.000**	-2.300	1.469	n = 2500
	within		0.000	0.000	0.000	T = 10
Similarity of projects	overall	0.000	1.000	-1.742	1.757	N = 25000
	between		**1.000**	-1.742	1.757	n = 2500
	within		0.000	0.000	0.000	T = 10

4.2. Regressions indicating the effects of contingency factors on the outcomes

In this section, the results of regressions that indicate the effects of the contingency factors on *average consensus* are presented in subsection 4.2.1. Subsections 4.2.2 and 4.2.3 present the effects of the contingency factors on *average accuracy of knowledge* and *percentage of project completed*, respectively.

4.2.1. Average consensus as the outcome variable

In this subsection, table 4.2.1 presents the results of the time-invariant set of contingency factors on *average consensus*. Then, tables 4.2.2 and 4.2.3 present details of regressions of the time-variant set of contingency factors on *average consensus.* Finally, subsections 4.2.1.1 – 4.2.1.3 provide interpretations of the results.

The results of the between-effects (time-invariant) regressions of the contingency factors on *average consensus* indicate that the contingency factors *average propensity to share*, *average self knowledge* , *time in training-phase*, *probability of exchange of information about a non-task-specific-knowledge area*, number of agents and number of locations have significant effects. All of the above contingency factors belong to the category *characteristics of the workgroup.*

Table 4.2.1. Between-groups effects of the time-invariant set of contingency factors on *average consensus*

Between regression (regression on group means)	**Number of obs**	**=**	**25000**
	Number of groups	**=**	**2500**
R-sq: within = 0.000	**Obs per group: min**	**=**	**10**
between = 0.688	**avg**	**=**	**10.0**
overall = 0.682	**max**	**=**	**10**
	F(14,2485)	**=**	**392.04**
sd(u_i + avg(e_i.))= 0.557	**Prob > F**	**=**	**0.0000**

Average Consensus	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95% Conf.**	**Interval]**
Average propensity to share	**0.704**	**0.011**	**62.820**	**0.000**	**0.682**	**0.726**
Average knowledge level	0.005	0.011	0.430	0.668	-0.017	0.027
Average Openness to change	-0.002	0.011	-0.140	0.885	-0.024	0.020
Average self-knowledge	**-0.024**	**0.011**	**-2.140**	**0.032**	**-0.046**	**-0.002**
Average maximum number of failed tries	0.016	0.011	1.440	0.151	-0.006	0.038
Time in training-phase	**0.295**	**0.011**	**26.380**	**0.000**	**0.273**	**0.317**
Probability of non-specific exchange	**0.169**	**0.011**	**15.140**	**0.000**	**0.147**	**0.191**
Average richness of chat	-0.010	0.017	-0.580	0.561	-0.043	0.023
Average richness of email	-0.006	0.017	-0.350	0.724	-0.039	0.027
Average richness of phone	0.001	0.013	0.070	0.940	-0.024	0.026
Average proportion of common knowledge	-0.015	0.011	-1.370	0.171	-0.037	0.007
Average turnover	-0.015	0.011	-1.330	0.185	-0.037	0.007
Number of agents	**0.220**	**0.018**	**12.500**	**0.000**	**0.186**	**0.255**
Number of locations	**0.072**	**0.018**	**4.090**	**0.000**	**0.038**	**0.107**
Constant	0.000	0.011	0.000	1.000	-0.022	0.022

4.2.1.1. Effects of characteristics of the workgroup on average consensus

Higher values of *average propensity to share* (0.704, p-value = 0.000) imply that agents are more likely to share knowledge, and information about the presence of knowledge, with other members of their workgroup. Greater the amount of *time in training phase* (0.295, p-value = 0.000), greater is the number of interactions the agents will have with each other in the pre-project phase, and hence greater is the likelihood that they would develop perceptions about others' areas of knowledge. The higher the value of *probability of exchange of information about non-task-specific-knowledge area* (0.169, p-value = 0.000), the greater is the likelihood that, in a given time period, two agents that are involved in a knowledge-transmission transaction are also going to exchange information about areas of knowledge not related to the current task on which the recipient is working.

A greater geographic spread implies a greater likelihood of agents relying on an electronic medium for communication. As described in chapter 3, equation 3.7, the use of an electronic medium adds to the time required to transmit knowledge or instructions based on knowledge. Since the the exchange of information about knowledge in a non-task-related area is associated with the transmissions of knowledge or instructions based on the required knowledge, a longer duration of knowledge transmissions implies a greater number of exchange of information about knowledge in a non-task-related area. Hence, *number of locations* (0.072, p-value = 0.000) positively affects *average consensus*.

Average self-knowledge (-0.024, p-value = 0.032) was shown to have a negative and significant effect on the development of TMS in a workgroup. Two aspects of the simulation are pertinent in explaining this result: (1) the agents in the simulation would contact other agents for obtaining knowledge only when the focal agents believe that they lack the knowledge required for completing the subtasks assigned to them, and (2) beyond the training phase, agents would participate in the exchange of non-task-related information only if they are part of an on-going transmission of knowledge. Reason (1) implies that the higher the value of *average self-knowledge,* the lower is the likelihood that an agent would seek knowledge from another agent. Consequently, agents with higher self-knowledge would be less-likely to participate in the exchange of information about the existence of knowledge in other agents. Given that every project terminates after a certain number of time periods (which is not determined directly, but results when all the subtasks of the project have been either completed or abandoned), workgroups that have agents with higher levels of self-knowledge (that is agents which have beliefs that they possess knowledge in a greater proportion of areas), would have fewer number of exchanges of information leading to development of TMS (note that TMS only reflects the beliefs that the workgroup, as a whole, has about its members' areas of knowledge – these beliefs need not be true). Hence, workgroups with higher values of *average self-knowledge* are less likely to have higher values of TMS.

In terms of absolute numbers, large workgroups are more likely to have a greater number of exchanges that are related to exchanges of knowledge and exchange of information about the presence of knowledge. This is because, in smaller workgroups, an agent that is seeking knowledge that is needed to complete a subtask assigned to it has fewer number of sources to contact. Since an agent would participate in a non-task-related exchange of information only after it has begun obtaining (or providing) knowledge, a fewer number of agents in the workgroup would imply that there is smaller likelihood that an agent can find a source agent that would meet its knowledge requirements. Hence, in such workgroups, tasks are more likely to be abandoned before they are completed, because an agent can attempt to complete a subtask only a limited number of times before abandoning it for failing to obtained the required knowledge. Since the likelihood of a non-task-related exchange is predicated on the initiation and continuation of a task-related-knowledge transmission, fewer agents in a workgroup implies a smaller likelihood of knowledge-related exchanges, and thus a smaller likelihood of non-task-related exchanges. Therefore, smaller workgroups are at a disadvantage, both with respect to their agents finding the knowledge they need and with respect to their agents forming beliefs about other agents' areas of knowledge. Hence, *number of agents* (0.220, p-value = 0.000) has a positive effect on *average consensus*.

Next, the effects of contingency factors that have both between-groups and within-group effects are obtained via the respective regressions. As described in section 3.3, the within-group effects of a set of predictors on an outcome can be determined using either fixed-effects or random-effects regressions. To determine the appropriateness of fixed-effects regression over random-effects regression, in the case of each outcome variable, both regressions were run, and their models were used as input to the *xtoverid* routine (Schaffer and Stillman, 2006) in Stata. In the case of all three sets of regressions, the Sargan-Hansen statistic indicated that the fixed-effects regression was appropriate. Here, for the sake of brevity, results from only the fixed-effects regressions are presented.

The results indicate that the time-variant effects of only *average project intensity* (category: *interface between workgroup and tasks*) and *average task priority* (category: *characteristics of the tasks*) are significant. The results of the between-effects (time-invariant) regressions of the contingency factors on *average consensus* indicate that the contingency factors *average knowledge-intensity of subtasks* (category: *characteristics of tasks*), *average task intensity* (category: *characteristics of tasks*), *average project intensity* (category: *interface between workgroup and tasks*) and *average number of tasks per agent* (category: *interface between workgroup and tasks*) have significant effects.

Table 4.2.2. Within-group effects of the time-invariant set of contingency factors on *average consensus*

Fixed-effects (within) regression		**Number of obs =**	**25000**
		Number of groups =	**2500**
R-sq:	**within = 0.002**	**Obs per group: min =**	**10**
	between = 0.048	**avg =**	**10.0**
	overall = 0.008	**max =**	**10**
		F(8,22492) =	**5.32**
corr(u_i, Xb) = -0.0935		**Prob > F =**	**0.0000**

Average Consensus	Coef.	Std. Err.	t	P>|t|	[95%Conf.	Interval]
Average direction time	0.000	0.001	0.480	0.630	-0.001	0.002
Average stickiness time	0.000	0.001	0.320	0.746	-0.001	0.002
Average project intensity	**0.002**	**0.001**	**2.140**	**0.032**	**0.000**	**0.004**
Average knowledge intensity of subtasks	-0.001	0.001	-0.820	0.413	-0.002	0.001
Average task intensity	0.001	0.001	1.880	0.060	0.000	0.003
Average task priority	**0.002**	**0.001**	**3.200**	**0.001**	**0.001**	**0.003**
Tasks per agent	0.002	0.001	1.320	0.188	-0.001	0.004
Task connectedness	0.000	0.001	0.490	0.624	-0.001	0.002
constant	0.000	0.001	0.000	1.000	-0.001	0.001

sigma_u = 0.996

sigma_e = 0.099

rho = 0.990(fraction of variance due to u_i)

F test that all u_i=0: F(2499, 22492) = 940.16 Prob > F = 0.0000

Table 4.2.3. Between-group effects of the time-variant set of contingency factors on *average consensus*

Between regression (regression on group means)		**Number of obs**	**=**	**25000**
		Number of groups	**=**	**2500**
R-sq:	**within = 0.001**	**Obs per group: min**	**=**	**10**
	between = 0.085	**avg**	**=**	**10.0**
	overall = 0.047	**max**	**=**	**10**
		F(8,2491)	**=**	**28.75**
sd(u_i + avg(e_i.))= 0.954		**Prob > F**	**=**	**0.000**

Average Consensus	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95%ConfInterval]**	
Average direction time	0.009	0.047	0.200	0.844	-0.082	0.101
Average stickiness time	-0.011	0.047	-0.230	0.818	-0.103	0.082
Average project intensity	**0.182**	**0.070**	**2.620**	**0.009**	**0.046**	**0.319**
Average knowledge intensity of subtasks	**0.129**	**0.046**	**2.790**	**0.005**	**0.038**	**0.220**
Average task intensity	**-0.084**	**0.037**	**-2.300**	**0.021**	**-0.156**	**-0.013**
Average task priority	-0.024	0.061	-0.400	0.688	-0.143	0.095
Average number of tasks per agent	**-0.391**	**0.027**	**-14.510**	**0.000**	**-0.444**	**-0.339**
Task connectedness	-0.098	0.067	-1.460	0.145	-0.229	0.034
constant	0.000	0.019	0.000	1.000	-0.037	0.037

4.2.1.2. Effects of characteristics of the tasks on average consensus

The contingency factors, *average knowledge-intensity of subtasks* (0.129, p-value = 0.005), and *average task intensity* (-0.084, p-value = 0.021), have a positive and negative significant effect, respectively. Other factors being constant, in workgroups whose agents work on tasks with relatively-higher *average knowledge-intensity of subtasks*, the agents are more likely to be in need of knowledge that they do not possess and therefore are more likely to seek knowledge from others. A likely consequence of this situation is that agents in such workgroups are more likely to learn about the knowledge areas of other agents, who are their knowledge sources, which implies that such workgroups tend to have a higher level of *average consensus*. Higher values of *average task intensity* imply that a greater number of subtasks are associated with each task. This implies that knowledge is needed in a greater number of areas. The within-group negative effect of *average task intensity* implies that as the number of subtasks associated with a given task is high. When a task has more subtasks, individuals are less likely to have consensus on knowledge related to all, or most of, the subtasks. This results in a lower likelihood of high levels of consensus, thereby reducing average consensus.

Average task priority also was found to have a positive significant effect (0.002, p-value = 0.001) on *average consensus*. Between two tasks that have satisfied their precedence constraints, agents are going to choose the task with higher priority. The positive result implies that, within a workgroup, such a choice would lead to greater number of interactions involving an exchange of knowledge related to these (i.e., high-priority) tasks as compared to other (i.e., low-priority) tasks. Therefore higher values of *average consensus* are attained with respect to high-priority tasks.

4.2.1.3. Effects of the interface between the workgroup and the tasks on average consensus

The contingency factors in this category that were found to have a significant effect are *average project intensity* (0.002, p-value = 0.032) and *average number of tasks per agent* (-0.391, p-value = 0.000). Higher values of average project intensity implies that the number of tasks, and consequently subtasks that an agent has to complete would be higher. Therefore each agent has to possess knowledge in a greater number of areas, leading agents in such workgroups to interact with other agents in their workgroups for a relatively greater number of times. As explained earlier, in section 4.2.1.2, greater interactions with other agents lead to greater likelihood of developing perceptions of other agents' areas of knowledge, and thus to a greater level of agreement between any given pair of agents about a third agent's areas of knowledge. Hence, higher values of *average project intensity* lead to higher *average consensus*. When individuals are assigned a greater number of tasks, they are less likely to have consensus on knowledge related to all, or most of, their tasks, thereby lowering the likelihood of high levels of consensus, and therefore reducing average consensus.

Figure 4.1 summarizes the results of regressions on *average consensus*. It indicates that (a) only three of the five categories of contingency factors have significant effects on *average consensus* and, (b) within each category, some factors have a positive significant effect while others have a significant negative effect.

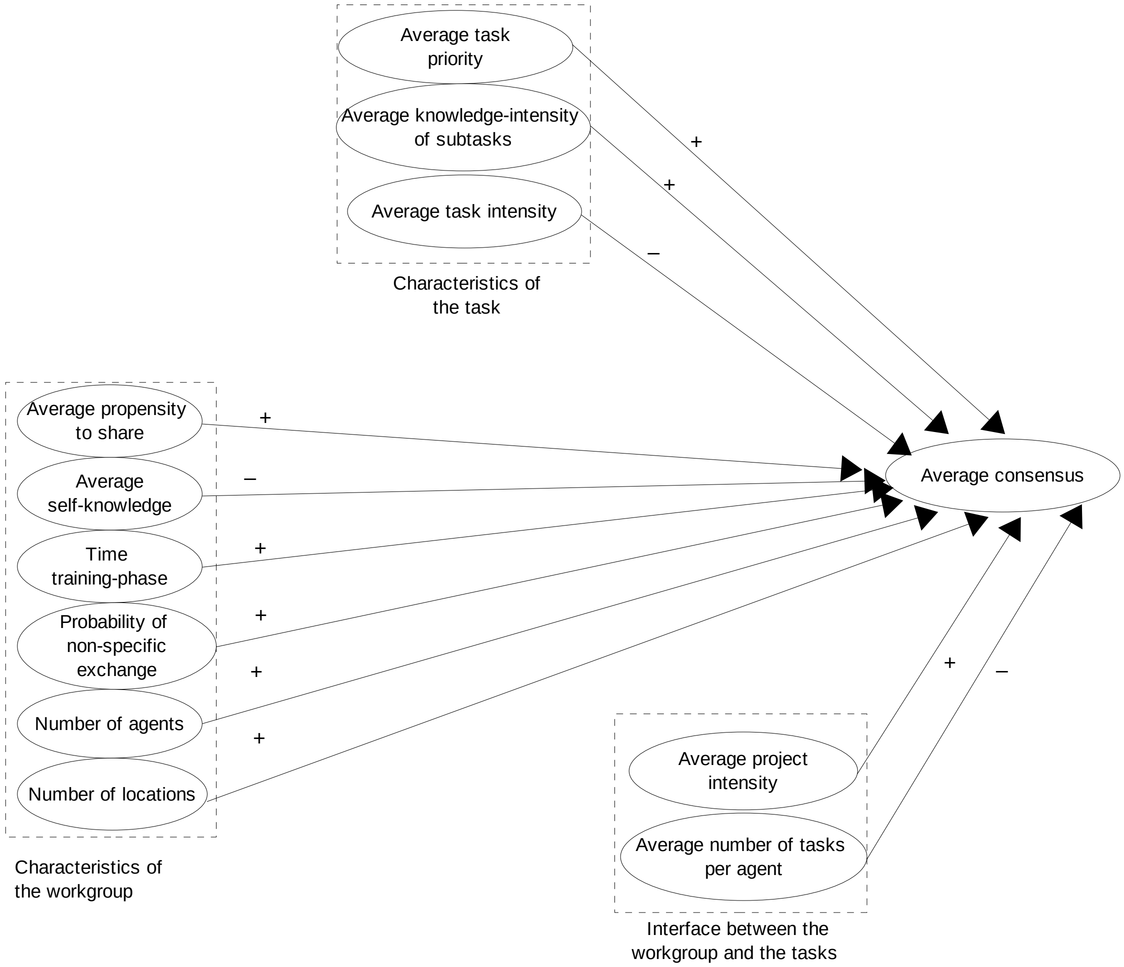


Figure 4.1. Effects of contingency factors on *average consensus*.

4.2.2. Average accuracy of knowledge as the outcome variable

In this subsection the results of the regressions of the contingency factors on *average accuracy of knowledge*. First, the results of the between-effects regression results are presented in table 4.2.4. Then, tables 4.2.5 and 4.2.6 presents results of within-effects and between-effects regressions of those contingency factors that were found to have both types of standard-deviation components, as reported in table 4.1. Finally, the results of the regressions are interpreted in sections 4.2.2.1 – 4.2.2.3.

Table 4.2.4. Between-groups effects of the time-invariant set of contingency factors on *average accuracy of knowledge*

Between regression (regression on group means)		**Number of obs**	**= 25000**
Group variable: expno		**Number of groups**	**= 2500**
R-sq:	**within = 0.000**	**Obs per group: min**	**= 10**
	between = 0.875	**avg**	**= 10.0**
	overall = 0.875	**max**	**= 10**
		F(14,2485)	**= 1239.49**
sd(u_i + avg(e_i.))=	**0.355**	**Prob > F**	**= 0.000**

Average Accuracy	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95% Conf.**	**Interval]**
Average propensity to share	-0.001	0.007	-0.200	0.839	-0.015	0.013
Average knowledge level	**0.936**	**0.007**	**131.260**	**0.000**	**0.922**	**0.949**
Average Openness to change	0.001	0.007	0.170	0.862	-0.013	0.015
Average self-knowledge	-0.003	0.007	-0.360	0.723	-0.016	0.011
Average maximum number of failed tries	-0.006	0.007	-0.900	0.368	-0.020	0.008
Time in training-phase	0.006	0.007	0.820	0.410	-0.008	0.020
Probability of non-specific exchange	-0.012	0.007	-1.720	0.086	-0.026	0.002
Average richness of chat	-0.007	0.011	-0.620	0.536	-0.028	0.014
Average richness of email	-0.002	0.011	-0.160	0.871	-0.023	0.019
Average richness of phone	0.013	0.008	1.630	0.103	-0.003	0.029
Average proportion of common knowledge	0.001	0.007	0.090	0.925	-0.013	0.015
Average turnover	0.003	0.007	0.370	0.713	-0.011	0.017
Number of agents	**-0.026**	**0.011**	**-2.310**	**0.021**	**-0.048**	**-0.004**
Number of locations	0.008	0.011	0.700	0.484	-0.014	0.030
constant	0.000	0.007	0.000	1.000	-0.014	0.014

4.2.2.1. The effect of characteristics of the workgroup on average accuracy of knowledge

Average accuracy of knowledge is operationalized as the proportion of the number of areas of knowledge that are correct, that is, can be applied to completed a subtask, in each agent, averaged across all agents. Workgroups whose members have higher levels of knowledge, initially, tend to retain their high levels of knowledge throughout the lifetime of the workgroup, after accounting for factors such as attrition of knowledge due to forgetfulness of the workgroup's agents, and obsolescence of knowledge (as modeled by March 1991; Kane and Alavi, 2007, in the form of environmental turbulence), which has not been included in the simulation, and turnover. Hence, high values of *average knowledge level* (0.936, p-value = 0.000), which indicates the average level of knowledge at the beginning of a workgroup's lifetime, have a positive significant effect on *average accuracy of knowledge*.

The effect of *number of agents* on *average accuracy of knowledge* was found to be negative (-0.026, p-value = 0.021). This may be explained as follows: in workgroups with higher number of agents, indicated by the contingency factor *number of agents*, a greater number of agents would need to be knowledgeable for the average knowledge to be high, and it might take more time for a greater number of individuals to acquire knowledge, even if they may be able to learn from more people. Therefore, the average knowledge accuracy at a given point in time would be relatively lower in relatively larger groups.

Next, the regressions of the contingency factors which have both time-variant and time-invariant components of standard-deviation on *average accuracy of knowledge* are shown in tables 4.2.5 and 4.2.6.

Table 4.2.5. Within-group effects of the time-variant set of contingency factors on *average accuracy of knowledge*

Fixed-effects (within) regression		**Number of obs =**	**25000**
		Number of groups =	**2500**
R-sq:	**within = 0.002**	**Obs per group: min =**	**10**
	between = 0.048	**avg =**	**10.0**
	overall = 0.008	**max =**	**10**
		F(8,22492) =	**5.32**
corr(u_i, Xb) = -0.094		**Prob > F =**	**0.000**

Average Accuracy	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95% Conf. Interval]**	
Average direction time	0.000	0.000	0.120	0.907	-0.000	0.000
Average stickiness time	0.000	0.000	1.010	0.312	-0.000	0.000
Average project intensity	0.000	0.000	-0.330	0.740	-0.000	0.000
Average knowledge intensity of subtasks	0.000	0.000	0.050	0.960	-0.000	0.000
Average task intensity	0.000	0.000	1.280	0.200	-0.000	0.000
Average task priority	0.000	0.000	1.630	0.103	0.000	0.000
Average number of tasks per agent	-0.000	0.000	-0.490	0.625	-0.000	0.000
Task connectedness	-0.000	0.000	-1.050	0.294	-0.000	0.000
constant	0.000	0.000	0.000	1.000	-0.000	0.000

sigma_u = 1.000

sigma_e = 0.002

rho = 0.999 (fraction of variance due to u_i)

F test that all u_i=0: F(2499, 22492) = 1.7e+06 Prob > F = 0.0000

Table 4.2.6. Between-groups effects of the time-variant set of contingency factors on *average accuracy of knowledge*

Between regression (regression on group means)			Number of obs	=	25000	
			Number of groups	=	2500	
R-sq: within = 0.001			Obs per group: min	=	10	
between = 0.085			avg	=	10.0	
overall = 0.047			max	=	10	
			F(8,2491)	=	28.75	
sd(u_i + avg(e_i.))= 0.954			Prob > F	=	0.0000	

Average Accuracy	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95%Conf.Interval]**	
Average direction time	-0.003	0.049	-0.060	0.953	-0.099	0.093
Average stickiness time	0.020	0.050	0.410	0.679	-0.077	0.118
Average project intensity	-0.016	0.073	-0.220	0.828	-0.159	0.127
Average knowledge intensity of subtasks	0.031	0.049	0.640	0.523	-0.064	0.126
Average task intensity	-0.004	0.038	-0.100	0.922	-0.079	0.072
Average task intensity	0.005	0.064	0.070	0.942	-0.120	0.130
Average number of tasks per agent	0.030	0.028	1.070	0.283	-0.025	0.086
Task connectedness	-0.055	0.070	-0.790	0.432	-0.193	0.083
constant	0.000	0.020	0.000	1.000	-0.039	0.039

The results of the regressions indicate that none of the contingency factors that have within-group and between-group effects are significant. Figure 4.2 summarizes the overall results of the different contingency factors on *average accuracy of knowledge.* It illustrates the fact that of the five categories of contingency factors, only a subset of contingency factors belonging to *characteristics of workgroup*, viz., *average knowledge level* and *number of agents*, have a significant positive effect on the outcome *average accuracy of knowledge.*

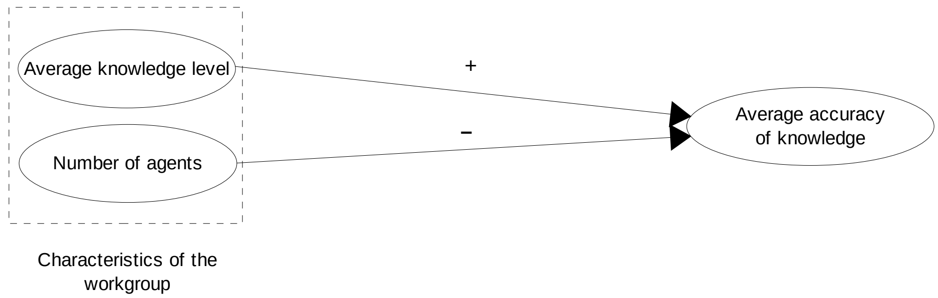


Figure 4.2. Effects of contingency factors on *average accuracy of knowledge*

4.2.3. Percentage of project completed as the outcome variable

The outcome variable *percentage of project completed* indicates the performance of a workgroup; it describes the extent to which a workgroup's members were successful in completing the tasks (and their associated subtasks) assigned to them. It measures the proportion of subtasks that are completed. In table 4.2.7, the between-effects regressions of the time-invariant set of contingency factors on *percentage of project completed* are reported. Tables 4.2.8 and 4.2.9 present details of the regressions of the time-variant set of contingency factors. Finally, subsections 4.2.3.1 – 4.2.3.5 provide interpretations of the results.

Table 4.2.7. Between-group effects of time-invariant set of contingency factors on *percentage of project completed*

```
Between regression (regression on group means)  Number of obs      =     25000
                                                Number of groups   =      2500
R-sq:  within  = 0.000                          Obs per group: min =        10
       between = 0.516                                         avg =      10.0
       overall = 0.335                                         max =        10
                                                F(14,2485)         =    188.88
sd(u_i + avg(e_i.))=  0.563                     Prob > F           =    0.000
```

Percentage of project completed	Coef.	Std. Err.	t	P>|t|	[95% Conf.Interval]	
Average propensity to share	**0.138**	**0.011**	**12.210**	**0.000**	**0.116**	**0.160**
Average knowledge level	**0.277**	**0.011**	**24.460**	**0.000**	**0.254**	**0.299**
Average Openness to change	0.000	0.011	0.020	0.981	-0.022	0.022
Average self-knowledge	**0.468**	**0.011**	**41.370**	**0.000**	**0.445**	**0.490**
Average maximum number of failed tries	0.011	0.011	1.000	0.317	-0.011	0.033
Time in training-phase	**0.091**	**0.011**	**8.080**	**0.000**	**0.069**	**0.113**
Probability of non-specific exchange	**0.084**	**0.011**	**7.440**	**0.000**	**0.062**	**0.106**
Average richness of chat	0.030	0.017	1.770	0.077	-0.003	0.063
Average richness of email	**-0.037**	**0.017**	**-2.210**	**0.027**	**-0.070**	**-0.004**
Average richness of phone	-0.016	0.013	-1.220	0.224	-0.041	0.010
Average proportion of common knowledge	-0.003	0.011	-0.250	0.799	-0.025	0.019
Average turnover	0.008	0.011	0.670	0.505	-0.015	0.030
Number of agents	0.024	0.018	1.340	0.180	-0.011	0.059
Number of locations	0.010	0.018	0.580	0.561	-0.025	0.045
constant	0.000	0.011	0.000	1.000	-0.022	0.022

The above results indicate that a subset of the contingency factors belonging to the category *characteristics of the workgroup* have significant effects on the outcomes variable *percentage of project completed.* An interpretation of these effects is presented next.

4.2.3.1. The effects of characteristics of the workgroup on *percentage of project completed*

Average propensity to share indicates the likelihood of a workgroup's agents to share knowledge or information about knowledge in a particular area with other agents of the workgroup. The effect of this contingency factor was found to be positive and significant (0.138, p-value = 0.000). Higher values of *average propensity to share* affect a workgroup in two ways: (a) an agent looking for a source for knowledge in a given area is more likely to locate a source, since agents are more likely to let other agents know about their areas of knowledge during their encounters; (b) once an agent has located a source, the likelihood of the source providing accurate knowledge or correct directions based on the appropriate knowledge is higher. Reasons (a) and (b) imply a greater likelihood of an agent completing its assigned tasks (subtasks), and therefore a higher value of *percentage of project completed.*

At the agent-level *knowledge level* indicates the proportion of knowledge areas in which an agent has correct knowledge that would help it complete the subtasks with which the knowledge areas are associated. At the workgroup-level, *average knowledge level* indicates the aggregate ability of the workgroup to complete the assigned tasks and their associated subtasks. Hence, *average knowledge level* was found to have a positive significant effect (0.277, p-value = 0.000) on *percentage of project completed*.

The effect of *average self-knowledge* was found to be positive and significant (0.468, p-value = 0.000). When compared to agents in workgroups with lower levels of *average self-knowledge*, agents of workgroups with higher levels of *average self-knowledge* that agents are less likely to contact other agents, in instances when they have high levels of knowledge, and more likely to contact other agents in instances when they have low levels of knowledge. However, if the distribution of knowledge is skewed, then despite having high levels of average knowledge, a workgroup's agents which have higher than within-group average level of knowledge, might be unavailable to act as a source-agent to newer requests, if they are acting already as a source. Hence, the specific effect of *average self-knowledge* appears to be contingent upon other within-group factors. However, since agent-level data were not collected, such contingent effects could not be examined in this study.

The contingency factor *time in training* has been shown to have a positive significant

(0.091, p-value = 0.000) effect on *percentage of project completed*. Higher values of *time in training phase* imply that, before a workgroup begins working on its projects, agents will have obtained information about (a) greater number of agents' areas of knowledge, and (b) greater number of sources for each knowledge area. Hence, during the project phase, the likelihood of an agent being able to locate a source for the knowledge in the areas it requires for completing its assigned tasks and subtasks is higher, thereby resulting in a higher value of *percentage of project completed.*

The effect of *probability of exchange of information about a non-task-specific knowledge area* was found to be positive and significant (0.084, p-value = 0.000). Higher values of *probability of exchange of information about a non-task-specific-knowledge area* imply that during a given time period, during the training-phase, and during the project-phase, if an agent is involved in a knowledge-transmission-related transaction, the likelihood of learning about another agent's areas of knowledge is higher. Thus, in such workgroups, agents are more likely to be able to locate a source agent for the knowledge in the area they require, by relying on their relatively greater amount of information about sources of knowledge in different areas. Hence, this contingency factor was shown to have a positive significant effect on *percentage of project completed*.

4.2.3.2. The effect of characteristics of information technologies

The regression results show that richness of email is the only significant contingent factor from the category characteristics of information technologies and that it has a negative effect (-0.037, p-value = 0.027) on percentage of project completed. The richness, synchronousness and locality attributes of email as a communication medium imply that (as described in section 3.2.1.16) email should be chosen for communication when the priority of the task is relatively low, and synchronicity of transmission is not required. Given the relatively low level of richness of email, its use leads to relatively longer transmission times, both for direction and transfer. Email should be used for low-priority tasks, but if individuals perceive it to be a rich medium, they may start using it for high-priority tasks as well, leading to lower knowledge transfer and consequently lower project completion. Perceived richness of email might have had a negative effect on knowledge accuracy but for the fact that average knowledge level dominates that regression. Once a workgroup member feels that he/she needs to acquire knowledge from another individual, and identifies a potential source, the member would be more likely to use email for knowledge transfer if he/she perceives email to be a rich medium, but would fail to actually acquire knowledge because email would not actually transfer the needed knowledge. As a result, over time, the likelihood of agents having the knowledge needed for completing all the subtasks would be lower when perceived richness of email is higher, and therefore subtasks would be more likely to be abandoned. Hence, higher values of average richness of email lead to lower project completion as seen in lower values of percentage of project completed.

Next, the regressions of the contingency factors that have both within-group and between-groups components of standard deviations are presented.

Table 4.2.8. Within-group effects of time-variant set of contingency factors on *percentage of project completed*

Fixed-effects (within) regression		**Number of obs**	**= 25000**
		Number of groups	**= 2500**
R-sq:	**within = 0.076**	**Obs per group: min**	**= 10**
	between = 0.005	**avg**	**= 10.0**
	overall = 0.030	**max**	**= 10**
		F(8,22492)	**= 230.43**
corr(u_i, Xb) = -0.009		**Prob > F**	**= 0.000**

Percentage of project completed	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95%Conf.**	**Interval]**
Average direction time	-0.003	0.004	-0.790	0.429	-0.011	0.005
Average stickiness time	-0.005	0.004	-1.180	0.240	-0.013	0.003
Average project intensity	-0.007	0.006	-1.190	0.233	-0.019	0.005
Average knowledge intensity of subtasks	**-0.178**	**0.004**	**-42.740**	**0.000**	**-0.186**	**-0.170**
Average task intensity	0.007	0.005	1.510	0.131	-0.002	0.016
Average task priority	0.003	0.004	0.860	0.389	-0.004	0.011
Average number of tasks per agent	**0.020**	**0.008**	**2.590**	**0.009**	**0.005**	**0.035**
Task connectedness	-0.001	0.004	-0.140	0.891	-0.009	0.008
constant	0.000	0.004	0.000	1.000	-0.007	0.007

sigma_u = 0.805

sigma_e = 0.599

rho = 0.64348269 (fraction of variance due to u_i)

F test that all u_i=0: F(2499, 22492) = 17.98 Prob > F = 0.0000

Table 4.2.9. Between-group effects of time-variant set of contingency factors on *percentage of project completed*

Between regression (regression on group means)		**Number of obs =**	**25000**
		Number of groups =	**2500**
R-sq: within = 0.042		**Obs per group: min =**	**10**
between = 0.014		**avg =**	**10.0**
overall = 0.023		**max =**	**10**
		F(8,2491) =	**4.39**
sd(u_i + avg(e_i.))= 0.802		**Prob > F =**	**0.000**

Percentage of project completed	Coef.	Std. Err.	t	P>|t|	[95%ConfInterval]	
Average direction time	**-0.092**	**0.039**	**-2.350**	**0.019**	**-0.169**	**-0.015**
Average stickiness time	0.044	0.040	1.110	0.266	-0.034	0.122
Average project intensity	0.024	0.058	0.400	0.687	-0.091	0.138
Average knowledge intensity of subtasks	**-0.167**	**0.039**	**-4.290**	**0.000**	**-0.243**	**-0.091**
Average task intensity	-0.026	0.031	-0.840	0.404	-0.086	0.035
Average task priority	-0.058	0.051	-1.130	0.257	-0.158	0.042
Average number of tasks per agent	**-0.059**	**0.023**	**-2.600**	**0.009**	**-0.103**	**-0.015**
Task connectedness	0.066	0.056	1.160	0.245	-0.045	0.176
constant	0.000	0.016	0.000	1.000	-0.031	0.031

Next, interpretations of the results of regressions involving those contingency factors that were shown to have within-group and between-groups standard-deviation components are presented.

4.2.3.3. The effect of characteristics of the tasks

Average knowledge intensity of subtasks was found to have a significant negative within-group (-0.178, p-value = 0.000) and between-groups effects (-0.167, p-value = 0.000) on percentage of project completed. As the number of areas in which knowledge is required to complete a subtask, on an average, increases, the likelihood that each agent is able to obtain the knowledge or instructions based on the knowledge, in a fixed amount of time (if the agent does not have the required knowledge) decreases. Additionally, for a fixed value of *knowledge level* for an agent, as the knowledge requirements increase, through an increase in the *average knowledge intensity of subtasks*, agents are less likely to have all the required knowledge. Given the above two reasons, higher values of *average knowledge intensity of subtasks* lead to lower values of *percentage of project completed,* since the proportion of subtasks that agents can complete during a fixed amount of time reduces. Hence, *average knowledge intensity of subtasks* was found to have a significant negative effect on percentage of project completed.

4.2.3.4. The effect of characteristics of knowledge

While obtaining knowledge required to complete a given subtask, agents can obtain knowledge either via transfer or via direction (as described in subsection 3.2.1.7). Knowledge in a given area that was obtained by transfer implies that an agent does not have to seek it once again, in the future, if knowledge in the same area is associated with another subtask. *Average direction time* was found to have a negative significant effect (-0.092, p-value = 0.019) on *percentage of project completed*. If agents in a workgroup have to spend more time giving others direction (which does not lead to knowledge transfer), fewer members of a workgroup would receive knowledge over time. Hence, over time, the likelihood of agents having needed knowledge for completing all their assigned subtasks would be lower, thereby leading to lower values of *percentage of project completed*.

4.2.3.5. The effect of the interface between the workgroup and the tasks

In this category of contingency factors, *average number of tasks per agent* was found to have a significant positive effect (0.020, p-value = 0.009), within a workgroup, on *percentage of project completed*; it was found to have a significant negative effect across workgroups. This implies that within a workgroup, as the average number of tasks per agent increases, some of the agents would be working on a greater number of tasks than others. As Littlepage et al. (2008) found, if such skewness in the assigned workload is such that those members of a workgroup who are best qualified (in terms of knowledge and skills) in working on a set of tasks are assigned those tasks, then the performance of the workgroup, as a whole, improves. Littlepage et.al (2008) explain that this occurs because of a greater degree of match between tasks, the expertise they require and the persons to whom the tasks are assigned. This observation is consistent with the task-expertise-person fit described by Brandon and Hollingshead (2004). While the observed result is consistent with the explanation of Littlepage et al., in order to investigate and verify the underlying mechanism by which the effect has resulted, further investigation is needed.

The between-groups effect of *average number of tasks per agent* (-0.059, p-value = 0.009) was found to be negative. While comparing workgroups, those that have higher levels of *average number of tasks per agent*, tend to have a greater number of knowledge requirements for each agent, resulting in a lower likelihood of each agent completing its assigned set of tasks and subtasks. Hence, in workgroups that are "overworked"[37], the proportion of subtasks that agents can complete is lower. Consequently, across workgroups, *average number of tasks per agent* negatively affects *percentage of project completed.*

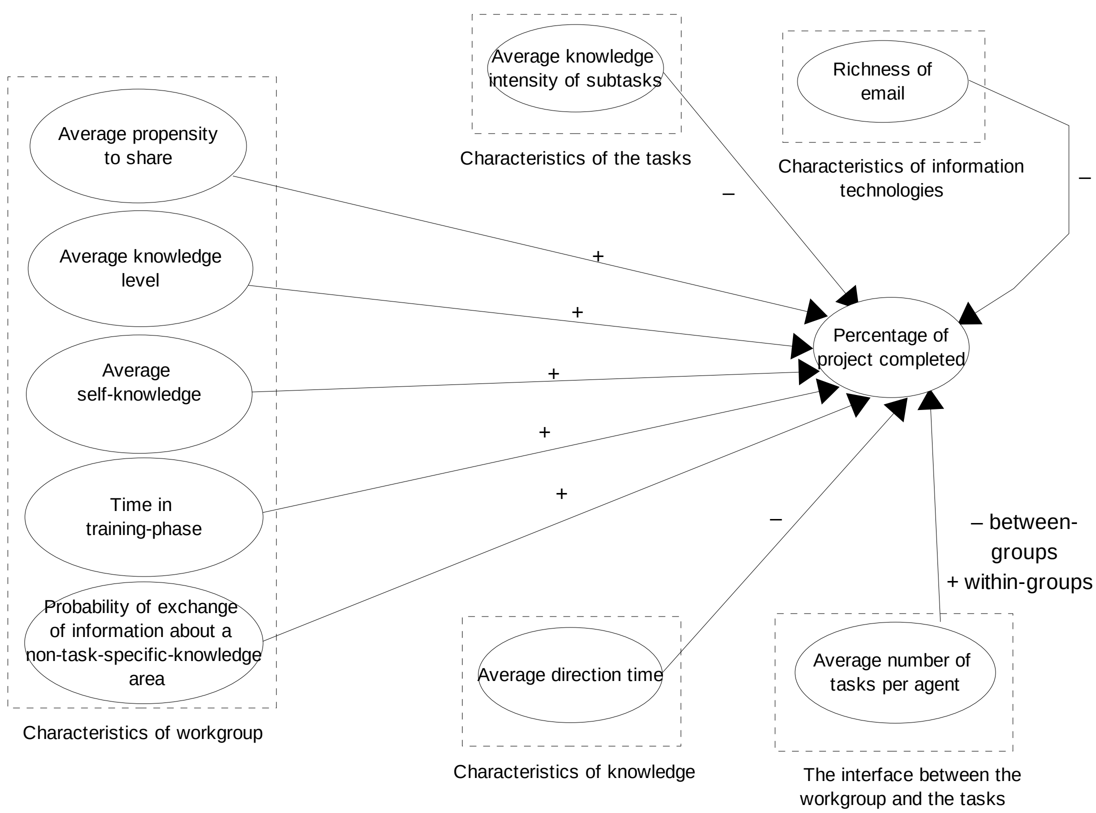


Figure 4.3. Effects of contingency factors on *percentage of project completed.*

37 As defined in the subsection 3.2.1.6.

The results of the effects of the contingency factors that have a significant effect on *percentage of project completed* are shown in figure 4.3. The figure illustrates the observation that subsets of contingency factors, which are drawn from every category, are significant. Those belonging to *characteristics of the tasks*, *characteristics of information technologies* and *characteristics of knowledge* were observed to have a negative effect, while those belonging to *characteristics of workgroup* were observed to be positive; the contingency factor drawn from the category *the interface between tasks and workgroup* was observed to have a positive within-group effect and a negative between-groups effect.

Next, tables 4.3 – 4.5 provide summaries of the significant effects of various contingency factors on each of the three outcome variables respectively[38].

38 In the following tables, BE stands for between-groups effects and WE stands for within-group effects

Table 4.3. Summary of significant effects of contingency factors on *average consensus*

Average Consensus	Coef.	Std. Err.	t	P>|t|	[95%Conf.	Interval]
Average propensity to share (BE)	0.704	0.011	62.820	0.000	0.682	0.726
Average self-knowledge (BE)	-0.024	0.011	-2.140	0.032	-0.046	-0.002
Time in training-phase (BE)	0.295	0.011	26.380	0.000	0.273	0.317
Probability of non-specific exchange (BE)	0.169	0.011	15.140	0.000	0.147	0.191
Number of agents (BE)	0.220	0.018	12.500	0.000	0.186	0.255
Number of locations (BE)	0.072	0.018	4.090	0.000	0.038	0.107
Average project intensity (WE)	0.002	0.001	2.140	0.032	0.000	0.004
Average task priority (WE)	0.002	0.001	3.200	0.001	0.001	0.003
Average project intensity (BE)	0.182	0.070	2.620	0.009	0.046	0.319
Average knowledge intensity of subtasks (BE)	0.129	0.046	2.790	0.005	0.038	0.220
Average task intensity (BE)	-0.084	0.037	-2.300	0.021	-0.156	-0.013
Tasks per agent (BE)	-0.391	0.027	-14.510	0.000	-0.444	-0.339

Table 4.4. Summary of significant effects of contingency factors on *average accuracy of knowledge*

Average Accuracy	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]	
Average knowledge level	0.936	0.007	131.260	0.000	0.922	0.949	(BE)
Number of agents	-0.026	0.011	-2.310	0.021	-0.048	-0.004	(BE)

Table 4.5. Summary of significant effects of contingency factors on *percentage of project completed*

Percentage of project completed	Coef.	Std. Err.	t	P>|t|	[95%ConfInterval]	
Average propensity to share (BE)	0.138	0.011	12.210	0.000	0.116	0.160
Average knowledge level (BE)	0.277	0.011	24.460	0.000	0.254	0.299
Average self-knowledge (BE)	0.468	0.011	41.370	0.000	0.445	0.490
Time in training-phase (BE)	0.091	0.011	8.080	0.000	0.069	0.113
Probability of non-specific exchange (BE)	0.084	0.011	7.440	0.000	0.062	0.106
Average richness of email (BE)	-0.037	0.017	-2.210	0.027	-0.070	-0.004
Average knowledge intensity of subtasks (WE)	-0.178	0.004	-42.740	0.000	-0.186	-0.170
Average number of tasks per agent (WE)	0.020	0.008	2.590	0.009	0.005	0.035
Average direction time (BE)	-0.092	0.039	-2.350	0.019	-0.169	-0.015
Average knowledge intensity of subtasks (BE)	-0.167	0.039	-4.290	0.000	-0.243	-0.091
Average number of tasks per agent (BE)	-0.059	0.023	-2.600	0.009	-0.103	-0.015

4.3. Summary

In this chapter results were presented of the regression analyses that were conducted to determine the effects of the contingency factors on the three outcomes. Section 4.1 provided an overview of the regression analyses and described the rationale for dividing the contingency factors into two groups, those that have a between-groups effect and those that have a within-groups effect. Section 4.2 and its subsections provided details of the regressions and interpretations of the results. They also included figures (4.1, 4.2, and 4.3) and tables 4.3 – 4.5 that summarize the results.

In the next chapter, Discussion, the findings of the study are presented via a synthesis of the regression results with findings from extant literature. Conclusions, limitations and directions for future research that can be based on the current work are also presented.

5. Discussion

The goal of this dissertation was to understand how different contingency factors affect the knowledge outcomes, specifically *average consensus* and *average accuracy* of knowledge, and performance of a workg5.1. Interpretation of the resultsroup, measured as *percentage of project completed*. *Average consensus* indicates the degree to which the members of a workgroup agree with each other about each other's areas of expertise. It represents the aspect of transactive memory that was operationalized and studied in this dissertation. *Average accuracy* of knowledge is an aggregate measure that indicates the average ability of a workgroup member to complete the tasks assigned to her/him without having to seek help from others.

The motivation to study the two knowledge-related outcomes and the performance of the workgroup were the three issues: (1) an increasingly-transient workforce; (2) ever-increasing levels of data, and the associated increase in the knowledge required to process the data; and, (3) the large proportion of knowledge staying in individual's minds, rather than being converted into knowledge artifacts. Understanding the effects of the contingency factors would help organizations better manage their knowledge-related processes in the light of the three issues identified above. In addition, by understanding which of the contingency factors that affect the knowledge outcomes also affect the performance of a workgroup, organizations can decide to focus on the most important (significant) subset of contingency factors to maximize the returns of their investments in the human and technological resources. The effects of the contingency factors were investigated by analyzing the results obtained from simulations an agent-based model using the methodology of panel data analysis. The results were presented in the previous chapter, *Results*.

In this chapter, in section 5.1, the results are discussed in the light of findings from prior literature, where appropriate, with a goal of providing answers to the research question that was presented in chapter 1. Section 5.2 identifies the limitations that circumscribe the current study. In section 5.3, implications for practitioners are presented. This is followed by section 5.4, where implications for theory are discussed. Then, in section 5.5, directions for future research that can be based on the current work are described. Finally, section 5.6 presents an overall summary of this dissertation. Next, the results of this study are discussed alongside findings from extant literature.

5.1. Interpretation of the results

The results of the regression analyses indicate that not all contingency factors have significant effects on the outcome variables. Specifically, in the case of *average consensus*, only subsets of contingency factors belonging to the categories *characteristics of the workgroup*, *characteristics of the tasks* and *the interface between workgroup and tasks* were found to have significant effects. In the case of *average accuracy of knowledge*, only a subset of contingency factors belonging to the category *characteristics of the workgroup* were found to be significant, while in the case of *percentage of project completed*, subsets of contingency factors from all five categories, viz., *characteristics of the workgroup*, *characteristics of the tasks*, *the interface between workgroup*, *characteristics of knowledge* and *characteristics of information technologies* were found to be significant. The details of the significant results mentioned above are discussed next.[39]

5.1.1. Contingency factors affecting average consensus

A workgroup's *average consensus* indicates its members' overall agreement about each others' areas of knowledge. Consensus among a workgroup's members is seen as one of the metrics by which the workgroup's transactive memory can be measured (Austin, 2003). Not all categories of contingency factors were found to have significant effects on *average consensus*. Figure 5.1a presents an overview of the results.[40] In each of the three categories of contingency factors, *characteristics of the workgroup*, *characteristics of the tasks* and *the interface between the workgroup and the tasks*, some contingency factors were found to have a positive effect, while others, a negative effect. The specific effects of each of the contingency factors that were significant are explained in the rest of this section.

39 It should be noted that in this section's subsections, unless specified explicitly, all results that are reported as significant are between-group effects.

40 To ensure consistency across all the figures in this chapter, the same legend is used to depict the specific effects of the categories of contingency factors on the individual outcome variables.

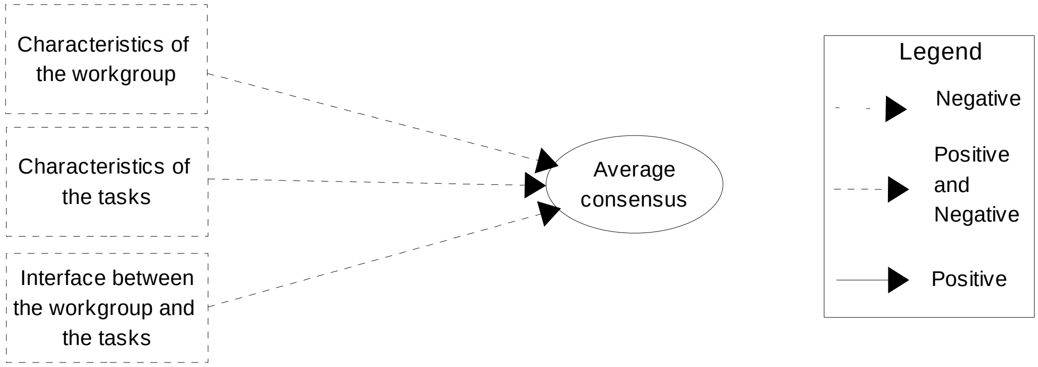


Figure 5.1a. Categories of contingency factors that have significant effects on *average consensus*

The contingency factor *average propensity to share* (category: *characteristics of the workgroup*) was found to have a positive significant effect. This is consistent with the theorizing and findings of researchers (e.g., Austin, 2003; Brandon and Hollingshead, 2004; Wegner et al. 1995), according to whom, openness in communication is essential for members of a workgroup to learn about others' areas of knowledge. Additionally, the contingency factor *probability of exchange of information about a non-task-specific-knowledge area* (category: *characteristics of the workgroup*) was found to have a positive significant effect on *average consensus*. The non-task-specific exchanges are the ones where a workgroup's members discuss topics that are not related to work, described by Davenport and Prusak (1998), as “water-cooler conversations”. Davenport and Prusak (1998) explain that such informal conversations provide opportunities for members of organizations to learn more about each other's skills, experiences and knowledge, which might subsequently lead to collaborations involving exchange of knowledge. Hence, if the likelihood of such exchanges is higher, the members of a workgroup would be more likely to know about their colleagues' areas of knowledge, thereby leading to a more developed transactive memory. Additionally, if the members are more willing to share information about their areas of knowledge, other members of the workgroup are more likely to develop perceptions of their areas of knowledge. Hence, the likelihood of informal conversations and the likelihood that members would share information about their knowledge during the informal (non-task-related) conversations have a positive effect on the outcome *average consensus*.

The contingency factor *time spent in training* (category: *characteristics of the workgroup*) was also found to have a positive significant effect. This is consistent with the findings of Liang et al. (1995), Lewis et al. (2005) and Ren et al. (2006), who state that during the training that a workgroup's members undergo together, several opportunities are provided during which they can interact with each other, with a consequence of learning about each other's areas of knowledge and thereby developing a perceptions of each other's areas of knowledge.

The effect of the contingency factor *number of agents* (category: *characteristics of the workgroup*) was found to be positive and significant. This result suggests that under the conditions of this study, where exchange of information is predicated on the transmission of knowledge related to a task, larger workgroups tended to produce a greater number of transmissions related to knowledge, and thus, a greater number of exchanges of information about non-task-related knowledge. Hence, members of a workgroup that is larger in size, and where the search for a knowledge source is not curtailed due to organizational factors, are more likely to locate a knowledge source and, while obtaining the needed knowledge, are also likely to develop a better understanding of each other's areas of knowledge, thereby resulting in a higher value of *average consensus*.

The contingency factor *number of locations* (category: *characteristics of the workgroup*), also was found to have a positive significant effect on *average consensus*. As described in subsection 3.2.1.16, if two members of a workgroup are geographically separated, they need to use an electronic medium for communications. Additionally, the use of electronic media tends to increase the time required for transmitting both knowledge and, in the case of direction, the time required for transmitting instructions that are based on the appropriate knowledge. Thus, to the extent that knowledge-related communications lead to the exchange of information about non-task-related areas of knowledge, a pair of members using an electronic medium for communicating, are more likely to have a greater number of exchanges of information about non-task-related areas of knowledge, since the duration of transmission is longer. Hence, the greater the geographical spread of a workgroup's members, as indicated by the contingency factor *number of locations*, greater is their *average consensus*.

The contingency factor *average self-knowledge* (category: *characteristics of the workgroup*) was found to a negative significant effect. In workgroups, where all exchanges between any pair of members are predicated on one of the member's need for knowledge, members are less likely to contact and be involved in an exchange with other agents if they (the focal members) believe that they already possess the knowledge they need. Consequently, members of such workgroups that have higher levels of self awareness of their own areas of knowledge, are going to focus mainly on completing their tasks and are less likely to seek help from others, unless they lack the knowledge needed to complete their own work. Hence, in such groups, members are less likely to develop an understanding of others' areas of knowledge, which results in a relatively lower consensus in the workgroup about the members' areas of knowledge.

The contingency factor, *average task intensity* (category: *characteristics of the tasks*), was found to have negative significant effect on *average consensus.* This is because it adds to the time taken to complete a task. Higher values of *average task intensity*, imply that a greater number of subtasks are associated with each task, implying further a need for knowledge in a greater number of areas. Thus, other factors being constant, when *average task intensity* increases, members of a workgroup are less likely to have opportunities of serving as sources for other members. This results in a fewer number of exchanges of information related to each other's areas of knowledge, and, consequently, to lower *average consensus*.

The contingency factor, *average knowledge-intensity of subtasks* (category: *characteristics of the tasks*) was found to have a positive significant effect on *average consensus*. This contingency factor indicates the degree of difficulty associated with a subtask; it is measured in terms of the amount of knowledge that a workgroup member has to apply in order to complete a subtask. Higher the values of this contingency, the greater is the need for knowledge, which further implies that, other factors being constant, members of a workgroup are less likely to have the needed knowledge and are more likely to seek knowledge from others. Hence, there is a greater likelihood of such members learning about their source's knowledge in other areas as well. Thus, such workgroups are more likely to have a higher level of *average consensus.*

Average task priority (category: *characteristics of the tasks*) was found to have a positive significant within-group effect on *average consensus*. In terms of their completion, higher-priority tasks receive preference over lower-priority tasks, assuming that they meet their precedence requirements. Therefore, tasks of higher priority are more likely to lead to interactions with other agents, thereby increasing the likelihood of a member, who is assigned a higher-priority task, interacting with other members. This leads to the consequence of a higher *average consensus*.

The contingency factor, *average project intensity* (category: *the interface between the workgroup and the tasks*), was found to have a positive significant effect both within and across workgroups. A greater number of tasks implies a greater number of task-related interactions involving exchange of knowledge, which increases the likelihood of exchange of information about non-task-related areas. Thus, a greater number of tasks leads to a higher level of workgroup transactive memory.

The effect of *average number of tasks per agent* (category: *the interface between the workgroup and the tasks*) was found to be negatively significant. This result is consistent with the finding of Littlepage et al. (2008) to the extent that an increase in average workload, which is indicated by *average number of tasks per agent*, results in conditions where the members with increased workload are matched with tasks for which they have the required knowledge. In such situations, the workgroup members with increased workload are more likely to work on their own tasks, and less likely to act as a source of knowledge in a given time period. Since they are assigned a relatively larger share of the work, they are unavailable for a relatively more number of time periods, implying that they are less likely to serve as sources of knowledge. Given this situation, at the workgroup-level, there are likely to be fewer number of knowledge-related, and consequently non-task-related exchanges of knowledge and information about knowledge. Hence, higher levels of *average number of tasks per agent* leads to lower levels of *average consensus*, to the extent that the workgroup's composition of members, their areas of knowledge and assignment of tasks, match the conditions described above.

In summary, those contingency factors that were found to have a positive effect on *average consensus*, do so by increasing the number of interactions among members of a workgroup, thereby increasing the likelihood of sharing information about each other's areas of knowledge. Those contingency factors, which were found to have a negative effect on *average consensus,* do so by reducing the likelihood of knowledge-sharing (and their associated non-task-related) interactions.

Of the contingency factors that affect *average consensus*, those that are relatively more controllable by the project manager are *time in training*, *project intensity*, *task* *intensity*, *task priority*, and *number of tasks per agent*. Those that are relatively less-controllable by the project manager are *number of agents*, *number of locations, average propensity to share*, *average self knowledge*, *probability of exchange of information about a non-task-specific-knowledge area*, and *average knowledge-intensity of subtasks*.

To the extent it is possible, the *task intensity*, *task priority* and *project intensity* must also be increased. This ensures members have a better chance of being assigned tasks that match the knowledge they have, have less ambiguity related to scheduling of tasks and have a greater number of tasks to complete, respectively. The results indicate that doing so would lead to an increase the number of interactions with others, thereby leading to greater *average consensus*.

Since, the number of human resources available, and their geographic distribution might be out of control of a project manager, these contingency factors, while contributing positively to *average consensus*, cannot be easily manipulated by the project manager. Hence, to compensate for these, *time in training* should be assigned a maximum possible value, so that during the *training* phase, members become more familiar with each other's areas of knowledge. It is also possible that *number of agents* is more easy to control than the geographic locations of the members of a workgroup. Hence, to the extent possible, this contingency factor should also be maximized, so that members have a greater number of knowledge sources to contact, and learn about their areas of knowledge. Through appropriate incentives (see Kankanhalli et al. 2005), members of a workgroup can be motivated to share knowledge with others. However, their propensity to share cannot be fully and directly controlled. Hence, by increasing the number of members in a workgroup, the effect of relatively less control a project manager has over the workgroup members' propensity to share can be partly mitigated. Figure 5.1b presents a categorization of the contingency factors into more-controllable and less-controllable groups, based on the above explanation.

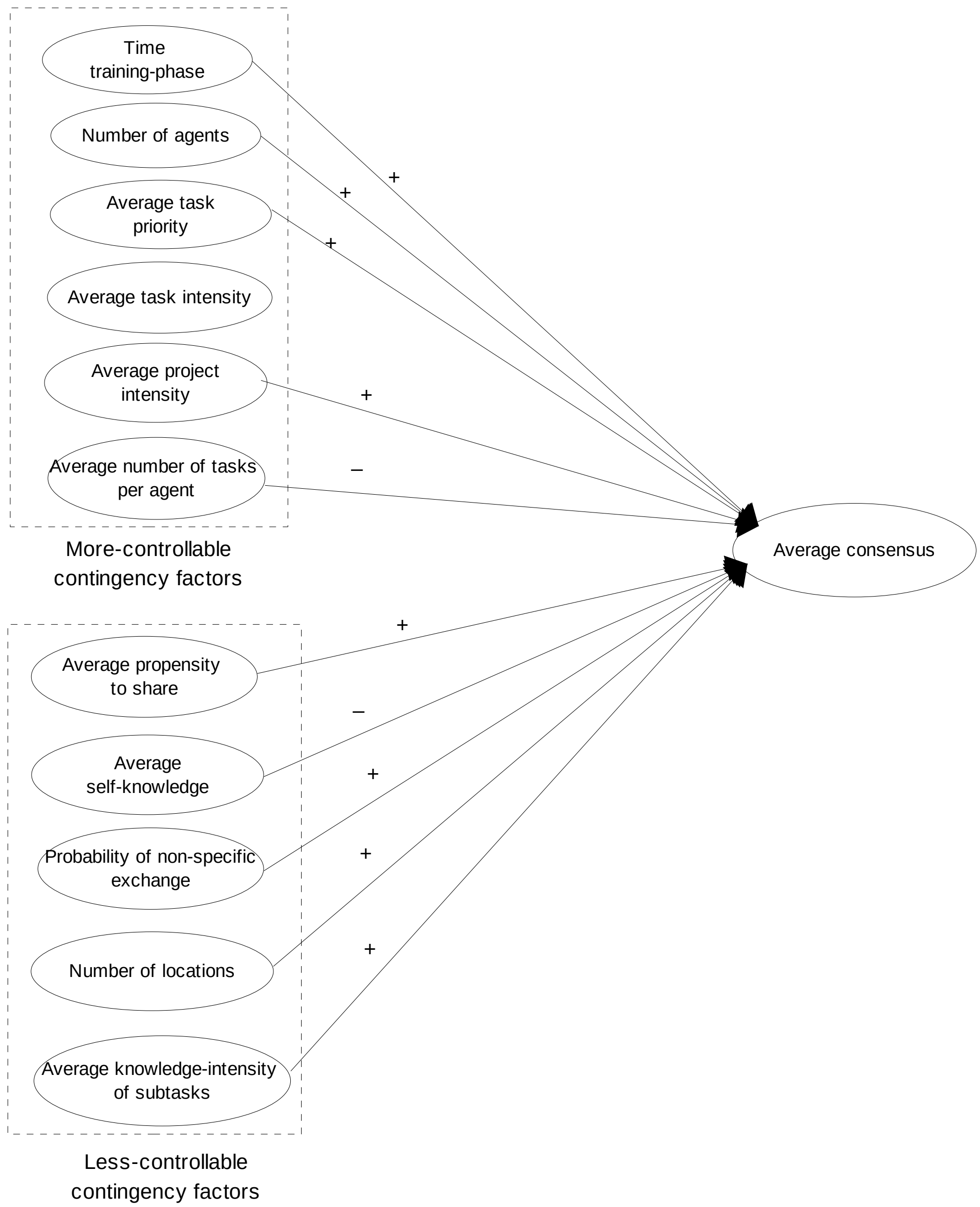


Figure 5.1b. Contingency factors that have significant effects on *average consensus*, categorized into "more-controllable" and "less-controllable" groups.

5.1.2. Contingency factors affecting average accuracy of knowledge

Compared to the outcome variables *average consensus* and *percentage of project completed*, relatively fewer number of contingency factors were found to have a significant effect on *average accuracy of knowledge*; both contingency factors belong to the category *characteristics of the workgroup*. *Average propensity to share* and *probability of exchange of information about a non-task-specific-knowledge area* are the two contingency factors that were found to have a positive significant effect on *average accuracy of knowledge*. Consistent with the approach described in section 5.1.1, the following figure (figure 5.2) depicts a high-level view of the effects of the category of contingency factors that affect *average accuracy of knowledge*. Following the figure, details of the contingency factors that significantly affect *average consensus* are presented.

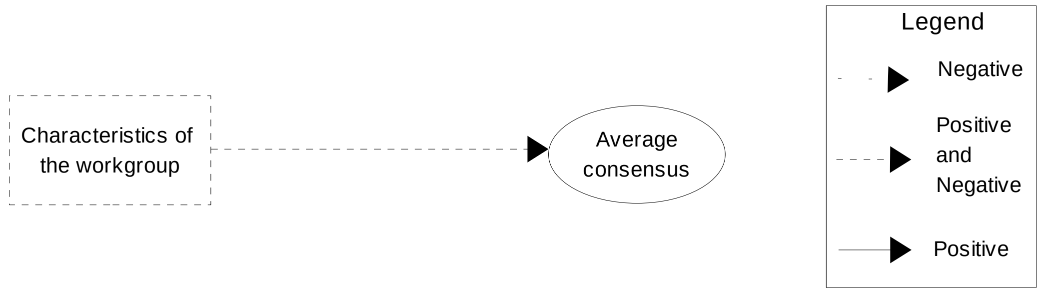


Figure 5.2a. Categories of contingency factors that have significant effects on *average accuracy of knowledge*

A member of a workgroup seeking knowledge in an area where he/she does not possess it is more likely to succeed if he/she knows, based on past interactions with other members of the group, the areas of knowledge of other members of the workgroup. Additionally, even if a potential source was found, the likelihood that the recipient will indeed obtain the required knowledge depends on whether the source is willing to provide the requested knowledge accurately. Hence, in workgroups where agents are more likely to (a) share information about their areas of knowledge, and (b) provide the requested knowledge accurately, the knowledge level of the workgroup as a whole, indicated by *average accuracy of knowledge*, goes up. This result finds support in the findings of Shen et al. (2008) who found that when a workgroup's membership fragments, thereby leading to decrease in the sharing of knowledge among the workgroup's members, the workgroup's members can no longer learn about each other or also learn new knowledge. Consequently, both the workgroup's average consensus and accuracy of knowledge are lowered.

In summary, the two contingency factors, *average propensity to share* and *probability of exchange of information about a non-task-specific-knowledge area*, both of which belong to the category *characteristics of the workgroup*, have a significant positive effect on *average accuracy of knowledge.* They do so by increasing the frequency, and the likelihood during each time period, of the workgroup members receiving accurate knowledge from their colleagues.

Of the two contingency factors, *number of agents*, which indicates the number of workgroup members, is more controllable, and can be used to compensate for the relatively lesser amount of control possible on the *average knowledge level*. This is because the knowledge possessed by individual members of a workgroup is affected several factors, such as prior on-job experience, educational background, etc. Training can be used to increase the amount of project-relevant knowledge possessed by each workgroup member. However, given the time and resource constraints, the amount of project-related knowledge that can be transferred to each workgroup member is limited. Hence, to ensure that the aggregate amount of knowledge available to a workgroup is maximized, the size of the workgroup has to be increased to the extent allowed by the constraints on human resources in the organization. It should, however, be noted, that while increasing the absolute number of agents increases the aggregate amount of knowledge available, for the knowledge to be useful, it should be accurate. The negative effect of *number of agents* on *average accuracy of knowledge* implies that in order for the average level of knowledge that can be applied (that is, knowledge that is accurate) an increase in the size of the workgroup should result in an addition of individuals who have high levels of accurate knowledge, in order for them to be useful additions to the workgroup. The following figure (figure 5.2b) illustrates the above explanation.

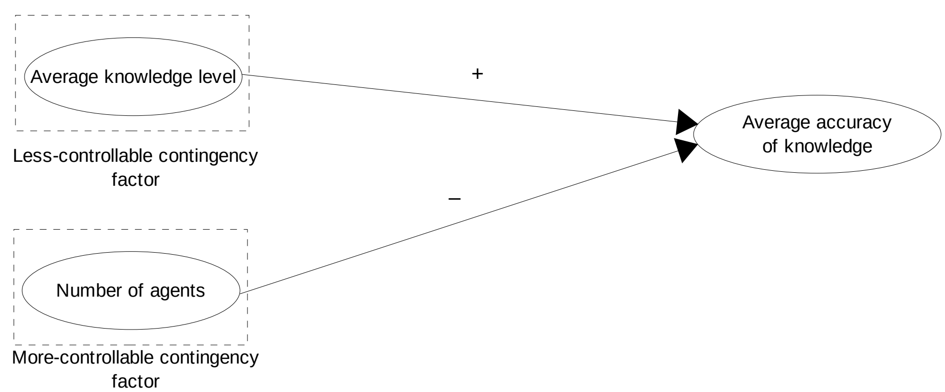


Figure 5.2b. Contingency factors that have significant effects on *average accuracy of knowledge*, categorized into "more-controllable" and "less-controllable" groups

5.1.3. Contingency factors affecting percentage of project completed

The outcome variable *percentage of project completed* represents the performance of a workgroup – it indicates the proportion of work, which was assigned to a workgroup, that is completed by the workgroup. Subsets of contingency factors belong to all five categories were found to have significant effects on *percentage of project completed*. Figure 5.3 depicts a high-level view of the effects of the categories of contingency factors that affect *percentage of project completed.*

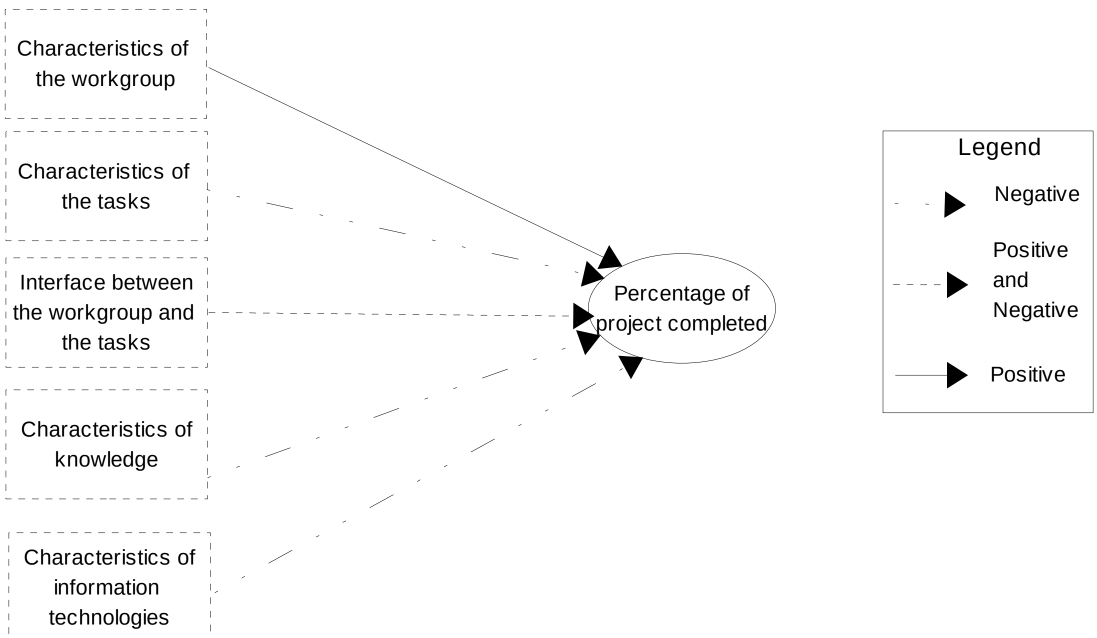


Figure 5.3a. Categories of contingency factors that have significant effects on *percentage of project completed*

In the category *characteristics of the workgroup*, the contingency factors *average knowledge level*, *average propensity to share*, *average self-knowledge*, *time in training phase*, and *probability of exchange of information about a non-task-specific-knowledge area* were all found to have a positive significant effect on *percentage of project completed.*

Workgroups with higher initial value of *average knowledge level* are more likely to complete a greater proportion of the work assigned to them, because their members are better equipped to complete the tasks. This result is consistent with the notion that the performance of a workgroup depends on the degree of skills and knowledge possessed by the members of the workgroup (Larson and Lafasto, 1989).

Workgroups whose members are more likely to share accurate knowledge with each other, as indicated by *average propensity to share*, are more likely to complete a greater proportion of the work assigned to them, since a greater proportion of the knowledge-seekers are going to receive the knowledge needed for completing their assigned tasks. This is partly consistent with the findings of Cummins (2004), who found that knowledge-sharing had a positive influence on a workgroup's performance and that this result was moderated by the structural diversity of the workgroup.

The observed positive significant effect of *average self-knowledge* on *percentage of project completed* is partially consistent with the results reported by Katz-Nevon and Erez (2005) who found that, under conditions of low task interdependence, self-efficacy (used to indicate awareness of one's strengths, which is similar to the operationalization of *self-knowledge* in this study) was a significant positive predictor of the performance of individuals of a workgroup.

The two contingency factors, *time in training phase* and *probability of exchange of information about a non-task-specific-knowledge area*, increase the likelihood of a member of a workgroup learning about his/her colleagues' areas of knowledge. Hence, both contingency factors were found to have a significant positive effect. The observed effect of the *time in training phase* is consistent with the findings of Lewis et al. (2005), Liang et al. (1995) and Ren et al. (2006), who observed a positive significant relationship between the duration of the training period of a workgroup and its performance. The observed effect of *probability of exchange of information about a non-task-specific-knowledge area* is consistent with the positive effects of informal conversations on the spread of knowledge and consequently on the workgroup's performance that was described by Davenport and Prusak (1998).

In the category *characteristics of the task*, *average knowledge intensity of subtasks* was found to have a negative significant effect, both within-groups and between-groups, on *percentage of project completed.* Within a workgroup, those subtasks that are associated with a larger number of knowledge areas are less likely to be completed. This is because, for a given set of workgroup conditions, including the knowledge level of workgroup's member, the likelihood of an agent (a) having the required knowledge in all knowledge areas associated with the given subtask, or (b) being able to locate an agent or a set of agents that has knowledge in all the required areas associated with the given subtask, decreases with an increase in the number of areas in which knowledge is required. Hence, within a workgroup, subtasks with a greater number of knowledge area requirements are less likely to be completed.

In addition to having a negative significant effect within a workgroup, *average knowledge-intensity of subtasks* was found to have a *negative effect on percentage of project completed* across workgroups. Using the reasoning presented in the previous paragraph, it can be inferred that, across workgroups, those workgroups whose subtasks have a lower knowledge requirements are more likely to be completed. This result is consistent with Grant's (1996a; 1996b) exposition that organizations succeed in dynamic environments when they have the ability to acquire and apply the needed knowledge. Furthermore, an organization's ability to successfully compete, which involves meeting its goals that are derived based on a strategy that is meant to help the firm compete, is limited by the knowledge of its employees. Thus, firms or workgroups that try to complete projects which require a bigger repertoire of knowledge, are less likely to be successful when their members lack all the required knowledge or are unable to distribute the needed knowledge among themselves such that the needed knowledge is made available to the employees in charge of completing a work unit.

In the category, *the interface between the workgroup and the tasks*, the contingency factor, *average number of tasks per agent*, was found to have a positive significant effect within workgroups and a negative significant effect between workgroups. The differential effects of workload on workgroup performance may be explained by the observation made by Littelpage et. al. (2008). According to Littlepage et al. (2008) the context in which the workgroup performs its tasks, along with the details of the tasks themselves and how well they match with the skills and expertise of the workgroup's members, are relevant in determining the effect of workload on a workgroup's performance. They also found that an increase in the workload of an employee may lead to a better performance of the workgroup, if those members of the group that have superior skills and knowledge are assigned the work that they are most proficient (compared to other group members) of doing. This was explained as resulting from a reduction in inter-personal communication, with a consequence of reduction in wastage of time spent on non-task-related activities. However, an additional consequence is that it would decrease the likelihood of the development of transactive memory in the workgroup. In the agent-based model used in this dissertation, the allocation of tasks was random, not based on a match between an agent's knowledge and the knowledge requirements of the subtasks associated with the task. The collection of agent-level data, which indicates the degree of match between an agent's knowledge areas and the knowledge requirements of the task assigned to it, was not programmed into the simulation. Hence, the similarity of this study's conditions to those described by Littlepage et al. (2008) could not be ascertained.

In a study with results that are opposite to those reported by Littlepage et al. (2008), Urban et al. (1995) have found that an increase in workload leads to a decrease in the workgroup's performance. It should be noted, however, that the study by Urban et al. (1995) did not include the construct of transactive memory. This result is consistent with the between-groups negative significant effect of *average number of tasks per agent* on *percentage of project completed.* Thus, to summarize, workload per member tends to have a positive effect within a workgroup, if it leads to a better match between members' knowledge and the knowledge requirements of the tasks assigned to them. Across workgroups, however, greater workload per member tends to decrease the performance of the workgroup (presumably, if it does not lead to a corresponding increase in the match between members' knowledge and the knowledge requirements of the tasks assigned to them).

In the category, *characteristics of knowledge needed to complete the tasks*, the contingency factor, *average direction time*, was found to have a significant negative effect (both within-group and between-groups) on *percentage of project completed*. Higher *average direction time* values imply that members of a workgroup are more likely to spend time giving others direction, an activity which does not lead to knowledge transfer. Hence, over time, fewer members of a workgroup would receive the knowledge needed to complete all the subtasks assigned to them. Thus, over time, relatively fewer number of subtasks would be completed, leading to lower *percentage of project completed*.

In the category, *characteristics of information technologies*, the contingency factor *characteristics of email* was found to have a negative effect on *percentage of project completed*. Email as a communication medium has the lowest absolute richness among all communication media. However, it is possible that some members of a workgroup might perceive it to have a higher level of richness than other members do. This relatively high perception of the richness of email may be due to factors such as shared common knowledge among a pair/group of members who use email to convey richer messages via text, by relying on their shared background knowledge (Ngwenyama and Lee, 1997). In the agent-based model, the above rule was taken into account in determining (a) the perceived richness of email, and, (b) how it affects the expected time for transmitting knowledge via direction and transfer (equation 3.7, subsection 3.2.1.15). Based on equation 3.7, it can be seen that use of email for communications can lead to an increase in the transmission time. In terms of actual richness, email ranks the lowest among the three communication media and as such should be used for low priority tasks. But, if individuals perceive it to be a rich medium, they may start using it for high-priority tasks as well, leading to lower knowledge transfer and, consequently, to a lower level of project completion. Once a member of a workgroup feels that he/she needs to acquire knowledge from another individual, and identifies a potential source, the member would be more likely to use email for knowledge transfer, if the member perceives email to be a rich medium. However, the member would fail to acquire knowledge, actually, because email would not transfer knowledge. As a result, over time, the likelihood of workgroup members having the knowledge needed to complete all their assigned subtasks would be lower when perceived richness of email is higher. Therefore, subtasks are more likely to be abandoned, resulting in lower project completion rates, as seen in *percentage of project completed* values.

In summary, the contingency factors that have a positive significant effect on *percentage of project completed*, are those characteristics of the workgroup that enable the members of the workgroup in locating sources of knowledge within a limited time period, encourage them to share the requested knowledge and information about knowledge, and provide them with knowledge in several areas that is required by them to complete their assigned tasks. The effects of the contingency factors that were observed to have a negative effect on *percentage of project completed* are due to the fact that they decrease the likelihood of transfer of knowledge that is needed for completion of tasks, thereby leading to a greater proportion of subtasks being abandoned.

Of the contingency factors affecting *percentage of project completed*, the contingency factors *average number of tasks per agent* and *time in training-phase* are more-controllable, while the contingency factors *average propensity to share*, *average knowledge level*, *average self-knowledge*, *probability of exchange of information about a non-task-specific knowledge area*, *average direction time* and *average knowledge intensity* *of subtasks* are less-controllable. This is illustrated in the following figure (figure 5.3b).

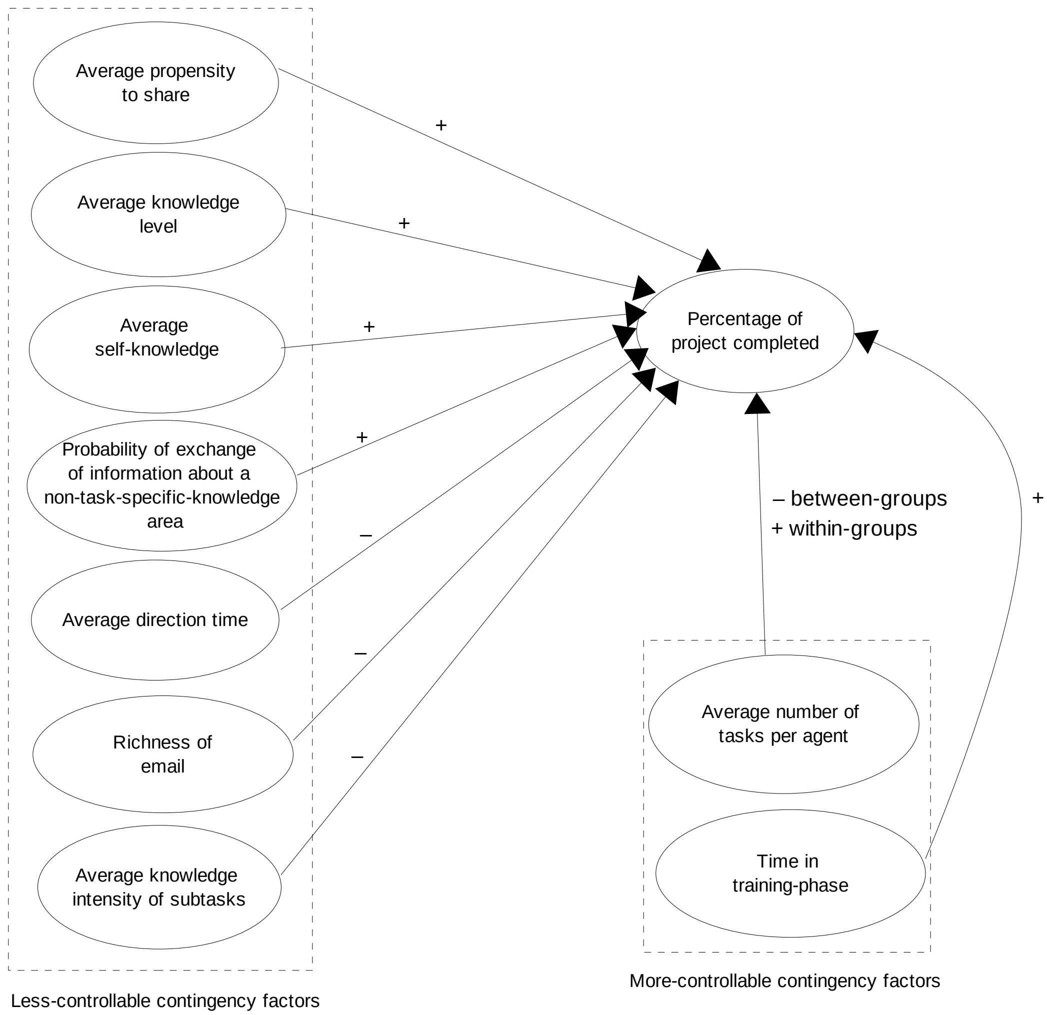


Figure 5.3b. Contingency factors that have significant effects on percentage of project completed, categorized into "more-controllable" and "less-controllable" groups.

As illustrated in figure 5.3b, a majority of factors affecting *percentage of project completed* are less-controllable, while only two, *time in training-phase* and *average number of tasks per agent* are more-controllable. The effect of *time spent in training-phase* complements the effects of *average knowledge level* and *average self-knowledge.* Workgroup members succeed in completing their assigned tasks if they have the required knowledge. However, if they lack the required knowledge and need to obtain it from their colleagues, the effect of *time spent in training-phase* comes into play – they are more likely to develop a transactive memory and thus more likely to succeed in locating a source for the knowledge they require. This implies that the time spent in the training phase is a crucial determinant and, given that is controllable, it should be maximized.

The effect of *average number of tasks per agent* indicates that within a workgroup across time, increasing the workload of members leads to better performance. However, while there is no direct evidence from this study, prior work (see Littlepage et al., 2008) suggests that increasing the work-load of members of a workgroup has a positive effect on the workgroup's performance if such an increase results in a better match between the knowledge requirements of the tasks and the knowledge possessed by the members of the workgroup who are assigned a greater amount of work. On the other hand, across workgroups, those with higher *average number of tasks per agent* were shown to have lower performance. This could result, if, in workgroups with higher overall workload per member, there is a poorer fit between members and the tasks assigned to them. Data at the workgroup member-level (agent-level) were not collected in this dissertation, so the exact reason why *average number of tasks per agent* has a differential effect could not be ascertained substantively.

The overall effects of the contingency factors are summarized in table 5.1. It can be seen that of the effects of all the contingency factors, the effects of (a) *average self-knowledge*, (b) *average knowledge intensity of subtasks*, and (c) *average number of tasks per agent* on *average consensus* and *percentage of project completed* are contradictory. Specifically, *average self-knowledge* has a negative significant effect on *average consensus* and a positive significant effect on *percentage of project completed*; *average knowledge intensity of subtasks* has a positive significant effect on *average consensus* and a negative significant effect on *percentage of project completed*; *average number of tasks per agent* has a negative significant effect on *average consensus* and mixed effects on *percentage of project completed*. Of the above three contingency factors, only *average number of tasks per agent* is more-controllable. Using the explanation provided by Littlepage et al. (2008), the positive effect of *average number of tasks per agent* is expected to occur when it results in a better match between the knowledge of members of a workgroup and the knowledge

Table 5.1. Summary of significant effects of contingency factors on the outcomes

Contingency Factors	*Controllability*[41]	*Effect on Outcomes*		
		Average consensus of the workgroup	*Average accuracy of knowledge of the workgroup*	*Percentage of completion of work*
Average knowledge level (characteristics of the workgroup)	Less-controllable	Positive	Positive	Positive
Average self-knowledge (characteristics of the workgroup)	Less-controllable	Negative	–	Positive
Number of agents (characteristics of the workgroup)	More-controllable	Positive	Positive	–
Number of locations (characteristics of the workgroup)	Less-controllable	Positive	–	–
Probability of exchange of information about a non-task-specific-knowledge area (characteristics of the workgroup)	Less-controllable	Positive	–	Positive
Time in training phase (characteristics of the workgroup)	More-controllable	Positive	–	Positive
Average propensity to share (characteristics of the workgroup)	Less-controllable	Positive	–	Positive
Average task intensity (characteristics of the tasks)	More-controllable	Negative	–	–
Average task priority (characteristics of the tasks)	More-controllable	Positive	–	–
Average knowledge intensity of subtasks (characteristics of the tasks)	Less-controllable	Positive	–	Negative
Average project intensity (the interface between the workgroup and the tasks)	More-controllable	Positive	–	–
Average number of tasks per agent (the interface between the workgroup and the tasks)	More-controllable	Negative	–	Positive (within-workgroup); Negative (between-workgroups)
Average direction time (characteristics of knowledge)	Less-controllable	–	–	Negative
Characteristics of email (characteristics of information technologies)	Less-controllable	–	–	Negative

requirements of the tasks assigned to them. The other "more-controllable" contingency factors that have positive effects on the outcomes are *number of agents*, *time in training phase*, *average task intensity*, *average task priority* and *average project intensity.* Hence,

41 Indicates the degree to which a contingency factor is controllable by a project supervisor/manager.

increasing the values of the above-identified contingency factors, to the extent it is feasible to do so, while ensuring that members of a workgroup are matched closely with to the tasks that they are equipped to complete, would result in an overall positive values for the workgroup outcomes.

More detailed recommendations to practitioners are provided in section 5.3. Next, limitations of the current dissertation are presented.

5.2. Limitations

The goal of this dissertation was to understand the effect of five categories of contingency factors on three outcome variables that represented a workgroup's transactive memory, the aggregate level of knowledge of its members and the performance of the workgroup. The approach of the study was exploratory, to determine how each of the contingency factors, result in different levels of each of the outcome variables, via a fixed, per-determined, set of processes that represent a stylized workgroup and its operations. Given the above aspects of the current study and the methodology, the following are key limitations of this dissertation.

The collection of data, via simulation, at the workgroup level meant that the effects of individual differences in workgroups on the outcome variables, via the fixed set of processes that describe individual members' behaviors, could not investigated. This precluded a more nuanced understanding of how interactions between the differences in individual characteristics on the one hand, and the five categories of contingency factors on the other hand, affect the outcome variables.

In realistic workgroups, project deadlines matter: prior work (e.g., Waller et al., 2001) showed that time pressure affects a workgroup's performance. Members of the workgroups in the current study experience time pressure only during their search for a source of knowledge; they do not experience any time constraints while obtaining knowledge or during the completion of the task. This specific operationalization was made to ensure that the complexity associated with the implementation of the agent-based model, particularly in the light of the exploratory nature of the study, was manageable. Hence, the effects of the interplay between contingency factors from the category *characteristics of knowledge* (specifically, *average stickiness time* and *average direction time*), the choice of the communication media, as determined by the contingency factors from the category *characteristics of the information technologies* and a suitably operationalized version of time pressure, could not be investigated. Consequently, the question “how does the workgroup's choice of information technologies for its communication affect each of the three outcome variables under varying levels of time pressure and varying levels of characteristics of knowledge?” while interesting, could not be answered.

Another simplifying assumption of the study was that the the change in the knowledge requirements across a series of projects would be constant. Changing knowledge requirements across a series of projects may act as another stressor on the workgroup's members. As Lewis et al. (2007) indicate, a change in the knowledge requirements might make old-timers' knowledge obsolete, thereby affecting the collective pool of useful knowledge available to a workgroup. Prior simulation-based studies, e.g., March (1991) and Kane and Alavi (2007), implemented the notion of environmental turbulence to represent obsolescence of a workgroup's knowledge. In the current study, effects of such factors as members' forgetfulness and environmental uncertainty, which results in a change of knowledge areas associated with tasks in a given project, were not implemented. Hence, the effect of obsolescence of members' knowledge on the outcomes could not be investigated.

Based on the interviews and support from prior literature, the current study implemented the effects of three electronic communication media. However, as literature in the area of knowledge management and organizational learning (e.g., Kane and Alavi, 2007; Kankanhalli et al. 2005) suggests, other technologies such as *knowledge repositories*, *electronic communities of practice*, are being used by organizations for managing their knowledge. Hence, while the functionalities of information technologies have been modeled, the variety of technologies investigated is smaller in comparison to the technologies reported in literature.

The model implemented in this dissertation assumes that members' propensity to share with others remains constant across time. However, literature (e.g. Kankanhalli et al., 2005) indicates that a member's propensity to share knowledge and information about presence of knowledge may change over time, as a result of interactions with other members of the workgroup. The assumption of constancy of members' propensity to share prevented an investigation of the patterns of change in members' propensity to share (which may be seen as another outcome) and how this change compares with the patterns of changes in the other outcome variables, across time.

In the current set of analyses, standard deviation values of the contingency factors and the outcome variables were not included in regressions because doing so would result in the following:

(a) an increase in the number of findings that need to be reconciled and synthesized in a meaningful fashion in the current study, and,

(b) a need for skewness data of the variables, so that the results of the standard-deviation- regressions could be explained in terms of the skewness in the values of the contingency factors

Based on the above reasons, it was deemed that the additional analyses would add further to the scope of results reported and to be synthesized, and are best left for future investigations.

Next, the implications of the current study for practitioners are identified.

5.3. Implications for practice

Based on the the effects of the different contingency factors that were investigated in this dissertation, the following recommendations are made to practitioners who are project managers, human-resources personnel and designers of information technologies used for communication and collaboration:

- Project managers should assign unambiguous priorities to each task in the project. This allows members of the workgroup in performing their own, internal scheduling of tasks assigned to them with a goal of maximizing the number of subtasks (and thus tasks) they can perform. Information about relative priority of tasks, in conjunction with their assessment of their own knowledge in the required areas and their perceptions of their colleagues' areas of knowledge, would help the members choose and work on tasks and subtasks that they believe have the highest probability of successful completion.
- Project managers should (a) partition the work into the basic units of work, the subtasks in the jargon used in this dissertation, and group them into tasks such that the decomposition of the entire amount of work into tasks is maximized while the number of subtasks per task is minimized, and, (b) make the assignment of tasks to members such that the three-way fit between tasks assigned to a person, the areas

of knowledge required to complete the tasks and the person's areas of knowledge is maximized. Suggestion (a) ensures that each member of the workgroup can determine the best order in which to complete the tasks and their subtasks assigned to her/him. Suggestion (b) ensures that each member is not overwhelmed with work, and can also take some time to help other members of the workgroup by providing them the needed knowledge.

- Project managers and their superiors must ensure that adequate incentives are provided to motivate the members of a workgroup to share the knowledge their colleagues request, so that, as a workgroup, they are more likely to complete a greater proportion of the project assigned to them. Of course, this suggestion must be employed in conjunction with the previous suggestion regarding the workload assigned to each member so that member have both the motivation and the opportunities to share their knowledge.
- Project managers and their superiors should also encourage socialization among the workgroup members. Socialization facilitates the workgroup's members learning about each other's areas of knowledge, and can thus know whom they may contact when they need knowledge to complete an assigned task.
- Project managers should recognize the positive effect of provide training. The training should be of adequate duration, and provide knowledge that is necessary for members of a workgroup to complete their assigned tasks. Additionally, the training period also provides opportunities for members to socialize and learn about each other's areas of knowledge, whose positive effect is mentioned in the previous point.
- Misconceptions about the usefulness of different information technologies in the exchange of knowledge might result in the choice of a technology, such as email, by members of a workgroup to exchange knowledge associated with a high-priority task. Results from this dissertation show that such a choice has negative consequences. Hence, human resources executives and project managers must ensure that workgroup members are provided appropriate training so that the right type of technology is used for exchanging knowledge of specific level of difficulty associated with tasks of specific priorities.
- Additionally, those responsible for designing and implementing information

technologies that are aimed towards communication should incorporate functionality that allows both synchronous and asynchronous communication[42]. This allows workgroup members to choose the most appropriate mode of communication using the same technology. Also, with continued use, workgroup members will develop a better understanding of the technology that allows them to use the technology in the best way possible, leading to better knowledge-exchange, and consequently, better performance, outcomes for the workgroup.

- In addition to being encouraged to share knowledge, members of a workgroup must also be encouraged to provide knowledge via transfer, rather than via direction, where possible. This ensures that in the future, a greater number of members of a workgroup working on a task that requires knowledge that they lacked prior to obtaining it from their colleague, can use the knowledge they obtained and complete the task immediately. This improves the performance of the workgroup.

Next, directions for future research that can build on the current dissertation are presented.

5.4. Implications for theory

This dissertation considered the simultaneous effects of a relatively large set[43] of contingency factors on each of the three outcome variables via a fixed set of workgroup processes. As described in chapters 2 and 3, the contingency factors that were included were drawn from the results and theoretical arguments presented in prior literature. The inclusion of the contingency factors in the current study implies, implicitly, that they are expected to have an effect on at least one the outcome variables. However, the results suggest that on each outcome variable, only a subset of contingency factors are significant. In this section, a brief explanation of what was surprising about the results and what these results imply for future development of theory is presented. “Surprise” associated with the effect of a particular contingency factor is defined here as a result that was unexpected based on the rationale used for the inclusion of the contingency factor in the model.

Based on literature related to transactive memory, specifically cognitive-interdependence, (e.g., Wegner, 1986) average knowledge level and self knowledge were

42 Presently, technologies such as Skype and Google's Gchat, incorporate functionality that facilitates synchronous and asynchronous communication, in addition to providing rich communication via video and voice chat.

43 Relative to the typical number of contingency factors considered in each study drawn from prior literature

both included in the model used in this dissertation. However, as described in chapter 4 and as summarized in table 5.1, the effects of these two contingency factors is different on different outcome variables. Only by explicitly considering the individual effects of these two factors, could their unique contributions to the outcome variables be understood. The different roles played by the perceptions of their own knowledge by members of a workgroup and the aggregated knowledge of the workgroup on a workgroup's outcomes was not hitherto studied in much depth. This lack of prior work, along with the different results of the two contingency factors, as reported in this dissertation, imply that their study, in conjunction, would lead to a greater understanding of their individual and joint contribution to workgroup outcomes.

Prior work (e.g., March 1991) indicated that *probability of turnover* can have a significant effect on a workgroup's knowledge and performance outcomes. However, in the model used in this dissertation, *probability of turnover* was shown to not have a significant effect. This was unexpected in the light of prior studies and their results. In future work, why this is so, that is, which other contingency factors mask the effect of this contingency factor, should be investigated to determine how the result found here can be reconciled with expectations based on other, prior studies. Associated with *probability of turnover*, is the contingency factor *average proportion of knowledge areas common with the replaced agent*. This factor was included to observe whether variations in it would affect any of the outcome variables. The results imply that, perhaps, the lack of significance of this factor is associated with the lack of significance of *probability of turnover*.

The expectation, based on the inclusion of the contingency factor *maximum number of failed tries* was that this factor would have a positive effect on the outcome *percentage of project completed*, since a greater number of tries were expected to lead to a higher likelihood of finding a source who can provide the knowledge, and therefore, a higher likelihood of obtaining the required knowledge, thereby completing the assigned subtask. The non-significant effect of this contingency factor, was therefore unexpected. In the future, the reasons why this factor was found to not have a significant effect and under what conditions, and, under what set of workgroup processes this contingency factor has a significant effect, would lead to a greater understanding of the interactions between this contingency factor and the workgroup processes and their consequential effect on *percentage of project completed.*

The contingency factor, *average openness to change*, was expected to affect *average consensus,* and, indirectly, *average accuracy of knowledge*. The lack of significant effect of this contingency factor implies that its effect is possibly masked by one or more other contingency factors. Future investigations to determine the exact mechanism by which this masking of the contingency factor's effect would lead to a greater understanding of the interplay between this and other contingency factors and the workgroup processes and their consequent effects on the outcome variables.

Of the characteristics of three electronic media, which are used for communication, in addition to face-to-face communications, only *richness of email* was found to have a significant effect, and only on the outcome variable *percentage of project completed*. This implies that under the set of workgroup processes considered in the current study, the use of email, which differed from the other media by supporting asynchronous communication, affects the workgroup processes, and consequently *percentage of project completed* differently from the other media. While chat and email, as implemented in the model, have similar ranges of richness values, email was shown to be the one with a significant effect. This was unexpected. In future work, the interplay between synchronousness (as defined in chapter 3) of communications, the nature of the knowledge and tasks, and the workgroup processes must be explored in greater depth to understand the specific effect, and to validate the explanation provided in subsection 5.1.3.

Prior work on the use of *expert-seeker* (e.g., Becerra-Fernandez, 2000) implied that its use would significantly reduce the time taken to search for knowledge sources, therefore leading to better performance outcomes. However, the lack of significant effects of the contingency factor *use of expert-seeker* implies that the effect of this factor is masked by other contingency factors and their interactions with workgroup processes. Understanding this phenomenon would help determine under which circumstances its usage will make a positive contribution to the workgroup outcomes.

Of the two characteristics of knowledge, *average stickiness time* and *average direction time*, only direction time was found to be significant in its effect (negative; only on *percentage of project completed*). Since agents in the modeled workgroups choose one of the two possible modes of knowledge transmission, transfer or direction, based on which of the two characteristics of knowledge results in faster transmission of knowledge, the two contingency factors' effects are not independent of each other. However, the specific effect could not be predicted, a priori, so, the result reduced the uncertainty regarding the specific effect of characteristics of knowledge.

Finally, the contingency factor, *connectedness of network of task-interdependencies* was included in the model to capture the notion of inter-dependencies among tasks; it was expected that those projects where the tasks are more inter-related might experience greater difficulties of completion and worse workgroup performance. However the lack of significant effect implies that, under the given set of workgroup characteristics and processes considered in this dissertation, this contingency factor does not affect the workgroup outcomes. This implies that further investigations, involving experiments with different sets of workgroup processes, might lead to discovery of situations where this contingency factor would have a significant effect on the workgroup outcomes.

Next, directions of future research are discussed.

5.5. Directions for future research

Future research, building on the current work can explore several avenues. An immediate extension to the current research would be the identification of a studies that use a subset of the contingency factors considered in this dissertation and at least one of the outcome variables. Then, panel data regressions of these reduced sets of contingency factors on the appropriate outcome variables would be run. The results of these regressions would be compared with those from the corresponding studies from which the contingency factors were identified to be included in the reduced models. The results of this dissertation (and indeed the implementation of the simulation) would be validated to the extent that the results of the regressions are consistent with those from the studies from which they were drawn. Inconsistency of results obtained from the regressions on the reduced set of contingency factors as regressors, and the results from the studies from which they were drawn, indicates that further investigation is necessary to understand the probable causes of the discrepancies. During the process of reconciliation of the two sets of results, points of commonality and points of divergence between the conditions of the current dissertation's model, and the conditions pertaining to the context in which the reported studies that were considered, could be used in synthesizing the findings of the two sets of studies. Such a synthesis could produce a more comprehensive theory explaining the effects of the considered contingency factors on the outcome variables.

In the future, the effect of time pressure can be studied in terms of how it interacts with the characteristics of knowledge and a workgroup's members use of information technologies for communication. Additionally, the set of information technologies used by a workgroup can be extended to include knowledge repositories and electronic communities of practice (Kane and Alavi, 2007) to determine how the emergence of transactive memory is affected by a workgroup's use of information technologies for communications and for storing, retrieving and utilizing knowledge captured in an electronic knowledge repository.

In the current work, the set of processes followed by agents representing members of a workgroup are fixed. Future work, by drawing on the current work's specifications a workgroup's behaviors, and literature detailing the knowledge-search and knowledge-exchange behaviors of workgroup members, can implement a collection of behaviors from which agents choose based on preferences for each type of behavior that evolve over time. Such work can also investigate the effects of using other members of a workgroup as references for knowledge sources. That is, in addition to the 'expert-seeker' database, members of a workgroup can refer to members of the workgroup with whom they have interacted before and seek their help in identifying sources of knowledge. Evidence for such work was indicated in interview 1 of this dissertation. Literature based on Granovetter's (1983) seminal work on the influence of an individual's social network of ties on the individual's ability to find the needed knowledge/information could inform the extensions to be made to the current model.

Another approach to the effects of contingency factors would involve considering the outcomes, *average accuracy of knowledge* and *average consensus* as the intermediary outcomes, being the result of the workgroup processes and the contingency factors, and, *percentage of project completed* as the final outcome, in the next time period. Stated differently, *average accuracy of knowledge* and *average consensus* in time period *t-1* would affect *percentage of project completed* in time period *t.* The suggested approach is as follows:

(a) the contingency factors with only between-groups components of standard deviation would be combined with the two knowledge outcomes as the regressors; the performance outcome would be the outcome variable; the two knowledge outcomes' values would be time-lagged, that is, the values of all the between-groups-varying contingency factors and the two knowledge outcomes at time *t-1* would act as the regressors in predicting the values of the performance outcome at time *t*

(b) the contingency factors with both within-groups and between-groups components of standard deviation would be combined with the two knowledge outcomes as the regressors; the performance outcome would be the outcome variable; the two knowledge outcomes' values would be time-lagged, that is, the values of all the contingency factors that vary between groups and within groups, and the two knowledge outcomes at time *t-1* would act as the regressors in predicting the values of the performance outcome at time *t*. These regressions would indicate whether the knowledge outcomes in a given time period, acting as intermediary outcomes, are significant predictors of the performance outcome in a future time period.

The above approach would lead to the development of richer process-based theory explaining the effects of the contingency factors on the intermediary and final outcome variables across time, as a function of a given set of workgroup processes.

Finally, future work can implement methods by which individual behaviors of the members of a workgroup are recorded at each time period, along with the changes in the workgroup-level variables and outcome variables. Such a dataset can be analyzed by mixed-effects methods (Frees, 2004) to reveal the role that different interactions between individual-level and workgroup-level attributes play across time in producing different levels of outcome variables.

5.6. Conclusions

There are three phenomena affecting the effectiveness of an organization's knowledge over time: (1) an increasingly-transient workforce; (2) ever-increasing levels of data, and the associated increase in the knowledge required to process the data; and, (3) the large proportion of knowledge staying in individual's minds, rather than being converted into knowledge artifacts. Considered together, the three phenomena imply that in order to successfully obtain, share and apply knowledge, organizations must understand how various contingency factors affect the processes by which knowledge, transactive memory and performance of workgroups result. These phenomena motivated the investigation carried out in this dissertation, which was guided by the following research question:

Which factors, from the five categories of factors (a) characteristics of the workgroup; (b) characteristics of the tasks assigned to the workgroup; (c) the interface between the workgroup and the tasks; (d) characteristics of the knowledge required to complete the tasks; and (e) characteristics of the information technologies, affect workgroup outcomes, including (i) average consensus among a workgroup's members about each other's areas of knowledge; (ii) average accuracy of knowledge; and (iii) performance of the workgroup, over time, and in what way?

Figure 5.4 (presented at the end of this chapter) presents an overview of the findings of this dissertation. A review of extant literature revealed that few studies exist that investigated the longitudinal effects of the salient contingency factors, simultaneously, on the three outcome variables. Hence, this dissertation used an exploratory approach that involved simulations of an agent-based model to answer the research question.

The agent-based model that was used to investigate the effects of the contingency factors was created by drawing on the findings reported in literature. Key aspects of the model were validated using data obtained from a series of qualitative, semi-structured interviews. The results of the simulations were analyzed using the statistical methodology of panel data analysis, which was deemed appropriate, given the longitudinal nature of the data. The results indicate that those contingency factors that were found to have a positive effect on *average consensus*, do so by increasing the number of interactions among members of a workgroup, thereby increasing the likelihood of sharing information about each others areas of knowledge. Those contingency factors, which were found to have a negative effect on *average consensus*, do so by reducing the likelihood of knowledge-sharing (and their associated non-task-related) interactions. The contingency factors, which have a significant positive effect on *average accuracy of knowledge*, do so by increasing the frequency and likelihood of the workgroup members receiving knowledge that is accurate from their colleagues. The contingency factors that have a positive significant effect on *percentage of project completed*, are those characteristics of the workgroup that enable the members of the workgroup in locating sources of knowledge within a limited time period, encourage them to share the requested knowledge and information about knowledge, and provide them with knowledge in several areas that is required by them to complete their assigned tasks. Those contingency factors that affect *percentage of project completed* negatively do so by decreasing the likelihood of workgroup members obtaining the needed knowledge over time.

Of the contingency factors that have a significant effect on one or more of the outcome variables, some are more-controllable, while others are not. The implications of what this means, to practitioners, are discussed (in section 5.3) so that practitioners can focus on the key subset of the contingency factors that would help them maximize their investments in human and technological resources.

By addressing the research question, this dissertation has made a few contributions. These are as follows:

1. The dissertation contributes to the literature by identifying the key contingency factors, from a group of contingency factors, which have significant effects on the outcome variables over time, via a fixed set of workgroup processes.
2. It acts as a guide to future studies by narrowing the set of contingency factors that need to be studied.
3. It provides a set of recommendations to practitioners whereby they can focus on those contingency factors that can help them maximize their investments in human and technological capital. By ensuring that their practices and policies increase the effects of the contingency factors that have positive effects on the outcomes and mitigate the effects of the contingency factors that have negative effects, practitioners can improve the performance outcomes of the workgroups to which they belong and/or supervise.
4. The agent-based model that was used in this dissertation, and the source code through which it was implemented, provide a foundation for more sophisticated models that could relax some of the assumptions made in this study, while adding more processes that would capture a greater amount of richness seen in real-life workgroups.

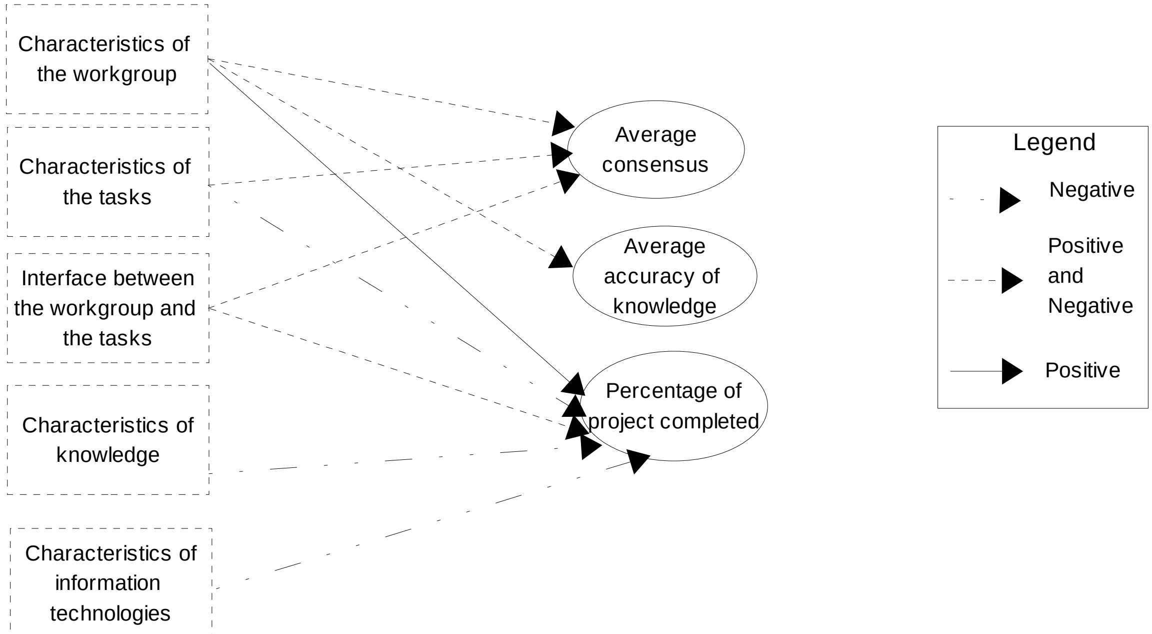


Figure 5.4. An overview of the findings

References

Ackoff, R.L., *Creating the Corporate Future: Plan Or be Planned for*, New York: John Wiley and Sons, 1981.

Akgun, A.E., J. Byrne, H. Keskin , G. S. Lynn and S. Z. Imamoglu, "Knowledge Networks in New Product Development Projects: A Transactive Memory Perspective," *Information and Management*, 42 (2005), 1105-1120.

Akgun, A.E., J. Byrne, H. Keskin and G. S. Lynn, "Transactive Memory System in New Product Development Teams," *IEEE Transactions on Engineering Management*, 53 (2006), 95-111.

Anand, V., C. C. Manz, and W.H. Glick, "An Organizational Memory Approach to Information Management," *Academy of Management Review*, 23 (1998), 796-809.

Austin, J.R., "Transactive Memory in Organizational Groups: The Effects of Content, Consensus, Specialization, and Accuracy on Group Performance," *Journal of Applied Psychology*, 88 (2003), 866-878.

Baum, C.F. *An Introduction to Modern Econometrics Using Stata*, (2007) College Station, TX: Stata Press.

Becerra-Fernandez, I. "Facilitating the Online Search of Experts at NASA using Expert Seeker People-Finder," *Proc. of the Third Int. Conf. on Practical Aspects of Knowledge Management, Basel, Switzerland*, 30-31 Oct. 2000, (U. Reimer, ed.), 3-1 – 3-7.

Becerra-Fernandez, I., A. Gonzalez, and R. Sabherwal, *Knowledge Management: Challenges, Solutions, and Technologies*, Upper Saddle River, NJ: Prentice Hall, 2004.

Brandon, D.P., and A.B. Hollingshead, "Transactive Memory Systems in Organizations: Matching Tasks, Expertise, and People," *Organization Science*, 15 (2004), 633-644.

Bureau of Labor Statistics "Employee tenure in 2008," downloaded from http://www.bls.gov/news.release/pdf/tenure.pdf, last accessed on July 25, 2010.

Butts, C.T., "sna: Tools for Social Network Analysis," (2010a), R package version 2.1.

Butts, C.T., *Software Manual for the R sna Package*, (2010b) R package version 2.1.

Campion, M.A., G. J. Medsker and A.C. Higgs, "Relations between work group characteristics and effectiveness: Implications for designing effective work groups," *Personnel Psychology*, 46 (1993), 823-851.

Carley, K. "Organizational Learning and Personnel Turnover," *Organization Science*, 3 (1992) 20-46.

Carley, K.M., "Computational Organizational Science and Organizational Engineering," *Simulation Modelling Practice and Theory*, 10 (2002), 253-269.

Cummins, J.N., "Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization," *Management Science*, 50 (2004), 352-364.

Dalton, D.R., W.D.Todor, M.J. Spendolini, G.J. Fielding, J. Gordon and L.W. Porter "Organization Structure and Performance: A Critical Review," *The Academy of Management Review*, 5 (1980), 49-64.

Daft, R.L., and R.H. Lengel, "Organizational Information Requirements, Media Richness and Structural Design," *Management Science*, 32 (1986), 554 – 571.

Davenport, T. H., and L. Prusak, *Working Knowledge: How Organizations Manage What They Know*, (1998), Cambridge, MA: Harvard Business School Press.

Davis, J.P., K.M. Eisenhardt, and C.B. Bingham, "Developing Theory through Simulation

Methods," *Academy of Management Review*, 32 (2007) 480-499.
Epstein, J.M. And R. Axtel, *Growing Artificial Societies: Social Sciences from the Bottom Up*. Cambridge, MA: MIT Press, 1997
Epstein, J.M., *Generative Social Science: Studies in Agent-Based Computational Modeling*, Princeton NJ: Princeton University Press, 2007.
Espinosa, J.E., and C. Pickering, "The Effect of Time Separation on Coordination Processes and Outcomes: A Case Study," *Proceedings of the 39th Hawaii International Conference on System Sciences – 2006*, (2006), 1-10.
Espinosa, J.E., S. A. Slaughter, R.E. Kraut and J.D. Herbsleb, "Familiarity, Complexity and Team Performance in Geographically Distributed Software Development," *Organization Science*, 18 (2007), 613-630.
Ewusi-Mensah, K and Z.H. Przasnyski, "On Information Systems Project Abandonment: An Exploratory Study of Organizational Practices," *MIS Quarterly*, 15 (1991), 67-86.
Faraj, S., and L. Sproull, "Coordinating Expertise in Software Development Teams," *Management Science*, 46 (2000), 1554-1568.
Fraidin, S.N., "When is One Head Better than Two? Interdependent Information in Group Decision Making," *Organizational Behavior and Human Decision Processes*, 93 (2004), 102-113.
Frees, E.W. *Longitudinal and Panel Data: Analysis and Applications in Social Sciences* (2004). Cambridge, UK: Cambridge University Press.
Glaser, B., and A.L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, (2009). Piscataway, NJ: Aldine Transaction.
Granovetter, M., "The Strength of Weak Ties: A Network Theory Revisited," *Sociology Theory*, 1 (1983), 201-233.
Grant, R.M., "Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration," *Organizational Science*, 7 (1996a), 375-387.
Grant, R.M., "Toward a Knowledge-Based Theory of the Firm," *Strategic Management Journal,* 17 (1996b), 109-122.
Gray, P.H., "The Impact of Knowledge Repositories on Power and Control in Workplace," *Information Technology and People*, 13 (2001), 368-384.
Hollingshead, A.B., "Retrieval Processes in Transactive Memory Systems," *Journal of Personality and Social Psychology*, 74 (1998), 659-671.
Hollingshead, A.B., "Cognitive Interdependence and Convergent Expectations in Transactive Memory," *Journal of Personality and Social Psychology*, 61 (2001), 1080-1089.
Hollingshead, A.B., "Perceptions of Expertise and Transactive Memory," *Group Processes and Intergroup Relations*, 3(2000), 257-267.
Hollingshead, A.B., and D.P. Brandon, "Potential Benefits of Communication in Transactive Memory Systems," *Human Communication Research*, 29 (2003), 607-615.
Hollingshead, A.B., and S.N. Fraidin, "Gender Stereotypes and Assumptions about Expertise in Transactive Memory," *Journal of Experimental Social Psychology*, 39 (2003), 355-363.
Jarvenpaa, S.L., and A. Majchrzak, "Protecting National Security: Role of Transactive Memories in Ego-Centered Knowledge Networks," *Organization Science*, 19 (2008), 260-276.
Jensen, M.C., and W.H. Meckling, "Specific and General Knowledge, and

Organizational Structure," In *Myers, P.S. (Ed.), Knowledge Management and Organizational Design*, Newton, MA: Butterworth-Heinemann (1996), 17-38.

Kanawattanachai, P., and Y. Yoo, "The Impact of Knowledge Coordination on Virtual Team Performance Over Time," *MIS Quarterly*, 31 (2007), 783-808.

Kane, G.C., and M. Alavi, "Information Technology and Organizational Learning: An Investigation of Exploration and Exploitation Processes," *Organization Science*, 18 (2007), 796-812.

Kankanhalli, A., B. C. Y. Tan, and K.K Wei, "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Quarterly*, 29 (2005), 113-143.

Kogut, B., and U. Zander, "Knowledge of the Firm, Combinative Capabilities and the Replication of Technology," *Organization Science*, 3 (1992), 383-397.

Kogut, B., and U. Zander, "Knowledge of the firm and the evolutionary theory of the multinational corporation," *Journal of International Business Studies,* 34 (2003), 516–529.

Kotlarsky, J., B. van den Hoof and M. Huysman, "Bridging Knowledge Boundaries in Cross-Functional Groups: The Role of a Transactive Memory System," *Proceedings of the Thirtieth International Conference on Information Systems*, Phoenix (2009), 1-16.

Kotlarsky, J., and I. Oshri, "Social Ties, Knowledge Sharing and Successful Collaboration in Globally Distributed System Development Projects," *European Journal of Information Systems*, 14 (2005), 37-48.

Kotlarsky, J., and I. Oshri, "Social Ties, Knowledge Sharing and Successful Collaboration in Globally Distributed System Development Projects," In *Kotlarsky, J., I. Oshri and P.C. Van Fenema (Eds.)*, 1-23, London: Palgrave, 2008.

Larson C.E and F.M.J LaFasto, *Teamwork: What must go right/what can go wrong*, (1989) Newbury Park, NJ: Sage.

Lesca, N. and M.L Caron-Fasan, "Strategic Scanning Project Failure and Abandonment Factors: Lessons Learned," *European Journal of Information Systems*, 17 (2008) 371-386.

Lewis, K. "Knowledge and Performance in Knowledge-Worker Teams: A Longitudinal Study of Transactive Memory Systems," *Management Science*, 50 (2004), 1519-1533.

Lewis, K., D. Lange and L. Gyllis, "Transactive Memory Systems, Learning and Learning Transfer," *Organization Science*, 16 (2005), 581-598.

Lewis, K., M. Belliveau, B. Herndon and J. Keller, "Group Cognition, Membership Change, and Performance: Investigating the Benefits and Detriments of Collective Knowledge," *Organizational Behavior and Human Decision Processes*, 103 (2007), 159-178.

Liang, D.W., R. Moreland, and L. Argote, "Group Versus Individual Training and Group Performance: The Mediating Role of Transactive Memory," *Personality and Social Psychology Bulletin*, 21 (1995), 384-393.

Littlepage, G.E., Hollingshead, A.B., L.R. Drake and A.M. Littlepage, "Transactive Memory and Performance in Work Groups: Specificity, Communication, Ability Differences, and Work Allocation," *Group Dynamics: Theory, Research, and Practice*, 12 (2008), 223–241.

Majchrzak, A., S. L. Jarvenpaa and A.B. Hollingshead, "Coordinating Expertise Among

Emergent Groups Responding to Disasters," *Organization Science*, 18 (2007), 147-161.

March, J.G., "Exploration and Exploitation in Organizational Learning," *Organization Science*, 2 (1991), 71-87.

Moreland, R.L. And L. Myaskovsky, "Exploring the Performance Benefits of Group Training: Transactive Memory or Improved Communication," *Organizational Behavior and Human Decision Processes*, 82 (2000), 117-133.

Ngwenyama, O.K., and A.S. Lee, "Communication Richness in Electronic Mail: Critical Social Theory and the Contextuality of Meaning," *MIS Quarterly*, 21 (1997) 145 – 167.

Nonaka, I., "A Dynamic Theory of Organizational Knowledge Creation", *Organization Science*, 5 (1994), 14-37.

Nahapiet, J., and S. Ghoshal, "Social Capital, Intellectual Capital and the Organizational Advantage," *Academy of Management Review,* 23 (1998), 242-266.

Pan, G., and S.L. Pan, "Examining the coalition dynamics affecting IS project abandonment decision-making," Decision Support Systems, 42 (2006), 639-655.

Pavitt, C. "Do Interacting Groups Perform Better Than Aggregates of Individuals?" *Human Communication Research*, 29 (2003), 592-599.

Polanyi, M., *The Tacit Dimension*, London: Routledge and Keoan, 1966.

Ren, Y., K.M. Carley, and L. Argote, "When Is It Beneficial to Know What Others Know?" *Management Science*, 52 (2006), 671-682.

Rentsch, J. R. and R. J. Klimoski, "Why do Great Minds Think Alike?: Antecedents of Team Member Schema Agreement," *Journal of Organizational Behavior*, (2001), 107-120.

Robey, D., *Designing Organizations*, Boston, MA: Irwin, 1991.

Sabherwal, R., and I. Becerra-Fernandez, "Integrating specific knowledge: insights from the Kennedy Space Center," *IEEE Transactions on Engineering Management*, 52 (2005), 301-315.

Schaffer, M.E. and S. Stillman, "XTOVERID: Stata module to calculate tests of over-identifying restrictions after xtreg, xtivreg, xtivreg2, xthtaylor," Statistical Software Components, (2006), S456779, Boston College Department of Economics, revised 07 Mar 2010.

Shen, M. Gallivan and X. Tang, "The Influence of Subgroup Dynamics on Knowledge Coordination in Distributed Teams: A Transactive Memory System and Group Faultline Perspective," *Proceedings of the Twenty Ninth International Conference on Information Systems*, (2008), Paris.

Simon, H.A., "Bounded Rationality and Organizational Learning," *Organization Science*, 2 (1991), 125-134.

Singley, M., and J. Anderson, *The Transfer of Cognitive Skill*, Cambridge, MA: Harvard Press, 1989.

Smith, R.G., and R. Davis, "Frameworks for Cooperation in Distributed Problem Solving," *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-11 (1981), 61-70.

Stata, *Cross-Sectional Time Series,* College Station, Texas: Stata Press, 2003.

Strober, M.H. "Human Capital Theory: Implications for HR managers," Industrial Relations, 29 (1990), 214-239.

Steiner, I.D., Group Process and Productivity, New York, NY: Academic Press, 1972.

Szulanski, G., “Exploring internal stickiness: Impediments to the transfer of best practice within the firm,” Strategic Management Journal, 17(1996), 27-43.

Szulanski, G., “Exploring internal stickiness: Impediments to the transfer of best practice within the firm,” *Strategic Management Journal*, 17(1996), 27-43.

Tesfatsion, L. “Verification and Empirical Validation of Agent-Based Computational Models”, available at http://www2.econ.iastate.edu/tesfatsi/empvalid.htm (retrieved on September 14, 2010).

Tsoukas, H., “The firm as a Distributed Knowledge System: A Constructionist Approach,” *Strategic Management Journal*, 17 (1996), 11-25.

Urban, J.M., C.A. Bowers, S.D. Monday, B.B. Morgan Jr., "Workload, Team Structure, and Communication in Team Performance," *Human Factors* 38(1996), 300-310.

von Hippel, E., “'Sticky Information' and the Locus of Problem Solving: Implications for Innovation,” *Management Science*, 40 (1994), 429-439.

Waller, M.J. J.M. Conte, C.B. Gibson and M.A. Carpenter, "The Effect of Individual Perceptions of Deadlines on Team Performance," *The Academy of Management Review*, 26 (2001), 586-600.

Wasko, M. M., and S. Faraj, “Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice,” *MIS Quarterly*, 29 (2005), 35-57.

Wegner, D. M., “Transactive Memory: A Contemporary Analysis of the Group Mind” *In B. Mullen & G. R. Goethals (Eds.), Theories of group behavior*, 185–205). New York: Springer-Verlag, 1987.

Wegner, D. M., “A Computer Network Model of Human Transactive Memory,” *Social Cognition*, 13 (1995), 319-339.

Wegner, D., R. Erber, and P. Raymond, “Transactive Memory in Close Relationships,” *Journal of Personality and Social Psychology*, 61 (1991), 923-929.

Wegner, D., T. Giuliano and P. T. Hertel, “Cognitive Interdependence in Close Relationships,” *In W.J. Ickes (Ed.) Compatible and Incompatible Relationships*, 253-276, New York: Springer-Verlag, 1995.

Wittenbaum, G. M., S. I. Vaughan, G. Stasser, “Coordination task performing groups,” in R. S. Tindale, L. Heath, J. Edwards E. J. Posavac, F. B. Bryant, Y. Suarez-Balcazar, E. Henderson King, and J. Myers, eds. *Theory and Research on Small Groups Social Psychological Applications to Social Issues*, 1998 (4), New York: Plenum Press.

Yuan, Y.C., J. Fulk, and P.R. Monge, “Access to Information in Connective and Communal Transactive Memory Systems,” *Communication Research*, 34 (2007), 131-155..

Zhang, Z-X., P.S. Hempel, Y-L. Han, and D. Tjosvold, “Transactive Memory System Links Work Team Characteristics and Performance,” *Journal of Applied Psychology*, 92 (2007), 1722-1730.

Appendix A1. Summary of prior literature

Table A1.1 Summary of prior findings related to transactive memory and teams

Study	***Findings***	***Methodology***	***Comments***
Akgun et al., (2005).	Data drawn from 27 Turkish companies relating to 69 new product development projects were used in the study. Team stability, team member familiarity, and interpersonal trust had a positive impact on the TMS and also had a positive influence on team learning, speed-to-market, and new product success. The impact of the TMS on team learning, speed-to-market, and new product success was higher when there was a higher task complexity.	Survey-based cross-sectional data analyzed via multiple regression models (of various paths) and confirmatory factor analyses (of scales).	Cross-sectional data: longitudinal nature of the relationships could not be studied. Path analyses of the predicted relationships were not conducted via structural equation models.
Akgun et al., (2006)	The study tests the effects of TMS on new product development outcomes including mediating and moderating factors, i.e., the collective mind and environmental turbulence, respectively. Data were collected from 79 Turkish new product development teams. The study found that: 1) the TMS has a positive impact on team learning and speed-to-market; 2) the collective mind (i.e., team members' attention to interrelating actions) mediates relations between the TMS, team learning, and speed-to-market; and 3) team learning and speed-to-market mediates relations between the TMS and new product success. The study also found that 1) the impact of the TMS on speed-to-market is negative when market and technology turbulence associated with the environment is high and 2) team learning changes quadratically with respect to the	Survey-based cross-sectional data analyzed via multiple regression models (of various paths) and confirmatory factor analyses (of scales).	Cross-sectional data: longitudinal nature of the relationships could not be studied. Path analyses of the predicted relationships were not conducted via structural equation models.

Study	Findings	Methodology	Comments
	market and technology turbulence.		
Austin (2003)	The relationship between transactive memory and group performance in mature, continuous groups was examined. A group's transactive memory was found to be positively related to group goal performance, external group evaluations, and internal group evaluations. This relationship was true both for task as well as external relationships-based transactive memory.	Multiple surveys of 27 groups (a total of 263 respondents, with between 8 and 11 members per each group) in a large sporting goods and apparel company. Surveys were conducted over a period of 5 months. Posited relationships were tested using multiple regressions.	The measurement of group's TMS as a combination of knowledge stock, knowledge specialization, transactive memory consensus, and transactive memory accuracy, is unique and informative. Also, the study of continuing groups using multiple (longitudinal) surveys provides better insights into the TMS processes over time. The use of multiple measures to operationalize the outcome (group performance in terms of internal and external evaluations and an objective, proportion-of-goals-attained measure lend more credibility to the measures). Transactive memory included both members' knowing who knows what and group members' knowing who knows whom to contact for knowledge regarding something that no one in the groups knows.
Espinosa et al., (2007)	The interaction between task and team familiarity on the one hand with task and team coordination complexity on the other hand was studied in terms of the effect of the interaction on team performance. Results show that the beneficial effects of task familiarity decline when tasks are more structurally complex and are independent of task size. The benefit of team familiarity for team performance is enhanced when team coordination is more challenging, i.e., when teams are larger or geographically dispersed. Task and team familiarity were found to be more substitutive than complementary in their joint ef-	Hypotheses were tested on archival data of software projects performed at a large telecommunications firm.	Provides an interesting model that includes both knowledge about what others know, and knowledge of the task. To make my study more inclusive, I need to capture task and team familiarity along with task and team coordination complexities in my simulation model.

Study	Findings	Methodology	Comments
	fects on team performance. Task familiarity was found to improve team performance more strongly when team familiarity is weak and vice-versa.		
Faraj and Sproull (2000)	Studied a 69 software development teams and found that the ability of a team to coordinate expertise plays a significant role in predicting the team's performance outcome and that this effect remains significant even after accounting for the contributions made to the team's performance by team input characteristics, presence of expertise and administrative coordination.	Sample consisted of software development teams drawn from the application development division (with > 100,000 people) of a high-technology firm. 333 respondents from 69 teams spread across 13 geographic locations in the US provided survey responses that were analyzed.	This study adds to the literature by identifying some important attributes that complement a TMS in its effect of team performance.
Fraidin (2004)	This study found that when a pair of unfamiliar individuals had to work on a decision-making task, the effects of cognitive load were offset but the formation of transactive memory, when the individuals considered sharing information with each other. The effect of transactive memory on decision quality was moderated by whether related pieces of information were all given to one individual or were distributed between the two individuals: dyad decisions were more likely to be accurate when each pair of interdependent items was allocated to a single member (a ‘‘connected’’ hidden profile) than when each pair of interdependent items was separated between the two members (a ‘‘disconnected’’ hidden profile).	The participants were 413 undergraduates (198 males, 215 females) taking introductory psychology classes. Students participated in the study in partial fulfillment of a course requirement. A 3 X 2 X2 design was used. The first factor manipulated the distribution of information to dyad members (all shared, connected hidden profile, or disconnected hidden profile). The second factor manipulated cognitive load (low or high). The third factor varied whether pre-discussion importance ratings were made (ratings or no ratings). Dyad members in the hidden profile conditions were assigned specializations to create TMSs. The tasks were the solving of a murder mystery and the hiring of a job candidate.	
Hollingshead (1998)	The study consists of two experiments that examined retrieval processes in TMSs. The results of the	A 2 X 2 X 3 factorial design was employed, manipulating the relationship of the dyad (strangers vs.	Only one aspect of the transactive memory phenomenon was studied; the setting cannot be generalized to individuals

Study	Findings	Methodology	Comments
	two experiments indicate that both nonverbal and paralinguistic communication play an important role in the retrieval of knowledge in TMSs.	intimate couples) and the mode of communication (computer-mediated vs. face-to-face), with the administration of the knowledge test as a three-level repeated measures factor. Videotapes were made of the face-to-face interactions, and transcripts were recorded of the interactions occurring over the computer network. Participants received $10 for their participation in the experiments. They were mostly junior and senior undergraduate students, with an average age of 21 years and who have been in a relationship for at least 6 months.	working in a group in a realistic business setting.
Hollingshead (2001a)	The study found that the nature of transactive memory that develops in a dyad depends on the incentives that the partners have: incentives that favor common knowledge tend to produce integrated transactive memories whereas incentives that favor unique knowledge tend to produce differentiated transactive memories. The amount of transactive memory formed was moderated by prior experiences or knowledge that a partner has about the other's knowledge at the beginning of a cognitive task.	The design was a 2 X 4 factorial that controlled expectations about the partner's knowledge (similar or different from the participant's) and cognitive interdependence, the degree to which participants' outcomes depended on whether they recalled the same or different information as their partner (defined by 4 incentives). Students in an introductory psychology course (116, with 12-17 participants per each combination of factors.	The study provides a spectrum-type measure of the quality of transactive memory, with the two ends being differentiated and integrated. Also, prior experience with and/or knowledge of the partner's expertise affects the “quantity” of transactive memory formed and performance on the task. These should provide valuable inputs in programming the agents' cognitive make-up in the simulation.
Hollingshead (2001b)	The study found that a) people learn and recall more information in their own areas of expertise when their partner has different rather than similar work-related expertise; and (b) this effect reverses for recall of information outside work-related expertise. The experiment was conducted in a work setting.	44 office staff members drawn with little prior interaction, drawn from 24 departments in an academic setting.	The findings from this and the previous study (Hollingshead, 2001a) together indicate the 'mental shortcuts' that people take in deciding how much and what new knowledge they are going to acquire, based on their perceptions of others' knowledge levels.

Study	Findings	Methodology	Comments
	The author states that information about one's partner's role leads individuals to determine how much common and differentiated (specialized) knowledge they are going to learn by themselves and how accurately they meet their task requirements that rely on an individual's knowledge.		
Hollingshead and Brandon (2003)	Review prior literature and critique another review article (Pavitt, 2003) to show that while not all communication can lead to the formation of 'better' transactive memory, those communications that provide accurate information about other members' knowledge lead to better transactive memory (better in terms of its positive effect on the outcomes of tasks that members of a group were required to perform in the experiments that were part of the prior studies that were reviewed).	Review and critique of prior articles.	Not all communication leads to the formation of 'accurate' transactive memory.
Hollingshead and Fraidin (2003)	The results of this study showed that both male and female participants shared similar gender stereotypes across knowledge domains. Participants with opposite-sex partners were more likely to assign categories based on gender stereotypes than were participants with same-sex partners. Participants with opposite-sex partners learned more information in categories consistent with those stereotypes. These findings suggest that transactive memory systems may perpetuate gender stereotypes.	A 2X2 factorial design-based lab experiment, where the subjects were undergraduate students.	In the absence of information about one's partners' areas of knowledge, members of a group tend to use their prior assumptions/biases/knowledge (beliefs) to guess their partners' areas of knowledge – the association relating a person of a particular gender with specific areas of knowledge that is strongest gets triggered, leading members to prejudicially assign areas of knowledge to their partners (and assuming ignorance in certain other areas, again based on a prejudicial reaction).
Jarvenpaa and Majchrzak (2008).	The study reports that in the case of an inter-organizational network, where the participants have mixed motives, the perceived	Survey of 104 professionals working in the area of IT security for private firms and government agencies. Hypothesized relationships	The notion of incongruence of goals and motives of agents participating in a network where a transactive memory develops over time was not considered in

Study	Findings	Methodology	Comments
	levels of formation of TMSs depends on the participants' trust in the credibility of other sources. Their contribution to the common pool of knowledge depends on the amount and type of information they are permitted to divulge themselves (based on their organizational policies) and their organization's approach towards inter-organizational learning.	tested via structural equation modeling	prior literature, given that prior literature focused primarily on romantic relationships and work groups where all members have the same goal. The study points out key factors such as benevolence-based trust, organizational mandate towards sharing knowledge with members of other organizations that can affect the formation of a TMS in a inter-organizational network.
Kanawattanachai and Yoo (2007).	In virtual teams, the volume of communication during the initial phase of a team project is positively related to the formation of transactive memory. During the mature phases of the project, transactive memory, particularly as it relates to task-expertise coordination among team members, has a positive effect on the team's performance	38 online-MBA student teams were studied (using email logs and performance on project tasks) over a period of 8 weeks. All team members communicated in virtual space, with no face-to-face contact.	The role of communication in determining information about who knows what, and who can do what, is shown.
Kotlarsky and Oshri (2005)	In addition to rapport and trust, collective knowledge and transactive memory are key determinants of successful collaboration in globally-distributed software development projects.	Interviews of globally-distributed teams from two organizations, SAP and LeCroy (the companies had operations in several countries and the team-members interviewed were geographically and temporally distributed).	
Kotlarsky et al. (2008)	Found that the strength of a group's transactive memory, whose creation is aided by technologies such as codified and personalized directories, lexicons, etc., aids in the reduction of gaps that exist in the use of a shared jargon (termed *syntactic knowledge boundary*), in a commonly-shared deep understanding of the terminology applied (termed *semantic knowledge boundary* and commonly-understood and shared goals, practices and interests (termed *prag-*	Survey within a large Dutch governmental organization of employees drawn from several departments whose employees had to work together on department-spanning projects. The departments were located in different geographic regions. 150 respondents (41%) completed the survey.	Provides support to the inclusion of geographic location-based diversity of workgroup members in the simulation. The key contribution of this study is in delineating the effect of usage of technology, such as directories and lexicons that provide the workgroup members a common language to use while communicating with their workgroup members, on the development of a TMS in the group.

Study	Findings	Methodology	Comments
	matic knowledge boundary).		
Lewis (2004)	The study found that teams with initially distributed expertise and familiar members are more likely to develop a TMS. Frequent face-to-face communication led to TMS emergence. However, communication via other means had no effect. Teams with more established TMSs later benefited from face-to-face communication, but they were less helped by frequent communication via other means, suggesting that transactive retrieval processes may have been triggered during face-to-face communication and suppressed during other types of communication. TMSs were positively related to team viability and team performance, suggesting that developing a TMS is critical to the effectiveness of knowledge-worker teams.	Student/consultant teams were required to work on a project proposal during the course of 13 weeks. Three surveys (first: demographics, the rest: details of the team processes) were conducted to obtain data from 35 teams and 36 teams (two datasets were generated, one per semester). There were no common participants between the two surveys.	The nature of communication was different in different phases of the project. During the initial phase, the communications included exchanges for obtaining information regarding each others skillets, knowledge, etc. During the latter phases of the project, the communications were mainly related to refinement of members' perceptions of others' expertise and allocation and re-allocation of knowledge-specific tasks to members who were most qualified to fulfill them, indicating the formation and refinement of task-expertise-person triadic links in the minds of the team members (the author does not explicitly use the task-expertise-performance construct). The author points towards some interesting research questions: "Some questions to be addressed by future research include the following: (1) At what points in a team's development should members focus on TMS emergence? (2) At what point is a TMS mature enough to facilitate knowledge retrieval and integration, and how is this maturity manifested? (3) Does the efficacy of a TMS diminish over time? If so, what are the markers and outcomes of this decline? Answers to these questions would help researchers define, in a better fashion, how the structure of a team and its tasks influences TMSs and their impact on performance."(pg. 1530)
Lewis et al., (2005)	The study found that groups with a prior TMS and experience with two tasks in the same domain were more likely to develop an abstract understanding of the principles relevant to the task domain. There was no support for the au-	Longitudinal experiment, comprising of training in an assembly task (week 1), performance of the task (week 2), performance of a different assembly task (week 3), a knowledge task (week 4). Subjects were undergraduate stu-	The findings are limited to "short term", laboratory settings. However, they provide insights into learning and transfer of knowledge, related both to tasks as well as others' expertise, again indicative of the possible applicability of the TEP unit as a measure of TMS as it

Study	Findings	Methodology	Comments
	thors' contention that that a TMS facilitates learning transfer after experience with only a single task. The extent to which members maintained expertise across tasks influenced the degree of learning transfer, especially for groups whose members had previously developed a TMS with another group. The study demonstrates that TMS influences group learning and knowledge transfer.	dents (300 participants; 100 groups of 3 each) who participated for extra credit. To test for expertise stability some groups retained membership throughout the 4-week period, while other groups had some members swapped with members from other groups.	evolves over time. However, the authors have not measure TMS in terms of TEP units.
Lesca and Caron-Fasan (2008)	In their exploratory study of information systems development projects at multiple sites, the authors have found that the following factors are primary contributors to the abandonment of projects: (1) stakeholders' qualifications and experience; (2) the management and organization of the project system; (3) strategic alignment and changes in the organization's internal structure; (4) poor project impetus; (5) uninvolved management; (6) unqualified people (6) inaccurate expectations, project mismanagement, strategy misalignment; (7) poor participation; (8) hostile culture; (9) insufficient budget; (10) conflating technical and managerial problems; (10) previous project trauma, and (11(underestimated complexity.	The authors acted as consultants at the time of collecting data via the action-based research methodology. The set client companies with whom the authors worked was geographically diverse. The authors participated in 39 projects, drawn from 32 companies.	Provide a basis for some of the "subtask abandonment rules" - specifically, the factors that the authors have identified: complexity of the project and lack of competent workers can be operationalized to determine the maximum amount of time to be spent on a subtask, before abandoning it. Abandoned subtasks lead to partially-completed projects.
Lewis et al., (2007)	The results of this study show that groups that experience partial membership change tend to rely on the TMS structure that old-timers developed in their original group, and that doing so is ultimately detrimental to performance because it creates inefficient	Laboratory experiment involving undergraduate business students.	Indicates the role that old-timers in a group can play in facilitating or inhibiting the accuracy of a group's TMS. What specific characteristics of old-timers facilitate/inhibit the accuracy of TMS? It depends on the fidelity of old-timers' knowledge with what is required to complete a task: if the old-

Study	Findings	Methodology	Comments
	TMS processes. Results from a supplemental study indicate that these TMS process inefficiencies can be avoided when old-timers are instructed to reflect upon their collective knowledge prior to task execution.		timers' knowledge is irrelevant or incorrect, then accuracy of TMS and consequently task performance deteriorates.
Liang et al., (1995)	This study found that groups consisting of members who trained together performed better on group tasks than those groups that consisted of members who trained individually, before joining a group. The authors indicate that TMS was a mediator between task and team characteristics and team performance on the other. The state that in groups where the members trained together, the TMS was higher, resulting in a better team performance.	Laboratory experiment where the subjects were undergraduate students.	Ostensibly, during collective training, a team's members can obtain information cues from other members of the team whereby they can ascertain their partners' knowledge and skills. These cues could be verbal and/or non-verbal, such as observing someone perform a specific sub-task. This study also highlights the importance of information cues in determining others' expertise which is crucial in the development of TMS.
Littlepage et al. (2008)	Found that the application of transactive memory-based allocation of work is most effective when the allocation is done based on specifics of a task, rather than the knowledge domain to which it belonged. Additionally, they also found that the overall performance of the group was better when its members differed in terms of their abilities associated with performing tasks in a given domain and that allocating a greater proportion of work to those that are more proficient tends to improve the overall performance of the group, as doing so reduce the amount of inter-personal communication. This was said to occur because greater communication requirements tend to lower the efficiency of the workgroup.	Thirty-six secretarial staff members at a large university in the southeastern United States(18 pairs) participated in an lab experiment; all were female, and the mean age was 47.83 (SD = 6.72).	Provides support to the decomposition of a task into subtasks, since the findings of the study suggest that the question of how agents seeking other agents' knowledge at for completing individual subtasks (analogous to the items used in the experiment employed in this study) are worth investigating.

Study	Findings	Methodology	Comments
Majchrzak et al., (2007)	The authors apply the theory of TMS to emergent response groups and explain that such an application requires extending the theory in three ways: 1)the role of expertise in task assignment, 2) how groups function when credibility in member expertise cannot be validated, and 3) how expertise is coordinated.	Examine publicly-available information pertaining to three catastrophic situations: Hurricane Katrina and Rita in the US, the tsunami in South-East Pacific and Indian oceans and the recent earthquake in Pakistan.	Suggest that in order for TMS to be applicable in a non-traditional situation such people dealing with the aftermath of an environmental disaster, the assumptions of TMS must be reexamined and key constructs such as expertise, location of expertise, credibility of expertise, and task and expertise coordination must be redefined to take into account the expedient nature of work that emergence response teams perform.
Moreland and Myaskovksy (2000)	This study found that for TMS to develop in a group communication among members is not essential. The transmission of information about others' areas of expertise/skills to the members of a group was essential for the development of TMS. This could be (as shown in the study) instructions to group members at the beginning of a task informing them about the expertise of their group members.	126 undergraduate students who participated in a laboratory experiment, where the factors manipulated were the presence/absence of team communication and presence/absence of instructions to the team (used synonymously with group) members about other members' skills.	This study highlights the importance of information that group members need to have about other members' areas of expertise in order for transactive memory to develop in a group. While communication among team members might provide the information needed, it is not essential and is not sufficient (if it does not provide the right information).
Ren et al., (2006)	The results of this study show that TM decreases group response time by facilitating knowledge retrieval processes and improves decision quality by informing task coordination and evaluation. The results also suggest that the effects of TM are contingent upon group characteristics, such as group size and environment, as well as the dimension along which group performance is assessed. TM seems to be more beneficial to small groups using quality as the dependent variable, but more beneficial to large groups, groups in a dynamic task environment, and groups in a volatile knowledge environment using time as the depend-	Virtual (computer simulation-based) experiments using a special software suite ORGMEM developed at Carnegie Mellon University.	Identifies important team attributes that could be used in the proposed simulation.

Study	Findings	Methodology	Comments
	ent variable.		
Rentsch and Klimoski (2001)	The study found that demography, team experience, team member recruitment, and team size were significantly related to team member schema agreement, which was significantly related to team effectiveness.	The study was conducted in a naturalistic setting involving 315 individuals who comprised 41 teams	Of interest is support for the authors' hypothesis that *team size* negatively affects *team member schema agreement .*
Shen et al. (2008)	Describe how the formation of smaller subgroups due to differences in attributes of its members such as gender, ethnicity, age, etc., leads to fragmentation, which the authors describe as appearance of *faultlines*. The fragmentation of a workgroup was found to be negatively correlated with the formation of a TMS in the group and consequently to lower coordination of knowledge-sharing activities and lower group performance. The authors also found that lower TMS resulted in greater dissatisfaction of the group members with their work in the group. The cross-sectional nature of their study precluded an investigation of the specific processes that lead to the appearance of *faultlines* and how the appearance of *faultlines* reduces in group performance.		Provide support to the inclusion of distinguishing a feature in the agent: location. While other features such as gender, ethnicity have been reported in the study as contributing to the development of fault-lines, including these attributes and the associated mechanisms of how they affect inter-agent dynamics of knowledge-exchange is deemed beyond the scope of the simulation.
Yuan et al., (2007)	This research tested a transactive theory model of how individuals allocate and retrieve task-related information in work teams. It extended prior research by exploring the role of communal information repositories in the context of human information resources. It was found that the usage of information repositories was significantly related to individual access to information. However, the rela-	The proposed model was tested on data collected from 179 people in 15 project teams in organizations from a variety of industries.	In the article, the authors use *information* and *knowledge* interchangeably. Knowing who knows what can be obtained through direct interactions (as shown in previous studies) or through instructions provided about others' expertise (e.g., in some of the experimental studies reviewed earlier) or via the use of “expert locator” software. Also, those who have a higher IT competency are more likely to use a

Study	Findings	Methodology	Comments
	tionship between individual direct information exchange with team members (the human repositories) and individual access to information was significant only among average-level users of organizational information repositories. The development of individual expertise directories significantly influenced individual direct information exchange with team members. The perceived usage of organizational information repositories by team members significantly influenced actual usage. Finally, technology-specific competence in using intranets significantly influenced the actual usage of intranets as organizational information repositories.		technology-based solution for seeking out the specific experts in their team.
Zhang et al., (2007)	The chief goal of this study the establishment of the ecological validity and generality of TMS research findings. The study examined the relationships between team characteristics, TMS, and team performance. The results indicate that task interdependence, cooperative goal interdependence, and support for innovation are positively related to work teams' TMS and that TMS is related to team performance. TMS was found to mediate the team characteristics–performance links.	Data were collected through multiple respondent surveys in high-technology firms in China, belonging to the following industries: information technology, telecommunications, electronic engineering, biological engineering, and related fields. Teams consisting of 3–9 members who had been together for 3 months or more were surveyed.	The study lends to the credibility of the TMS construct in a different culture.

Appendix A2. Content analysis of the interviews

The purpose of the interviews is to obtain information regarding the types of IT used for communication among the members of a workgroup and to obtain support for some of the assumptions made about the behavior of agents. The following tables present a content analysis of the interviews, and identify the information provided regarding the use of IT and support for specific points in the proposed simulation[44]. In the column titled “Specification item for which support provided”, the items from the previous (as given in the dissertation proposal document) and the current version of simulation specifications are identified by explicitly by prefixing each item with “previous” and “current” respectively. Also, in the case of specifications from the current version, the first word from the name of the subsection, along with the specific number of a point, if a specification is given as a numbered point in a subsection, is used to identify the category of the specifications that are supported by the response to a question.

Table A2.1. Interview of M

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
a)	S: *I am here with M, doing the first interview. Thank you M for spending time with me.* M: Glad to help	Introduction		
b)	S: *To begin, I will start with question number 0. Could you please describe a recent project in which you were involved that involved team work.* M: We are still in it. It is a project to implement and integrate a new loan origination system, I work at YYY mortgage. We have a new loan origination system called CORE and I work in the capital markets section of YYY, which is the group that takes the originated mortgages and sells them in secondary markets to Freddie Mac and Fannie Mae and Federal Govt. through FHA. My involvement in this project from capital markets perspective is to integrate capital market systems oriented around secondary markets in with this new CORE origination system. We are currently in user-acceptance testing and it is a challenge to say the least. I work with a virtual team users and other technology people, both in capital markets, technology and technology integration group of CORE. TIG is the acronym we use at WF to describe technology integration IT functions. That's I think is changing to TOG Technology and Operations Group. Right now we call ourselves TIG	Description of the project and team composition		

44 To ensure privacy of the interviewees, they and their employers have not been identified; instead, their names and the names of their employers have been substituted with initials and acronyms, respectively.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
c)	S: *So when you say that you work as part of a virtual team, does that it mean that some of your team members are in a different location?* M: Pretty much everyone I work with is in a different location. My manager is in XXX, YY. We have team members, developers in XXX, QQ, we have a developer in XXX, ZZ, we have business users in RR FF., XXX, HH, FG, HH. We have project team members in the CORE side that I deal with in all of those cities + SF bay area and XXX, ZZ, and we also have developers in XXX, WWW. And there's one developer in AAA, besides myself, but pretty much everybody else is not local. So very virtual.	Description of the project and the composition of the workgroup		• The workgroup members can be geographically separated – this requires that each agent have “location” as an attribute. • The location of the agents would determine, though not solely, the choice of IT used by them for communication.
d)	S: *That brings up the second question: when you collaborate with your team members, what types of communication technology do you use?* M: We have, pretty much everybody has, the same set of technologies. First time we come into the firm they give you a wireless telephone headset. So basically you have a wireless boom mike and headset so you can pretty much be hands-free. Of course, everybody gets a PC. For somebody like myself, my role is a Business Systems Consultant – an analyst designer. That role typically is provided with a laptop computer, whereas hardcore developers are provided with a desktop computer for doing production support. The software suite is MS Office suite. Up until recently, everyone was AOL IM client, but now we are switching to MS Communicator Client. We extensively use MS Netmeeting Client. We extensively utilize a product called Centra that is actually an educational website hosting package that we use for meeting facilitation and meeting tracking. On Centra we have people who are meeting leaders, presenters and facilitators. We have centra to actually record voice calls and record the actual screen shots as the meeting progresses and then team members can go into centra system and recover those at any time. We are with MS Office Communicator and MS Office Live to	Description of the IT used by the workgroup for communications.	Subsection 3.2.1.4: 4	• The IT used should facilitate synchronous communication, e.g., Phone-calls, IM, group-chat with voice (and possibly with video), and asynchronous communication via email and document repositories e.g., MS Sharepoint.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	replace Centra. I don't know how that is going to play out. From a knowledge sharing perspective, we have MS Sharepoint. Most teams have an MS Sharepoint site or two for sharing documentation. Our Sharepoint site is limited to the technology side for the developer side. For the business side we also have a ...I am sure it's packaged software that we have implemented but I don't know what that is. There is another website that we use for publishing documents and meeting minutes and things like that for the business team. We are looking at Wikis. We have not had a great experience with Sharepoint. Sharepoint has some limitations and I don't know if they are Sharepoint limitations per se or if they are those that are exposed to us in terms of capabilities. We are looking at Wikis, of course we all use file sharers. We have a lot of software that is oriented around the idea of virtual teaming, sharing, capturing knowledge.			
e)	S: *Do you also use any software for synchronous document edition (should be editing)?* M: Not really. We have used things like … you know with Excel you can save a Workbook with shared mode? We have used that. The main documentation that we tend to share is … I call it metadata. Our function at Capital Markets is a very ETL function. So we have four origination systems, plus various other systems that we bring into loan data everyday and we are sending loan data in various levels of summarization out everyday. So what we have is quite a bit of metadata, keeping track of source to target mapping and data transformations and things like that. We tend to use Access for that. Pretty much everyone who has MS Office on their PC has Access, so we are lousy with MS Access databases. So we have a lot of databases that were written to provide these metadata mapping function, documentation function, and those can be shared. We have spaces on our shared drives somewhere so people can open and access the metadata. But as far as using a Word document or anything like that, we don't do a lot of collaboration in terms of multiple people editing at the same time, but we will see a lot though is over the course of a meeting somebody either through NetMeeting or Centra somebody who will have a document open will be driving the document and capturing changes and actually editing the document as other people are observing.	Description of the IT used by the work-group for communications.	Subsection 3.2.1.4: 4	
f)	S: So is it similar to keeping minutes? M: Well, it may be keeping minutes but it also could be, for instance, for this particular project in the next stage we are doing the systems requirements spe-	Obtain details of how the IT used for group-	Subsection 3.2.1.4: 4	While the use of “group chat” as a multi-way communication method is

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	cification for a new function and we actually have a .. several weeks of meetings to work out the system requirements, SRS is what we call it, and the facilitator was actually documented changes to the SRS, writing the document, and editing the document, as we were doing the meetings. So, through Netmeeting, we have a discussion, the analyst will capture some verbiage, and we'd actually edit that and..no, I didn't actually quite say that I, change it to say this, add a bullet...even though we aren't all collectively keying in at the same time, we are through the NetMeeting and Centra doing a joint...	chat is used.		described here, implementing it is considered to be out of scope of the dissertation.
g)	S: *So it's somewhat similar to a Whiteboard?* M: Similar to, yes, exactly. We have one person who is driving, and the others are modifying by giving the instructions... S: *But it's always one person who does the keying...* M: Yes S: *Others critique and..* M: Yes, exactly...	Obtain details of how the IT used for group-chat is used.	Subsection 3.2.1.4: 4	
h)	S: *So, as you know tacit knowledge is knowledge that is difficult to articulate, to verbalize, to explain to others. In your daily routine if you need to exchange tacit knowledge, either to obtain tacit knowledge or to provide tacit knowledge, do you have a preference for a particular type of technology?* M: I think it is situational, Srikanth. I think it depends on the purpose and the audience. Let me run your through a couple of examples. As I mentioned , we are in user-acceptance testing and my role is ...its complicated. Probably the best I need to do is to draw a quick picture. (begins drawing on a sheet of paper) If you imagine, I have an online system that I call CORE and CORE has a DB of its own..and CORE is replicating..down to another DB that we call ODS. The ODS has slices that we call Lock, Stage and Static. Data moves through these stages at different points in the processing cycle. And then out of this ODS DB we have feeds to a legacy origination system and then we have feeds out of that to downstream systems like Capital Markets, which is my function and then we also have some direct feeds that come this way. So I am Happy Harry sitting down here and the business user in Frederick calls me and says “you know this field isn't right. I expected a value of A in there and it has a value of C. And this is looking at a table in Capital Markets. So my role here is to determine where it broke.	The reasoning behind the choice of a specific medium of communication when the knowledge exchanged is “tacit”	Subsection 3.2.1.7: 5, 6	Stickiness of knowledge results in a difficulty in exchanging and transferring it. In the literature, tacit knowledge is described as being hard to articulate and verbalize and hence, is difficult to transfer. Hence, knowledge that has a high level of tacitness is also highly sticky . So, the response to this question provides a basis to determine how, given a need to obtain knowledge of a certain perceived difficulty, a work-group member decides on which communication medium needs to be used. The functional form would be: *probability of choos-*

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	S: *So you retrace your steps?* M: Exactly. Now the good news is that I have SQL access to here and I can write SQL queries against Static and I have SQL access to here. So I can see two different views. This is black box. I can see what data is loaded into it. So, I can see this but I cannot see that. And so then, the challenge is, well let's see I will use the field ABC, that's the name. So I will go to my metadata, I look at my metadata to see what the source of ABC is right, so I know what its source is, coming in here, I know what its source is supposed to be here and my metadata tells me how to find it here. So, very typically, the first thing I'll do is I'll go look here and I'll take the transformation rules that I find in the metadata and go find the data here and go see what it looks like. S: *OK* M: And if the data looks good there, then I know that the problem is downstream, I know that somewhere between here to here it broke. If the data isn't good here, then I'll do some more inspection to find out why and then its either going to be a problem in here or there may be a problem upstream. Alright? Now, I am working with a very sophisticated business user in Frederick who also queries this. And so we do lots of knowledge sharing when it comes to finding things here. Very often what will happen is he'll have problem or something that he doesn't think is right and he'll say "Can we talk about this?" We'll start a NetMeeting an informal meeting and we'll start looking at his query and result-sets and things like that and my queries and result-sets and start querying data and one of us will be driving, to find out what might be breaking and why things aren't working. So there's some tacit knowledge sharing because one of us is helping the other understand how to do outer joins and other things like that and it gets very informal. Another area where you may tacit knowledge sharing in the same vein is I got a community of users who are interested in the same piece of data and one of them has reported a defect and now I need to communicate to all of them what is broken and we are going to do about it and in that case the most likely method of knowledge sharing would be me writing some sort of a document with lots of examples showing why it broke, showing what I found, what I think is broke and what I am going to do to fix it and sending it via email. Lots of times you get into these things and you don't really know where it's broke, so, what you end up doing is scheduling what we call a topic specific meeting and it might involve people across this whole range		Subsection 3.2.1.7: 5, 6	*ing a medium for communication = f(stickiness of the knowledge to be exchanged, richness of the medium).* The above equation would be used in lieu of the probability values of choosing a medium that are set *a priori* . That is, each time an agent needs to communicate with another agent, it chooses a specific medium (face-to-face) or IT based on how rich the medium is (governed by the *rich_ftof, rich_tr, rich_email* set of parameters) and the tacitness of the knowledge to be exchanged (this information is obtained from the knowledge-stickiness vector.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	because we don't know where it is broke. So we'll get people in through all this and we'll start working through here's what I saw, what did you send and things like that. So it is very situational.			
i)	S: *But each time you need to collaborate or do things at the same time, you use NetMeeting...* M: Or Centra or something like that. It almost always evolves to that. So a Tim or (unclear)... the one I was talking about who is a really savvy user I was talking about.. typically what will happen is that I will receive an IM “You there?” “Yes” “Got a minute?” “Can we call into a NetMeeting?” So typically that's a reach out to call someone. If it is something simple we'll just use IM back and forth but if it is anything more difficult, then we'll end up with a NetMeeting and a phone call before it's all over. S: *So when you first walk into your office in the morning, do you sign on automatically?* M: Oh, absolutely. S: *Is it a company-given ID that you use?* M: Yes. And it's mostly single-sign-on. We use...they call it ADENT, I don't know if that's a YYY term or industry term, but we have an ADENT ID and when we sign on to that ID... that ID and password is generally the keys to everything. There are multiple places where you have to sign on, but that ID and password is generally that you use to sign on to everything. The other tool we use, we Rational Clearquest Defect Tracking System – DTS – that's another example of a place where you'd try to code tacit knowledge into.. you know try to codify tacit knowledge. You can imagine in this environment (pointing to the sheet of paper) I find something wrong here, it needs to be fixed here... we have challenges where users try to write to write their defects, in fact we want the users to write their defects rather than us trying to be in the middle and that will assign them to some developer over here and based on some rule given over here the developer will fix something and send it back. They'll look at it, they won't like it because they won't get the right results, so they'll send it back and it'll bounce back and forth about three or four times and then they'd come back to me and say, “I can't get this fixed, will you help me?” and so I'll go through that, look at all these points, figure out what is broke and I will very explicitly document in the defect record, what I found, why I think it is broke and what needs to be fixed. I'll update the defect that way and then I'll typically send email or contact the developer and say, or the data-modeler and give them more specific instructions or	Obtain information about what type of IT is used to facilitate synchronous communication	Subsection 3.2.1.7: 5, 6 Subsection 3.2.1.7: 5, 6 Subsection 3.2.1.7: 5, 6	This point indicates that some types of communication media facilitate synchronous communication while others do not. Therefore, the choice of a communication medium is a function of how well the medium facilitates synchronous communication. Hence the previous equation now becomes *probability of choosing a medium for communication = f(stickiness of the knowledge to be exchanged, richness of the medium, synchronous)*. In addition to the richness parameter, a communication medium should have an attribute called *synchronous* which is boolean: true indicates that it facilitates synchronous communication, false indicates that it does not facilitate synchronous communication. It should be noted that email can be used, technically, to simulate an IM session. However, in this interview there is no indication that email was used as a replacement for IM (Netmeeting, whose

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	feedback or...it's a multiple step process..very time-consuming, very painful sometimes but necessary. S: *Based on in which location you are in this graph, you choose the technology accordingly would it be just an IM or would it be something more involved based on the nature of the task..* M: Exactly. It's situational based on the nature of the communication need. My first preference would be IM because I can see [whether] somebody is online or not..so if I have an issue, first what I am going to do is IM the developer or the business user "are you there?" or my manager...pretty much anybody and it really more goes to email if you want to make sure that you have a documented record, you want to share the information with more people, if it's important for you to have the history of what's going on. S: *Do you also share logs of your IMs and NetMeeting conversations?* M: Not normally. But now with IM we weren't really logging with...now with the communicator, it does log....So every conversation you have with your Communicator is saved in your Outlook mailbox. So I don't know how we'll use that yet. S: *But, if need be, you could go and search the logs* M: Right. And I mentioned that we also use Centra. And when we use Centra, we use Centra specifically because we want to record the meeting and have people look back. So we have formal user meeting where formal UAT defect management meetings, staff meetings, we try always to run those on Centra so that who needs or wants to can go back to the meetings and listens to them.			logs could be recorded and imported into the Outlook email client, is considered a chat application). Hence, email would have a value *false* for the *synchronous* parameter.
j)	S: *When you first started with this project in which you are working, how did you come to come to know each team member's areas of expertise. Given a particular type of problem or situation, how did you know whom to contact?* M: I don't know … I never thought about it that way. I can tell you one of my pet peeves when it comes to technology and business is IT people absolutely need to understand the business. And so I've always made it part of my core being that I understood the business. When I started at YYY, other than having a mortgage of my own, I didn't know much about the mortgage business. So I really set about becoming an expert in the mortgage business. And even when I was successful in it. I learned a lot about secondary marketing for mortgages. Over that period of time, I	Obtain information about the process the interviewee used for developing an understanding of others' areas of knowledge and selecting his source of	Subsection 3.2.1.7: 3a; 3b	This question provides support for the use of an 'expert-seeker' type of IT. Additionally, acquaintances of a knowledge seeker can be a good source of the knowledge areas of other sources with whom the knowledge seeker does not have a direct relationship. So, the

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	had a lot of questions, I contacted a lot of people who may have the answers and learned over a period of time who has what knowledge. And so the mortgage business in the secondary markets business we've got people who actually sell me assets, we've got people who are delivering the assets, we've got people who work with the GSCs ...the govt... like Fannie and Freddie and FHA and they work with them on a contract.. You learn over time who those people are because of their expertise because people tell you through word of mouth...mostly it's word of mouth I think. Or we go to a colleague whom you know pretty well and ask – I have this problem, whom would you recommend...sure. When you first start on a project or one of your coworkers and say, I need to find more about this and they might say "You need to talk to Bobby, because Bobby knows everything." So you go over to Bobby, your introduce yourself to Bobby and ask some questions and ask Bobby if she has any documents that you could see or you ask Bobby for a demonstration of the system. Of course you make sure that you understand enough of the mortgage business upfront so you don't totally... S: *Say you have a common ground, so to speak* M: Right, right. And you know, if you have a mortgage, then you know something about it...same thing with the IT side...you find people in your team that have the specific skills you need the DBA, the data-modelers, the ETLs...mostly what YYY takes is they take a lead approach where for a project there will be a lead ETL developer and may be more multiple developers underneath that person and there would be a lead BSC – business systems consultant – and I am the lead BSC in this Capital Markets and the other guy, Leroy, who lives in Atlanta, is the lead developer and if I need something for...if I need a developer to look at a problem, the formal approach would be..Leroy might just look at a problem and he'd sign it and the informal approach that he and I have..is..because of the trust he and I have. The trust relationship here is I know what developer developed it, so I'll send it to that developer and copy Leroy ...if I don't know then I'll call Leroy, email Leroy or IM Leroy and say "I need someone to look at this" whom would you recommend...if I do know, I'll send it right to that developer same thing with the data modelers, same thing with the DBAs. I'll just reach out to specifically the person I need. You just learn more with time. S: *Is there also some technology that, for instance, that is like corporate white pages or something like*	knowledge.	Subsection 3.2.1.7: 3a; 3b Subsection 3.2.1.7: 3a; 3b; Subsection 3.2.1.10.	simulation must include, in addition to direct relationships that are formed during the training phase of the simulation and the use of 'expert-seeker', the idea of a network. That is friends introduce a knowledge seeker to their friends. Friends-of-friends can then form the potential set of sources of knowledge. The exchange of knowledge about others' areas of expertise can happen when the source approaches his friends directly, while trying to complete a task or during the non-task related exchange of knowledge. The above idea can be implemented as follows: • based on the interactions formed during the training phase, form a network • each agent, in addition to using the expert-seeker and their friends also use the friends-of-friends in

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	that where you have a listing of folks and their areas of expertise that you refer to? M: Yeah, there are...on Outlook there's something we call the GAL – the global address list – and so you can find people on there ...it's a little bit overwhelming though, and a little bit difficult. The reason I say it's overwhelming is...last I saw there were over 200,000 employees in all of YYY and its companies and so finding the person that you need there is hard. And the second problem is their titles and how they are listed in GAL isn't always useful. YYY has a lot of people with the title Risk Manager and so that's not always useful to understand what they do, but tied to the GAL and I don't know if this is just an Outlook function or something that YYY built in, they have organization charts and so if to find somebody...a typical problem by the way is we have a lot people with the same name. If you can find somebody, you can click on their org-chart and follow their organizational chain up to see if you can find out where they report; that might help you find the right person. Another thing they have is..they have a fairly expensive intranet called TeamWorks. And TeamWorks has quite a few tools on it. One of the things it has is a team-member look up and say you can key in a fragment of a name, first-name or a last-name and it will start showing you all the possible matches. It also has an A-Z look-up where you can get to a reference list of different sites intranet sites within the company that might be useful to you. So there are very good, lots of tools for doing, for providing assisted information through websites and things like that.			determining whom to contact for knowledge. • As described in point 48.j of the original simulation specification, knowledge not related to a task can also be exchanged. In the current simulation, when a source is contacted by a recipient, the source can indicate the presence/ absence of certain (randomly-determined) areas of knowledge of the source. This specification could be extended to include the possibility that, governed by a parameter, the source can provide information about the source's friends' areas of

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
				expertise to the recipient of knowledge. This probability could be governed by a parameter called *p_friends_ expertise,* whose range of values will have to be determined based on a survey of literature. However, using the criterion of simplicity, this will not be implemented in the current dissertation.
k)	S: *So based on what you told me, you use different types of methods, to get know each others' areas of expertise. Now once you know that, once you've been told that such and such person is an expert or someone who can answer your question in a certain area, how do you figure out if the person is in fact knowledgeable and can in fact help you with a particular question that you have?* M: that one's kind of tricky. A lot of the people I deal with, in fact primarily I deal with the business side and the business people I deal with know a lot about business. So even though their specific area is this little niche, they know the whole process and they are recognized for knowing the whole process. So, it's not like I am going to doubt them. And mostly they'll be very upfront if they are not sure. So here's what I think happens based on my past experience I think to answer your question, I think it happens this way, this way, but who you really should talk to is Sally over in closing group because Sally really knows this stuff and she can probably give you a better answer. So if you really need that answer, you'd go to Sally, if you are content with the level information that Bobby had for you, then you'd go ...but you learn who those people are over time and I guess part of it boils to	To determine the process by which a person identifies a source for the knowledge that he/she requires	Subsection 3.2.1.7: 3a Subsection 3.2.1.7: 3a	Indicates that notion of perceived trust that a source would share the knowledge that is requested by the recipient who contacted the source based on *a priori* information that the source has knowledge in the specific area that is needed by the recipient. The response to the final question in this set implies that the history of interactions is a factor in determining the perceived value of the suggested parameter *propensity to share.* It can be

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	what your default case is, so if your default case is to be skeptical or to be accepting and you know for the most part, my default case is..if somebody pointed at that person saying that that person has the answer, then I'd generally believe them unless what they tell me doesn't ring true consistently with what I have seen. Then I'll have to dig some more. S: *So, when you contact someone for help, for knowledge, how do you determine the person is really willing to share knowledge?* M: it's assumed. S: *So it's always a trust-based thing, when you ask, you trust that the person is not withholding anything* M: It's assumed. There are some people who I am skeptical about, who I don't trust. And it's a matter of my perception of their agenda is not that they have the same agenda that...what we are trying to achieve with the project. S: *So it's based on prior interactions with that person?* M: Yes		Subsection 3.2.1.8: 1, 2, 3	computed as follows: *Perceived propensity to* $share_{task\ n}$ = *f(accuracy of knowledge received during task n, perceived propensity to* $share_{task\ n-1}$*)*
l)	S: *Could it also be based on something you heard about them from someone else?* M: It could be that, it could be past experiences. I can think of one example. I won't name a name but a person who was instrumental to the first project I had who really didn't want the system that we were developing. He wasn't the principal user but he was impacted by the system and did not want to implement the system. There were things that his business system could have done to make the system more successful that he would not do because it meant additional things for his function to do. And that person, to this day, I am wary of, because I am not, I don't trust, I don't know what his motivations are, trust that he is motivated by the same goal that other people have. But it's something that you'd learn in the school of hard knocks more than anything else. So you run into resistance and you kinda learn who is resistant and who doesn't necessarily has the same goal.	Seek information on sources of knowledge about others' areas of expertise	Subsection 3.2.1.7: 3b	Supports the notion that acquaintances can serve as a source for others' areas of expertise. This point supports the need for including the notion of a network of connections among the workgroup members, where each member's network of connections can act as a source of knowledge about the areas of expertise of the non-connected (directly) members' areas of expertise. However, to keep the implementation simple, this phenomenon will not be implemented.
m	S: *So how do you, how is your decision to seek know-*	Seeks sup-	Subsection	The perceived value

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	ledge from a member affected by the person's tendency to make the knowledge public? This is building on what you just described in that if you have had prior experience, that turned out to be good or bad then you'd probably be more or less willing to contact him. M: The person that I mentioned, I am least likely to go to that person. If I need something from that department, I will try to find somebody else in that department who could answer that question . I will not go to that specific person unless I absolutely have to. S: *So, OK, so you won't go back to the same person twice because you've had a bad experience?* M: I won't know if I would say because of a bad experience but if I've seen a pattern. S: *So it takes more than one experience for you to decide?* M: Sure. If there is a pattern of experience that that person is going to be difficult and my word, my interpretation, my perception..but if there's a pattern of behavior that would lead me to believe that that person is not going to necessarily cooperate, I will try to find a different way.	port for determining whether a person would use the same member as source on more than one occasion, despite receiving less-than-accurate knowledge.	3.2.1.7: 3a; 3b Subsection 3.2.1.8: 2, 3	of a source's *propensity to share* is updated after each interaction and when a new knowledge-seeking need arises, the updated value of the perceived *potential to share* is used in determining whether a particular source can be approached once again for knowledge.
n)	S: *OK. Now assuming that you do go ahead with contacting someone and that the person has given you some knowledge about which you had no prior conception or idea, how do you go about judging the accuracy of the knowledge given by that person?* M: it sort of depends. Sometimes.. often...I am trying to find the right modifier...often the knowledge that I have collected is something that I can verify in data . So mortgage data is pretty finite. Other than what we have to add because of government regulation that are coming down every week, what we need to know for mortgages for the secondary market is finite. So, if you are asking for information that is going to be about something that exists already, or exists in some form in data, you can actually go into the data and find it. So, if I ask somebody a question about calculation or how to decide if a loan meets certain program requirements or whatever, and they provide enough information, typically what I would is I would go to the data to see if I could confirm that in the data. Then I'd be pretty comfortable that I understand it. If I can't, then I would collect the data that is confusing to me and I would go back to that person and try to triangulate it to find out what I misunderstood. S: *So you try to validate and if there are some incon-*	Understand how the accuracy of a source's knowledge is determined. Understand how the causes of inaccuracy are ascertained (in this and	Accuracy is determined based on whether the knowledge provided can be applied to solve the task at hand. Subsection 3.2.1.4: 3 This set of responses also can be	An interesting phenomenon worth considering is the deliberate provision of incorrect knowledge. The knowledge vector should now be modified to hold three, instead of two, states: 0 indicating absence of knowledge, 1 indicating presence of correct knowledge, -1 indicating presence of wrong knowledge. Transmission of inaccurate knowledge therefore means that a -1 is passed by the source to the recipient instead of a 1. the probability of sending a -1 instead of a 1 is governed

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	sistencies, you'll go back. M: Go back... S: *So it's like an iterative loop...* M: Absolutely S: *Unless you...until the time you have it resolved.* M: Right. Absolutely. S: But what about those inst... have you had instances where you did not have such data to validate a person's... M: Oh sure, new requirements... S: *OK* M: And there all you can do is take.. I mean...If you are going to a business user for their expertise, you have no other way to validate it...all you can do there...let me relate it as a situation. Typically, in that scenario we go to the user to get requirements. Get requirements to make a design decision to meet a requirement. So out of that interaction there will be something produced, typically a requirements document, a change request, a detailed design spec and through our process those things would be reviewed by more people. So I would take the requirements person A gave me, document it, send it out to the whole business unit and all the key participants..we call them SMEs (subject matter experts). So in this project we have subject matter experts from each critical business function. So a requirements document or a business document would be sent to all of them. And we would actually facilitate a review through NetMeeting or a Centra session to review that document all of them and that would probably take a couple of passes..right..because the documents are never right the first time.. you modify them and again you are doing them in real-time so people see what you are keying, and you send them out for review again and you may have another review and then they send they send them out for formal sign off. So there is a back-end documentation, review sign-off process that occurs where everybody at least agrees that the information collected from the expert is right. Now in the end eventually comes through and the data might not support the rule and we'd have a defect or a change request or something else. S: *So when such a thing happens, is it because even the SMEs did not understand the requirement correctly?* M: Our largest challenge is the data. So if we go back	the following questions in this set of questions). Also, to determine the factors that the interviewee thinks are the causes for receiving inaccurate knowledge from the source and to determine whether a single instance of receiving inaccurate knowledge would lead the recipient to refrain from contacting the same source again in the future.	interpreted as providing support for the phenomenon of stickiness of knowledge Subsection 3.2.1.7. Subsection 3.2.1.4: 3	by the value of *pts*. For example a value of *pts* = 0.6 implies that there is a 40% chance that a -1 is sent instead of a 1 to the recipient by the source. The reason for providing inaccurate knowledge can be time pressure, misunderstanding, deliberately misleading etc. Furthermore, misunderstanding can be seen as resulting from a lack of common/shared knowledge. To avoid complicating the simulation, the following simplifying assumption will be made: the recipient of inaccurate knowledge does not distinguish, qualitatively, the three different types of causes for being provided inaccurate knowledge. All that the agent knows in order to make a rational (that is, goal driven, based on the conception of rational behavior by Simon (cf. 1995)) is that it received inaccurate knowledge from the specific source and that it should therefore update its perceived value of *pts* for that source. In the future, when trying to determine whom to contact for knowledge, the updated

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	to this picture up here (pointing to the sheet of paper). I have field ABC and it has values 1, 2 and 3 and 1 means something and 2 means something and 3 means something. And we are going to work with the people up here because the data is being originated at this new origination system and we are going to have some mapping...some data mapping...and so we are going to say “hey, we need this information, its definition is X, the values we are looking for are 1, 2 and 3 what do you have?” They'll tell us. They'll give some construct, some table names, some column names...things like that and tell us what values they will have. It might be a relationship, it might be a series of tables...we don't know. But they will find someway...the data-modelers out here will find some way to get this piece of information to us. So then what happens is the data-mapping plus transformation rule. And let's say that up here the field is...well...X, Y and Z and its values are a, b, x and 2.. you know I am just making that up. And so the transformation will say that “if you get A that will equal to 3, if you get B that will equal to 1, two Ss equal to 2 and that's where your problem occurs and actually in this system, they use very verbose domain values and so they might have a domain value that something like this (writes something on the sheet of paper). So we will have a transformation rule that says if funsperse [not clear] then move 1, if return a float then move 3 and so on and where this breaks down is when, well, that's not the value, if it is upper-case or has an embedded space or it doesn't exist. And so, even though all of the experts agree that this was the right thing, in the end it was wrong because they implemented something different out there and that's where it's really gray because the data-mapping is occurring at the requirements-level, at the design-level, not at the running code-level, so you are subject to those kinds of issues. S: *So when you discover that SMEs or the folks whom you contact, the knowledge that they give you turned out to be inaccurate, how does it affect your perception of the SMEs' knowledge?* M: In this context, it does not because it is outside their control. If I found an SME was consistently incorrect, then I would probably try to find another SME....as another voice S: *Would you still consider that person to be an SME?* M: Yes, I would. And the reason I say that is because so much of it is contextual. So one of the business functions I have is loan delivery. And these are people		Subsection 3.2.1.8: 3 Subsection 3.2.1.8: 2, 3 Subsection 3.2.1.8: 3	value of the perceived *pts* will be used and the agent seeking knowledge would therefore choose to contact, with a greater probability, those agents who are more likely to provide it with accurate knowledge that it needs, irrespective of their motives for providing the right knowledge.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	who actually take the mortgage and where it is originated...so you sign your paperwork, you got your house and now a big packet of information a folder comes floating through the system and we also are going to sell the mortgage to Freddie Mac or Fannie Mae or FHA and the loan delivery people are the people who take and prepare the mortgage to be delivered to the downstream from upstream. They deal with all kinds of different loan products. And so much of their answer is going to be very contextual to the specific product. So the project that I am describing right now is purely to deliver to FHA – Federal Housing Administration loans, first mortgages. And so the data requirements and the procedure requirements for FHA are manageable. And so they are going to give you information specific to FHA but sometimes its gray. Because there are shared processes, we do the same things for all loans, for FHA loans, we do this. S: *So, certain exceptions...* Certain exceptions. And you know people, they generally try to tell you if they are not sure. And sometimes they don't realize that they're not sure. So if you have a good SME, you don't get bad information because you do not...you're talking to, that's just bad information [not clear]. If it is consistently bad, you look for a different SME. S: *So what do you think..I guess you have already alluded in that sometimes they don't realize that they are wrong. Can you think of other factors that make people give you wrong knowledge; wrong information?* M: It's the whole communication gap, right? So, you ask a question that you think is clear and comprehensive and they give you an answer that they believe is clear and comprehensive and it's kind of like answering comps questions, you know..you answer the questions you are asked and sometimes the answer is so compelling that...you know...sometimes the question is so ..sometimes the question can lead you in the wrong direction, sometimes your answer to the question could be interpreted in the wrong way. So you've got a communications gap. That's the biggest challenge. S: *Do you think that people might also give you wrong information on purpose, to mislead you, for instance?* M: I have not experienced that. I have experienced people who may give me a quick or a snap answer without a lot of thought S: *Possibly because they don't have time?*			

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	M: Time pressures, maybe they don't think the question is important to them. I can't say that I felt ever that someone's intentionally misled me. S: *So it's because they have more important priorities and things like that?* M: Sure. The biggest danger is when you ask a question that appears simple and you get an answer that is simple but the context of the question and the answer are entirely different. So it takes a lot of energy and work to frame the question to get the right answer and it is not always obvious. S: *So you have to bring in the whole context into picture and the other person is on the same page...* M: Exactly. It takes a lot of work to know that the context is the same. At face value you look at a question that is straightforward and the person answering the question thinks that it is very straightforward, they answer it from their context, you receive the answer from your context. Everybody goes off and you do your thing and then you get down to testing it and you have a defect because..”oh that's what you meant” S: *So, that again is the communication gap that you described?* M: Exactly. Exactly. And the SMEs, they know their business cold and sometimes they make wrong assumptions about the level of understanding of the questioner. S: *Is it because the amount of shared knowledge is low?* M: Yes. That's a good way to say it. The shared context is just not there or apparent. I was working with this user in our Secondary Market Accounting Control and I guess I have been working with her for two years and she'll say things or give answers to things that we both think that we understand. And it's not until we get a defect that we realize that we weren't... S: *It's because you have a different...* M: “Oh that's what you meant”... S: *It's because you have a different mental picture...* M: Right. And a lot of times you cannot find it until you have real data and you can say “this isn't working the way I thought it should. This is what I see. Why is that? Because it should be like this but it's not.” S: *So the data provide a kind of litmus test, so to speak...*			

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	M: It really does.			
o)	S: *OK, I guess question 10 leads in a slightly different direction. Do you revise and if you do, on what basis do you revise your judgment of others honesty when it comes to sharing knowledge. This is slightly different from your previous response in that where you said people don't deliberately mislead you. But..* M: Is this more in the area where they withhold information? S: *Yeah, or if you believe that they have, for some reason, misled you on purpose, then how do you revise your judgment of them? Or would you continue to think that they probably did not want to mislead you? They just...* M: Well, I explained my basic premises, my default cases... S: *Yes, always trust...* M: They weren't malicious in their intent. But again if you figure out that it happens enough times, you start to take on a more protective stance. So, if you received information from that person, make sure...make doubly sure that you validate with the data, that you have somebody else you can vet it against. Make sure that it is good. I mentioned a lot of times, or most of times, this information ends up in design documents or in requirements documents and so in those cases you really do make mental notes of those areas where you relied on that person exclusively for information and you try to make sure that it gets a real sound review. So I go to that person and I make sure that we go over this very well and walk through and make sure that people are very clear, that they understand. S: *Have there been any instances at all where you have gone through the whole thing and you came to the realization that that person probably misled you on purpose? Because maybe the person is not honest or forthcoming..you know various such reasons* M: No examples come to mind, at least not my experience at YYY. S: *OK. Which is probably a good thing to have in a company, a very trusting culture* M: It is a very trusting culture. I know I am an individual contributor and I work with individual contributors. If you get up...I don't know what might happen if you get up and say look we are afraid of managers and business leaders and things like that.	To obtain further support to the notion that no matter what the putative cause is (according to the interviewee or the recipient of knowledge), the source would still be contacted in the future if a need to do so arises.	Subsection 3.2.1.8: 1	

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
p)	S: *This builds on all the things we have discussed so far. What makes you ..or what do you think affects a person's tendency to share knowledge with others. Or, may be I should be more specific – not just sharing knowledge but also making all of his/her areas of expertise known to others.* M: Well, I am in the ..I suppose the classic would be job protection. Job retention. You know, if I am an expert in something that makes me valuable S: *So you'd like to let it be known to others that …* M: Yeah, if people need this information they need to come to me. To me that would be a ...probably a more explainable reason. Of course, there is a need to know. I don't deal much with that in my role. I mean there's not anything that is extremely sensitive. There would be possibly, the keeper of the information didn't think I didn't know, didn't think I would understand it, doesn't understand why it's needed. You know there are a number of reasons...doesn't have the time to tell me. So there are a number of reasons why people might withhold information. Probably job protection would be the one that would come to mind. S: *Do you think the person's prior experiences as a receiver or seeker of knowledge also affect the way the person chooses to either provide or not provide information?* M: Yeah, I can see that, I would think so. You have, maybe you have some people who would just hate to have to ask for information and want to know it all, find it all themselves and … S: *Do you think such a person let others know that he/she is knowledgeable or would he/she just keep his/her knowledge to himself/herself?* M: I don't know. I mean, it depends what their motivations are. Some people are just inquisitive and want to know . Others want to know, they want people to know for egotistical reasons, right? S: *Right* M: So, if somebody wants to know for egotistical reasons then they make sure that everybody knows that they know and they are the expert, whereas if someone is internally motivated to know then they are not likely to be so outspoken about what they do know.	To identify some of the factors that lead to a person's willingness or unwillingness to share knowledge with others.	Provide general support to the notion that inaccurate knowledge could be provided to the recipient of knowledge. Subsection 3.2.1.4: 3, 4. While not directly supporting a specification, these responses identify the factors that affect the tendency to share accurate knowledge, implying support for the inclusion of the parameter *propensity_to_share*, which encapsulates the effects of all the	While this does not directly inform any particular point in the simulation specification, it does provide evidence to determine whether (based on the set of interviews) there are common factors that drive a source of knowledge to be unwilling to share knowledge. Therefore, it substantiates the inclusion of the parameter *propensity to share* by showing that the tendency to share knowledge with others is not constant across all members of a workgroup and across different workgroups. That is, there is a variation in the tendency to share knowledge across individuals and the workgroups and organizations where they work and therefore this variation should be captured explicitly in the simulation via an appropriate parameter.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	S: *They might still share when they are approached but they don't go...they don't put any extra effort into making others know that they are experts* M: Right. And to me, you get into number of inter-personal dynamics. Some are introverted/extroverted, are they sure of themselves/unsure of themselves. So someone may have all the information but if they aren't sure of themselves in terms of communicating, being able to defend what they know if it's challenged, people might not be sure of what they know just because their friends might challenge them and they'll be uncomfortable having to defend what they know. There are people who are very risk averse. They could have a document that says 100% what it is, but they might not share it because they don't want to do anything that can come back and get them later. I think it's ..you can explore the whole realm of personal-interpersonal motivations. S: *And also the personality itself..the person might be introverted.* M: Yes, exactly. And then there is the time factor.		relevant factors in determining the likelihood of sharing accurate knowledge by an agent.	
q)	S: *The final question: in your judgment, how does a person's tendency to accurately share knowledge change over the course of interactions with others?* M: So you are talking in a general sense or in a more specific pairs. S: *Yeah, well, pairs as well as part of a work-group* M: Obviously, I can answer from my own interactions. As I mentioned my default case, I am going to tend to trust and share as much as I can, if I feel that somebody is open and honest and sharing, I am going to continue to develop and take advantage of that and share. If I feel that somebody is going the other way where they are are least likely or not likely to be open and honest then I will try to find other sources of information. I deal with a lot of people where it's sort of a training rule and there are people who...you are....it's clear that they get it, or that they don't get it and you start to reach the perception or the conclusion that they won't ever get it and so it probably tends to span ...it's like a self-fulfilling prophecy that when you probably don't spend enough time trying to impart knowledge on those people because .. S: *Because you believe that they'll never get it* M: Yes, that's probably not fair but it happens, right? So as you interact with people, they develop confidence in that interaction and they tend to share more, you'll build on contexts, you'll share more information,	To determine whether the parameter, which captures a source's tendency to share knowledge, is a relatively stable trait.		A person's tendency to share can be deconstructed into two components: a general tendency to share and a more specific dyadic-level tendency to share with a specific person. Incorporating this deconstructed version of a source's propensity to share would add to the complexity of the model. Hence, it is assumed that this trait remains constant throughout the simulation, while recognizing that in reality the two component approach might be more meaningful.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	you'll receive more information, and so you'll develop.. so it'll become the whole being bigger than the sum of its parts and as...if on the other hand if you see that interaction isn't productive, then they tend to not follow that interaction. And so it's back to there are some people that you'll go to if you absolutely have to. You'll provide the information that they need, if they need it. My approach is never to deny anything if somebody needs it then they need to know. But, you'll spend less energy on those kinds of relationships, those kinds of knowledge exchanges and you'll want one that is productive. S: *So you'll want who's willing to share, one that's more productive..* M: Right, right. That's my sense. S: *Thank you very much for spending so much time with me.* M: Sure. Glad to help.			

I

Table A2.2. Interview of K

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
a)	S: *I am here with K, CTO of SSCC. Thank you K for being generous with your time.* K: You are quite welcome.	Introduction		
b)	S: To begin, please describe the nature of a project with which you were involved recently. K: You must be a C programmer because you started at zero. OK. Basically, it was a project to centralize the storage and set up our remote replication and give them the ability to have disaster recovery for a company that happened to have two locations, separated by about 15 miles. They did not have to have real-time replication. So, replication could be behind a bit and it involved both standard file system data that was served and database data. S: *What was the size of your project team.* K: OK The project team...there were two people from the customer who were involved. One of the them was a manager and he also did hands-on work with that we needed to know how to do it and change it. Another person was more or less...he did administrative stuff and needed to do it. He was an employee with	Description of the project and the team's composition		

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	them but also consulted with them. Which was interesting. Two people from the company, myself and one other person, and other people if we needed to call them in, which we did. So it was a total of about, really, outside a salesperson we added another person, so you could say 7 people at a maximum. Normally, it was 4.			
c)	S: *OK. When you were working with other members of your group, what types of technologies did you use for communicating?* K: We did face-to-face meetings, we emailed, we WebExed and we talked on the phone a lot. I would say primarily, for touching base, when we were not on-site, we were making phone calls to each other. And my style, is more of a phone-call style...a lot of people are email-style, which is nice because you have nice chain of events with email and somehow, somewhere you will get email. Even if you are reading from a phone, you'll eventually get it. But my personal was, I am a phone person. It's probably because I am older. S: *Is it also because you wanted a more immediate feedback?* K: Yes. A lot of times, if you are not immediate feedback, you can get bogged down waiting. It can be something that you cannot ask and the customer has to ask the question for themselves for you and if you cannot get hold of people for meetings or for any other reason, it can knock your timeline off. So all of a sudden they are complaining to you saying you're not reaching your goal, you are past your deadline. But some hinged on a question "yes", "no" that took about 30 seconds to answer. So, I do depend on the phone definitely for the immediacy. S: *So the technology that you use, is based on the context?* K: Yes, it is definitely based on the context. If it is not that critical and it is general information and it is not critical, I will email it first. If they are general questions and, you know, they are not timely, I'll email it. If it is timely, I would go for either calling them on their phone or texting them on their phone, if their phone will accept text. A lot of time you can get by, by texting people in meetings and getting information from them and getting by. People don't necessarily like meetings but it is something they accept, particularly in large corporate situations.	Identify the different types of IT used and the reasoning behind the choice of specific types of IT under different circumstances.	Subsection 3.2.1.4: 4	• The workgroup members can be geographically separated – this requires that each agent have "location" as an attribute. • The location of the agents would determine (not solely) the choice of IT used by them for communication. • The priority of the task of the knowledge seeker also determines the choice of technology This set of responses provides

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
				support for the functional form from points h) and i) of table A2.1 *probability of choosing a medium for communication = f(stickiness of the knowledge to be exchanged, richness of the medium, synchronous)*
d)	S: *Moving on to the next question, third question, question number 2, tacit knowledge is something that is described as difficult to articulate, verbalize, things that you learn over a period of time and cannot be put into words immediately to explain to someone else. So, essentially, it is difficult to transmit to others. If you were dealing with such knowledge, such tacit knowledge, how do you decide on which medium to use, which type of technology to use?* K: You use tacit knowledge a lot when you are hitting a problem. If things are flowing smoothly, what you know and how you are doing it is fine. When you are getting into a problem, your standard knowledge is used up and you are hitting tacit areas. So, depending on first time you've done it or the 100th time you have done it, you'll have a feel. Just like you were saying, it is very difficult to articulate things about this problem. It feels like it's a network problem; it feels like it's a hardware problem. And that's really, coming from a computer science background, where it is science and not a social science background where you are worried about feelings. But, the more I have dealt with computers, the more I feel like a computer psychologist in terms of dealing with vague situations. It is not that cut and dry many times. But what I will first do, when I am hitting something like this, is I will tie into the people immediately around me. As you work with people you'll know what their strengths and weaknesses are. You generally talk to people, you talk about this before you go into a project or a meeting to have them help you with some aspects or give you their feelings, if you will. If that doesn't work, you happen to work with people that happen to have greater knowledge with the subject or more experience with it, you tap them. After that, depending on the problem, simultaneously do things like Google. Also depending on the product, you might have a secondary website	The reasoning behind the choice of a specific medium of communication when the knowledge exchanged is “tacit”	In the simulation tacitness of knowledge is captured by the notion of stickiness. Subsection 3.2.1.7 Subsection 3.2.1.7: 1, 2	The comments made in the item h) of table A2.1 are applicable here. The functional form from item c) above is relevant here, too: *probability of choosing a medium for communication = f(stickiness of the knowledge to be exchanged, richness of the medium, synchronous).* The responses to the last two questions in this set imply that the preference of IT of the source and the requester of knowledge are relevant. Inclusion of this aspect into the simulation would add to the complexity of the simulation by expanding the set of parameters. The inclusion would require also the creation of rules for determ-

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	that people who are partners to the company have access to and not the general public. So these are higher technical resources you might have, so I would look at the ...everything you can find on Google, I am also simultaneously looking at the manufacturer's site, because since it is a tacit knowledge problem, the description of what the problem might be shown as many different ways vague problems. Then you may be put this up as a question on the Web. This is a specific thing that you are looking at this angle of it and you might look at several different things depending on who did it. But it can all be the same problem and you are feeling different symptoms depending on what you are doing. So that would be it. S: *But, when dealing with problems that require tacit knowledge, do you also seek help from your colleagues?* K: Yes, absolutely. S: *When you are dealing with colleagues, do you have a preference for a particular type of technology, communication technology?* K: generally, it is going to be... when you are getting into a problem where it's tacit knowledge, then it's a timely thing, so I would try to do the phone, IMing and then email, in that order. That's basically how I do it. Because normally when things at your end are tight, where you are getting to a point and something is going to take you off your deadline, that's what I would do in that order. S: *Is your decision to choose one particular technology also based on what you know about the other person's preference?* K: First, it is my preference, then it is theirs. Because I am a phone guy, I would like to call them.			ining when the preference of the source dominates the preference of the recipient (and vice-versa). Hence, unless there is further evidence from the remaining two interviews and/or literature, which explains how to create the above-mentioned rules and that the preference for IT is important from a theoretical-richness perspective, it is assumed that the preference of the source for a specific type of IT will not be considered by the requester while making a choice of IT for communicating with the source.
e)	S: *When you need to contact someone else, some other colleague, a distant colleague, how do you decide on...given a choice of three or four colleagues, who potentially have the knowledge you need, how do you pick the one you do pick?* K: OK, if I know that of the four colleague, the one colleague is not busy on something, I would call them first. And if everybody's busy, then I would call the colleague that I have the best relationship with. You will run into people that are extremely technically knowledgeable, but they... but from a personality standpoint, they don't like to be bothered. Or, possibly, they don't like to share their knowledge. That actually	Obtain information about the process the interviewee used for selecting his source of knowledge.	Subsection 3.2.1.4: 3, 4	Indicates that the level of activity (how busy the person is) of the source of knowledge should be considered by the knowledge seeker before contacting them. Given that the goal of dissertation is to model transactive

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	comes up as a power thing. “I know this, I know what your problem is, but I am not going to tell you.” So, that's something that you run into. S: *OK. Do you also use things like recommendations from your previous contacts, friends...* K: Yes. Or even customers. Depending on the thing you might call the customer and when you are working as a company to do installation and troubleshooting, you often don't get the long-term time administering or seek time in front of it, that makes you have much more in-depth knowledge of its behavior, if you will, or things that you need to do to make it work properly, that are not documented or not documented to public. So, I have a problem with something and I barely touched it, and that's what the customer does, then I would call them and ask them. S: *OK. Were there instances where you had to contact complete strangers based on just the information that that person is knowledgeable?* K: Yes, absolutely and I would contact them and say “I am sorry to bother you. I spoke to this person and they asked me to go ahead and call you and ask you, because you happened to deal with this”. S: *So, you came to know about an expert through another expert, based on recommendation?* K: Yes. Based on recommendation. S: *How do you figure out the areas of expertise of your workgroup members?* K: Basically, if you have been around them for a while, you'll know their areas of expertise. If you don't, what you do is, before you embark on a project or before you embark on a meeting, every body will introduce themselves and say what areas of things they are experts at or what areas of things they have worked in more. So, basically, someone telling you “these are things I have experience with” that's primary way. And then what you will find is in certain cases, you will have people that have capabilities or experience because of certain problems. But you have people “I have solved this one narrow problem involving this and doing this” and you have other people, just you dealt more with it...but talking to them, that's how you learn. S: *Previously, you have mentioned that you will also get recommendations from your existing network of friends.* K: That is correct.		Subsection 3.2.1.18: 1	memory, that is a group-level memory of “who knows what”, the specific mechanics of determining who among the given set of potential knowledge sources should be contacted is irrelevant. Hence, while this information is interesting, it does not add to the simulation and would be excluded. The last five responses to questions in this set provide support for the ideas of including notion of a network in the simulation and of having acquaintances acting as source of information about others' areas of expertise. Additionally, there is support for the “training” phase portion of the simulation in the response to second-to-last question “...before you embark on a project or before you embark on a meeting, every body will introduce themselves and say what areas of things they are experts at or what areas of things

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
				they have worked in more...”
f)	S: *Moving on to the next question, which is number 5, how is your decision to seek from a member of your group affected by that person's tendency to make his/her knowledge public?* K: When you have people who are not afraid to share their knowledge publicly, it does make it easy to know what kinds of things they have dealt with and helps you quickly say “hey, you were working on Exchange the other day” or “you were working on this big clustered UNIX server, what do you think of this problem?” So actually, that's good. Some people are completely quiet, they don't share knowledge and they might be the smartest person in the room and you'd never know it. You won't be able to do things as quickly because you don't know that they know, but where someone who talks about it, you'll at least have some idea that there is someone to talk to. S: *In what context would someone talk about their areas of expertise?* K: just talking about some things...just talking. S: *A general conversation?* K: Yes, a general conversation. S: *At a water cooler or somewhere..?* K: That would be correct, yes. S: OK. Did you also have a technology similar to corporate whitepages of sorts where different people's areas of expertise are listed? K: Yes. Quite often when they start getting smart, you might set up a wiki or something like that to have everybody write about how they have installed something or done something... it's kind of like a global, corporate knowledge-base. And also this could also potentially integrate into trouble tickets and things like that if your system allowed it. If you have a customer that had this problem, this was what the problem was, this was how you solved it. These....may be you had links to versions of the software if you had to upgrade the firmware on the switch or they needed a new version of the client and may be have the versions of the software on your site or link to their original site. Sometime it's better to have a redundant copy in your control so if you needed a specific version or feature you can get it, but that can make it much better and you have sort of positive feedback. “Someone's had this similar problem before, what date was it, what was	Obtain information that provides support for the inclusion of the parameter *propensity to share* and the inclusion of *perceived propensity to share* as an attribute of the seeker of knowledge and support for 48.j. Also obtain support for the use of “expert seeker” type of IT	Subsection 3.2.1.18: 1f Subsection 3.2.1.7: 10 Subsection 3.2.1.10.	The information obtained provides support for the proposed *propensity to share* parameter. Support also obtained for “expert seeker” functionality.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	it" and it can be very useful.			
g)	S: *Let's say over the course of a project during your interaction with various members, you come to know that a person is an expert in five areas, he's made his knowledge public in only two areas, how does that affect your impression of that person's tendency to share?* K: If they are sharing about something, I figure they'll figure other stuff too, and they'd not have a reason to not share the other stuff. So, generally if you have someone who's willing to talk and help you on one thing, even if it's something they haven't talked about and they know about, they'll share it with you. I should say the converse of it would be where some people, if they do not help you with one thing, they're not going to help you with another either. S: *OK. So it is based on your prior interactions that you decide that this guy was more forthcoming therefore he would be more forthcoming in other areas as well?* K: Correct.	Support for the applicability of the *propensity to share* parameter.	Subsection 3.2.1.4: 7	The information provided appears to support the notion described after table A2.1, where a person's propensity to share is described as composed of a generalized tendency to share and a more person-person, that is, a dyad-level tendency to share. However, as described in that section, the deconstruction of the *propensity to share* might not be practical from the perspective of its implementation. Hence, even though this set of responses can be interpreted as supporting the above description, the deconstructed version of *propensity to share* will not be implemented.
h)	S: *OK. How do you form the, now we are on question number 6, how do you form your judgment of others' areas of expertise based on the accuracy of knowledge they share with you?* K: You do judge the accuracy, it kind of depends on the problem and how long ago they solved the problem? Because as you get things like firmware changes, driver changes, things like that, sometimes you have a problem solved and a version or two later, the problem creeps back up and it could be that the way to solve it, how they solved it previously no longer works. So it's one of those things that, depending on how you diagnose it or they help you diagnose it or the remediation or fixing of the problem is not working. You do judge what they gave you. But you also take the knowledge they give you depending on how long	Understand how the accuracy of a source's knowledge is determined.	48.k, to the extent that the possibility of inaccurate knowledge being given by the source to the seeker is recognized. Accuracy is determined based	The description of modeling inaccurate knowledge being transmitted and retained will be implemented, since there is evidence that accuracy of knowledge transmitted affects the recipient's perception of source's level of expertise in the specific area of knowledge that was received by

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	ago they solved it and how close the close the versions were. Information could be very accurate. It's just that things have changed so much that you have to come up with a new solution but at least it was something. S: *Following what you said, how do you judge the accuracy of the knowledge that was given to you?* K: Somethings, if it's ...you're very lucky and they have the exact same problem and what they tell you instantly fixes the problem then you know that it is accurate. Other times, you can tell if they are close or not, on what they,... part of the problem would be how you explain what you see to someone else. You might see that there are entirely different views on the same thing. S: *So, you need to be first on the same page, with respect to the problem you describe...* K: Yes. And there are some people who, for whatever reason, who will knowingly spread misinformation to you and may be they will come along and they will "look, I can fix this in 2 seconds" and they will fix it in 2 seconds. They knew how to fix it, but they were deliberately trying to make you look bad. That can also happen. That's a totally different thing. That's dealing with jerks. That's not a normal thing. S: *OK. If it turns out that the knowledge that was give to you was not accurate and less accurate, would you contact that person again in the future when you need knowledge?* K: Probably would. And you can probably can also tell "well, that stuff you asked me to do didn't work, but this stuff worked" and you'd feed the information back to them. So you probably try them anyway. S: *What if it turned out to be...?* K: What if they were trying to make your life harder? Then you wouldn't, because they are like, they want to get this person. S: *But how can you tell what the intent is?* K: I would say again, it's a gut feeling. Which might ...generally you'd find that the kind of people who would do that kind of a thing to be jerky anyway. So you'll develop gut intuition if someone is trying to make you look bad. S: *So, it is not that particular instance on which you base your judgment.* K: No. Correct. It's other evidence.		on whether the knowledge provided can be applied to solve the task at hand. Subsection 3.2.1.4: 2 Support for Subsection 3.2.1.7: 3a to the extent that the recipient is	the recipient. The description of accuracy of knowledge as being determined by whether it can be applied to solve the task is consistent with the interpretation in point n, table A2.1. Further evidence is also provided to the change suggested in point m) and in the section immediately following table A2.1 about a recipient seeking knowledge from a source despite having prior experience(s) where the knowledge provided by the source proved to be inaccurate. It is

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	S: *It's a pattern of behavior?* K: Yeah.		aware of a source's tendency to provide inaccurate knowledge and (implicit in the response) chooses a source that the recipient believes as having a relatively greater tendency to provide accurate knowledge.	evident from the response “I would say again, it's a gut feeling. Which might ...generally you'd find that the kind of people who would do that kind of a thing to be jerky anyway. So you'll develop gut intuition if someone is trying to make you look bad...” that determining the actual reason for being provided inaccurate information is a “tacit” skill and hence modeling it would not be practical. Hence, just representing the phenomenon of transmission of inaccurate knowledge to a recipient suffices for practical (implementation-related) reasons.
i)	S: *Question 7 overlaps with what you said, in that it asks you how do you evaluate the accuracy of knowledge that others share with you.* K: Just basically, when you are using their knowledge to help you solve the problem, you see if indeed that looks like what they were saying is correct. Is it a real close match, an exact match, or completely off-base. You have have judge each incident separately. S: *So, it's based on whether it fixes your problem or not?* K: Yes. Yes.	Understand how the accuracy of a source's knowledge is determined.	Subsection 3.2.1.14:2	The explanation provided in point n) of table A2.1 while not directly supported, is not contradicted.
j)	S: *Moving on to question 8, when and how do you revise your judgment of the accuracy of others' knowledge? This, as you can tell, builds on the points you made previously.* K: Yeah. You can have someone that used to have to work on a particular thing a lot but it has been several years and when it's several years, may be if it's a real	Understand how the perceived value of the source's areas of expertise	Provide general support to the notion that inaccurate knowledge	Questions 1 and 2 in this set provide interesting information: the decay in the accuracy of source's knowledge with time.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	general problem they are very accurate, but in newer, version-specific things which may be a Windows issue or a driver issue, their knowledge would be downgraded because they did not have to touch it in a long time. But, at the same time and you sit there and say "well, this is what I have seen before" so you will still trust what they say, but will not worry about dead accuracy. S: *OK. So, it's because you perceive that the person's knowledge is not up-to-date...* K: Correct. It could be that their knowledge is not up to date, but accurate enough. At other times, it could be nowhere near accurate. So it depends on the product.	changes in a given area, if the knowledge provided by the source in that area turns out to be inaccurate.	could be provided to the recipient of knowledge. Subsection 3.2.1.6. Subsection 3.2.1.7. Subsection 3.2.1.4: 3	For the sake of simplicity, it is assumed that the knowledge does not decay.
k)	S: *OK. Moving on to question 9, in addition to the things that you described previously about a person being a jerk, for instance, not wanting to share knowledge, can you think of other factors that affect a person's tendency to share information about his/her areas of expertise with others?* K: Generally I found that a majority of people will help you and freely provide you information. And there are others who don't want you to, or they will, if they are being forced to. So I think it's a personality thing. It might be that they get bothered by a lot of people because they have a know things and they want their work done. It could also be a power thing – a personality trait. S: *OK. So, it essentially, comes down to a person's context and the person's personality.* K: Yes.	Understand the factors that the interviewee thinks are determinants of a source providing inaccurate information.	Provide general support to the notion that inaccurate knowledge could be provided to the recipient of knowledge.	While this question does not support any specific points in the simulation specification, it obtains evidence to the phenomenon that a source of knowledge can and does provide inaccurate knowledge to the recipient and this can be due to several reasons.
l)	S: *On what basis do you revise a person's honesty regarding sharing knowledge with you?* K: On how well their knowledge helped you? S: *No, this is slightly different from the previous set of questions. The previous set was about how many areas a person was willing to make public. Does he know X, Y, Z and he tells you that he knows X, Y , Z or he tells you that he only knows X. This is about how sincere are they in reporting what they know and being accurate about what they convey to others.* K: Most people you talk to will tell you how long it has been since they've done something and they will also tell you, it's like "I've done this, this and this" and I go " Eh, I am fuzzy about that, but here's what it sounds like". They also give you their own judgment of how good their information is. So, generally you get the two things. The information and how sure they are. And, you know, they are generally pretty honest.	Obtain information about the process by which the perceived value of a source's *propensity to share* is modified by a recipient.	Subsection 3.2.1.8.	The responses do not provide the required information. The next set of questions – point m) – contain the required information.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
m)	S: *Have you had any instances where people give you information but they do so for some ulterior motive. That this, they are not completely honest with...* K: Yes, absolutely. S: *Would you care to...* K: Yes. Basically, it's when they are angry at people, they're jealous at people or they're basically jerky individuals. They will give you certain pieces of information and withhold critical pieces of information and will in some way attack you later on with "well any idiot knows something" or blah blah blah. When they give you information they knew all of this but they make it, they made it available publicly afterwards to have greatest impact – show they are better than you, they know more than you, or something like that. But it is basically a personality issue. I have seen with not just me but with other people. It is generally a personality thing. Personally, if I find people like that, I would just fire them, because not a whole lot of good can come out of it. S: *What do you think affects a person's tendency to be honest. Would they change their tendency to be honest as a result of their interactions with others or do they choose to be honest or dishonest all the time? Do they remain consistent, in your view? If they don't, what makes them change their behavior?* K: Most people are consistent. If you find someone who is a dishonest person, you are not going to ask them any more because they are going to waste your time on purpose.	Obtain support for the phenomenon where the perceived value of a source's *propensity to share* is modified by the recipient	Subsection 3.2.1.8.	The information obtained indicates that the recipient's perceived value of the source's *propensity to share* does remain constant, supporting the assumption made in point q) of table A2.1.
n)	S: *OK. Do you think a person who is dishonest, for example, gives you wrong information on purpose, is it because that person had bad experiences in his/her own life, because of which he turned out to be the way he is?* K: I don't know what caused them to be they way the are. I have feeling that it has to do with just them growing up. S: *OK*. K: It's separate from work. It's just personality. S: *It developed over a period of time?* K: Yes. Yes. And I am willing to bet that it extends beyond work. Just a generalized personality trait. S: *It's not something that appears/disappears in the*	Follow-up to the previous set of questions to find out the interviewee's perception of the causes that make a source unwilling to share accurate information.	Subsection 3.2.1.8: 2, 3	

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	work environment? K: No. It's like this: people who are like that [laughs]..I do not want to find out other aspects of their life. They could be absolutely great people at home, to their family, but at work they are abysmal human beings. S: *Is it because they need to play company politics.* K: May be. There definitely are people who are better at playing politics than others.			
o)	S: *I think I have gone over all the questions. Thank you once again.* K: You are welcome.	Conclusion of the interview.		

Summary of interview 2
Support: The second interview provided support for the following aspects of the simulation specification

- Geographical separation of workgroup's members
- The use of different types of IT based on the location, stickiness of knowledge, synchornicity of communication, and the richness of the communication medium
- The network of relations between a knowledge-seeking agents and other members of the workgroup can help a knowledge-seeking agent in determining who among the "friends-of-friends" can be a potential source for knowledge.
- The pre-project phase (described in the "training phase" portion of the simulation specification) where workgroup members interact with each other and find out about each other's areas of knowledge.
- 'Expert-seeker' type IT
- The use of *generalized propensity to share* to indicate a base-level tendency to share knowledge or information about the presence of knowledge in a particular area, averaged across all agents in a workgroup.
- Accuracy of knowledge defined as whether it can used in completing a task.
- *Perceived propensity to share* values, which are updated for each agent with respect to the specific agent that it contacted for knowledge, are used in determining whom to contact in the future. This makes it possible for an agent, which provided inaccurate knowledge during one of the previous instances when it was contacted for knowledge by a knowledge-seeking agent, can be chosen as source of knowledge in one of the future tasks, if it turns to be the agent with the required knowledge and has highest *perceived_pts* value associated with it
- The inclusion of the provision of inaccurate knowledge by a source agent to a knowledge-seeking agent is neither contradicted nor confirmed

Changes suggested: The following changes are suggested to the simulation specification:

- the probability of choice of IT depends, in addition to the stickiness of knowledge, richness of the communication medium and synchronicity of communication, on the priority of the task that is assigned to the knowledge-seeking agent

Table A2.3. Interview of R

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
a)	S: *I am here with R. Thank you R for giving me this opportunity.* R: You're welcome.	Introduction		
b)	S: *I will start with a background question. Please describe a project in which you were involved recently.* R: Recently...I have to go back a few years because I changed careers. But I will probably pick one of the projects I worked on at my last employer, that would be M. I worked on a number of different projects there. If I were to pick one, there's one where I did the analysis, design and development of a time-entry system. When I was there, the position I was in...I was the team lead at the company. There were fifteen, sometimes between ten and fifteen systems and we would support. I would do a lot of projects in the company. I am just picking a software development one that I did personally. S: *You said your role was that of the team lead.* R: Yes, yes. I was in the role of a team lead. I was actually assigned that role by the Vice President of the department. But this particular project was something I was doing on my own. On this particular project, I did analysis, design, development of a time entry system. S: *The question was to jog your memory and also to provide a context for the following set of questions.* R: OK. OK.	Description of the project and the team's composition		
c)	S: *While collaborating with your team members or members of your workgroup, when you are working on a single project, what different types of communication media and technologies did you use?* R: As far as media technologies for communication, email is probably the big one, as far as discrete electronic media. Most of the communication was done in person – face-to-face. That's by and large the number one way of how we collaborated.	Identify the different types of IT used and the reasoning behind the choice of specific types of IT under different circumstances	Subsection 3.2.1.4: 4	
d)	S: *If I understand you correctly, were all your team members in the same physical location?* R: Yes, yes. We did have some offshore. We did have some offshore components going on. That happened at different phases of the project. At one	Follow-up to question in b) to determine wheth-	Subsection 3.2.1.4: 2	Provides support to the comment made in c) of tables A2.1 and A2.2 regarding having a location

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	point we had a lot of work being done offshore. That didn't seem to work out really well, so we cut back on that. At one point, we had no offshore. When I left, we had a little bit of offshore work going on. It depended on the needs of team. In the team I worked on, we did not have one big project. We are part of maintenance, support and development. We had a lot of projects going on, all kinds of different phases. Just based on what was needed, we did some off-shoring.	er the team was geographically dispersed		attribute.
e)	S: *Was there a reason for you to choose email. Was it because it was asynchronous?* R: Yeah, especially for something like offshore. We would do...it just seemed to be most convenient. We did a lit of chat, very tiny amount. We also used Windows Sharepoint Services. That was, I thought that was critical. Checking in documents, identifying changes...we used Sharepoint quite a bit. These all pale in comparison to face-to-face. S: *When you say "chat", did you use text-based chatting or video, too?* R: Just text-based. We did not do any video conferencing, I don't think. S: *How about telephone conferencing?* R: Yes, we did telephone...still a minor role. Email ...here's how it worked: if it was emergency, we used telephone or face-to-face. Other than that we probably relied more on email. S: *OK. Based on how immediate you wanted your response to be, you chose...* R: Yes. That drove how we treated it. Yeah.	Description of the various types of IT used	Subsection 3.2.1.4: 4 Subsection 3.2.1.7: 5 Subsection 3.2.1.7: 5	Provides support to the points made in d), h) and i) of table A2.1, and c) of table A2.2.
f)	S: *As you know from our Ph.D. Studies, tacit knowledge is knowledge that is difficult to articulate, to verbalize, share with others. So, if there were instances where you had to share tacit knowledge, did you choose a specific type of communication medium?* R: That would be face-to-face. When there's something important, something hard to communicate, it was definitely face-to-face. We had a lot of that on our team. Some of the systems we were developing, we were supporting...they were fairly complicated. To had somebody go through and read documents on our system. I didn't think that was very fruitful, it would take them forever and then as we both know, tacit knowledge cannot be captured in a document. So we were training face-to-face and then we would apply what we called system rotation. We would share system support among various team members	The reasoning behind the choice of a specific medium of communication when the knowledge exchanged is "tacit"	Same as described in h) of table A2.1 and d) of table A2.2. Subsection 3.2.1.7.	The assumption made in d) of table A2.2 is not contradicted, hence the assumption can be considered reasonable.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	and the whole team could come up to speed on all the systems. S: *OK. So, different members were exposed to the systems at different times?* R: Yes. It was actually my idea and we actually carried it through. Yeah. S: *OK. But were there any instances when you had to share tacit knowledge with folks offshore?* R: Yes. In that one...we would actually use phone calls, sometimes in the morning, sometimes in the evening. That's the best we could do. We ran along with offshoring. At one point we discontinued it. Here's something interesting. We discontinued it..but once we reinstated it, it was with people who had already been onsite and knew the systems. Then they could go back offshore and they could resume. S: *So they had the contextual knowledge.* R: Yeah, yeah. In fact, one of my team members, was visiting his family in India and he was there for a few months. It was almost automatic – how he dealt with problems quickly. I would send an email saying "this needs to be done with the system". I will notify him at the end of the day and then I come back the next morning and it would be fixed. There was very little communication. S: *Because he knew what was needed?* R: Yes, because he knew what was going on. I thought it was great. It seemed to work really well. S: *That's because he had the tacit knowledge?* R: Right. All that stuff had been transmitted, so at that point we could communicate in small amounts of information. S: *So email sufficed?* R: Yeah, email sufficed. Sometimes I would send him an email and he would know exactly what was going on. It was quite nice. S: *If you needed knowledge to solve a particular problem and you did not have that knowledge already with you, how did you decide on whom to contact for that knowledge?* R: For a particular system or a business problem? S: *Both*. R: OK. For finding the knowledge needed when something comes up, I first try to identify "is this a		This response indicates that a pre-project phase where each member learns about at least a few of the areas in which other members of the workgroup have knowledge is a reasonable specification. Subsection 3.2.1.18: 1 Knowledge 4, 5	While the response to this question does not necessarily imply a "training" phase as described in subsection 3.2.1.18:1 it does imply that there is a pre-project phase whose outcome is that the members of the workgroup develop a perception of at least some of the others' areas of knowledge. The functional form presented in d) of table A2.2 should be modified to include the amount of knowledge the seeker of knowledge believes is common with the source. So, the functional form becomes *probability of choosing a medium for communication = f(stickiness of the knowledge to be exchanged, richness of the medium, synchronous, perceived amount of common knowledge)*.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	business problem? Is this a system problem?” in some cases it would be both. May be I wouldn't know. My first step, if it is a system problem, is to go to whomever created the system. Because they are the ones who have the most knowledge about the system. If that's not available, sometimes I would look through code; I did that a lot. Sometimes, I look through the code first, rather than go to whomever created the system. But usually, that's the fastest way to get problems resolved. If they are not around, I would go to the main support person for the system, who wasn't always me, obviously. Sometimes in the company people move around. So you might actually need to go to a different department to track down whomever is most knowledgeable about that. For business needs, obviously there's an owner of the application. So you have a system owner and a business owner. A system owner is obviously technical. A business owner is more usually on the business side of the software. And I might contact them as well.			
g)	S: *OK. How do you know the whereabouts of the person now, that such and such person was responsible for the system?* R: There are a couple of ways to track them down. One is that I just worked there so long, so I kind of know who owned the systems. Sometimes you could ask some people. Also, we tracked the system owner in one of the systems that we created. We created a system called “centralized infrastructure database” CID for short, and you can look up the system owner in there. S: *So it is like a corporate database with “person and responsibility” pairs?* R: Yes, yes. Here's something funny, a funny note, an aside. Sometimes at the companies where I worked, I am not saying it is at this particular company, you have systems in place and nobody would know who the owner was [laughs]. No body would know who created a system. S: *It it because the persons who created the system left the company?* R: That's most likely the case, yes. And there was one place where I worked where we actually had to go through and try to track down the systems to see if systems were running and who would support them and in some cases we would find systems that were running may be for years without support. They weren't assigned to anybody. It's just something that	Support for the use of 'expert seeker' Obtain more information about how a person with the required knowledge, when he/she is not part of the workgroup, is	Subsection 3.2.1.10. Subsection 3.2.1.7: 3b	The “centralized infrastructure database” is similar to the 'expert-seeker' described in the simulation specification (34.b) and is consistent with the findings from the previous two interviews as identified in point j) table A2.1 and point f) in table A2.2. The response to this question implies the possibility where no member of the group has the knowledge required to complete

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	somebody created years ago and was using and that's... it's surprisingly common for such systems...there are not many such systems. If there's a big company, there might be a system out there that has been created as may be something small and nobody knows who created it, owns it or supports it. S: *Would you also look at the source code to find the author of a particular system?* R: Yeah, we used Visual Source Safe. You can use it to see who checked in the code originally. Sometimes there would be comments in the code. Not always. Lot of code that I have seen wasn't well commented.	identified.		a task or the possibility that a seeker of knowledge does not have the knowledge required to complete a specific task and does not know of any other member of workgroup that has the required knowledge. This creates a situation where the specific task cannot be completed as a result of a lack of requisite knowledge or as a result of ignorance on the part of the knowledge seeker of another agent who has the required knowledge. The latter phenomenon can be interpreted as a case of poor transactive memory. This phenomenon can be implemented as follows: if an agent cannot obtain the knowledge to complete the task after trying for a pre-specified number of time periods, the task is abandoned. Consequently, the project also is abandoned. While the interviews do not provide direct evidence to corroborate this observation, a review of literature is expected to provide the required evidence.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
h)	S: *Going on to question #4, what process did you go through to develop an understanding of the areas of expertise of various members of your group?* R: That, I would actually start before they were hired. I looked at their resume pretty thoroughly. Check references, ask what kind of a person they were. I may have already known some of these individuals.	Obtain information about the process the interviewee used for developing an understanding of others' areas of knowledge	Subsection 3.2.1.18: 1 Subsection 3.2.1.18: 1 Subsection 3.2.1.7: 8	The response to this question is consistent with the observation made in point e) table A2.2 about the pre-project interaction phase described as "training phase" in the simulation specification.
	S: *Personally, or through acquaintances?* R: Both. But I found that usually when somebody started on a team, I would spend a couple of days with them because I was the most knowledgeable. Working with them very closely. Might chat with them for a little while and then we will make assignments to them on things they could do. Based on how quickly they turned those assignments around, I can gauge the quality of the work pretty quickly. If if assigned somebody a task that I estimate should take them an hour to do, even for somebody new, it should not take them more than an hour to do and it took them 4 hours, then red flags would go off. Other indications might be something I would assign, and think that it might take a little while and they might solve it pretty fast. That would be my first indication of their knowledge level, one of the first indicators of their abilities. As time will go on, you start to see a trend. You know this person is turning around problems really quickly, coming to me after identifying flaws in the code that I had missed before … you know those types of things.	Obtain further support for the notion of learning of others' areas of knowledge through mutual acquaintances	Subsection 3.2.1.18: 1 Subsection 3.2.1.7: 8	The response supports the inclusion of the "network" notion and the use of one's personal network as a source for learning about non-relationship-having-members' areas of knowledge, as proposed in point j) table A2.1 and supported via evidence in point e) table A2.2.
	S: *OK. So, it's mostly through personal interactions that you develop a feel for the other person's expertise?* R: Some yes and some no. If somebody sent me an email, you'd count it as a personal interaction, "yes, I've completed this problem, I checked it out", you know I would verify what the changes were. Someone would develop a kind of reputation for a team, I would know what their ability was. S: *Based on their output?* R: Yeah, based on their output and I will check to see what they've done. On the other hand, I really.. sometimes it might just be a matter of work ethic that is causing problems. You know, may be I might go and	Obtain support for the notion of learning about areas of knowledge of the source that are not related to the knowledge needed to		Given the interviewee's role as a team-leader, the dynamic he shared with his workgroup members is different from that that the previous two interviewees had with members of their workgroups. Hence, the current interviewee's process of knowing his workgroup's members' areas of ex-

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	visit one of the developers, stop in and see how is this problem going, and see that they are surfing the web instead of working on a problem... not that it's a problem to do that every now and then. But if I come by a lot, if I visit your cube a lot and ask how a problem is .. and you're surfing the web every single time I visit, then it's probably more of an ethical question than one of ability.	solve a specific task		pertise is based on whether they complete the task they were assigned. However, the interviewee could determine whether each member of his workgroup did indeed have knowledge that is needed to perform the specific tasks assigned to him/her. Thus, the interviewee can develop an understanding of each member's areas of knowledge and thus, the interviewee can be part of the workgroup's transactive memory system. Even though the distinction of the roles is recognized in this case, using the criterion of simplicity, it is assumed that the simulation would not implement multiple roles e.g., worker, supervisor.
i)	S: *OK. How is your decision to seek knowledge from a member of your group affected by that person's tendency to make that knowledge public? To elaborate, when you try to seek knowledge from someone, how is your tendency to talk to that person to get knowledge affected by that person's tendency to make his/her knowledge public?* R: Oh, it's a critical factor. Critical. Interestingly enough, a lot of times, in the corporate world, their reputation will precede them. So if I need a piece of knowledge about a system, it's frequent …if someone's been at a company for any length of time and if he's hard to work with and doesn't want to share knowledge, then that's pretty well known by a	Obtain support for the construct *propensity to share* whose need and meaning are cited in table A2.1, point k and explained in	Subsection 3.2.1.4: 7 Subsection 3.2.1.7: 3a	The response indicates that the “network” can be used by a seeker of knowledge to not only gain information about who knows what but also about the *propensity to share* of each person who is connected to the person that the knowledge-seeker

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	lot of people. It's not something that's a secret. You might hear from a person that such and such a person is hard to work with or I have worked with that person before and he's difficult to work with... or on the other side it might happen that he's very easy to work with, very approachable, then, too their reputation precedes them. So when I need knowledge, I go to somebody about whom I already have preconceived ideas...about what they were like. It will definitely have an impact. If somebody is difficult to work with, then I might just say heck with it, I don't care...I will just figure this out on my own. I would say it's a critical thing. At the same time if there's something that is urgent, I might approach somebody and say "I really need this, I need some help here". And they might be willing or might be difficult to work with. But at least you might give it a chance. It depends on the urgency of what you need to do.	the section following table A2.1.		contacts. Implementing this phenomenon would add to the complexity of the model. However, it would also add to the richness of the simulation and help answer questions such as 1) how does a group's consensus on each of its member's *propensity to share* affect the accuracy of the group's TM 2) how does a group's consensus on each of its member's *propensity to share* affect the accuracy of the group's performance, as measured in the time taken to complete the set of projects assigned to the group. The phenomenon described earlier can be implemented as follows: each time an agent contacts one of the other agents for knowledge about another agent's areas of expertise, then in addition to receiving information about the third agent's areas of expertise, the recipient agent would also receive information (perceived by the source of such information) about the third agent's

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
				propensity to share. The implementation of the situation where a seeker agent contacts one of the other agents for information about other agents that might have the required knowledge, which it believes both itself and the agent that it has contacted lack, is beyond the scope of the current work.
j)	S: *OK. Have you come across instances where it turns out that a person knows more than what he/she lets others know about what he/she knows?* R: I've definitely seen that. There are people like that and their reputations are known as well. There are some people who do not want anyone to know anything about the systems they support because they are using that as a form of job security and/or some type of control.	Obtain support for the idea where an agent corrects its perception of a source upon discovering that the source has knowledge in a particular area when it was believed that the source does not have such knowledge.	Provides implicit and overall support to the inclusion of the parameter *perceived_propensity_to_share* – Subsection 3.2.1.8: 1, 2, 3	
k)	S: *Would you still go to such a person? If so, under what conditions?* R: Probably more. It depends upon how approachable and friendly they are. There are people who are doing that for a couple of different reasons. A large number of people are probably doing that out of insecurity. And if they are nice and they are insecure, I might still go to them and let them know that they do	Obtain support for the points m) in table A2.1 and h) in table A2.2.	Subsection 3.2.1.18.	The response provided here supports the idea that an agent, after revising its *perceived propensity to share* of a source agent, can potentially con-

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	not need to be threatened and then still try to get the information. If they are mean and unfriendly, then I might just stay away from them altogether. S: *OK. That makes sense.* R: Because otherwise, you go to them and you won't get anywhere and all you do is, you'll get upset and you've lost time.			tact that source, despite receiving incorrect knowledge in an earlier encounter.
l)	S: *How do you form your judgment of others' areas of expertise based on the accuracy of knowledge they share with you?* M: Based on how quickly they can answer my question or address my issues. That's one. But sometimes they might be very knowledgeable but not be able to answer my question very quickly. But if somebody is able to answer a lot of random questions very quickly, that tells me an awful lot, very quickly, about them. They obviously know what's going on here. Especially if I ask off-the-wall questions. It's kind of like sampling. It's my kind of sampling. S: *Do you think they are positively correlated? The more accurate their responses are, in your judgment, the more expert they are?* R: Yes. Absolutely. Absolutely. S: *Would you say the converse is also true? Would you say that because someone's answers are inaccurate, they have low expertise in the areas in which you question them?* R: I would think that it's true as well. I would think it works both ways.	Determine the process used by the interviewee to determine others' areas of knowledge	Subsection 3.2.1.4: 3 Subsection 3.2.1.8.	The interactions described in this set of responses can be interpreted as support for interactions that result in a) obtaining information about others' areas of knowledge and, b) modifying those perceptions during the pre-project phase and project phase
m)	S: *OK. This leads to the next question. How do you evaluate the accuracy of knowledge that others give you?* R: Well, I first see if it meets the need. That would probably be the best thing available. For example, if I approach somebody and say "I have this problem" and they say "there's this solution", and if they or those around them, or those around me say "Yes, that works", then that tells me pretty much everything I need to know. S: *Do you apply the knowledge that you receive and then determine whether it's accurate?* R: Yes, that's how I do it, yeah. S: *Is there any other basis to judge the accuracy of the knowledge you've been given?* R: Yeah, there is one another. That is to match it against my knowledge. If there is something that I	Understand how the accuracy of a source's knowledge is determined.	Subsection 3.2.1.4: 3 to the extent that the possibility of inaccurate knowledge being given by the source to the seeker is recognized. This response provides support to the notion that the accuracy of knowledge is determined	The description of how the inaccuracy of knowledge and its transmission will be modeled, as described in point n) in table A2.1 and in point h) in table A2.2 applies here. The description of accuracy of knowledge as being determined by whether it can be applied to solve the task is consistent with the interpretation in

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	think is going to solve a problem and if they give me the same one and if consider myself to be an expert on that then … if I approached them to verify what I know ...that would be another way to judge the accuracy – compare it with what I know and/or what others that I think are experts know. S: *So you look for consistency between what they tell you and what you know or between what they tell you and what someone else, whom you consider an expert and trust, knows...* R: Yes. Or look at it in a logical chain sense and see if it makes sense. And then then I might go and implement the solution.		based on whether the knowledge can be applied to solve the task at hand.	point n, table A2.1 and point h) in table A2.2. In addition to whether the knowledge provided was useful in completing a task, the accuracy of the knowledge received is determined by comparing it with what one knows and by comparing it with what "experts" know. For simplicity, neither of these two would be implemented, even though they are recognized as valid ways of determining the accuracy of knowledge received.
n)	S: *OK. In your judgment what factors affect a person's tendency to provide information about his or her areas of expertise?* R: I would say probably security is a big one; another one is ability to communicate. Security can have a lot of sub-categories. Security in terms of job security, control factor ...some people just like to have control. What's going on around them...if they give out too much, that can affect their security. S: *So, they want to make themselves indispensable?* R: Yes. Artificially. If somebody is a poor communicator, they might be hesitant to share information. That's another one. If somebody questions your authority, they may not share information. If they say "I don't even know if you work at this company, that could be a problem. If they are coming from another department, they may have information, but they might not want to share it with your department. That's another factor. Let me think … a tendency to avoid conflict might be one. Or just to have control. Or, here's a big one, fear of the unknown. If I don't know what's going happen...that ties back to understanding authorizations. Somebody approaches one individual, let's say individual A is approached by individual B and individual A says "I know you and I	To identify some of the factors that lead to a person's willingness or unwillingness to share knowledge with others.	Provide general support to the notion that inaccurate knowledge could be provided to the recipient of knowledge. Subsection 3.2.1.4: 3	The responses to the questions in this set support the arguments provided in point p) in table A2.1 and point k) in table A2.2.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	know you are in the company, but you know, I don't know if it's OK for me to share this information with you. There's a lot of such stuff that can go on in the corporate world. This thing called PIC – payment industry certification – and it's becoming more and more important not to share data if you don't have to. And it even happens here at L. You're not supposed to share data unless you absolutely have to. Just because it increases liability, so they are locking things down. That may be another driver of people's tendency to share or not share something. Another one is “projection of self”. Someone may not want to share information because that may make them look bad, or their department look bad, or their manager look bad or their team look bad. It could be anything, data or code. If they've written a bunch of code and you need to know something about their system, they may not want you to see their badly-written code, just because they may have had a legitimate reason but have had to throw it together quickly just to have a system up and running. They may not be proud of that. I've seen that in many places. S: *Just to avoid embarrassment?* R: Yes. I have thrown together some code quickly, and I've said to someone “I am not really proud of this code. We did this when we were really under the gun.” And I think I may have already said this, but I want to repeat this just to make sure, aside from that, maybe someone is really good at what they do anyway. So you bring in a developer who has no background, no experience and who can't even do their job. They may be really hesitant to share anything with you because they don't know what they're doing. And I've seen some people that are actually pretty good at what they do, but they may be a highly paid consultant. I've been a consultant. I've worked with other consultants from rival consulting firms. And at one consulting firm, we would notice a rival consulting firm. May be we were working on something in the same company and if we needed a solution, we would ask one of the consultants from the rival consulting firm, how to solve a problem and it evaded all of us and they were very reluctant to show their ignorance of the topic. There was a joke among our firms consultants that one of us should tell them that “it's OK to say that you don't know the solution. We understand that you don't know everything”. So there may be that factor as well – rival relationship between two firms – more so than the reluctance of one consultant towards showing his ignorance. He might be thinking that he's a representative of the firm and he			

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	wouldn't want to make the firm look bad.			
o)	S: *OK. Going on to the next question. Have you ever revised your impression of the honesty of someone regarding sharing knowledge and if you have, on what basis, under what conditions did you revise?* R: Yeah, I've definitely done that. Probably, if it is revised, it's probably done earlier than later in a relationship. It's probably because I misassessed somebody. I've evaluated somebody, had a preconceived idea or early idea of their ability. I've been dead wrong both ways – about somebody's inability and about somebody's ability. Early in the relationship, I may have thought that they are not good, but over time they have proven themselves as quite competent. I've seen the opposite to be true. This is seen in interviews. You interview somebody, you bring them in and you'd think "this person is going to do a great job" and they can't get anything done. They may have a great looking resume but cannot get a darned thing done. S: So, it's based on your direct experience? R: Yes. And I revise it based on the output of the person. Or it could be based on some third party that "this person is really great, he's struggling in this little area", or "this person isn't really very good".	Determine whether the interviewee revised his perception of the source's *propensity to share* and if so, under what conditions.	The response provides support to the description presented in the section following table A2.1. Subsection 3.2.1.8.	Supports the need for modeling a change in *perceived propensity to share*
p)	S: *OK. Going on to the next question: how does a person's tendency to share knowledge with others change as he/she interacts with others?* R: I'd go back to security. You know, somebody is secure in a relationship, that really opens up the channels a lot. Secure about themselves, secure about the relationship. If they understand that if I am trying to get something from somebody, and they know that I am not just out to get them, you know, make them look bad, it's a trust issue. This is probably the best way to describe it. S: *So as time goes by and trusts builds up, people would be more forthcoming...* R: Yes, some people are already willing to share any knowledge they have. I generally try to do that. When someone comes up to me and even if I don't know anything about how ...especially in companies where I work, I generally share whatever knowledge I have. If I get let go of the company because somebody else is more knowledgeable about the system, I don't care. I am confident I can find work somewhere else. S: *Has the reverse also appeared to be true? Somebody has been open initially and they, for some reas-*	To determine whether the parameter, which captures a source's *propensity to share* knowledge, is a relatively stable trait.		The response obtained is consistent with the response presented in point q) table A2.1, but is contradictory to the response presented in point m) table A2.2. Therefore, the assumption of constancy of *propensity to share* should be relaxed. As described in point q) table A2.1, *propensity to share* can be modeled as consisting of two components: a generalized propensity and a person-specific dyad-level propensity. The two components can be imple-

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	on, become less forthcoming as time goes by? R: Yeah, that can definitely happen. If a consultant might come on to the team and the consultant is really doing the best, and them make some mistakes, even if it is one or two small projects, if the manager – I have seen this happen – if the manager really comes down hard on them, they will be less likely to share information later. S: *So, even though, as a person, they are more open and honest, if extraneous factors, such as manager or somebody forces them to...* R: Or even a peer who says "how can you be so stupid?" they will close up a little bit. S: *The next question builds on the previous one. How is a person's tendency to accurately share knowledge change as he/she interacts with others? The previous one was about if they are sharing, how forthcoming they are, here it's about they are going to share knowledge with you but the knowledge can be accurate or inaccurate. That's done on purpose. So how does the tendency to be accurate or inaccurate....* R: So you are asking about misinformation, which is intentionally wrong. S: *Yes. In response to the previous questions you described how somebody would open up as trust builds, or somebody would close up because somebody was critical. So here, what I am asking about how somebody would be more accurate with what they tell you or less accurate, like misinformation.* R: Situations drive it one way or the other. One would be like the situation we just talked about, like where somebody is... something negative was said and that would drive things one was or the other. I had another one but it's eluding me...another example...more accurate...[pauses] give me just a minute...oh OK, here we go. This is done to make someone intentionally look bad. If you wanted to make somebody intentionally look bad, you'd share bad information with them intentionally and that would be one situation. On the flip side, you must really like someone, I've done this and I've seen this done...you want to make someone look better...I've done that...for example, I've received emails emails before where they've written wrong things and sent it out. And then I reply back to their email saying that we should go down this track and then correct what they put wrong in the original email chain to make them look better and they replied back saying, that was an error, thank you doing it, I appreciate it. Those would be the two situ-			mented as follows: at the time of initialization of the simulation, each agent is assigned a value of *propensity to share* – this is the generalized value. The values assigned to the agents are drawn from an N(avg_pts, avg_pts/3) distribution where avg_pts represents the workgroup-level value of generalized pts. This generalized tendency forms the basis for agent-agent *pts* values. Initially, for each given agent, the value of *propensity to share* is the same with respect of all the agents. However, over the course of interactions, the value of agent-agent *pts* changes. This change in *pts* can be modeled in the same way as the change in *perceived pts*, which is described in the section following table A2.1. However, despite the above explanation, to keep the implementation of the simulation reasonably simple, it is assumed that the value of *propensity_to_share* of an agent remains constant across a simulation

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	ations. S: *So it depends on the nature of the relationship between the source and the recipient.* R: I think so, yes. If it's a positive relationship then they'd probably try to be as accurate as possible. Even in a negative relationship you could still try to be as accurate as possible, but I am acknowledging that there are times when that isn't the case. S: *People tend to do that...* R: Oh yeah, share bad information, definitely, definitely happens. S: *It's based on the person's experiences that they do that?* R: Yes, it's probably is.			run.
q)	S: *I have gone over all the questions.* R: Oh, good. Good. S: *Thanks once again, R.* R: Sure. Anytime, Srikanth.	Conclusion of the interview.		

Table A2.4. Interview of J

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
a)	S: *I am here with J, VP of Sales at SSCC. Thank you, J, for giving me this opportunity.* J: You're welcome.	Introduction		
b)	S: *I'll begin with question # 0. Please describe a project in which you were involved recently.* J: The project was a property development project that involved the rehabilitation of 178 apartment buildings in the NSL region. It included tax-credits for affordable housing in an antique building. So essentially it was a large construction project to do a gut-rehab where we had to bring 178 apartments to a brand-new status.	Description of the project and the team's composition		
c)	S: *When you were working with other members of your workgroup, during the course of the project, what were the different media and technologies that you used for communication?* J: Well, if you think about it, we had to communicate all different levels from construction companies to state government to architects to staff, office person-	Identify the different types of IT used and the reasoning behind the choice of specific types of	Subsection 3.2.1.4: 4	Support for the comments made in d) of table A2.1, c) and d) of table A2.2,

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	nel, leasing agents and a property management company. So there were a lot of different parties involved in the workflow if you will. S: *OK. Did you have a specific type of technology that you used more often than the others, for instance, did you prefer face-to-face, or email, or video chat? Which are the different types of technology that you used?* J: Probably the majority, which would not be uncommon in any type of business, would be email, which we used on a mobile computer, a Blackberry, a desktop computer to get on to the Web at any convenient location which could be the State Capitol or downtown or whatever. For the most part it was email, lot of document exchanges via attachments.	IT under different circumstances		
d)	S: *Going on to next question, in literature, tacit knowledge is defined as knowledge that is difficult to articulate, to explain, to formalize to others. It includes things like intuitions, hunches, etc. When you had to collaborate with others and had to exchange tacit knowledge with others, did you choose a particular type of technology for communication?* J: In a lot of cases it was visual, so the technology would be...we would not be ashamed to do stand-up Powerpoint presentations, you know kind of illustrating our ideas, showing photographs of the current property etc. At other times we had to show the design via architectural drawing, which was also very visual, using AutoCAD systems and printing on those huge HP blueprint type printers. And that was very visual, so it was mostly via design software along with Powerpoint. S: *Were these used when you were communicating with clients or even among your group members?* J: It's a good question. Those were used both internally and externally. It could be a bank one day, politicians the next day, could be construction companies the following day. Internally, from an ideas standpoint, we would utilize the drawings and putting together our strategy via Powerpoint, trying to build support, and then taking it to the road. To build consensus you had to have unanimous support among all the stakeholders, so we used those tools.	The reasoning behind the choice of a specific medium of communication when the knowledge exchanged is "tacit"	Same as described in h) of table A2.1 and d) of table A2.2. Subsection 3.2.1.4: 4	The response provides new information: use of AutoCAD and Powerpoint to enhance the conveyance of knowledge. Since these two technologies were used when the medium of communication was face-to-face, for a simplicity perspective, they will be subsumed under the *ftof* medium in the simulation specification.
e)	S: *If you needed knowledge in a certain area for getting something done, whom did you contact and how did you contact a particular person?* J: Normally, as part of introductions and getting acquainted, when you are working together, you get a sense of roles and responsibilities. Of course, you	Obtain information about the process the interviewee used for developing	Subsection 3.2.1.18:1	The response to this question is consistent with the observation made in point e) table A2.2 about the

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	ask them outright, what they do for the company and for all intents and purposes, what that means to us and how we can work with that individual. Normally, it's based on experience. At a more tactical level, you need to know whom to engage. A lot of times executives volunteer that information... "if you need this, if you need that"... help you navigate through their organization as to whom you need to get in touch with and why, if you need to get something done.	an understanding of others' areas of knowledge		pre-project interaction phase described as "training phase" in the simulation specification.
f)	S: *OK. Could you please describe the process that you used to identify the areas of expertise of the members of your workgroup?* J: Yeah. The architect I worked with, the common denominator was that we both had lots of business experience and a lot of political or quasi-political influence, we were comfortable working in that space – high risk, high reward. Then individually we all had separate sets of experiences and expertise such as in my case, building team, building the organization, procedure flow, invoicing, accounts payable, accounts receivable, and then all the service incidents, service call flows, etc. They all came my way. One of my other partners was...he got his degree from Columbia University. He managed all the finances, all the performance. He took the ideas and broke them down into financial models to see if it was viable. Looked at different income sources and whatnot. And another key player was, still is, an architect with an architectural firm. The way it worked out was that we were all able to leverage core competencies as part of that partnership. We never really had any revenue streams except for the rent receipts and the deposits on the rent. And when the income sources came in from the tax credits, there was that. And we had an investment stream from a hedgefund investor. So we spent enough time lining up the different financial sources. I think we knew enough about each other to have a pretty good feel. There was a little conflict in the space of who was going to be the big boss in the office and who was going to direct the maintenance personnel and the leasing office. Everybody has different ideas about how that needs to be done and how people need to be dealt with but that was pretty much my responsibility. One of things about formulating partnerships is that once you have known a person for a good amount of time you have a feel for what they are going to bring to the table, their circles of influence, and how that ties back into how you are trying to accomplish your goals, creating a new business or making something brand new	Obtain information about the process the interviewee used for developing an understanding of others' areas of knowledge	Subsection 3.2.1.18: 1	The response to this question is consistent with the observation made in point e) table A2.2 about the pre-project interaction phase described as "training phase" in the simulation specification.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	happen, etc.			
g)	S: *How is your decision to seek knowledge from the member of a group affected by that person's tendency to make his knowledge public, that is make it known to other members of the group?* J: You know that is a pretty interesting question because it depends on the personality profile of an individual. Some folks are very guarded. They hold their, if you will, their copyright and their ideas on methodology and work-flows very close to their vests. They look at that as their copyright, as their patent and their space and don't really want to give access to that. If they do, it's probably in the form of a read-only document like a PDF or something and you cannot see the co-creation of may be an Excel spreadsheet or something like that. It might be highly sensitive or may be have some real thinking that went into that. Whereas at my end, I am more sales and marketing driven, so, I had to really ratchet that down a little bit. As far as property development was concerned, find out early that you don't want to talk a whole about the projects that you are working on due to the fact that you don't want someone to undermine your effort. S: *Even within your group?* J: Within a group...no. We pretty much trusted each other but believe it or not there was no transparency within the group as to who was touching what. There was like group interest and there was individual interest and there was a lot of conflict on both sides of that. S: *Would you describe your group to be very cohesive?* J: No. It was cohesive with respect to what were trying to get done. Other than that we were somewhat dysfunctional. We did not meet and agree very often...not nearly as often as we should have. At the end of the day we did pull it off, we were successful, but it could have gone much smoother if people were a bit more honest with each other and just basically answer questions in a timely fashion or been forthcoming with financials etc. That really eroded the trust and the bond between... S: *What was the size of your group?* J: There were five of us in total, of which three of us were active. Two parties were on the peripheral. They...one was with more political inclination and the other really outback...a gentleman that was man-	Obtain support for the construct *propensity to share* whose need and meaning are cited in table A2.1, point k and explained in the section following table A2.1.	Subsection 3.2.1.4: 7	Provides evidence to the existence of a variance in people's tendency to share knowledge with others, that is, to the inclusion of *propensity to share* in the simulation, as argued in point p), table A2.1. However, no evidence is available to either support or to invalidate the argument made for the relaxation of the assumption regarding the constancy of *propensity to share*, which was made in point p) of table A2.3. Hence, *propensity to share* is assumed to vary, as described in point p) of table A2.3.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	aging a hedgefund out in ZZZ. So, his visibility was more at a macro-level where you see the financials and debriefs and so forth about what we were doing. S: *But for project-related activities, it the group of the three of you?* J: Yeah, the other two were involved more in terms of decision-support and debate ideas and interaction, etc. But permanently, it was the three of us on the ground: the property-development individual, the architect and myself. I drove really all the lobbying effort and... again because you have to have unanimous agreement – state, city, local, municipality levels of government.			
h)	S: *OK. Moving to question # 6, how do you form your judgment of others' areas of expertise based on the accuracy of knowledge they share with you?* J: Well, you know, in the US it's a performance-based economy. So it's pretty easy to measure whether people are doing what they say they are going to do or they have the ability to do what they advertise they can do. So, normally, in most business projects with which I have been involved, things happen pretty rapidly and if people cannot do what they say they bring to the table, the whole team suffers because there's a break in the work-flow. Without really spending a lot of time analyzing where it broke, we kind of know in what camp things have really slowed down, back-logged. For instance, I am waiting for the drawings of the designs and I cannot finish my job until I have that product in hand. S: *OK. If they were to give you some knowledge that you didn't have, but you needed to get something done, how do you evaluate the person's expertise based on how accurate that knowledge was?* J: I guess it's subjective, on a day-to-day basis. One day you might be extremely happy with their performance and then you peel back the onion a couple of layers and you'd find that there were a couple of flawed variables in the judgment that went into the decision that you thought was brilliant two days before. So that's kind of hard to answer because you are ...people's credibility is like a roller-coaster at times … not a constant, whereas the other....i would say in the engineering space I would say that it was more reliable, constant and believable story that came out of that camp where they are dealing with absolute, things that are measured to the millimeter.	Understand how the accuracy of a source's knowledge is determined.	Subsection 3.2.1.4: 3; Subtask and Subsection 3.2.1.8: 3 to the extent that the possibility of inaccurate knowledge being given by the source to the seeker is recognized.	The description of how the inaccuracy of knowledge and its transmission will be modeled, as described in point n) in table A2.1 and in point h) in table A2.2 applies here.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	S: *If I understand you correctly, you judge the person's expertise based on whether the knowledge given to you is given to you is useful or not, whether it gets the job done or not.* J: Right, or if it is lopsided due to selfishness or some ulterior motive. Sometimes they hand you information that is convenient for them for their own selfish purpose so that's where it really gets complex because they are qualified, very qualified in that particular space to provide you with the right answer, but not being to able to tell whether that is the right answer ...becomes very difficult as far as judging character.			Agreement with the description of accuracy of knowledge as being determined by whether it was useful.
i)	S: *So how do you revise your judgment? When something happens or… on what basis do you revise your judgment of others?* J: Actions speak louder than words, so you really have to measure their results on their action side: yes or no did they do this, yes or no did they do what they said they were going to? Stay on plan once they've advertised what the plan would look like or did they deviate from the plan and not get consensus and go off and do a lot of things in a more of an independent mode... and probably violate the trust bond in the partnership. It gets really tricky believing everything you hear even from your own colleagues and partners. S: *So you revise your judgment based on whether it is consistent with your perception or whether it delivered the results?* J: Probably based more on the perception rather than being able to do a point-to-point comparison of what someone else may say in their same field who offered a separate opinion or tried to define it. You can really tell if someone is trying to be sneaky or deceptive. And this is kind of a gut skill... S: *That you develop over a period of time?* J: Yeah, over a length of time of being lied to. Most people don't do what they say they're going to do in business. That's the vast majority that do not. And if they do, that's not when they say they were going to deliver or at the same level in which they said that they were going to do. They alter their delivery and justify in their own mind as to why they modified it or shortened it or didn't include you in the communication or whatever the case might be.	Determine whether the interviewee revised his perception of the source's *propensity to share* and if so, under what conditions.		The role that the interviewee played had a greater proportion of "management/supervision" than obtaining knowledge to perform his own activities. In the light of this contextual information, revision of others' areas of expertise arises on a judgment of how well they performed their tasks and informed the interviewee that they completed their task. Hence, the accuracy of knowledge in this case is determined whether the interviewee's supervisees have indeed completed the tasks assigned to them to the satisfaction of the interviewee. That is, whether the interviewee can use this knowledge (of the tasks that he delegated as being

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
				completed) to perform his next task, which depended on his supervisees' completion of their assigned tasks. Hence, the response provided by the interviewee can be construed as providing support for the definition of accuracy of knowledge that was used in in the previous three interviews.
j)	S: *What factors do you think affect a person's tendency to share knowledge with others? Let me rephrase that. What factors do you think affect a person's tendency to let others know that she/he is knowledgeable in, say, X number of areas? Let's say, for instance, I am knowledgeable in five areas, but I choose to reveal that I know things in only three areas, what do you think made me behave that way, as in withhold knowledge?* J: Because it's capital, it's your brain-trust. That's your monopoly. What you bring to the table. I have found over time in business that a lot of times business people look at those thoughts, ideas, different documents...they are little copyright on how things need to be done, so they are very reluctant to...well the other part is that most people try to size up their audience … if I told you all five areas, chances are you'd only get three, and the other two don't really apply or if they do apply, you're going to use those areas of expertise for your own self and not share.	To identify some of the factors that lead to a person's willingness or unwillingness to share knowledge with others.		The response provides support to the notion that a source of knowledge might not be willing to let potential knowledge-seekers know about all the areas in which the source is knowledgeable. Hence, this provides support for the inclusion of the parameter *propensity to share.*
k)	S: *OK. On what basis would you revise, if you were to revise, your judgment of others' honesty?* J: Character is one of the hardest things in the world to define. Normally, character really surfaces after six months or a year after you've known somebody. So over a length of time, their real self will come forward, and whatever they advertise up front fades, so it's really hard to maintain over a length of time, unless you are actually doing the work and you are who you say you are and so forth. S: *Do you think your answer also applies to honesty with respect to sharing knowledge?*	Determine whether the interviewee revised his perception of the source's propensity to share and if so, under what conditions.	The response provides support to the description presented in the section following table A2.1. Subsection 3.2.1.8.	Supports the need for modeling a change in *perceived propensity to share*

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	R: Probably. Again, it's ...a lot of business people have been burned in a lot of different ways, so, they hate to over-share their knowledge. They taught somebody the ropes, shown them the least path, real trade secrets, contacts, whatever they could ..and those could be actually used in the future to compete against them. So, for the most part, senior business people are pretty guarded about exactly their reach, all the things they can do and are willing to do.			
l)	S: *How does a person's tendency to share knowledge with others, change as he/she interacts with others in his/her group over a period of time?* J: Hmm...obviously you get annoyed with the BS of not getting concise answers to closed-end questions where they ramble on aimlessly and you get frustrated in a way, get short with them, rude and cut them off and say "I am tired of the BS and why can't you just tell the truth?" also a lot of candor has to pop up and a little bit of courage, you know...just call somebody out. But that's not every project. Sometimes partnerships come together that appear to be great and appear to be not so good. S: *OK. Can you think of reason why someone is less honest initially would turn out to be more honest?* J: Probably because their honesty is being closely inspected by a couple of smart people that will basically say "the other day you said this and now you're saying this. Which is it?" And when their support starts eroding and they start losing momentum as far as their partnership with the workgroup is concerned, then people have a tendency to be less-than-awful and start coming clean more often, participating and being more honest. S: *Do you think people can also go the other way, that is, they are more honest initially...* J: Sure. It could be from a slacker mentality. May be they were not working that hard, may be they perceived everyone in the group as being beneath them, may be not as smart as them, may they started out a little strong but... S: *So even when sharing knowledge, do you think initially they are very honest and open about sharing knowledge but over a period of time, their tendency to honestly share knowledge goes down?* J: Yeah, I mean part of it is just "do they really need to know?" "do we really need to meet?" kind of just talk about stuff that's really low impact. Some of it is	To determine whether the parameter, which captures a source's propensity to share knowledge, is a relatively stable trait or if it changes		The argument made in point p) of table A2.3 regarding the relaxation of the assumption pertaining to the constancy of *propensity to share* is supported. Note: Using the criterion of keeping the implementation of the simulation relatively simple, the parameter *propensity_to_share* is assumed to remain constant for the duration of a simulation run.

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	just convenience. The other of it is just kind of like political lines being drawn within a small company or a small group. They may tell one of your colleagues but not you to kind of pit two guys against each other. I've seen things like that happen in business. S: *OK. Let's say two people have been very open with each other, sharing knowledge. Do you see the honesty in their knowledge-sharing relationship go down over a period of time?* J: Sometimes not. I've got a lot of colleagues that I am brutally honest with and they're brutally honest with me back. That's one of the reasons we are that close. The reason we enjoy working together is due to the fact that we can trust each other and there's not a lot of overhead required to think about things that are out of your control any how. S: *But would such honesty with respect to sharing knowledge go down? Has it ever been the case? And if so, what do you think was the cause?* J: If it's been, it's because of some other personal distractions popping up and little areas of conflict where you choose not to participate in a meeting or an event for any particular reason. Getting along with people on a personal and professional basis is always a challenge. Not everybody sees something the same way...that's normally the rub. S: *OK. Question # 13 is somewhat similar but not quite the same. Based on your experiences, do you think, a person's tendency to accurately share knowledge as in giving the right type of knowledge, would that tendency go down? Earlier it was about being honest, that is, for instance saying that I don't know something when in fact I do. This questions is about saying that I do know this but I give you the wrong information instead of the right one, i.e. deliberately misleading.* J: Oh, yeah, I can see when he/she interacts with others...and again that's the company politics where they are trying to perhaps look good themselves by undermining some co-worker or someone else who's involved with the project thus making them look lacking. Because they are compelling enough to get everybody else in the group with whom they are interacting to focus on the shortcomings of the particular individual. That would obviously change with the person being in the room or may be there in a certain interaction with a couple of people but when a third person comes into the room, they take a different posture, with the new person in the room.			

	Question/Response	Purpose	Specification item for which support provided	Comments/ changes needed to the specification
	S: *Let's say there are two people, A and B. If B was forthcoming with A, but A turned out to give inaccurate information to B, do you think that would lead B to provide inaccurate information to A in the future?* J: Again, if it was something they were trying to do selfishly for themselves, and mislead B who was operating in good faith, if I understand you correctly, normally, there's some selfish reason behind doing that to have them take their eye off the ball, or distract them from some shenanigans that they're up to. Always hard to tell but... S: *Do you think it's also a function of atmosphere or culture in the group? If the folks are not very trusting among themselves, that would also affect the one-one relationships between various group members?* J: Absolutely. You would get the odds of normal people being more open with each than, may be, engineering or finance type people. If you build a culture of certain like-minded individuals, there's a high likelihood that the whole team effort, the whole entrepreneurial spirit will come to the forefront, and they're going to put their best effort forward and really contribute to the team and try to make something special. Or, in a larger company, it's really hard for such things to take place because they're departmentalized and their efforts are really hard to...the employees are going to have a really difficult time trying to understand how their efforts fold into the big picture and whether or not it really makes a difference. S: *So that would also lead them to pull back...* J: Sure. Again you would find in a lot of businesses that people like to act real busy and that's just a busy smokescreen. If you get right down to it, they are failing at their job or failing at their roles with regards to what they are asked to do for the company, or partnership for that manner.			
m)	S: *OK. That completes the set of questions. Do you have any recommendations to improve the questions? Were they clear?* J: Yes, I think so.			

Appendix A3. Follow-up interviews

This appendix provides details of the data obtained from the follow-up interviews. The goals of the follow-up interviews were to:

(a) obtain information about the interviewee's tendency to use a communication medium under various situations provided in the questionnaire, and

(b) understand the rationale the interviewees used while making a choice of a communication medium under the conditions provided in the questionnaire.

Follow-up interview of M

The following questions seek information on your choice of various communication media for exchanging knowledge with members of your workgroup

1. If you are communicating with a workgroup member who is at the same location as you, how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call			x		
Via a chatroom/video conferencing software				x	
Via email		x			
Via instant messaging (IM)	x				

Please provide a short description of your rationale for the above response

Many of my immediate work group members are located on the same floor, and it is sometimes easier to just walk over to their cubicles to communicate. I added instant-messaging (IM) as a method because I view this as a separate technology from chatrooms and video conferencing. If I cannot easily communicate with someone face-to-face, I will try IM as the next best choice because of its immediacy. I believe I mentioned that we spend lots of time on conference calls, and IM offers an ability to get a quick answer while someone is otherwise busy on a call. I try to avoid email where possible because our email inboxes are usually overwhelmed with message traffic, unless it is important to maintain a record of the communication.

2. If you are communicating with a workgroup member who is at a different geographic location (not in your office building), how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face					x
Via a phone call		x			
Via a chatroom/video conferencing software		x			
Via email	x				
Via instant messaging (IM)	x				

Please provide a short description of your rationale for the above response
The IM tool we currently use (Microsoft Communicator) provides real-time visibility into a correspondent's calendar availability, and is usually the fastest way to reach someone offsite. Given that we're all constantly on conference calls, it's rare to call someone on the phone and actually have them answer. Email is a good second choice to IM, but again the inbox clutter becomes a deterrent. When I do send an email, I often identify a specific action required in the subject line, and use the Microsoft Outlook Follow-up feature to set a reminder for the recipient to ensure he/she is prompted to act on my message. When I do resort to a phone call, I will often use a web meeting facility (video conferencing software) such as Windows NetMeeting or Saba Centra to review a supporting document.

3. If you are communicating with a workgroup member on a task of medium priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face			x		
Via a phone call			x		
Via a chatroom/video conferencing software					x
Via email	x				
Via instant messaging (IM)		x			

Please provide a short description of your rationale for the above response
When immediacy is of lower importance, I am usually more inclined to send email instead of IM, especially if Microsoft Communicator shows my correspondent's status is 'away' or 'offline'.

4. If you are communicating with a workgroup member on a task of high priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software			x		
Via email	x				
Via instant messaging (IM)	x				

Please provide a short description of your rationale for the above response
Immediacy is very important for high priority activities. IM provides the most immediate way to contact someone, so unless Microsoft Communicator shows my correspondent's status is 'away' or 'offline', I'll use IM first. If the correspondent is not available on Communicator, then I'll resort to a phone call, then an email for an offsite correspondent, or face-to-face, then an email if the correspondent is onsite. When I do resort to a phone call, I will often use a web meeting facility (video conferencing software) such as Windows NetMeeting or Saba Centra to review a supporting document.

5. If you are communicating with a workgroup member on a task of low priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face					x
Via a phone call					x
Via a chatroom/video conferencing software					x
Via email	x				
Via instant messaging (IM)			x		

Please provide a short description of your rationale for the above response
When immediacy is not important, I am most inclined to send email instead of IM, especially if Microsoft Communicator shows my correspondent's status is 'away' or 'offline'. Face-to-face and phone calls are overkill if there is no urgency. By its non-interruptive nature, email best fits the immediacy requirements for low priority communications.

6. If you are communicating with a workgroup member to exchange knowledge that is highly tacit (difficult to articulate, verbalize and convey), how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software		x			
Via email		x			
Via instant messaging (IM)					x

Please provide a short description of your rationale for the above response
Tacit knowledge exchange requires multiple channels for success, in my view. If face-to-face is a possibility, I'll likely meet in a conference room with a white board, where I will draw diagrams or share a presentation deck to support the knowledge exchange. For a phone call, I'll likely have either sent a document by email, or will use a web meeting facility (video conferencing software) such as Windows NetMeeting or Saba Centra to review a supporting document. IM is of limited use in this context because of its transient nature.

7. If you are communicating with a workgroup member to exchange knowledge that is easy to articulate, verbalize and convey, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face				x	
Via a phone call			x		
Via a chatroom/video conferencing software					x
Via email	x				

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Via instant messaging (IM)	x				

Please provide a short description of your rationale for the above response
IM followed by email are great vehicles for communicating this type of information. The choice between the two would depend on any perceived need for persistence of the information. If the information is transient in nature, then IM is the best choice. If there is need for the information to be persisted for later reference, then email makes more sense.

Follow-up interview of K

The following questions seek information on your choice of various communication media for exchanging knowledge with members of your workgroup

1. If you are communicating with a workgroup member who is at the same location as you, how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call		x			
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
It's fastest to speak face-to-face followed by email.

2. If you are communicating with a workgroup member who is at a different geographic location (not in your office building), how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face					x
Via a phone call	x				
Via a chatroom/video conferencing software			x		
Via email	x				

Please provide a short description of your rationale for the above response
Phone is fastest, followed by email for me. If video conferencing is available, I would use it.

3. If you are communicating with a workgroup member on a task of medium priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Whichever way is fastest for that day.

4. If you are communicating with a workgroup member on a task of high priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Same as above.

5. If you are communicating with a workgroup member on a task of low priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call		x			
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Same as above.

6. If you are communicating with a workgroup member to exchange knowledge that is highly tacit (difficult to articulate, verbalize and convey), how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call		x			
Via a chatroom/video conferencing software					x
Via email		x			

Please provide a short description of your rationale for the above response
Face-to-face with a white board would be fastest.

7. If you are communicating with a workgroup member to exchange knowledge that is easy to articulate, verbalize and convey, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software					x

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Via email	x				

Please provide a short description of your rationale for the above response
Just the fastest way to get things done.

Follow-up Interview of R

The following questions seek information on your choice of various communication media for exchanging knowledge with members of your workgroup

1. If you are communicating with a workgroup member who is at the same location as you, how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call					x
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
I choose the face-to-face means when something is urgent or important. In addition, I sometimes choose face-to-face when I simply want to see someone (i.e., friendship). Phone calls and chatrooms/video conferencing seem to be more trouble than they're worth: it's easier to use the other media. Email is one of my primary means of communication, due to its asynchronous ability.

2. If you are communicating with a workgroup member who is at a different geographic location (not in your office building), how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face					x
Via a phone call		x			
Via a chatroom/video conferencing software			x		
Via email	x				

Please provide a short description of your rationale for the above response
Email is my first means of communication, due to the asynchronous ability of that medium. I'm unlikely to setup a face-to-face meeting, simply because I'm required to be in my office most of the time (i.e., when I'm not teaching). I don't use chatrooms/video conferencing software simply because it's not readily available (i.e., it's more trouble than it's worth). Phone calls are the most convenient means to speaking with someone over long distances (2^{nd} only to email).

3. If you are communicating with a workgroup member on a task of medium priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face			x		
Via a phone call				x	
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Email is my first choice, unless I simply want to pay somebody a visit (i.e., face-to-face due to friendship). All other media are much less important.

4. If you are communicating with a workgroup member on a task of high priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
For high-priority items, I use all media except one. I use whatever it takes to communication the most information in the shortest amount of time. I simply don't have chatroom/video conferencing readily available, or I might use it. It's more appropriate for me to say that *urgency* drives the means of communication. Higher urgency leads me to media richness.

5. If you are communicating with a workgroup member on a task of low priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call					x
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
For low-priority items, I tend to use email due to the asynchronous capability (i.e., convenience), unless I simply want to visit a coworker who is a friend.

6. If you are communicating with a workgroup member to exchange knowledge that is highly tacit (difficult to articulate, verbalize and convey), how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Via a phone call		x			
Via a chatroom/video conferencing software					x
Via email				x	

Please provide a short description of your rationale for the above response
If at all possible, I will use the most media-rich environment available (i..e, face-to-face). If face-to-face is unavailable, then I will most-likely use the phone. If I use email, it would be to setup a face-to-face meeting or a phone call.

7. If you are communicating with a workgroup member to exchange knowledge that is easy to articulate, verbalize and convey, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face		x			
Via a phone call					x
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Email is my first choice, unless I simply want to pay somebody a visit (i.e., face-to-face due to friendship). All other media are much less important.

Follow-up Interview of J

The following questions seek information on your choice of various communication media for exchanging knowledge with members of your workgroup

1. If you are communicating with a workgroup member who is at the same location as you, how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call		x			
Via a chatroom/video conferencing software				x	
Via email		x			

Please provide a short description of your rationale for the above response
First, go and try to find the person, second, give them a quick call if you fail to find the person, third, shoot them an email/text that will end up in the PDA/phone.

2. If you are communicating with a workgroup member who is at a different geographic location (not in your office building), how likely are you to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face			x		

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Via a phone call	x				
Via a chatroom/video conferencing software			x		
Via email	x				

Please provide a short description of your rationale for the above response
Probably need a quick question answered or status regarding a project or proposal, etc.

3. If you are communicating with a workgroup member on a task of medium priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face		x			
Via a phone call	x				
Via a chatroom/video conferencing software			x		
Via email		x			

Please provide a short description of your rationale for the above response
Medium priority drove the need to probably meet face-to-face preferably. Then most immediate form of communication afterward.

4. If you are communicating with a workgroup member on a task of high priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call	x				
Via a chatroom/video conferencing software		x			
Via email	x				

Please provide a short description of your rationale for the above response
Whatever it takes, if it's very important to have near-term discussion.

5. If you are communicating with a workgroup member on a task of low priority, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face			x		
Via a phone call		x			
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Being a lower priority, typically it's not as important to have immediate interaction. Probably can wait for an email response.

6. If you are communicating with a workgroup member to exchange knowledge that is highly tacit (difficult to articulate, verbalize and convey), how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face	x				
Via a phone call				x	
Via a chatroom/video conferencing software		x			
Via email					x

Please provide a short description of your rationale for the above response
Face-to-face would be the best choice. It's important to witness body language towards judging whether or not they are getting the points. Visual illustrations always help so could utilize video conferencing.

7. If you are communicating with a workgroup member to exchange knowledge that is easy to articulate, verbalize and convey, how likely are you to to communicate

	Very Likely	Somewhat Likely	Neutral	Somewhat Unlikely	Very Unlikely
Face-to-face				x	
Via a phone call		x			
Via a chatroom/video conferencing software					x
Via email	x				

Please provide a short description of your rationale for the above response
Should not require any major thought for questions to be asked, so make a call to simply tell someone something or send an email so that it's documented is typically what is done.

Appendix A4. Additional aspects related to the implementation of the model

Table A.4.1 provides the results of the t-tests related to the validation of the simulation. Then, the implications of these results in terms of their support for the validity of the implementation of the model are discussed.

Table A4.1 Results of Welch's t-tests[45]

Contingency factor	Average consensus	Average accuracy of knowledge	Percentage of project completed
Average knowledge level		-376.08	-34.7
Average openness to change	-4.12		3.38
Average self-knowledge			-53.91
Probability of turnover	3.55	-9.38	-3.42
Average proportion of knowledge areas common with the replaced agent	3.24	2.04	
Number of agents	-37.98		-6.55
Number of locations	-22.54	-2.04	
Probability of exchange of information about a non-task-specific-knowledge area	-1219.66	-7.81	
Time in training phase	-39.85		-10.77
Average propensity to share	-62.31		
Maximum number of failed tries	-5.65	6.63	-4.19
Average task intensity			
Average task priority			
Average knowledge intensity of subtasks			6.67
Average project intensity			
Average number of tasks per agent	8.97		
Number of projects per workgroup			
Similarity of projects			
Connectedness of network of task-interdependencies			
Average direction time			
Average stickiness time			-2.53
Use expert-seeker			
Telephone			
Email	254.52	-193.34	-8.91
Text-based chat	257.68	-195.84	-8.96
Face-to-face			

45 The t-values are reported in the table. All t-values are significant at $p < 0.001$ level.

Average knowledge level

Those workgroups that have higher initial levels of accurate knowledge are more likely to have higher final values of accuracy of knowledge and also more likely to complete a greater percentage of tasks assigned to them. The results are consistent with this expectation.

Average openness to change

Those workgroups whose members are more open to changing their views about each other's areas of knowledge are likely to have greater consensus about each other's areas of knowledge. They are also more likely to find sources of knowledge when needed, thereby completing a greater percentage of tasks assigned to them. The results are consistent with this expectation.

Average self-knowledge

This contingency factor captures the notion of correctness of perceived self-knowledge. However, the proportion of areas about which an agent has correct perceptions of its knowledge is distinct from the proportion of such knowledge areas in which the agent has accurate knowledge. Hence, in the case of each agent, seeking help from others, and obtaining correct knowledge from others are dependent on both the number of areas in which it has correct knowledge and the number of areas in which its perceptions of its knowledge are accurate. Therefore, it is difficult, based on intuition and reasoning, to predict how the workgroups belonging to the 'maximum' and 'minimum' categories for this contingency factor would differ on each of the three outcome variables.

Probability of turnover

In those workgroups where there is a greater likelihood of turnover, members are less likely to reach an accurate consensus about each other's areas of knowledge. They are also less likely to be able to find the most reliable source of knowledge they need and thus are less likely to complete the tasks assigned to them, since they are less likely to obtain the knowledge they need. Consequently, the group differences with respect to each outcome variable are expected to be significant. The results are consistent with the above-described expectations.

Average proportion of knowledge areas common with the replaced agent

Based on intuition and reasoning, it is hard to predict whether higher number of matches between a replacing and a replaced agent's areas of knowledge would lead to significant differences with respect to each of the three outcome variables. This is so because, the effect of this factor on the knowledge transmissions occurring in a workgroup, is tied to the effect of the probability of turnover. Hence, no predictions are made about differences between the “maximum” and “minimum” workgroups with respect to each of the three outcome variables.

Number of agents

Greater number of agents would result in a significantly more number of potential interaction combinations, thereby increasing the likelihood that in such groups, members are less likely to know about each other's areas of knowledge. Its consequence is that members of such workgroups are less likely to obtain the knowledge they need, leading to lower values of average accuracy of knowledge and percentage of project completed. The results are consistent with the above-described expectations.

Number of locations

More locations imply a greater likelihood of usage of electronic media, thereby increasing the time taken to transmit knowledge. With an increase in the transmission time, the likelihood of exchange of information about non-task-related knowledge increases correspondingly. Therefore, this factor is expected to affect significantly all three outcomes. The results indicate consistency in two out of three cases, thus supporting the expectations.

Probability of exchange of information about a non-task-specific-knowledge area

In workgroups where the likelihood of exchange of information about non-task-specific knowledge is high, members are more likely to exchange information about each others areas of knowledge. Consequences of such exchanges are (a) an increase in the aggregate knowledge level of the workgroup, and, (b) greater likelihood of completing a higher proportion of the project's tasks. Hence workgroups are supposed to differ significantly on all three outcomes. The results are consistent in the case of two out of the three expected outcomes.

Time in training phase

Greater the amount of time spent in the training phase, greater is the likelihood that members of such workgroups would learn about each others areas of knowledge. Additionally, in such workgroups, members are more likely to find reliable sources of knowledge and complete the tasks assigned to them. Hence, workgroups are expected to differ significantly with respect to this contingency factor on all three outcome variables. The results are consistent with the above-described expectations in two out of three cases.

Average propensity to share

In workgroups where members are more likely to share knowledge and information about the areas in which they have knowledge, members are more likely to learn about each other's areas of knowledge, and consequently have a greater aggregate level of accurate knowledge and complete a greater proportion of their tasks. The results indicate consistency in one of the three cases.

Maximum number of failed tries

In workgroups where members are allowed to try to obtain knowledge for a relatively a greater number times, they are more likely to obtain the needed knowledge, learn about other members' areas of knowledge and complete a greater proportion of tasks assigned to them. Hence workgroups are expected to differ significantly with respect to this contingency factor, on all three outcomes. The results are consistent with the expectations.

Average task intensity, average task priority, average knowledge intensity of subtasks, average project intensity, similarity of projects and *connectedness of network of task-interdependencies*

These contingency factors were included to determine whether workgroups differ on the outcome variables, with respect to these factor, without having any prior expectations of differences. This is because few studies exist that used the same, or even similar, operationalizations of these contingency factors. Hence, the results of these factors' t-tests are not considered in the validation process.

Average number of tasks per agent

As the workload on each member of a workgroup increases, it is expected that the each member is less likely to have the knowledge needed and therefore is more likely to contact other agents, thereby increasing the likelihood of learning about others' areas of knowledge. However, it is harder to argue that the increased workload would lead to a greater sharing of knowledge, hence, while contacts will be made, seeking knowledge, such contacts may not lead to actual transmissions of knowledge. Thus, it is expected that the in workgroups with higher number of tasks per agent, the values of average consensus would be higher. However, the values of the other two outcomes may not necessarily be higher. The results support the above-described expectations

Number of projects per workgroup and *richness of the face-to-face communication medium*

Both these values were fixed across all workgroups and hence were not part of the validation process.

Average direction time and average stickiness time

These contingency factors are inter-related: in the model, an agent would choose either transfer or direction as the mode of transmission based on the specific values of these two parameters associated with the specific knowledge area. Hence, the specific effect of each of the contingency factors cannot be delineated, independent of the effect of the other. Therefore, it cannot be stated with certainty that workgroups with high values would differ significantly from workgroup with low values, for each of the two contingency factors, on all three outcome variables. However, it is expected workgroups in the “maximum” category would differ significantly from workgroups in the “minimum” category with respect to the effect of at least one of the two contingency factors on at least one of the three outcome variables. The results support the above-described expectation.

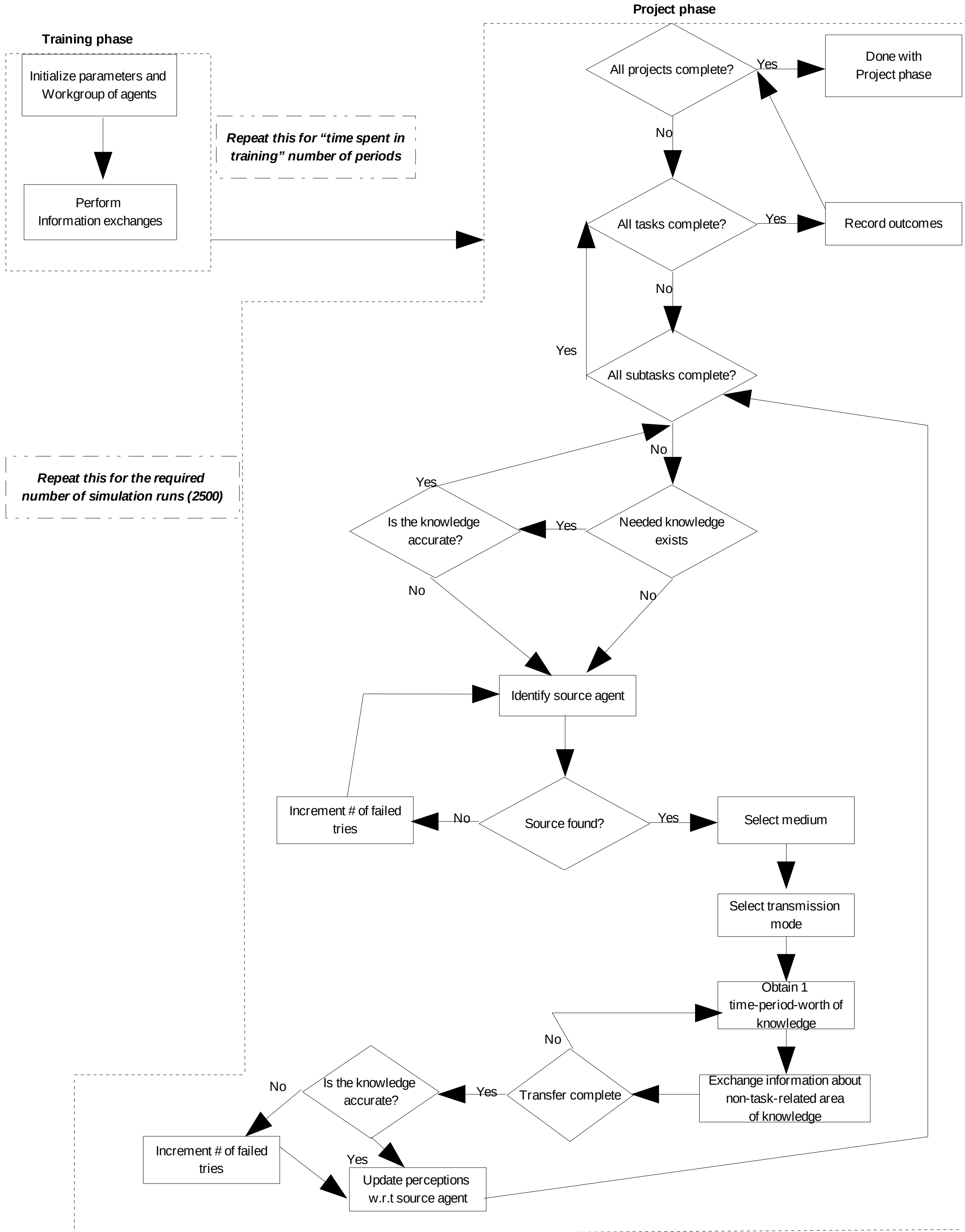
Richness of telephone, richness of email and *richness of text-based chat*

The three electronic communication media were included in the model based on prior literature. It was expected that, since there is a non-zero probability of the agents in the workgroups using electronic media while communicating, workgroups corresponding to the “maximum” value group would differ significantly from workgroups corresponding to the “minimum” value group in the case of the effect of at least one of the three contingency factors on at least one of the outcome variables. However, since the perceived richness of each medium changes over time, as a function of other contingency factors' values (which themselves change as a function of specific sets of other contingency factors) and workgroup processes, and it is hard to expect with high certainty the specific workgroup differences in terms of the effects of each of the contingency factors on the outcome variables. Hence, the only reasonable expectation is that the “maximum” and “minimum” groups will differ in the case of at least one of the three contingency factors associated with the richness of the electronic media in terms of their effects on each of the three outcomes. The results indicate that workgroups belonging to “maximum” and “minimum” groups differ in the case two of the three electronic media, in terms of the effects of the contingency factors on the outcome variables, implying a consistency with the expectations described above.

Use expert-seeker

The “expert-seeker” mechanism of locating sources of knowledge was included in the model based on prior literature. Hence, it was expected that workgroups in the “used” category would differ significantly from workgroups in the “did not use” category in the case of at least one of the three outcome variables. The results indicate that the above-described expectation was not met. However, data from the analysis of log files indicated that there have indeed been workgroups in the simulation experiments representing both conditions. Hence, while the results of the t-test do not provide support, evidence from the log files indicates that both conditions (“used” and “not used”) have occurred in the simulation runs.

Appendix A5. Flow-chart of the simulation



Appendix A6. Source-code of the implementation of the simulation

```
import math
from random import random, randint
import traceback
from diss_sim_simrun_v11_11 import SimRun
from diss_sim_helpers import _get_random, get_positive_avg
import psyco
psyco.full()

class Supervisor:
    """
    Controls the main set of processes that deal with
    (a) initialization of the simulation environment
    (b) running the simulation
    (c) recording the output
    """

    def __init__(self):
        """
        Initializes the supervisor object
        """
        self.dict_comm_media_richness =  {"ftof": {"synch": "yes",
                                                  "richness": 1.0,
                                                  "locality": "yes" },
                                           "phone": {"synch": "yes",
                                                   "richness": 0.45,
                                                   "locality": "maybe" },
                                           "chat": {"synch": "yes",
                                                  "richness": 0.25,
                                                  "locality": "maybe" },
                                           "email": {"synch": "maybe",
                                                   "richness": 0.25,
                                                  "locality": "maybe" } }

    def get_media_richness_vals(self):
        """
        Return a dictionary with randomly-determined values for richness of
        various communication media.
        """
        d = self.dict_comm_media_richness
        d["phone"]["richness"] = _get_random(d["phone"]["richness"])
        d["chat"]["richness"] = _get_random(d["chat"]["richness"])
        d["email"]["richness"] = _get_random(d["email"]["richness"])
        return d


    def initialize_run_parameters(self):
        """
        Initializes parameters for a single run. Yields a dictionary.
        """
        avg_turnover = random() # [0, 1)

        min_projs_run, max_projs_run = 10, 10 ##1, 10
        no_projects_per_run = randint(min_projs_run, max_projs_run)
        rng_task_priority = (1, 5)
        rng_know_intensity_stask = (1, 5)
        similarity_projects = random() #allow for zero similarity!
        min_proj_intensity, max_proj_intensity = 1, 12## 10, 30
        rng_proj_intensity = (min_proj_intensity, max_proj_intensity)
```

```
min_task_intensity, max_task_intensity = 1, 5
rng_task_intensity = (min_task_intensity, max_task_intensity)
# determine the size of organizational repertoire of subtasks
size_org_rep_subtasks = (int(math.ceil(2 - similarity_projects)) *
                         max_task_intensity *
                         max_proj_intensity)

## size_org_rep_kas = size_org_rep_subtasks * max_know_intensity
time_training_phase = randint(1, 10)

rng_no_locations = (1, 5)
dict_comm_media_richness = self.get_media_richness_vals()

min_no_agents, max_no_agents = 2, 10 ##2, 15
no_agents = randint(min_no_agents, max_no_agents)

use_expert_seeker = random() < 0.5

## avg. openness of an agent to change its perceptions of other agents'
## pts, KAs, etc.
avg_openness_change = get_positive_avg()
## interactions needed by one agent to know about other agent's KAs
rng_num_interactions = (1, 3)
# avg. amt. of KAs common to the replacing agent and original agent
# when turnover is executed
avg_common_know_repl = randint(1, 9)/10.0 ##get_positive_avg()
# avg. amt. of knowledge carried to a successive project
avg_carryover_exp_lvl = 0##random()
# avg. of initial amount of self knowledge
avg_self_knowledge = randint(1, 9)/10.0##get_positive_avg()
# agent's potential to share; remains constant during a sim. run
avg_generalized_pts = get_positive_avg()
# the average amt. of knowledge possessed by each agent
# same as saying the avg. proportion of beliefs held that are correct
avg_knowledge_level = randint(1, 9)/10.0 ##(1,9)

def get_incorrect_k_level():
    """
    Returns a value for proportion of knowledge areas that are
    incorrect
    """
    while True:
        v = random()
        if 0 < v < 1 - avg_knowledge_level:
            return v
## indicates, what proportion of beliefs held by an agent are incorrect
avg_incorrect_know_level = get_incorrect_k_level()

## parameter that determines exchange of information about
## non-task-specific knowledge
p_nonspec_exchange = _get_random(0.5)

## assign an upper bound on the number of failed tries for finding a
## knowledge source

max_failed_tries = randint(6, 10)##(5, 15) #1, 5
rng_half_life = (5, 10)
rng_stickiness_time = (1, 10) ## (5, 10)
rng_direction_time = (1, 10) ## (5, 10)

## While creating a DAG of the tasks, initially, a digraph with
## cycles is created. The number of edges for a given number of nodes
```

```
        ## is needed for the initial version of the graph (with cycles). This
        ## value (essentially graph-density), is determined by a parameter.
        ## The range of this value is [0, max_no_edges], which, for 'e'
        ## edges is e*(e-1)/2. Given that each project's tasks are determined
        ## after the simulation has begun, the number of edges between the
        ## nodes representing tasks in a graph will also be determined at
        ## run-time.

        ## a full-factorial design will not be used; instead, data would
        ## be sampled from the domain of each parameter's value
        ## create a dictionary with the various parameters and return it

        return {"avg_carryover_exp_lvl": avg_carryover_exp_lvl,
                "avg_common_know_repl": avg_common_know_repl,
                "rng_direction_time": rng_direction_time,
                "rng_half_life": rng_half_life,
                "avg_incorrect_know_level": avg_incorrect_know_level,
                "avg_knowledge_level": avg_knowledge_level,
                "rng_num_interactions": rng_num_interactions,
                "avg_openness_change": avg_openness_change,
                "avg_self_knowledge": avg_self_knowledge,
                "rng_stickiness_time": rng_stickiness_time,
                "dict_comm_media_richness": dict_comm_media_richness,
                "avg_generalized_pts": avg_generalized_pts,
                "rng_know_intensity_stask": rng_know_intensity_stask,
                "no_projects_per_run": no_projects_per_run,
                "rng_proj_intensity": rng_proj_intensity,
                "rng_task_intensity": rng_task_intensity,
                "no_agents": no_agents,
                "rng_no_locations": rng_no_locations,
                "p_nonspec_exchange": p_nonspec_exchange,
                "rng_task_priority": rng_task_priority,
                "similarity_projects": similarity_projects,
                "size_org_rep_subtasks": size_org_rep_subtasks,
                ## "size_org_rep_kas": size_org_rep_kas,
                "max_failed_tries": max_failed_tries,
                ## "time_project_phase": time_project_phase,
                "time_training_phase": time_training_phase,
                "avg_turbulence": avg_turbulence,
                "avg_turnover": avg_turnover,
                "use_expert_seeker": use_expert_seeker}

    def check_subtask_uniqueness(self, project):
        """
        Returns True if, in the given project, no two Task objs share a subtask.
        """
        d_sts = {}
        for taskobj in project.lst_tasks:
            for stobj in taskobj.lst_tsubtasks:
                if stobj.obj_id in d_sts:
                    d_sts[stobj.obj_id] += 1
                else:
                    d_sts[stobj.obj_id] = 1
        for k in d_sts:
            if d_sts[k] > 1:
                return False
        return True

    def check_simrun_initialization(self, simrun):
        """
        Runs various tests to ensure that simrun object was initialized
        correctly.
```

```
        """
        ## ensure that every task has been assigned unique subtasks
        if not all([self.check_subtask_uniqueness(projobj) for
                    projobj in simrun.lst_projects]):
            raise Exception("Project %d failed the task-uniqueness-commonness"
                            + " test!")
        return

    def initialize_output_file(self):
        """
        Initializes the outputfile, by writing the header into it.
        """
        colnames = ['avg_common_know_repl',
                    'avg_generalized_pts', 'avg_incorrect_know_level',
                    'avg_knowledge_level', 'avg_openness_change',
                    'avg_self_knowledge', 'avg_turnover',
                    'max_failed_tries',
                    'no_agents', 'no_projects_per_run', 'p_nonspec_exchange',
                    'similarity_projects', 'size_org_rep_subtasks',
                    'time_training_phase',
                    'media_richness_ftof', 'media_richness_phone',
                    'media_richness_email', 'media_richness_chat']
        colnames2 = ['use_expert_seeker','avg_direction_time',
                     'avg_stickiness_time',
                     'avg_st_knowledge_intensity', 'no_locations',
                     'avg_proj_intensity', 'avg_task_intensity',
                     'avg_task_priority', 'avg_tasks_per_agent']
        proj_cols0 = ['usage_ftof_total', 'usage_ftof_avg', 'usage_ftof_std',
                      'usage_ftof_skew',
                      'usage_phone_total', 'usage_phone_avg', 'usage_phone_std',
                      'usage_phone_skew',
                      'usage_email_total', 'usage_email_avg', 'usage_email_std',
                      'usage_email_skew',
                      'usage_chat_total', 'usage_chat_avg', 'usage_chat_std',
                      'usage_chat_skew']
        proj_cols = ['percent_complete', 'avg_accuracy_knowledge',
                     'std_accuracy_knowledge', 'skew_accuracy_knowledge',
                     'avg_accuracy_consensus', 'std_accuracy_consensus',
                     'skew_accuracy_consensus', 'total_general_interactions',
                     'avg_general_interactions', 'std_general_interactions',
                     'skew_general_interactions', 'total_specific_interactions',
                     'avg_specific_interactions', 'std_specific_interactions',
                     'skew_specific_interactions']
        proj_cols2 = ['total_direction_uses', 'avg_direction_uses',
                      'std_direction_uses', 'skew_direction_uses',
                      'total_transfer_uses', 'avg_transfer_uses',
                      'std_transfer_uses', 'skew_transfer_uses']

        with open("output.csv",'w') as fout:
            fout.write("expno|projno|" + "|".join(colnames) + "|" +
                       "|".join(colnames2) + "|" +
                       "|".join(proj_cols0) + "|" +
                       "|".join(proj_cols) + "|" +
                       "|".join(proj_cols2) + "|" +
                       "no_tasks|task_inter_matrix\n")
        return

    def run_experiments(self):
        """
        Runs the required number of experiments by
        a) initializing the parameters
        b) creating a SimRun object
```

```
        c) executing the appropriate methods of the SimRun object to run
           the training and project phases
        d) dealing with any error conditions
        """

        from os import system
        print system('date')
        with open("errorlog",'w') as fout:
            self.initialize_output_file()
            for expno in xrange(1875, 2500):
                print "Running experiment: %d" % expno
                try:
                    dict_params = self.initialize_run_parameters()
                    simrun = SimRun(dict_params, expno)
                    self.check_simrun_initialization(simrun)
                    simrun.execute_training_phase()
                    simrun.execute_project_phase(expno)
                except Exception as e:
                    print "Exception ", e
                    print "Creating error log"
                    fout.write("Exception %s \n" %str(e))
                    fout.write("Experiment %d\n" %expno)
                    fout.write(" - " * 20 + "\n")
                    traceback.print_exc(file=fout)
                    fout.write("\n ------ end of log ----------\n")
                print "^*^" * 25
        return True
if __name__ == '__main__':
    supervisor = Supervisor()
    supervisor.run_experiments()


"""
Implements the simulation-run object and its associated methods.
A simulation-run object represents a single simulation run, that is, a single
experimental run.
"""
from random import randint, random, shuffle, sample, choice
import math
from diss_sim_classes_v11_11 import Task, Subtask, Expert_seeker
from diss_sim_project_v11_11 import Project
from diss_sim_classes_v11_11 import AgentKArea
from diss_sim_agent_v11_11 import Agent
from diss_sim_helpers import _get_random, _create_lst_kas, fetch_obj
from diss_sim_helpers import get_indiv_richnesses
from diss_sim_helpers import compute_avg_std_skew_KL_accuracy
from diss_sim_helpers import compute_avg_std_skew_consensus
from diss_sim_helpers import compute_percentage_proj_completed
from diss_sim_helpers import compute_no_interactions, compute_other_stats
from diss_sim_helpers import compile_usage_counts
from diss_sim_helpers import create_serialized_prec_matrix
from diss_sim_helpers import compute_transmission_mode_counts
from pygraph.algorithms.generators import generate
from pygraph.algorithms.cycles import find_cycle
from pygraph.readwrite.dot import write
import psyco
psyco.full()

class SimRun:
    """
    Represents the class that holds information about, and executes actions
    pertaining to, a single simulation run.
```

```
    """
    def __init__(self, run_params, expno):
        self.run_params = run_params
        ## the following are created once for each simulation run,
        ## essentially, just once.
        self.expno = expno
        ## create the required number of subtasks
        self.lst_subtasks = [Subtask(stid) for stid in
                             xrange(run_params['size_org_rep_subtasks'])]

        ## create and assign knowledge areas to subtasks
        rng_st = run_params['rng_stickiness_time']
        rng_dt = run_params['rng_direction_time']
        rng_hl = run_params['rng_half_life']

        ## the following procedure (in CL parlance) works via side-effects
        self._assign_kas_to_subtasks(rng_st, rng_dt, rng_hl)
        ## next, create a list of KAs by collating the KAs assigned to every
        ## subtask
        self.lst_knowledge_areas = self._get_list_kas()

        ## create the required number of agents
        self.lst_agents = self._create_lst_agents()

        ## assign knowledge areas to agents via a procedure using side-effects
        self._assign_kas_to_agents()
        ## create an empty Expert_seeker object, which will be updated at the
        ## beginning of each project.
        self.expert_seeker = Expert_seeker(self.lst_agents)

        ## initialize the project objects via side-effects
        self.lst_projects = [] ## will be initialized in initialize_projects()
        self.initialize_projects()

        ## the following two attributes store the reqd. baseline info
        self.baseline_accuracy_kl = () ## holds avg., std, skew
        self.baseline_consensus = () ## holds avg., std, skew

    def _assign_kas_to_subtasks(self, rng_st, rng_dt, rng_hl):
        """
        Assigns knowledge areas to subtasks
        """
        ## print "Inside _assign_kas_to_subtasks"
        rng_ka_intensity = self.run_params['rng_know_intensity_stask']
        last_kid = 0
        for stask in self.lst_subtasks:
            no_kas = randint(rng_ka_intensity[0], rng_ka_intensity[1])
            stask.lst_kas = _create_lst_kas(no_kas, rng_st,
                                            rng_dt, last_kid)
            last_kid += no_kas
        return

    def _get_list_kas(self):
        """ - """
        l = []
        for st in self.lst_subtasks:
            for ka in st.lst_kas:
                l.append(ka)
        return l

    def _create_lst_agents(self):
        """
```

```
        Initialize the required number of agent objects and return them via
        a list.
        """
        ## first create a list of partially-complete agents

        l = [Agent(agid, randint(self.run_params['rng_no_locations'][0],
                                 self.run_params['rng_no_locations'][1]),
                   _get_random(self.run_params['avg_generalized_pts']),
                   get_indiv_richnesses(self.run_params),
                   _get_random(self.run_params['avg_openness_change']),
                   self.run_params['use_expert_seeker'])
             for agid in xrange(self.run_params['no_agents'])]
        return l

    def get_kl_statuses(self):
        """
        Returns a tuple with counts of wrong, absent and correct areas of
        knowledge and a count of number of areas in which the agent
        perceives correctly that it has knowledge.
        """
        ## print "Inside get_kl_statuses"
        while True:
            ps = int(math.ceil((_get_random(self.run_params
                                            ['avg_knowledge_level'])
                               * len(self.lst_knowledge_areas))))
            ms = int(math.ceil(_get_random
                              (self.run_params
                               ['avg_incorrect_know_level'])
                              * len(self.lst_knowledge_areas)))

            if ps + ms < len(self.lst_knowledge_areas):
                zs = len(self.lst_knowledge_areas) - (ps + ms)
                break
        while True:
            perc_true = int(math.ceil(_get_random
                                     (self.run_params
                                      ['avg_self_knowledge'])
                                     * len(self.lst_knowledge_areas)))
            if perc_true <= len(self.lst_knowledge_areas):
                break
        return (ms, zs, ps, perc_true)

    def _assign_kas_to_one_agent(self, agobj):
        """
        Assign knowledge areas, i.e., AgentKArea objects, to one agent
        after initializing each AgentKArea object appropriately.
        """
        ## first, create the required number of AgentKArea objects
        l_agkas = [AgentKArea(ka_obj.obj_id, ka_obj.stickiness_time,
                              ka_obj.direction_time)
                   for ka_obj in self.lst_knowledge_areas]

        ## next, set the reqd. number of statuses to -1, 1 and zero


        lst_agobj_ids = [ka.obj_id for ka in l_agkas]
        minuses, zers, pluses, perc_statuses = self.get_kl_statuses()
        shuffle(lst_agobj_ids)

        for minuscount in xrange(minuses):
            fetch_obj(lst_agobj_ids.pop(), l_agkas).status = -1
```

```
        for pluscount in xrange(pluses):
            fetch_obj(lst_agobj_ids.pop(), l_agkas).status = 1

        for zerocount in xrange(zers):
            fetch_obj(lst_agobj_ids.pop(), l_agkas).status = 0

        for kaobj in sample(l_agkas, perc_statuses):
            kaobj.perceived_status = 1

        agobj.lst_ka_details = l_agkas
        return

    def _assign_kas_to_agents(self):
        """
        Assigns knowledge area objects to agent objects by calling
        _assign_kas_to_one_agent for each agent.
        """
        ## print "\nInside _assign_kas_to_agents"
        for agentobj in self.lst_agents:
            self._assign_kas_to_one_agent(agentobj)
        return

    def execute_training_phase(self):
        """
        Perform various initializations and simulate initial (training-phase)
        exchanges of information about various knowledge areas.
        ""”

        def get_kid(srcobj):
            """
            """
            lst_pot_ids = [kobj.obj_id for kobj in srcobj.lst_ka_details
                           if kobj.perceived_status == 1]
            if len(lst_pot_ids):
                return choice(lst_pot_ids)

        def swap_info(source, recipient):
            """
            Simulates the interactions and mutual exchange of information
            between various agent pairs. The exchanges are about the
            areas in which each agent perceives itself as having knowledge.
            """
            ## works through side-effects
            if source.obj_id in recipient.dict_relations:
                recipient.dict_relations[source.obj_id][
                    'cnt_interactions_general'] +=1
                if random() <= source.pts:
                    kid = get_kid(source)
                    if (kid not in recipient.dict_relations[source.obj_id]
                        ['lst_ka_ids']):
                        recipient.dict_relations[source.obj_id][
                            'lst_ka_ids'].append(kid)
                        ## print "Added a new ka ID", kid,
                        ## print "to", source.obj_id, "'s dict_relations",
                        ## print "from", recipient.obj_id
            else:
                if random() < source.pts:
                    d = {'location': source.location,
                         'perceived_pts': recipient.pts, ## should be self's pts
                         'lst_ka_ids': [get_kid(source)],
                         'cnt_interactions_as_src': 0,
                         'cnt_interactions_as_rcpt': 0,
```

```
                         'cnt_interactions_general': 0 }
                    recipient.dict_relations[source.obj_id] = d
                    recipient.dict_relations[source.obj_id]
['cnt_interactions_general'] += 1
            return

        for period in xrange(self.run_params['time_training_phase']):
            ## perform exchange of information about randomly-chosen
            ## areas of knowledge between two randomly-selected agents
            ## every interaction between a pair of agents counts!
            lst_agids1 = [ag.obj_id for ag in self.lst_agents]

            shuffle(lst_agids1)
            for ag1id in lst_agids1:
                lst_agids2 = [ag.obj_id for ag in self.lst_agents
                              if not ag.obj_id == ag1id]
                shuffle(lst_agids2)
                ag1 = fetch_obj(ag1id, self.lst_agents)
                for ag2id in lst_agids2:
                    ag2 = fetch_obj(ag2id, self.lst_agents)
                    if (random() <=
                        self.run_params['p_nonspec_exchange']):
                        if random() <= ag1.pts:
                            swap_info(ag1, ag2)
                            ## cnt_fswap += 1
                        if random() <= ag2.pts:
                            swap_info(ag2, ag1)
                            ## cnt_rswap += 1
                    ## else:
                        ## cnt_nswap += 1
        ## update the Expert-seeker database
        ## later updates would happen in the project phase
        self.expert_seeker.update_reported_KAs(self.lst_agents)
        self.baseline_accuracy_kl = compute_avg_std_skew_KL_accuracy\
                                    (self.lst_agents)
        self.baseline_consensus = compute_avg_std_skew_consensus\
                                  (self.lst_agents)
        return

    def assign_tasks_to_agents(self, projobj):
        """
        Assigns tasks to agents; keeps this information in the
        dic_agent_tasks dictionary. {agentid: [taskid1, taskid2...]}
        """
        lst_ag_ids = [agobj.obj_id for agobj in self.lst_agents]
        lst_task_ids = [taskobj.obj_id for taskobj in projobj.lst_tasks]
        d = {}
        for aid in lst_ag_ids:
            if not len(lst_task_ids):
                break
            d[aid] = [lst_task_ids.pop()]

        shuffle(lst_ag_ids)
        shuffle(lst_task_ids)

        while len(lst_task_ids):
            for aid in lst_ag_ids:
                if len(lst_task_ids):
                    howmany = randint(0, len(lst_task_ids))
                    d[aid] += [lst_task_ids.pop() for i in xrange(howmany)]
                else:
                    break
```

```
    assert(len(projobj.lst_tasks) == reduce(lambda x,y: x+y,
                                        [len(d[v]) for v in d]))
    projobj.dic_agent_tasks = d
    return

def create_task_interdependence(self, projobj):
    """
    Create a DAG that represents inter-task interdependence.
    """
    no_tasks = len(projobj.lst_tasks)
    no_edges = randint(0, (no_tasks * (no_tasks - 1))/2)
    g = generate(no_tasks, no_edges, True)
    while True:
        ed = find_cycle(g)[0:2]
        if not len(ed):
            break
        g.del_edge(ed[0], ed[1])

    def get_preceding_nodes(given_node):
        return [tpl[0] for tpl in g.edges() if tpl[1] == given_node]

    for node in g.nodes():
        projobj.dic_task_precedences[node] = get_preceding_nodes(node)
    return

def get_no_subtasks(self,projid, dic_proj_counts):
    """
    Returns the size of the organizational repertoire of subtasks.
    """
    def add(x, y):
        return x+y

    return reduce(add, [reduce(add, [dic_proj_counts[projid][tid]
                                 for tid in
                                 dic_proj_counts[projid].keys()])])
def get_subtasks(self, probjj):
    """
    Returns a list of all the subtasks in the given project.
    """
    return [stobj for tobj in probjj.lst_tasks
            for stobj in tobj.lst_tsubtasks]


def get_reqd_no(self, prev_val):
    """
    Returns a value for the number of subtasks such that the
    constraints of avg_task_intensity and similarity of projects
    are satisfied.
    """
    while True:
        comp_val = randint(self.run_params['rng_task_intensity'][0],
                           self.run_params['rng_task_intensity'][1])
        if prev_val == 0:
            return comp_val
        else:
            if comp_val >= int(round(prev_val * self.run_params
                                 ['similarity_projects'])):
                return comp_val

def initialize_projects(self):
    """
```

```
Initializes a list of projects by:
(a) a list of project objects
(b) for each project, assigning a subset of subtask repertoire
    while ensuring that the constraint of similarity between two
    consecutive projects is met
(c) assigning the required number of tasks to each project
(d) assigning the subtasks to tasks in each project
(e) creating a DAG of all the tasks in each project
"""
## first, create project objects
self.lst_projects = [Project(i, self.expno) for i in
                     xrange(self.run_params['no_projects_per_run'])]


## create a dictionary that keeps an account of the number of tasks
## needed by each project and the number of subtasks required by
## each task in each project.
dic_proj_counts = {}
for projobj in self.lst_projects:
    ## no. of tasks
    no_tasks = randint(self.run_params['rng_proj_intensity'][0],
                       self.run_params['rng_proj_intensity'][1])
    dic_task_counts = {}
    for tcount in xrange(no_tasks):
        if tcount > 0:
            dic_task_counts[tcount] = self.get_reqd_no(dic_task_counts\
                                                        [tcount-1])
        else:
            dic_task_counts[tcount] = self.get_reqd_no(0)

    dic_proj_counts[projobj.obj_id] = dic_task_counts

for projobj in self.lst_projects:
    ## create the required number of task objects for each project
    projobj.lst_tasks = [Task(tid,
                              randint(self.run_params
                                      ['rng_task_priority'][0],
                                      self.run_params
                                      ['rng_task_priority'][1]))
                         for tid in dic_proj_counts[projobj.obj_id]]

    if projobj.obj_id == 0:## the first project in the sequence
        no_sts_needed = self.get_no_subtasks(projobj.obj_id,
                                             dic_proj_counts)
        lst_sub_sts = sample(self.lst_subtasks,
                                        no_sts_needed)
        l = [str(stobj.obj_id) for stobj in lst_sub_sts]
        l.sort()
        for taskobj in projobj.lst_tasks:
            taskobj.lst_tsubtasks = [lst_sub_sts.pop() for i in
                                     xrange(dic_proj_counts
                                            [projobj.obj_id]
                                            [taskobj.obj_id])]
    else:
        prev_proj_obj = self.lst_projects[projobj.obj_id - 1]
        lst_used_st_ids = [st.obj_id for st in
                           self.get_subtasks(prev_proj_obj)]
        no_common_sts = int(round(self.run_params
                                  ["similarity_projects"] *
                                  len(lst_used_st_ids)))
        lst_common_st_ids = sample(lst_used_st_ids, no_common_sts)
        lst_all_ids = [st.obj_id for st in self.lst_subtasks]
```

```
            lst_exclusive_st_ids = list(set(lst_all_ids).difference\
                                        (set(lst_used_st_ids)))
            lst_new_st_ids = lst_common_st_ids + lst_exclusive_st_ids
            lst_new_sts = [fetch_obj(stid, self.lst_subtasks)
                           for stid in lst_new_st_ids]
            ## next, assign the subtasks to each task
            shuffle(lst_new_sts)
            for tobj in projobj.lst_tasks:
                num = dic_proj_counts[projobj.obj_id][tobj.obj_id]
                tobj.lst_tsubtasks = [lst_new_sts.pop()
                                      for i in xrange(num)]

        return


    def record_stats(self, projobj, expno):
        """
        Compiles the following statistics at the end of a project:
        (a) average accuracy of knowledge at group-level
        (b) average consensus about others' expertise at group-level
        (c) time taken to complete the project: just records the time
        (d) percentage of project completed:
            no. of subtasks completed/total subtasks
        """
        colnames = ['avg_common_know_repl',
                    'avg_generalized_pts', 'avg_incorrect_know_level',
                    'avg_knowledge_level', 'avg_openness_change',
                    'avg_self_knowledge', 'avg_turnover',
                    'max_failed_tries',
                    'no_agents', 'no_projects_per_run', 'p_nonspec_exchange',
                    'similarity_projects', 'size_org_rep_subtasks',
                    'time_training_phase']
        ## note these names differ from those used in the header
        ## because these refer to keys in the dict. whose associated values
        ## are not atomic - they are tuples
        colnames2 = ['direction_time', 'stickiness_time',
                     'st_knowledge_intensity', 'no_locations',
                     'proj_intensity', 'task_intensity', 'task_priority',
                     'tasks_per_agent']
        proj_cols0 = ['ftof', 'phone', 'email', 'chat']
        proj_cols = ['percent_complete', 'avg_accuracy_knowledge',
                     'std_accuracy_knowledge', 'skew_accuracy_knowledge',
                     'avg_accuracy_consensus', 'std_accuracy_consensus',
                     'skew_accuracy_consensus', 'total_general_interactions',
                     'avg_general_interactions', 'std_general_interactions',
                     'skew_general_interactions', 'total_specific_interactions',
                     'avg_specific_interactions', 'std_specific_interactions',
                     'skew_specific_interactions']
        d2 = compute_other_stats(projobj, self.lst_agents)
        d_usages = compile_usage_counts(self.lst_agents)

        with open("output.csv", 'a') as fout:
            outstr = "|".join([str(expno),str(projobj.obj_id)]) + "|"
            outstr += "|".join([str(self.run_params[item]) for item in
                                colnames]) + "|"
            outstr += "|".join([str(self.run_params['dict_comm_media_richness']
                                    [item]['richness'])
                                for item in ('ftof', 'phone', 'email', 'chat')
                                ]) + "|"
            outstr += str(int(self.run_params['use_expert_seeker'])) + "|"
            outstr += "|".join([str(d2[item]) for item in colnames2 ]) + "|"
```

```
            outstr += "|".join(["|".join([str(val) for val in
                                          d_usages[colname]])
                                for colname in proj_cols0]) + "|"

            outstr += str(compute_percentage_proj_completed(projobj)) + "|"


            outstr += "|".join([str(item) for item in
                                compute_avg_std_skew_KL_accuracy(
                                    self.lst_agents)]) + "|"
            outstr += "|".join([str(item) for item in
                                compute_avg_std_skew_consensus(
                                    self.lst_agents)]) + "|"

            # the following code computes the KL and consensus values
            ## as a difference from the baseline
            d = compute_no_interactions(self.lst_agents)
            outstr += "|".join([str(item) for item in d['general']]) + "|"
            outstr += "|".join([str(item) for item in d['task_related']]) + "|"
            d = compute_transmission_mode_counts(self.lst_agents)
            outstr += "|".join([str(item) for item in d['direction']]) + "|"
            outstr += "|".join([str(item) for item in d['transfer']]) + "|"
            ## add the number of tasks
            outstr += str(len(projobj.lst_tasks)) + "|"
            ## append the sequential-version of the task-precedence matrix
            ## NOTE: ';' will be used to separate rows in the matrix
            outstr += ";".join(create_serialized_prec_matrix(projobj)) + "\n"
            fout.write(outstr)
        return


    def execute_project_phase(self, expno):
        """
        Execute the project phase of the simulation.
        """
        ## print "\nInside execute_project_phase"
        for proj_obj in self.lst_projects:
            self.assign_tasks_to_agents(proj_obj)
            self.create_task_interdependence(proj_obj)
            ret_val = proj_obj.execute_project(self.lst_agents,
                                               self.expert_seeker,
                                               self.run_params)
            ## update the Expert-seeker database
            self.expert_seeker.update_reported_KAs(self.lst_agents)
            self.record_stats(proj_obj, expno)


"""
Holds the Project class and its methods.
"""
from diss_sim_helpers import fetch_obj, _get_random
from random import random, sample, choice, randint
from copy import deepcopy
from diss_sim_agent_v11_11 import Agent
from diss_sim_helpers import get_indiv_richnesses
import psyco
psyco.full()
class Project:
    """
    Represents the template used to create project objects.
    """
```

```
    def __init__(self, obj_id, expno):
        """
        obj_id: identifies place in the sequence of projects that constitute
                a simulation run
        tpl_lst_tasks: a tuple that holds all the tasks that are assigned to
                       the current project
        """
        self.obj_id = obj_id # int!
        self.lst_tasks = [] ## holds references to Task objects
        self.max_failed_tries = -1
        ## taskid: agent dictionary,
        ## used to keep track of agent_id: [taskid, taskid...] assignments
        self.dic_agent_tasks = {}
        ## the following dict is redudant but will improve object-access speed
        self.dic_task_agentobjs = {}
        ## holds the DAG of tasks; {taskid: [taskid, taskid,...],...}
        self.dic_task_precedences = {}
        self.status = 'PROJ-NOT-STARTED' # PROJ-COMPLETED/PROJ-ABANDONED
        self.expno = expno

    def __repr__(self):
        return ("obj_id:%d\nlst_tasks:%s\n"
                %(self.obj_id,
                  " ".join([str(task) for task in self.lst_tasks])))

    def initialize_subtasks_failure_counts(self, failure_count):
        """
        Sets up the failure count values for each subtask - these values
        would be used by each agent to determine whether to abandon a task.
        """
        ## print "\nInside initialize_subtasks_failure_counts"
        for taskobj in self.lst_tasks:
            for stobj in taskobj.lst_tsubtasks:
                stobj.max_poss_failures = failure_count
        return

    def assign_agentobjs_to_tasks(self, lst_agents):
        """
        Fills the self.dic_task_agentobjs datastructure.
        """
        for agid in self.dic_agent_tasks:
            for tid in self.dic_agent_tasks[agid]:
                self.dic_task_agentobjs[tid] = fetch_obj(agid, lst_agents)
        return

    def has_completed(self):
        """
        Returns True if every task assigned to the project has been completed.
        """
        for taskobj in self.lst_tasks:
            if taskobj.status != 'T-COMPLETED':
                return False
        return True

    def agent_with_several_assigns(self, taskobj):
        """
        Returns True if a given task is assigned to an agent who is assigned
        another task that has begun and has not been completed.
        """
        for other_task_id in self.dic_task_precedences[taskobj.obj_id]:
            if (fetch_obj(other_task_id, self.lst_tasks).status == 'T-STARTED'
                and (self.dic_task_agentobjs[other_task_id].obj_id ==
```

```
                    self.dic_task_agentobjs[taskobj.obj_id].obj_id)):
                return True
        return False

    def is_ready_to_start(self, taskobj, lst_agents):
        """
        Returns True if the given taskobj is ready to start. A task can be
        started when (a) it has no predecessors; (b)all its predecessors have
        been completed.
        """
        if not len(self.dic_task_precedences[taskobj.obj_id]):
            return True
        for taskid in self.dic_task_precedences[taskobj.obj_id]:
            tobj = fetch_obj(taskid, self.lst_tasks)
            result = self.agent_with_several_assigns(taskobj)
            if (tobj.status in ('T-STARTED', 'T-NOT-STARTED', 'T-ABANDONED') or
                result):
                return False
            else:
                return True

    def start_task(self, taskobj, lst_agents, expert_seeker,
                   p_nonspec_exchange):
        """
        Checks to see whether a given task can be started, if so, it starts
        the given task by calling the complete_task method associated with the
        agent that has been delegated the current task.
        """
        if self.is_ready_to_start(taskobj, lst_agents):
            taskobj.status = 'T-STARTED'
            self.dic_task_agentobjs[taskobj.obj_id].complete_task(taskobj,
                                                                  lst_agents,
                                                                  expert_seeker,p_
nonspec_exchange)
        return

    def continue_task(self, taskobj, lst_agents, expert_seeker,
                      p_nonspec_exchange):
        """
        Continues the execution of the given task.
        """
        assert(taskobj.status == 'T-STARTED')
        self.dic_task_agentobjs[taskobj.obj_id].complete_task(taskobj,
                                                              lst_agents,
                                                              expert_seeker,
                                                              p_nonspec_exchange)
        return

    def run_schedule(self, lst_agents, expert_seeker, p_nonspec_exchange):
        """
        Based on the inter-task precedence, execute tasks by having the
        associated agents complete their respective tasks.
        """
        for taskobj in self.lst_tasks:
            if taskobj.status == 'T-NOT-STARTED':
                self.start_task(taskobj, lst_agents, expert_seeker,
                                p_nonspec_exchange)
            elif taskobj.status == 'T-STARTED':
                self.continue_task(taskobj, lst_agents, expert_seeker,
                                   p_nonspec_exchange)
            else:
                assert(taskobj.status in ('T-ABANDONED','T-COMPLETED'))
```

```
        return


    def perform_agent_swap(self, p_turnover, common_kl, lst_agents,
                           dict_run_params):
        """
        Performs the activities involved in swapping an existing agent with
        its replacement.
        """
        def get_replaceable_kaids(old_agentobj, kl):
            """
            Returns the IDs of the required proportion of KAs that should be
            retained in the new agent replacing an existing agent.
            """
            no_kids = int(round((1-kl) * len(old_agentobj.lst_ka_details)))
            return sample([kobj.obj_id for kobj in
                           old_agentobj.lst_ka_details],
                          no_kids)

        if random() <= p_turnover:
            ## choose an agent
            assert(len(lst_agents) > 0)
            assert(dict_run_params is not None)
            ## print "\nGoing to swap agents"
            agentobj_tbr = choice(lst_agents)
            #new agent's ID = max(existing agents' IDs) + 1
            new_id = max([agobj.obj_id for agobj in lst_agents]) + 1
            new_agentobj = Agent(new_id,
                                 randint(dict_run_params[
                                     'rng_no_locations'][0], dict_run_params[
                                             'rng_no_locations'][0]),
                                 _get_random(dict_run_params[
                                     'avg_generalized_pts']),
                                 get_indiv_richnesses(dict_run_params),
                                 _get_random(dict_run_params[
                                     'avg_openness_change']),
                                 dict_run_params['use_expert_seeker'])
            new_agentobj.lst_ka_details = deepcopy(agentobj_tbr.lst_ka_details)
            ## now, choose the replaceable KAs and flip their status to
            ## either -1 or 0, with a 0.5 probability; change perceived_status
            ## similarly
            for kid in get_replaceable_kaids(agentobj_tbr, common_kl):
                repl_kaobj = fetch_obj(kid, new_agentobj.lst_ka_details)
                if random() <= 0.5:
                    repl_kaobj.status = -1
                else:
                    repl_kaobj.status = 0

                if random() <= 0.5:
                    repl_kaobj.perceived_status = 0
                else:
                    repl_kaobj.perceived_status = 1
            return (agentobj_tbr, new_agentobj)
        return (None, None)

    def execute_turnover(self, p_turnover, common_kl, lst_agents,
                         dict_run_params):
        """
        Executes turnover, that is, change in the membership of the workgroup,
        via side-effects. Specifically, a single agent is replaced, with
        a given number of knowledge areas in common with the agent that it is
        replacing.
```

```
        """
        def identify_affected_agent(given_agentobj, lst_agents):
            """
            Returns reference to the agent that is affected by the swap.
            """
            for agentobj in lst_agents:
                if agentobj.obj_id == given_agentobj.obj_id:
                    return agentobj
            return None

        old_agentobj, new_agentobj = self.perform_agent_swap(
            p_turnover, common_kl, lst_agents, dict_run_params)
        if old_agentobj is not None and new_agentobj is not None:
            ## identify the agent that is currently obtaining knowledge
            ## from the old agent and replace its KA activities
            affected_agent = identify_affected_agent(old_agentobj, lst_agents)
            if affected_agent is not None:
                affected_agent.modify_knowledge_transfer(new_agentobj)
        return

    def reinitialize_interaction_counts(self, lst_agents):
        """
        Re-initializes the counts of interactions of all agent-pairs
        """
        for agentobj in lst_agents:
            for other_agentobj in [agobj for agobj in lst_agents if
                                   agobj.obj_id != agentobj.obj_id]:
                if other_agentobj.obj_id in agentobj.dict_relations:
                    agentobj.dict_relations[other_agentobj.obj_id][
                        'cnt_interactions_as_rcpt'] = 0
                    agentobj.dict_relations[other_agentobj.obj_id][
                        'cnt_interactions_general'] = 0
        return

    def reinitialize_media_usage_counts(self, lst_agents):
        """
        ...
        """
        for agentobj in lst_agents:
            for medium in agentobj.dict_media_usage_counts:
                agentobj.dict_media_usage_counts[medium] = 0
        return

    def reinitialize_transfer_mode_counts(self, lst_agents):
        """
        ...
        """
        for agentobj in lst_agents:
            for tmode in agentobj.dict_transfer_mode_counts:
                agentobj.dict_transfer_mode_counts[tmode] = 0
        return

    def execute_project(self, lst_agents, expert_seeker, dict_run_params):
        """
        Execute the various activities involved in a project.
        """
        failure_count = dict_run_params['max_failed_tries']
        p_nonspec_exchange = dict_run_params['p_nonspec_exchange']
        turnover = dict_run_params['avg_turnover']
        knowledge_repl_level = dict_run_params['avg_common_know_repl']
        self.initialize_subtasks_failure_counts(failure_count)
        self.assign_agentobjs_to_tasks(lst_agents)
```

```
        self.reinitialize_media_usage_counts(lst_agents)
        self.reinitialize_transfer_mode_counts(lst_agents)
        while True:
            self.execute_turnover(turnover, knowledge_repl_level, lst_agents,
                                  dict_run_params)
            expert_seeker.update_reported_KAs(lst_agents)
            self.run_schedule(lst_agents, expert_seeker, p_nonspec_exchange)
            if self.has_completed():
                return "PROJ-COMPLETED"
        return

"""
Contains various helpder methods used by various associated modules.
"""
from random import gauss, random, randint
from numpy import mean, std
from scipy.stats import skew
from itertools import combinations
import psyco
psyco.full()

def _get_random(mean):
    """
    - mean: the mean value is used to determine a value drawn from
              N(mean,mean/3) such that the value is > 0 and <= 2*mean
    - norm: True indicates that the returned value should be < 1.0

    NOTE: if norm is False, then the value returned is an integer;
          if norm is True, then the value returned is a float.
          Not a pure way of returning values, but this is pragmatic :/
    """
    while True:
        val = gauss(mean, mean/3.0)
        if 0 < val <= mean*2 and val < 1.0:
            return val
def get_positive_avg():
    while True:
        val = random()
        if val > 0:
            return val

def fetch_obj(objid, lst_objs):
    """
    Returns the object, from the given list of objects, which
    matches the given object id.
    Assumes that the requested object ID exists in the list of objects.
    """
    return [obj for obj in lst_objs if obj.obj_id == objid][0]

def _get_ka_times(st, dt, hlt):
    """
    Returns appropriate values for halflife, stickiness and direction.
    """
    halflife = randint(hlt[0], hlt[1])
    stickiness = randint(st[0], st[1])
    dt = randint(dt[0], dt[1])

    ## ensure that the condition of direction time exceeding
    ## transfer (stickiness) time does not occur
    def getval(s, d):
        """
        ---
```

```
        """
        while True:
            val = _get_random(d)
            if val <= s:
                return val

    direction = getval(stickiness, dt)
    return (stickiness, direction, halflife)

def _create_lst_kas(no_kareas, rng_stickiness_time,
                    rng_direction_time, last_kid):
    """
    Returns a list of knowledge-area objects (those held in reality).
    """
    from diss_sim_classes_v11_11 import KArea

    lst_kas = []
    for kid in xrange(no_kareas):
        stime = randint(rng_stickiness_time[0], rng_stickiness_time[1])
        dtime = randint(rng_direction_time[0], rng_direction_time[1])
        lst_kas.append(KArea(last_kid + kid, stime, dtime))
    return lst_kas

def get_indiv_richnesses(dict_run_params):
    """
    Assign the perceived richness values to the four media individually
    to each agent.
    """
    ## create a copy-by-value
    ## this is a one-time-only hack, which is simple to read
    ## but is hard-coded to copy the structure of the original
    ## dictionary of richness values
    s = dict_run_params['dict_comm_media_richness']
    d = {"ftof": {"synch": "yes",
                  "richness": 1.0,
                  "locality": "yes"},
         "phone": {"synch": "yes",
                   "richness": _get_random(s['phone']
                                           ['richness']),
                   "locality": "maybe"},
         "chat": {"synch": "yes",
                  "richness": _get_random(s['chat']
                                          ['richness']),
                  "locality": "maybe"},
         "email": {"synch": "maybe",
                   "richness": _get_random(s['email']
                                                ['richness']),
                   "locality": "maybe"}}
    return d

def compute_avg_std_skew_KL_accuracy(lst_agents):
    """
    Returns the average level of accuracy of knowledge. This is the average
    value of the number of areas in which an agent has accurate knowledge,
    averaged across all agents in a workgroup.
    """
    def get_avg_kl(lst_kaobjs):
        """
        Returns the average KL for a given agent.
        """
        return float(len([kaobj for kaobj in lst_kaobjs
                          if kaobj.status == 1]))/len(lst_kaobjs)
```

```
    l = [get_avg_kl(agentobj.lst_ka_details) for agentobj in lst_agents]
    return (mean(l), std(l), skew(l))


def compute_avg_std_skew_consensus(lst_agents):
    """
    Returns the average level of consensus (representing the group's TM).
    For each knowledge area, the number of agreements are divided by the total
    number of agents having a non-zero perception. This gives the average
    agreements about the workgroup's agents about the focal agent. This
    average is computed for all agents.
    """
    dict_consesuses = {}## kaid: consensus_value
    assert(len(lst_agents))

    def compute_consensus(ref_agobj):
        lst_other_agents = [agobj for agobj in lst_agents if agobj.obj_id
                            != ref_agobj.obj_id]
        assert(all([len(ref_agobj.lst_ka_details) ==
len(otheragobj.lst_ka_details)
                    for otheragobj in lst_other_agents]))

        count_corrects, count_incorrects = 0, 0
        dict_consensus_counts = {}
        for kid in xrange(len(ref_agobj.lst_ka_details)):
            for oth_ag_obj in lst_other_agents:
                if (ref_agobj.obj_id in oth_ag_obj.dict_relations and
                    kid in oth_ag_obj.dict_relations[ref_agobj.obj_id
                                                     ]['lst_ka_ids']):
                    count_corrects += 1
                if (ref_agobj.obj_id in oth_ag_obj.dict_relations_negatives
                    and kid in
                    oth_ag_obj.dict_relations_negatives[ref_agobj.obj_id]):
                    count_incorrects += 1

            total = count_corrects + count_incorrects
            if total > 0: ## being explicit
                dict_consensus_counts[kid] = float(abs(count_corrects -
                                                       count_incorrects)
                                                   )/float(total)
            else:
                dict_consensus_counts[kid] = -999
        lll = [val for val in dict_consensus_counts.values() if val >= 0]

        return sum(lll)

    for ref_agent_obj in lst_agents:
        dict_consesuses[ref_agent_obj.obj_id] = compute_consensus(ref_agent_obj)
    ## overall consensus per knowledge area
    l = dict_consesuses.values()
    return (mean(l), std(l), skew(l))


def compute_percentage_proj_completed(projobj):
    """
    Returns the proportion of the subtasks in a given project that have
    been completed.
    """
    sts_completed, sts_total = 0, 0
    for taskobj in projobj.lst_tasks:
        for stobj in taskobj.lst_tsubtasks:
```

```
            sts_total += 1
            if stobj.status == 'ST-COMPLETED':
                sts_completed += 1
    return float(sts_completed)/sts_total


def compute_no_interactions(lst_agents):
    """
    Returns a dictionary with items containing the total, mean, std. dev. and
    skewness of number of non-task-related information exchanges and
    task-related knowledge exchanges.
    """
    d = {}
    ll = []
    for agobj in lst_agents:
        lc = [agobj.dict_relations[key]['cnt_interactions_general']
              for key in agobj.dict_relations]
        ll.append(sum(lc))
    d['general'] = (sum(ll), mean(ll), std(ll), skew(ll))

    ll = []
    for agobj in lst_agents:
        lc = [agobj.dict_relations[key]['cnt_interactions_as_rcpt']
              for key in agobj.dict_relations]
        ll.append(sum(lc))
    d['task_related'] = (sum(ll), mean(ll), std(ll), skew(ll))
    return d

def compute_other_stats(proj_obj, lst_agents):
    """
    Returns a dictionary containing the following values:
    avg. direction time, avg. knowledge intensity of subtask,
    no. of locations, project intensity (no. of tasks per project),
    avg. stickiness time of a knowledge area, avg. task intensity (no. of
    subtasks per task), avg. task priority, avg. number of tasks per agent
    """
    ## direction time is averaged across all knowledge areas.
    ## first, compile the list of all the tasks
    d = {}
    lst_direction, lst_stickiness = [], []
    lst_k_intensity_st, lst_task_st_intensity = [], []
    lst_task_priorities, lst_tasks_per_agent = [], []
    for taskobj in proj_obj.lst_tasks:
        lst_task_priorities.append(taskobj.priority)
        lst_task_st_intensity.append(len(taskobj.lst_tsubtasks))
        for stobj in taskobj.lst_tsubtasks:
            lst_k_intensity_st.append(len(stobj.lst_kas))
            for kaobj in stobj.lst_kas:
                lst_direction.append(kaobj.direction_time)
                lst_stickiness.append(kaobj.stickiness_time)


    lst_tasks_per_agent = [len(val) for val in
                           proj_obj.dic_agent_tasks.values()]

    d['direction_time'] = mean(lst_direction)
    d['stickiness_time'] = mean(lst_stickiness)
    d['st_knowledge_intensity'] = mean(lst_k_intensity_st)
    d['no_locations'] = len(set([agobj.location for agobj in lst_agents]))
    d['proj_intensity'] = len(proj_obj.lst_tasks)
    d['task_intensity'] = mean(lst_task_st_intensity)
    d['task_priority'] = mean(lst_task_priorities)
```

```
    d['tasks_per_agent'] = mean(lst_tasks_per_agent)
    return d

def create_serialized_prec_matrix(projobj):
    """
    Returns a sequential-version of the matrix representing
    task-interdependencies.
    """
    l = []
    for taskid in projobj.dic_task_precedences:
        for succeeding_task in projobj.dic_task_precedences[taskid]:
            l.append("(" + ",".join((str(taskid+1), str(succeeding_task+1), '1'))
+
                     ")")
    return l


def compile_usage_counts(lst_agents):
    """
    Returns the total number, avg., std, and skewness values of the
    number of times a particular medium is used by the agents. The media:
    ftof, phone, email, chat.
    """
    l_ftof, l_phone, l_email, l_chat = [], [], [], []

    for agentobj in lst_agents:
        l_ftof.append(agentobj.dict_media_usage_counts['ftof'])
        l_phone.append(agentobj.dict_media_usage_counts['phone'])
        l_email.append(agentobj.dict_media_usage_counts['email'])
        l_chat.append(agentobj.dict_media_usage_counts['chat'])

    return ({'ftof': (sum(l_ftof), mean(l_ftof), std(l_ftof), skew(l_ftof)),
             'phone':(sum(l_phone), mean(l_phone), std(l_phone), skew(l_phone)),
             'email': (sum(l_email), mean(l_email), std(l_email),
                       skew(l_email)),
             'chat': (sum(l_chat), mean(l_chat), std(l_chat), skew(l_chat))})


def compute_transmission_mode_counts(lst_agents):
    """
    Returns total, mean, std and skewness values for direction and
    transfer counts respectively.
    """
    l_direction, l_transfer = [], []

    for agentobj in lst_agents:
        l_direction.append(agentobj.dict_transfer_mode_counts['direction'])
        l_transfer.append(agentobj.dict_transfer_mode_counts['transfer'])
    return ({'direction': (sum(l_direction), mean(l_direction),
                           std(l_direction), skew(l_direction)),
             'transfer': (sum(l_transfer), mean(l_transfer), std(l_transfer),
                          skew(l_transfer)) })


import psyco
import random
psyco.full()

class KArea:
    """
    Represents a knowledge area object. Objects of this type are stored in
    'reality'. This is in contrast to objects of the class AgentKArea, whose
```

```
    objects are stored within agent and OthersKADetails, whose objects an agent
    uses to store its perceptions of other agents' knowledge areas.
    """
    def __init__(self, obj_id, stickiness_time, direction_time):
        """
        obj_id: knowledge area's identifier
        stickiness: stickiness (in time periods) associated with the KA
        direction_time: time taken to transmit directions associated with the
                         KA
        """
        self.obj_id = obj_id
        self.stickiness_time = stickiness_time
        self.direction_time = direction_time

    def __repr__(self):
        return ("obj_id: %s; stickiness: %d; direction_time: %d\n"
                "ka_halflife: %d\n"
                %(self.obj_id, self.stickiness_time, self.direction_time,
                  self.ka_halflife))

class Subtask:
    """
    Represents a subtask.
    """
    def __init__(self, obj_id):
        """
        obj_id: subtask's id
        lst_kas: list of KAs needed to complete the given task
        """
        self.obj_id = obj_id
        self.lst_kas = [] ## references to objects of KArea (not IDs)
        ## other values: 'ST-STARTED','ST-ABANDONED','ST-COMPLETED'
        self.status = 'ST-NOT-STARTED'
        self.time_spent = 0
        self.max_poss_failures = 0 ## no. of failed knowledge-seeking tries
allowed


    def __repr__(self):
        return ("obj_id: %s\nlst_kas:\n%s"
                %(self.obj_id, "\n".join([str(ka) for ka in self.lst_kas])))

class Task:
    """
    Represents a task.
    """
    def __init__(self, obj_id, priority):
        """
        obj_id: task id
        priority: priority assigned to the task. range: 'lo','med','hi'
        lst_subtasks: list of subtasks assigned to the task object
        Note: The task object will not contain information about the tasks that
              precede or succeed it; the precedence information is maintained by
              the supervisor object that controls the simulation.
        """
        self.obj_id = obj_id
        self.priority = priority ## 0: high; 1: medium; 2: low
        self.lst_tsubtasks = [] ## lst of subtask objects, not IDs
        self.status = 'T-NOT-STARTED' # 'T-STARTED'/'T-ABANDONED'/'T-COMPLETED'
        # will be set to an appropriate value in the create_task_interdependence
        # method of a SimRun object.
```

```
    def __repr__(self):
        return ("obj_id: %s\npriority: %d\n%s"
                %(self.obj_id, self.priority,
                  "\n".join([str(stid) for stid in self.lst_tsubtasks])))

class Expert_seeker:
    """
    Represents the expert-seeker database.
    """

    def __init__(self, lst_agobjs):
        """
        dict_agent_KAs: holds ka ids as the keys and lists of agent ids
        (only those agents that report that they have knowledge in a given area)
        as values.
        """
        self.dict_agent_KAs = {}
        self.update_reported_KAs(lst_agobjs)

    def update_reported_KAs(self, lst_agobjs):
        """
        Pings each agent and requests it to provide information about the
        various areas in which it has knowledge. The dict_agent_KAs
        datastructure is updated via side-effects.
        """
        for aob in lst_agobjs:
            for kob in aob.lst_ka_details:
                if random.random() <= aob.pts and kob.perceived_status == 1:
                    if aob.obj_id not in self.dict_agent_KAs:
                        self.dict_agent_KAs[aob.obj_id] = [kob.obj_id]
                    else:
                        self.dict_agent_KAs[aob.obj_id].append(kob.obj_id)
        return

    def get_ka_src_ag_ids(self, kid):
        """
        For a given knowledge_area id, return a list of all agents that are
        sources of that knowledge.
        """
        l2 = [key for key in self.dict_agent_KAs if
              kid in self.dict_agent_KAs[key]]
        return l2

class AgentKArea:
    """
    Used to represent knowledge area objects corresponding to each knowledge
    area within an agent. This is in contrast to KArea, whose objects represent
    knowledge areas in 'reality' and OthersKADetails, whose objects are used
    by each agent to represent its perceptions regarding the knowledge areas of
    other agents.
    """

    def __init__(self, obj_id, stickiness_time, direction_time):
        """
        status: -1: incorrect knowledge; 0: no knowledge; 1: correct knowledge
        perceived status: 0: absent; 1: present
        """
        self.obj_id = obj_id
        self.stickiness_time = stickiness_time
        self.direction_time = direction_time
        self.status = -9
        self.perceived_status = 0
```

```
    def __repr__(self):
        return ("obj_id: %s\nstickiness: %d\ndirection_time: %d\n"
                "status: %d\nperceived_status: %d\n"
                "ka_halflife:%d\n"
                %(self.obj_id, self.stickiness_time, self.direction_time,
                  self.status, self.perceived_status, self.ka_halflife))

"""
Implements the Agent class along with its associated methods.
"""
from random import random, shuffle, choice
import math
import psyco
from copy import deepcopy
from operator import itemgetter
from diss_sim_helpers import fetch_obj
psyco.full()


class Agent:
    """
    Represents the agent and the methods that are executed by the agent.
    """

    def __init__(self, obj_id, location, pts, dict_prm_richness, opench,
                 use_expert_seeker):
        """
        obj_id: unique identifier
        location: location identifier
        pts: propensity to share
        dict_prm_richness: dict. to hold perceived media richness values
        openness_change: openness to change
        lst_ka_details: list of AgentKArea objects
        dict_relations: dict. to hold values of the following structure
        **the key in the dictionary is an agent's obj_id**
            {'obj_id': {'location':,
                        'perceived_pts':.
                        'lst_ka_ids': [],
                        ##'cnt_interactions_as_src': 0,
                        'cnt_interactions_as_rcpt':0,
                        'cnt_interactions_general': 0
                        }
            }
        dict_relations_negatives: {agent's obj_id: [kaid1, kaid2,..]
        lst_agent_task_ids: list of tasks assigned to the agent of the form
                           [('obj_id': location in sequence)]
        curr_know_src: reference to the agent from whom knowledge is
                         being obtained
        curr_subtask: reference to the current subtask being completed
        sorted_network_tasks:  a topologically sorted network of task
                                 objects
        """
        self.obj_id = obj_id
        self.location = location
        self.pts = pts
        self.dict_prm_richness = dict_prm_richness
        self.openness_change = opench
        # lst of AgentKArea objects, one for each KArea object in SimRun
        self.lst_ka_details = [] ## item == reference to an AgentKArea object
        self.dict_relations = {} # ka_id: ag_id
        self.dict_relations_negatives = {}
```

```
        self.curr_stask_obj = None
        self.curr_ka_obj = None  ## current area of knowledge being obtained
        self.curr_ksrc_agent = None ## current src (agent) of knowledge
        self.curr_transfer_mode = -1 ## 0: direction; 1: transfer
        ## keeps track of how many units of time were spent in acquiring currKA
        self.dict_elapsed_times = {}
        self.curr_actual_transfer_time = -1
        self.dict_failure_counts = {}
        ## indicates whether Expert-seeker can be used.
        self.use_expert_seeker = False
        ## keeps track of no. of times each medium is used
        ## should be reset at the beginning of each project
        self.dict_media_usage_counts = {'ftof': 0, 'phone': 0, 'email': 0,
                                        'chat': 0}
        self.dict_transfer_mode_counts = {"direction": 0, "transfer": 0}

    def __repr__(self):
        return ("obj_id: %s\nlocation: %s\npts: %s\ndict_prm_richness: %s\n"
                "openness_change: %f\n"
                "st_ka_details: %s\ndict_relations: %s\n"
                %(self.obj_id, self.location, self.pts, self.dict_prm_richness,
                  self.openness_change, self.lst_ka_details,
                  self.dict_relations))

    def identify_src_agent(self, lst_agents, expert_seeker):
        """
        Use self's dictionary of relationships or Expert-seeker to obtain
        a perceived source of knowledge and assigns it to self.curr_ksrc_agent.
        """
        ## first, check if any of the 'known' agents has the required knowledge
        assert(self.curr_ka_obj is not None)

        needed_kid = self.curr_ka_obj.obj_id
        lst_pot_agids = [(ag_id, self.dict_relations[ag_id]['perceived_pts'])
                         for ag_id in self.dict_relations if
                          needed_kid in self.dict_relations[ag_id]
                          ['lst_ka_ids']]
        if not len(lst_pot_agids):
            ## otherwise, consult the expert-seeker, if allowed
            if self.use_expert_seeker:
                lst_ids = expert_seeker.get_ka_src_ag_ids(
                    self.curr_ka_obj.obj_id)
                lst_pot_agids = [(agid, self.pts) for agid in lst_ids
                                 if agid != self.obj_id]
            if not len(lst_pot_agids):
                self.curr_ksrc_agent = None ## just making sure
                return
        chosen_agid = sorted(lst_pot_agids, key=itemgetter(1), reverse=True)[0][0]
        self.curr_ksrc_agent = fetch_obj(chosen_agid, lst_agents)
        if self.curr_ksrc_agent.obj_id not in self.dict_relations:
            d = {'location': self.curr_ksrc_agent.location,
                 'perceived_pts': self.pts,
                 'lst_ka_ids': [self.curr_ka_obj.obj_id],
                 'cnt_interactions_as_rcpt': 0,
                 'cnt_interactions_general': 0 }

            self.dict_relations[self.curr_ksrc_agent.obj_id] = d
        return

    def perc_shared_knowledge(self):
        """
        Computes the perceived number knowledge areas shared
```

```
        between self and the source agent.
        """
        count_common = 0
        if self.curr_ksrc_agent.obj_id in self.dict_relations:
            for kaobj in self.lst_ka_details:
                if (kaobj.obj_id in
                    self.dict_relations[self.curr_ksrc_agent.obj_id]\
                    ['lst_ka_ids']):
                    count_common += 1
        return count_common

    def perc_richness_media_wrt_src(self):
        """
        Returns the perc. richness of various media w.r.t a source agent.
        """
        def bounded_val(value):
            """
            Ensures that the output is <= 1.00
            """
            if value > 1.00:
                return 1.00
            return value
        d_richness = {}
        for medium in self.dict_prm_richness:
            val = (self.dict_prm_richness[medium]['richness']
                   + (0.1 * math.log(self.perc_shared_knowledge() + 1,
                                     len(self.lst_ka_details)
                                     + 1)))
            d_richness[medium] = bounded_val(val)
        return d_richness

    def choose_medium(self, task_priority):
        """
        Choose a medium for receiving knowledge from an agent based on
        task priority, the source agent's location and perceived richness
        value of each medium.
        """
        ## based on task priority, determine synchronousness
        ## priority 0: synchronousness - yes
        ## priority 1: synchronousness - maybe
        ## priority 2: synchronousness - no
        c = task_priority
        if c == 0:
            lst_pot_media = ['ftof', 'phone', 'chat']
        elif c == 1:
            lst_pot_media = ['ftof', 'phone', 'chat', 'email']
        else:
            lst_pot_media = ['phone', 'chat', 'email']

        d = self.perc_richness_media_wrt_src()
        l = [(medium, d[medium]) for medium in d]

        for tpl in sorted(l, key=itemgetter(1), reverse=True):
            if tpl[0] in lst_pot_media:
                self.dict_media_usage_counts[tpl[0]] += 1
                return tpl[0]

    def get_transfer_time(self, task_priority):
        """
        Returns the time that would be used by the knowledge-seeking
        agent in determining when to end the transfer of knowledge.
        """
```

```
        medium = self.choose_medium(task_priority)
        return int(round(((float(self.curr_ka_obj.stickiness_time)) /
                          self.perc_richness_media_wrt_src()[medium])))

    def get_transfer_mode(self):
        """
        For a given transfer time, returns 0: direction; 1: transfer.
        """
        if self.curr_ka_obj.stickiness_time <= self.curr_ka_obj.direction_time:
            return 1
        return 0

    def transfer_is_complete(self):
        """
        Returns True if transfer is complete.
        """
        val = (self.dict_elapsed_times[self.curr_ka_obj.obj_id] >=
               self.curr_actual_transfer_time)
        return val

    def finalize_received_KA(self):
        """
        Sets self's received KA's status to that of the src agent's KA's value
        (which can be -1 or 1). Works through a side-effect.
        """
        obj = fetch_obj(self.curr_ka_obj.obj_id,
                        self.curr_ksrc_agent.lst_ka_details)
        if random() <= self.curr_ksrc_agent.pts and obj.perceived_status == 1:
            if obj.status in (-1, 0) :
                self.curr_ka_obj.status = -1
            elif obj.status == 1:
                self.curr_ka_obj.status = 1
            else:
                raise Exception("Inconsistent status of KA %d in src. agent %d"
                                %(obj.obj_id, self.curr_ksrc_agent.obj_id))

        return


    def reset_statuses(self, abandon=False, completed=False):
        """
        Resets the status values of various state-related agent attributes.
        """
        self.curr_stask_obj = None
        self.curr_ka_obj = None
        self.curr_ksrc_agent = None
        self.curr_transfer_mode = -1
        self.curr_actual_transfer_time = -1
        self.dict_elapsed_times = {}
        if abandon or completed:
            self.dict_failure_counts = {}
        return

    def reset_self_ka_perceptions(self):
        """
        Resets the perceived status of all self's KAs to zero
        """

        for kobj in self.curr_stask_obj.lst_kas:
            obj = fetch_obj(kobj.obj_id, self.lst_ka_details)
            if obj.status == 1:
                obj.perceived_status = 0
```

```
        return

    def knowledge_is_accurate(self):
        """
        Returns True if knowledge in all the areas is correct (all 1s). Of
        course, this assumes that incorrect knowledge transmissions involve
        the setting of -1 in the corresponding AgentKArea object's status (not
        perceived status - this will be set to 1) of self.
        """

        for kobj in self.curr_stask_obj.lst_kas:
            selfkaobj = fetch_obj(kobj.obj_id, self.lst_ka_details)
            if selfkaobj.status == -1:
                return False
        return True

    def update_src_pts(self, accurate):
        """
        Updates the perceived pts value associated with a source based on
        the whether the knowledge obtained was accurate or inaccurate.
        """
        assert(self.curr_ksrc_agent is not None)
        assert(self.curr_ka_obj is not None)
        if random() <= self.openness_change:
            if accurate:
                self.dict_relations[self.curr_ksrc_agent.obj_id
                                    ]['perceived_pts'] += 0.01
                ## add the curr_ka_obj.obj_id to the perceived
                ## list of areas in which curr_ksrc_agent is knowledgeable
                if self.curr_ksrc_agent.obj_id in self.dict_relations:
                    if not (self.curr_ka_obj.obj_id in
                            self.dict_relations[self.curr_ksrc_agent.obj_id
                                                ]['lst_ka_ids']):
                        self.dict_relations[self.curr_ksrc_agent.obj_id
                                            ]['lst_ka_ids']\
                                            .append(self.curr_ka_obj.obj_id)
                else:
                    self.dict_relations[self.curr_ksrc_agent.obj_id
                                        ]['lst_ka_ids'] = self.curr_ka_obj.obj_id


            else: ## accurate is False
                self.dict_relations[self.curr_ksrc_agent.obj_id
                                    ]['perceived_pts'] -= 0.01
                if (self.curr_ksrc_agent.obj_id in
                    self.dict_relations and
                    self.curr_ka_obj.obj_id in
                    self.dict_relations[self.curr_ksrc_agent.obj_id
                                        ]['lst_ka_ids']):

                    self.dict_relations[self.curr_ksrc_agent.obj_id
                                        ]['lst_ka_ids'
                                          ].remove(self.curr_ka_obj.obj_id)
                if (self.curr_ksrc_agent.obj_id not in
                    self.dict_relations_negatives):
                    self.dict_relations_negatives[self.curr_ksrc_agent.obj_id
                                                  ] = [self.curr_ka_obj.obj_id]
                else:
                    if (self.curr_ka_obj.obj_id not in
                        self.dict_relations_negatives[
                            self.curr_ksrc_agent.obj_id]):
                        self.dict_relations_negatives[
```

```
                    self.curr_ksrc_agent.obj_id].append(
                    self.curr_ka_obj.obj_id)
    return

def handle_failed_status(self):
    """
    Resets the current subtask's status to either 'ST-ABANDONED'
    or 'ST-NOT-STARTED' depending on the whether the number of failed tries
    exceeds the max. number of failed tries allowed for the given subtask.
    """
    if self.curr_stask_obj.obj_id in self.dict_failure_counts:
        self.dict_failure_counts[self.curr_stask_obj.obj_id] += 1
    else:
        self.dict_failure_counts[self.curr_stask_obj.obj_id] = 1
    if (self.dict_failure_counts[self.curr_stask_obj.obj_id] >=
        self.curr_stask_obj.max_poss_failures):
        self.curr_stask_obj.status = 'ST-ABANDONED'
        self.reset_statuses(abandon=True)
    else:
        self.curr_stask_obj.status = 'ST-NOT-STARTED'
        self.reset_statuses()

    return


def handle_completed_transfer(self):
    """
    Handles the events after a transfer is marked complete:
    (1) mark the transferred knowledge as right/wrong, based on the pts
        of the source agent.
    (2) verifies that the knowledge received is correct and,
        if it is, mark the subtask as 'ST-COMPLETED' and reset the state
        values; if it not, mark the subtask as either 'ST-ABANDONED' or
        'ST-NOT-STARTED'
    """
    assert(self.curr_ka_obj is not None)
    assert(self.curr_ka_obj.obj_id in self.dict_elapsed_times)

    self.finalize_received_KA()
    if self.knowledge_is_accurate():
        self.curr_stask_obj.status = 'ST-COMPLETED'
        self.update_src_pts(True)
        self.reset_statuses(completed=True)
    else:
        if (self.curr_ksrc_agent is not None and
            self.curr_ka_obj.status == -1):
            self.update_src_pts(False)
        self.handle_failed_status() ## handles reset_statuses...
    return

def assign_transfer_time_mode(self, task_priority):
    """
    Assigns current transfer actual time and transfer mode.
    """
    self.curr_actual_transfer_time = self.get_transfer_time(task_priority)
    self.curr_transfer_mode = self.get_transfer_mode()
    if not self.curr_transfer_mode:
        self.dict_transfer_mode_counts['direction'] += 1
    else:
        self.dict_transfer_mode_counts['transfer'] += 1
    return
```

```
    def handle_src_related_assignments(self, lst_agents, expert_seeker):
        """
        Handles the assignment of a source agent from whom the required
        knowledge may be obtained. Also handles the assignment of
        actual_transfer_time and transfer_mode.
        """
        if self.curr_ksrc_agent is not None:
            return
        assert(self.curr_ka_obj is not None)
        self.identify_src_agent(lst_agents, expert_seeker)
        if self.curr_ksrc_agent is None:
            self.handle_failed_status()
        return

    def handle_knowledge_transmission(self, p_nonspec_exchange):
        """
        Handles the transmission of one unit of knowledge, and the consequences
        if the transfer of knowledge is complete. Also, handles the transfer
        of information pertaining to an unrelated KA by calling
        handle_unrelated_knowledge_transfer (which, too, works via a
        side-effect).
        """
        assert(self.curr_ka_obj is not None)
        assert(self.curr_ksrc_agent is not None)
        assert(self.curr_stask_obj.status == 'ST-STARTED')
        if self.curr_ka_obj.obj_id not in self.dict_elapsed_times:
            self.dict_elapsed_times[self.curr_ka_obj.obj_id] = 1
        else:
            self.dict_elapsed_times[self.curr_ka_obj.obj_id] += 1
        ## increment the interaction count
        self.dict_relations[self.curr_ksrc_agent.obj_id]
['cnt_interactions_as_rcpt'] += 1
        self.handle_unrelated_knowledge_transfer(p_nonspec_exchange)
        if self.transfer_is_complete():
            self.handle_completed_transfer()
        return

    def handle_unrelated_knowledge_transfer(self, p_nonspec_exchange):
        """
        Implements the transfer of knowledge that may/may not be related to the
        current subtask. The exchange is bidirectional. The exchange is an
        exchange of only information related to knowledge areas, not an exchange
        of any actual knowledge from a knowledge area.
        Works through side-effects.
        """
        if random() < p_nonspec_exchange:
            if random() < self.curr_ksrc_agent.pts:
                ## recipient is self
                kobj = choice(self.curr_ksrc_agent.lst_ka_details)
                if kobj is not None:
                    kid = kobj.obj_id
                if (kid not in self.dict_relations[self.curr_ksrc_agent.obj_id]
                    ['lst_ka_ids']):
                    self.dict_relations[self.curr_ksrc_agent.obj_id]
['lst_ka_ids'].append(kid)
                    self.dict_relations[self.curr_ksrc_agent.obj_id]
['cnt_interactions_general'] += 1
                kobj = choice(self.lst_ka_details)
                if kobj is not None:
                    kid = choice(self.lst_ka_details).obj_id
                    if self.obj_id not in self.curr_ksrc_agent.dict_relations:
                        d = {'location': self.location,
```

```
                          'perceived_pts': self.curr_ksrc_agent.pts,
                          'lst_ka_ids':[kid],
                          'cnt_interactions_as_rcpt':0,
                          'cnt_interactions_general': 1 }
                     self.curr_ksrc_agent.dict_relations[self.obj_id] = d
        return


    def identify_ka_obj(self):
        """
        Identifies and assigns, via a side-effect, one of the knowledge areas
        required for completing the given subtask.
        """
        lst_needed_ka_ids = [kobj.obj_id for kobj in self.curr_stask_obj.lst_kas
                             if not fetch_obj(kobj.obj_id,
                                              self.lst_ka_details)\
                             .perceived_status]
        if len(lst_needed_ka_ids):
            shuffle(lst_needed_ka_ids)
            self.curr_ka_obj = fetch_obj(lst_needed_ka_ids.pop(),
                                         self.lst_ka_details)
        else:
            self.curr_ka_obj = None ## just to be sure
        return

    def handle_has_reqd_kas(self):
        """
        Handles the situation where the agent perceives that it has knowledge
        in all the required knowledge areas.
        """
        assert(self.curr_ka_obj is None) ## is a prerequisite
        if self.knowledge_is_accurate():
            self.curr_stask_obj.status = 'ST-COMPLETED'
            self.reset_statuses(completed=True)
        else:
            self.reset_self_ka_perceptions()
            self.handle_failed_status()## will handle specific resets..
        return

    def assign_curr_subtask(self, taskobj):
        """
        For the given taskobj (reference to a Task obj), identify a
        subtask whose status is 'ST-NOT-STARTED' and assign it to
        self.curr_stask_obj.
        """
        lst_pot_stasks = [stobj for stobj in taskobj.lst_tsubtasks
                          if stobj.status == 'ST-NOT-STARTED']
        if len(lst_pot_stasks):
            self.curr_stask_obj = lst_pot_stasks.pop()
        else:
            self.curr_stask_obj = None ## just making sure.
        return

    def handle_subtask_assignment(self, taskobj):
        """
        Handles the assignment of current subtask and the consequence if
        a subtask could not be identified.
        """
        self.assign_curr_subtask(taskobj)
        if self.curr_stask_obj is None:
            ## hardcoding is efficient :-P
            dic_st_statuses = {'ST-ABANDONED': 0,
```

```
                        'ST-COMPLETED': 0,
                        'ST-NOT-STARTED': 0,
                        'ST-STARTED':0}
        for stobj in taskobj.lst_tsubtasks:
            dic_st_statuses[stobj.status] += 1
        if (not dic_st_statuses['ST-NOT-STARTED'] and
            not dic_st_statuses['ST-STARTED']):
            taskobj.status = 'T-COMPLETED'
    else:
        self.curr_stask_obj.status = 'ST-STARTED'
    return

def complete_next_subtask(self, taskobj, lst_agents, expert_seeker,
                          p_nonspec_exchange):
    """
    Completes the current subtask. Once completed, it changes the status
    of the current subtask to either ST-COMPLETED or ST-ABANDONED.
    """
    assert(self.curr_stask_obj is not None)
    assert(self.curr_stask_obj.status == 'ST-STARTED')
    if self.curr_ka_obj is None:
        self.identify_ka_obj()
        if self.curr_ka_obj is None:
            self.handle_has_reqd_kas()
    else:
        self.handle_src_related_assignments(lst_agents, expert_seeker)
        if self.curr_ksrc_agent is not None:
            self.assign_transfer_time_mode(taskobj.priority)
            self.handle_knowledge_transmission(p_nonspec_exchange)
        else:
            pass ## handle_src_related_assignments already handles
    return

def complete_task(self, taskobj, lst_agents, expert_seeker,
                  p_nonspec_exchange):
    """
    Complete the task whose task id is given by completing each of its
    associated subtasks.
    """
    assert(taskobj is not None)
    assert(taskobj.status == 'T-STARTED')
    if self.curr_stask_obj is None:
        self.handle_subtask_assignment(taskobj)
        if (taskobj.status in ('T-ABANDONED','T-COMPLETED') or
            self.curr_stask_obj is None): ##should this be not None ?? - No
            return
    self.complete_next_subtask(taskobj, lst_agents, expert_seeker,
                               p_nonspec_exchange)

    return

def modify_knowledge_transfer(self, new_agentobj):
    """
    Models the activities that occur if the agent from whom self is
    obtaining knowledge has been replaced.
    """
    ## treat the transfer as a failure
    ## self.handle_failed_status()
    ## reset the current agent so that the transfer can recommence,
    ## and continue if it's possible.
    self.curr_ksrc_agent = None
```