Dynamics in organizational problem solving and the leveraging of social capital: An agent-based modeling (ABM) perspective

Yuan Lin

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:
Hans J. Scholl, Chair
Kevin C. Desouza
Michael E. Eisenberg
Hazel Taylor
Deanna Kennedy

Program Authorized to Offer Degree:

Information Science

University of Washington

Abstract

Dynamics in organizational problem solving and the leveraging of social capital:
An agent-based modeling (ABM) perspective

Yuan Lin

Chair of the Supervisory Committee:
Professor Hans J. Scholl, The Information School

Organizational problem solving – the pooling of individual members' efforts to solve a problem – includes knowledge exchange and knowledge creation, both of which are important for overall organizational performance. Interpersonal knowledge exchanges disseminate temporally good solutions across the organization, encouraging organizational members to create solutions based on these good ideas. The recombination of individual knowledge, however, leads to increasingly similar knowledge bases and a decline in organizational knowledge diversity, thus reducing the chance of creating even better solutions in the future. The success of knowledge exchanges or creation is significantly influenced by individuals' leveraging of social capital, which mostly resides in informal organizational structure. The outcome of social capital leveraging relies not only on the opportunity presented to an individual (determined by the individual's position in an organizational social network), but also on the individual's motivation and ability to seize that opportunity. Additionally, social capital goes through complex structural changes while being leveraged, thus affecting its subsequent impacts.

Considering all the above issues, this dissertation study investigated how overall organizational problem-solving performance would be affected by individual members' autonomous leveraging of social capital for knowledge exchange or creation purposes. The research method, agent-based modeling (ABM), provided a unique perspective on this question while other approaches cannot. It allowed to account for emergent collective outcomes by dynamic and decentralized individual interactions. The simulation results suggested non-linear relationships between organizational problem-solving performance and (a) individual members' motivation to leverage social capital, (b) individual members' preference on what social capital to leverage, and (c) the impact of existing social capital. This study advanced the understandings of organizational ambidexterity, organizational social capital, and organizational networks. The agent-based model developed in this study can benefit future research.

To my committee, parents, and friends, it has been quite a ride.
Thank you for keeping me grounded and focused along the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

Knowledge is an important asset and competitive advantage for organizations, regarding the various complex problems they encounter in today's highly diversified and ever changing business environment (Argote et al. 2003; Davenport & Prusak 1998; Nonaka & Takeuchi 1995). Due to the division of labor and the specialization of knowledge bases, now the most useful knowledge is owned by organizational members and scatters all over an organization. How to appropriate individual knowledge remains a subject of great interest to organizational researchers (Grant 1996; Nonaka & Takeuchi 1995). Outside organizational contexts, recently collective intelligence has shown unprecedented success. We see huge numbers of people from all over the world form various communities where everyone is active in creating and sharing knowledge to solve one complex problem. The achievements of these communities include Human Genome Map, Wikipedia, Mozilla Firefox, and so on. It is argued that, by developing similar social interaction environments inside, organizations can copy the success made by pure self-organized communities as mentioned above (Alavi & Leidner 2001; Brown & Duguid 2001; Hansen et al. 1999; Kane & Alavi 2005; McDermott & Archibald 2010; O'Dell & Grayson 1999).

A problem thus rises regarding the difference between an organization and a pure self-organized community: to what extent should organizations interfere with individual members' problem-solving activities such as independent knowledge creation and knowledge exchanges? This problem is worth looking into as some managerial efforts did not pay off or even backfired because of unexpected outcomes of individual behaviors (Cross et al. 2002; Haas & Hansen

2007; Hansen 2009; Hansen & Nohria 2004; Perlow 1999). In terms of organizational structure, the problem is about what formal structure an organization should impose on its members' problem-solving activities that continuously generate informal organizational structure – the emergent patterns of social interactions within an organization. To tackle this problem, the following questions need to be answered:

a) Comparing with self-organized communities outside organizational contexts, what are organizations' unique needs in terms of problem solving?
b) What kind of structure can coordinate organizational members' decentralized problem-solving activities to fulfill organizations' unique needs indicated in (a)?
c) How could organizational members' autonomous problem-solving activities collectively generate and dynamically maintain the structure indicated in (b)?

While previous researches have shed some light on the first and the second questions, few studies have been done and little has been known about the third one.

Organizations generally have fewer members and more stable membership than self-organized problem-solving communities, which means organizations have lower capabilities of knowledge creation or innovation. With this constraint, organizations are pressured to solve certain problems more quickly and economically, as they usually have specific agenda, timelines and budgets. In addition to being aligned and efficient in solving established problems, it is also important for organizations to be prepared and flexible for tackling emerging problems, so that they can survive the future. Thus, organizational members are expected, on the one hand, to refine existing knowledge and current paths of knowledge acquisition to improve efficiency and, on the other hand, to explore new knowledge and new paths of knowledge acquisition to ensure sustainability. These two types of the activity are generally referred to as exploitation and exploration respectively, and organizations perform the best when individual members' general exploitation and exploration activities are balanced at the organizational level (March 1991).

There are two major balancing approaches based on formal organizational structure (Eisenhardt & Martin 2000). The sequential approach is for organizations to shift strategies over time and to alternately assign exploitative or exploratory tasks to individual members. The simultaneous approach is to establish structurally separated subunits in charge of exploration and exploration respectively and to set up higher-level integration processes. Neither of these approaches, however, allows organizations to utilize their members' discretion and abilities to solve complex problems in today's turbulent business environments. This idea was picked up by a third approach. This so-called contextual approach proposed to build supportive organizational contexts (e.g., processes or structures) that allows and encourages individual members to divide and adjust their time between exploration and exploitation activities (Gibson & Birkinshaw 2004). No organizational design details have ever been concretely specified (O'Reilly III & Tushman 2011) though, except for findings on the impacts of organizational culture, individual attributes (Adler et al. 1999; Gibson & Birkinshaw 2004; Gulati & Puranam 2009), and most recently, the macro structure of interpersonal knowledge exchanges (Fang et al. 2010; Lazer & Friedman 2007; Mason & Watts 2012; Miller et al. 2006).

Interpersonal knowledge exchanges are both blessings and curses for organizational problem solving, with regard to the need of balancing exploitation and exploration. On the one hand, they allow organizational members to apply existing solutions instead of "reinventing the wheel" or repeating earlier mistakes, therefore improving the efficiency of organizational problem solving. Interpersonal knowledge exchanges also promote innovation when new solutions are created by integrating existing solutions. On the other hand, rapid dissemination of temporally good solutions distracts organizational members from further exploring their own ideas. Knowledge exploitation and recombination diminish organization-wide knowledge diversity, which keeps the organization from prematurely converging to suboptimal solutions. In order for organizations to maintain high problem-solving performances in the long run, interpersonal knowledge exchanges should be coordinated at the organizational level to make sure individual solutions will not spread too fast or too slowly (Fang et al. 2010; Lazer & Friedman 2007; March 1991; Miller et al. 2006). To this end, previous studies suggested confining organizational members' knowledge exchanges to a connected macro network that has both closure and brokerage structures (Fang et al. 2010; Lazer & Friedman 2007), a network referred to as **hybrid** hereafter in this dissertation.

In previous studies, the macro interaction network was assumed to be predefined, static, and exogenous. It was also assumed that organizational members repeatedly exchanged knowledge with their network neighbors (and no one else) in the same frequency. These rigid assumptions are appropriate only when the network is part of formal organizational structure. The current study was concerned with organizational members' autonomous knowledge exchanges, which are conducted more via informal social interactions than through formal avenues (Cross et al. 2001). Organizational members' social interactions generate and then continuously shape informal organizational structure, which, once established, structuralizes social interactions (Giddens 1984). The mutual influences between micro interactions and macro structure iterates:

previous micro interactions collectively and gradually generate current macro interaction structure, which subsequently impacts future micro interactions, and this process goes on and on. Thus, if the macro interaction network represents informal organizational structure, it should be emergent, dynamic, and endogenous. Moreover, it should coevolve with organizational members' autonomous problem-solving behaviors, including interpersonal knowledge exchanges and independent knowledge creation (i.e., no knowledge exchange). The current study tried to understand how such coevolution influence the overall organizational performance. The overarching hypothesis was that the longer a hybrid macro network stayed through the problem-solving process, the better the overall organizational problem-solving performance would be in the long run. The dynamic, cross-level view of the current study is fundamentally different from the view of previous studies (**Figure 1)**. This view is closer to organizational reality and may shed more light on the balance of knowledge exploitation and exploration as a process.

The preceding coevolution is driven by individual behaviors that deviate from the regulation of established social structure. To account for these behaviors, the current study applied the concept of social capital – actual and potential resources embedded within, available through, and derived from the structure of social relations (Nahapiet & Ghoshal 1998). Social capital influences knowledge transfer or exchange (Inkpen & Tsang 2005; Nahapiet & Ghoshal 1998; Wei et al. 2011) and knowledge creation (Fleming et al. 2007; McFadyen & Cannella 2004; McFadyen et al. 2009; Sosa 2011). While closely related to social networks and often measured by network statics, social capital has richer theoretical underpinnings. Network ties and structures only describe one dimension of social capital (Nahapiet & Ghoshal 1998) and opportunities for leveraging social capital (Adler & Kwon 2002). Whether and how social capital is actually leveraged also depend on individual motivations and abilities. Thus, the non-deterministic impact of informal organizational structure on individual members' autonomous problem-solving behaviors can be accounted for by individuals' leveraging of social capital. The defining components of a hybrid macro network – closure and brokerage structures – are network representations of two major types of social capital (Burt 2000b). Closure induced social capital originates from close strong social relations (called bonds) that support low-cost and high-quality knowledge exchanges. Brokerage induced social capital comes from distant weak social relations (called bridges) that provide diverse, novel knowledge to fuel innovation. The leveraging of social capital can thus be generally categorized as bonding or bridging (Adler & Kwon 2002; Putnam 2000; Reagans & McEvily 2008).

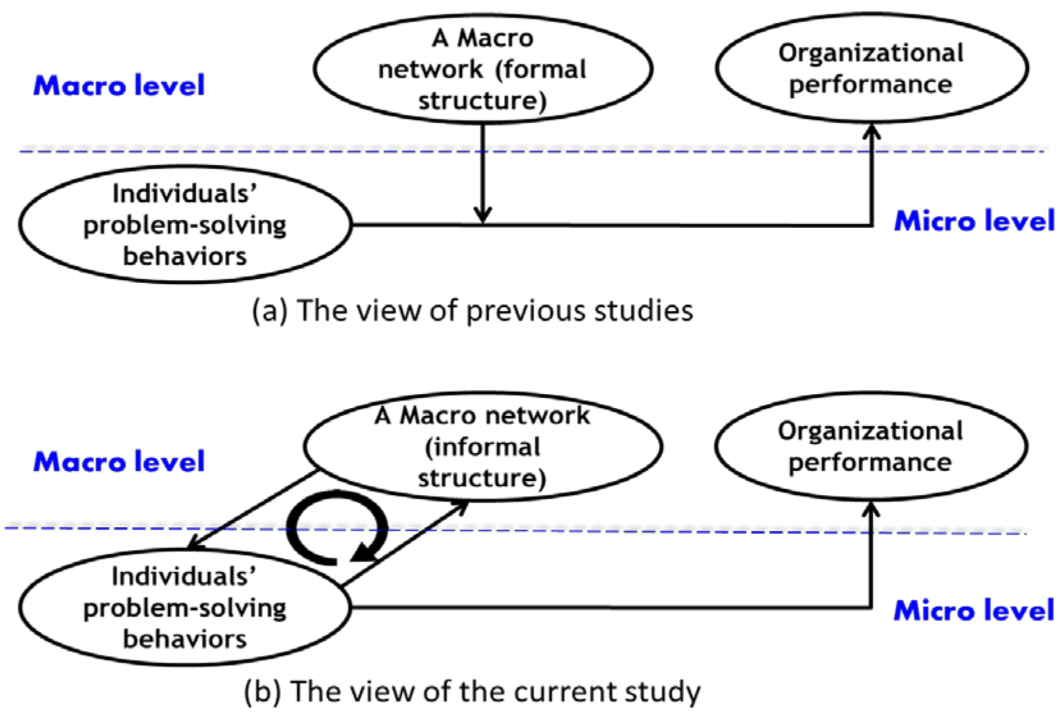


Figure 1. How individual members' problem-solving behaviors affect organizational problem-solving performance given a macro interaction network – the view of the current study vs. the view of previous studies

The current study asked the question: *how would organizational members' social-capital-based problem-solving behaviors, mainly knowledge exchange and creation, influence the overall organizational problem-solving performance?* Since the concomitant coevolution of micro interactions and macro structure is difficult, if ever possible, to capture in real organizations, computer modeling and simulation (Law & Kelton 1991) was applied as the research method with the following steps. First, the research problem was translated into an agent-based model (Epstein 1999; Epstein 2006; Gilbert 2008). Aligned with previous modeling studies, a simplified scenario of organizational problem solving known as parallel problem solving (Lazer & Friedman 2007) was simulated. The design of model variables and mechanisms was guided by the earlier hypothesis on generating and maintaining a hybrid macro network and drew on social networks and social capital theories as well as network formation models (Jackson 2010; Toivonen et al. 2009). Overall, the organization was modeled as a complex adaptive system, which consists of interacting agents, self-organized macro structure, and specific mechanisms jointly supporting the coevolution of micro interactions and the macro structure. Next, simulation experiments were conducted on a computer implementation of the preceding model to explore model behaviors, test theoretical hypotheses, and examine model validity. Since primary model inputs were expected to have non-linear and interactive effects on model outputs, experimental design and data analysis used Latin Hypercube Sampling and Multivariate Adaptive Regression Splines respectively. Results confirmed previous findings on the relationships between organizational problem-solving performances and factors that are unrelated to social capital and social networks, such as time, problem complexity, organizational size, and individual members' independent knowledge creating abilities. Results also revealed non-linear relationships between organizational performances and social capital or social network related factors, such as individual members' motivation to leverage social capital, their preferences on what social capital to leverage, and the influence of established social capital. These results, to the range of this study's examination, are valid and robust.

This dissertation has five chapters including this introduction. Chapter 2 reviews key concepts and relevant theories mentioned earlier, explaining their connections with the current study. Chapter 3 specifies conceptual design and computer implementation of the agent-based model developed to answer the research question raised above. Chapter 4 described experimental design and data analysis and presented the results. Chapter 5 discussed contributions and limitations of the current study.

Chapter 2

LITERATURE REVIEW

In this chapter, I review the theoretical and methodological foundations of the current study, including organizational exploitation and exploration, parallel problem solving, closure and brokerage structures, (related) network formation models, complex adaptive social systems, and agent-based modeling. Although each of these topics has a vast amount of literature, the review focuses on how they are connected with and fitted into the current study.

2.1. Organizational Exploitation and Exploration

There has been a long tradition of dividing organizational activities into exploitation and exploration (March 1991). Generally, exploitation relies on familiar sources or paths, attempting to leverage and refine extant knowledge; exploration seeks new sources or experiments with new paths, often inducing novel knowledge. Exploitation captures the ongoing benefits of established efficiency, but may lead to inertia in whatever has led to the success (Leonard-Barton 1992; Levinthal & March 1993; Tushman & O'Reilly III 1996). Exclusive exploitation makes an organization suffer from knowledge obsolescence and have difficulties in adapting to a new environment. In contrast, exploration can accelerate the renewal of organizational knowledge base, but too much exploration without exploitation may create chaos (Sastry 1997), trapping an organization in endless and unrewarding search (Levinthal & March 1993). Thus, organizations perform the best when exploitative activities, which maintain existing efficiency, and exploratory activities, which ensure future development, are balanced at the organizational level.

The literature suggests two major approaches to balancing exploitation and exploration: sequential or simultaneous. The sequential approach is based on a belief that exploration and exploitation are fundamentally conflicting activities. Advocates of this approach argue that an organization cannot pursue both activities simultaneously and succeed in both; thus, exploitation and exploration are better balanced by periodically switching attention between them (Brown & Eisenhardt 1997; Nickerson & Zenger 2002; Siggelkow & Levinthal 2003). However, the right timing is difficult to find, especially in today's turbulent business environment (Eisenhardt & Martin 2000). Thus, recent studies advocate the simultaneous approach also known as organizational ambidexterity (Gupta et al. 2006; O'Reilly III & Tushman 2008; Raisch et al. 2009). It is based on a belief that exploration and exploitation can be complementary and therefore should coexist within an organization.

The main challenge of the simultaneous approach is how to balance parallel exploration and exploitation activities (Gupta et al. 2006). Spatial separation seems to be an intuitive solution (Raisch et al. 2009): a business corporation can create two divisions, one in charge of exploration and the other in charge of exploitation; a team can assign individual members different roles. This solution, however, does not actually tackle the balancing problem but shift it to a higher level where the exploratory and exploitative units or roles need to be integrated and coordinated (Gupta et al. 2006), such as the corporation or the team level in the above examples. It is asserted that senior executives (Burgelman 2002; Jansen et al. 2008; Smith & Tushman 2005) or middle managers (Huy 2002) can be ambidexterity facilitators[1] at the higher level.

Another solution is to rely on organizational members' abilities to organize their own activities and to cope with the ensuing inconsistencies (Gibson & Birkinshaw 2004). Organizational behavior studies imply that non-managerial employees can act ambidextrously[2] (Leana & Barry 2000; Lewis 2000; Smith & Lewis 2011). Relying on the abilities of individual employees does not rule out managerial interventions. It is accepted that appropriate organizational design can make individual efforts collectively favorable for the organization. Earlier studies proposed a parallel formal structure that allows members to switch back and forth (Goldstein 1985; McDonough & Leifer 1983; Nonaka & Takeuchi 1995), while later studies advocated informal structure and organizational culture (Adler et al. 1999; Gibson & Birkinshaw 2004; Gulati & Puranam 2009). However, few practical guidelines have been provided so far, and most findings were obtained from mathematical or computational models.

March's seminal work (1991) focused on the interaction between individual members and the organization. In this model, a constant vector represented external environment. Individual knowledge of the environment (called individual belief) and organizational knowledge of the environment (called organizational code) were represented by vectors of the same length. Individual (organizational) performance was measured by the proportion of dimensions in the individual belief (organizational code) vector that match with the environment vector. Individual

---

[1]The management can determine the proportion of exploration or exploitation, coordinate the staff, resolve internal conflicts, and fulfill multiple roles. The goal is to maintain coherence and appropriate amount of inconsistency.

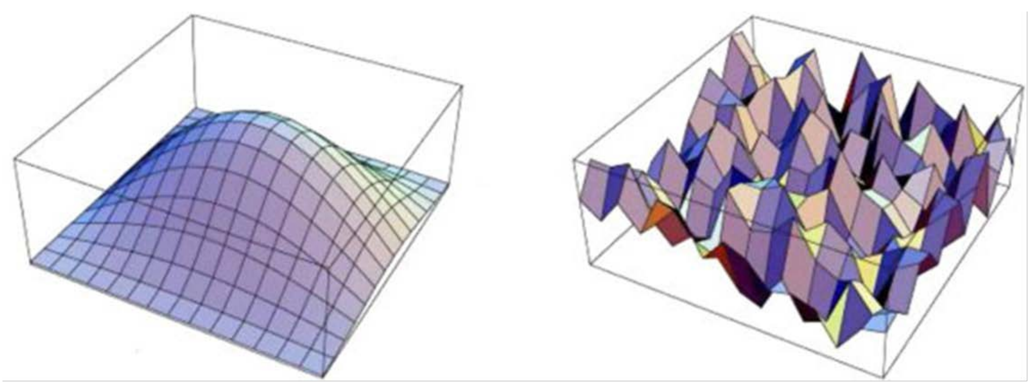
[2] To fulfill self-development and maintain job satisfaction, individual employees seek stable relations and dependable resources on the one hand, and new stimulation and variety in their work on the other hand.

and organizational performances improved through the interactions between individuals and the organization: on the one hand, individual beliefs changed to match the organizational code (a slow change leaves room for individual-level exploration); on the other hand, organizational code changed to match the majority of high-performing individual beliefs (a rapid change represented organizational-level exploitation). Because of the interactions, individual beliefs and the organizational code converged over time. When they finally became identical, neither individual nor organizational performance can further improve via interactions. Using this model, March found that the ultimate organizational performance is high when there is a balance between exploitation and exploration: the organizational code quickly changes to the best individual belief, while other individual beliefs slowly change to the organizational code. Miller (2006) extended March's model by allowing for direct interactions among individual members in addition to individual-organization interactions. He also distinguished explicit and tacit knowledge and the latter can only be exchanged via interpersonal interactions. The results showed that rapid interpersonal knowledge exchanges, like rapid convergence of individual beliefs to organizational code in March's model, negatively affect long-run organizational performance. Recently there have been two advances in modeling organization-environment interaction and the interactions among organizational members, which are reviewed in Section 2.2 and 2.3, respectively.

2.2. Parallel Problem Solving and the NK Landscape

March and Miller modeled the interaction between an organization and its environment from an organizational learning perspective – how fast and how accurate the organization and all its members learn about the external environment. Recent models (Fang et al. 2010; Lazer & Friedman 2007) applied a new type of organization-environment interaction known as parallel problem solving. In this scenario, the problems an organization encounters are complex and have more than one plausible solution. Also, the problem-solving process can yield valuable byproducts, so the organization assigns multiple members to work on the same problem independently and expects them to come up with diverse solutions. The parallel problem solving scenario differs from the organizational learning scenario used in March and Miller's models in two aspects with regard to their modeling. First, organizational performance depends on all individual solutions rather than a few superior solutions since there can be multiple solutions. Second, problem solving is path dependent – earlier solutions pave the way for later solutions, so the dimensions of a knowledge vector are not independent, at least with regard to the order of getting knowledge. In March and Miller's models, performances are measured by the proportion of matching dimensions, implying the independency of different dimensions. Thus, the models of parallel problem solving use a new representation of external environment and a new measurement of individual and organizational performances.

When modeling parallel problem solving, the external environment is often formalized as a NK landscape (Kauffman 1993). The NK landscape is a popular tool to model organizational environment (Felin et al. 2012; Foss 2011; Levinthal & Warglien 1999; Rivkin & Siggelkow 2003; Tiwana 2008) borrowed from complex systems research (Davis 2010). Every point in the landscape represents a potential solution that involves N knowledge areas. Each area contributes to the overall score of a solution equally and the contribution of each area is affected by K other areas. The level of problem complexity can be adjusted by changing the value of K: small value of K creates a smooth environment with simple problems, whereas large K creates a turbulent environment producing complex problems. This might be easier to explain by showing a 3D visualization of the NK landscape (**Figure 2**). When K = 0, the landscape looks like gently rolling ridges coming off a towering volcano. There is only one peak (called global optimum), and any move from that location will diminish performance. The performance scores of adjacent locations (i.e., solutions different in only one area) are so highly correlated that moving towards any location of higher performance will eventually lead to the global optimum. When $K = N - 1$, the landscape is extremely rugged with many peaks (called local optima) and valleys and the sides of ridges are precipitous. There is little correlation among adjacent solutions, as changing one area will affect the contribution of every other area. As K increases from 0 to $N - 1$, the number of peaks increases, the level of precipitousness increases, the correlation among moves decreases, and the height of the peaks decreases.



(a) Smooth NK landscape (K = 0) (b) Rugged NK landscape (K = N – 1)

Figure 2. Visualization of the NK landscape

The problem-solving process is modeled as agent(s) searching on the NK landscape; if there are multiple agents, they act in parallel. Individual performance is evaluated by the position of an agent: recall that each position corresponds to a scored problem solution. Organizational performance is measured by the average performance of all individuals. Kauffman (1993) distinguished two types of search conducted by single agents – adaptive walking and long jumping. Adaptive walking is to alter a random knowledge area of the current solution (i.e., explore one adjacent location) at a time. It simulates an incremental, inefficient, trial-and-error type of problem solving due to the bounded rationality of human beings: when people experiment on alternative solutions, it is difficult for them to anticipate and comprehend the consequence if too many changes are made at once. Agents who explore the NK landscape by adaptive walking are largely affected by their searching paths and present locations. Given a rugged landscape, an adaptive-walking agent tends to get stuck at the first local optimum it meets, because every possible next step results in performance decline.

Random long jumping, in contrast, allows agents to alter multiple knowledge areas at once and therefore to move beyond immediate neighborhoods and search paths. The number of areas an agent can alter at one time, denoted by a model parameter ω, controls how different an agent's new solution is from its last one, or metaphorically, how wide an agent can jump. As for a rugged landscape, a sufficiently big ω can improve the overall performance by protecting individual agents from the traps of local optima, but overly big ω might have just the opposite effect. As illustrated in **Figure 3**, an agent at the start point can follow any of the three searching paths. The green path has the widest span and leads to a local optimum, while the blue path has a moderate span and ends up at the global optimum. Further, the counterproductive effect of overly big ω is more severe on a smooth NK landscape (i.e., small K). While searching the landscape, an agent will move to a new location (i.e., adopt a new solution) only if that place has strictly better performance than the agent's current location. Otherwise, the agent will stay at its original location to explore other directions, which means the earlier trial fails. Since a smooth landscape has much fewer peaks than a rugged landscape, random long jumping is more likely to miss any peak and fail, as illustrated in **Figure 4**.

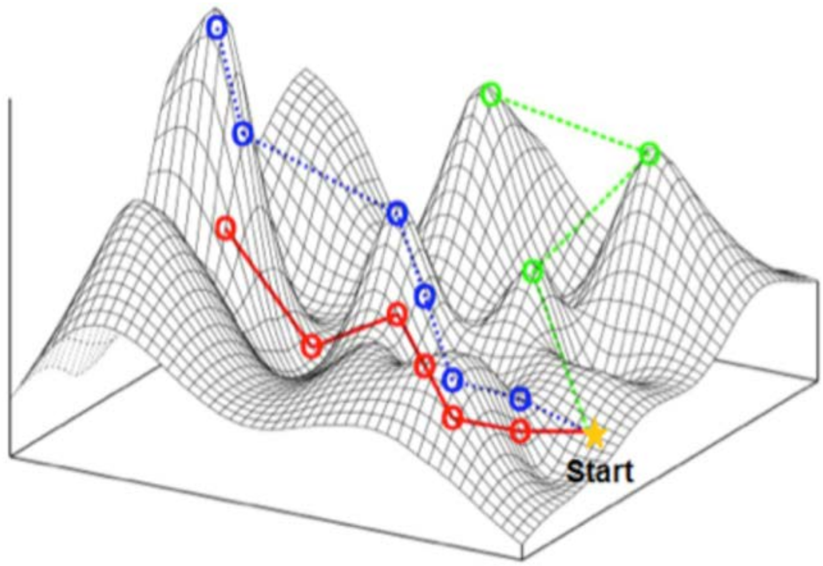


Figure 3. Searching paths from the same start with different strides and directions

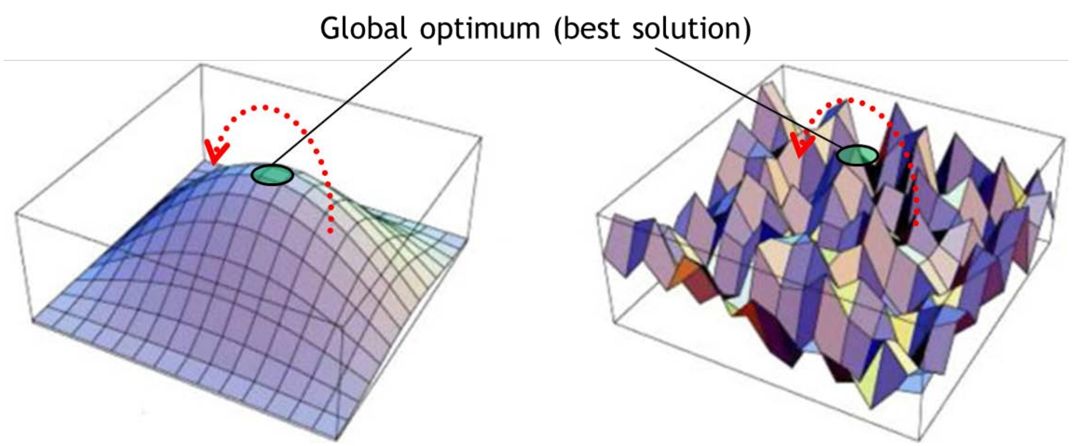


Figure 4. A miss-target random long jump on smooth and rugged landscapes

Adaptive walking and random long jumping are inefficient because they represent single agents' trial-and-error that are limited by individual abilities and disturbed by random chance. Fortunately, we can (and should) make agents collaborate when modeling a social system. In the real life, human beings always seek for advice from and offer advice to each other. In the face of a complex problem, it is very common for us to combine individual insights into a collective solution. Human beings' collective intelligence is created and made use of through social interactions. In the context of parallel problem solving, social interactions take the form of interpersonal knowledge exchange: individual members are allowed to learn from each other's solution at any point of time. In other words, while agents search on the NK landscape, they propagate signals about the areas they have covered through their observable performances. These signals promote long jumping, as people feel safer and are more likely to make big changes if they see others do so and succeed. When Agent A sees that Agent B's solution performs better (i.e., Agent B locates at a higher spot on the NK landscape), it may try to imitate B's solution (i.e., jump to where Agent B is or the neighborhood). This "guided jump" represents an adaptive behavior potentially more efficient than random arbitrary long jumps mentioned earlier in finding local peaks, as an agent can exploit other agents' knowledge about the NK landscape. Thus, an increasing number of guided jumps improve short-run organizational performances. However, rapid settling on the same local peaks prohibits individual agents' further exploration, undermining the organization's long-run problem-solving ability. In terms of the NK landscape metaphor of parallel problem solving, a balance between exploitation and exploration thus means having some agents to search the landscape by guided jumping while having others do so via adaptive walking or long jumping[3] at the same time. Guided jumping is enabled by interpersonal knowledge exchanges, which are essentially social interactions. Thus, an organization's macro social interaction structure can affect the combination of guided jumping and adaptive waling/long jumping by influencing who exchange knowledge with whom

[3]Whether it is an adaptive walk or long jump depends on individual agents.

at what time in what frequency[4]. In other words, an organization's macro social interaction structure impacts the outcome of parallel problem solving by enabling and constraining micro interactions. The next section reviews the development of this idea.

2.3. A Hybrid Macro Network – Closure and Brokerage

Some modeling studies formalized the macro interaction structure as a grid (Levine & Prietula 2011; Miller et al. 2006) and distinguished the micro interactions on the grid as local or distant. Distant interactions lead to rapid convergence of organization-wide knowledge; local interactions only cause the convergence of local knowledge (i.e., knowledge possessed by proximate individuals), preserving global knowledge diversity. Thus, a proper combination of local and distant interactions can enable an intermediate rate of knowledge dissemination across the organization and thus contribute to organizational ambidexterity. Recently, some researchers applied a network representation of the macro interaction structure and investigated its impact on organizational performances (Fang et al. 2010; Lazer & Friedman 2007). This advance is important as it opens a door to incorporating the vast literature on organizational social networks (Zollo & Winter 2002) and organizational social capital (particularly the structural dimension).

A network includes a set of nodes somehow connected by a set of ties. In the network topology of a system, nodes represent system units and ties represent the relations of these units; the way nodes interact with each other represents system functions or behaviors. Network ties can be

[4] For example, if an engineer team in the R&D department would like to test the usability of a new product, whom in the R&D department and other departments (e.g. Marketing and Legal departments) they would collaborate with and how often their meetings would be.

directed or undirected (indicating the direction of interactions) and the values of ties can be dichotomous or weighted (indicating the presence/absence or strength of interactions). Temporal aspects of ties, such as interaction frequency, can be represented by time sequence (Holme & Saramäki 2012) or aggregated into tie weights (also known as tie strength[5]). A network has multiple levels. The lowest level is a node. The highest level is the entire network topology. Between them there are dyads[6], triads[7], and larger sub-structures. Researchers can easily zoom back and forth to focus on different levels (Crozier 1972). Various measures and mechanisms have been developed to characterize network structures and dynamics (Albert & Barabási 2002; Barrat et al. 2004; Boccaletti et al. 2006; Butts 2009; Costa et al. 2007; Newman 2003; Strogatz 2001). Thus, the network topology provides a promising avenue to the study of structural and dynamical complexity of a system.

When the system being studied is a social system such as an organization, the macro network is a social network whose nodes and ties represent individual actors in a social system and their social interactions or relations (Wasserman & Faust 1994). Lazer and Friedman (2007) are the first ones who describe the macro interaction structure of an organization as a social interaction network. In their model of parallel problem solving, individual agents interact with each other via their network to exchange solutions, which leads to guided jumps on the NK landscape, as shown in **Figure 5**. By comparing four types of network topologies, the authors found that the macro network's efficiency in disseminating knowledge influenced short-term organizational problem-solving performances positively but long-term performances negatively. The reason,

---

[5]The two terms will be used in my dissertation interchangeably.
[6]a pair of nodes with or without a tie between them
[7]a triple of nodes and any ties among them

they argued, is that an inefficient macro network can better preserve knowledge diversity, which allows for better exploration in the long run. Fang and colleagues (2010) proposed a special inefficient network, in which the nodes are clustered into multiple subgroups and the connectivity is high inside each group but low between different groups. By systematically changing the degree of subgroup isolation and intergroup connectivity, the authors found that semi-isolated subgroups, with a moderate number of random intergroup links, were associated with the highest organizational performance, whereas very low and very high levels of intergroup connectivity were both associated with lower performances. Densely connected subgroups facilitate local dissemination or exploration of knowledge by providing multiple pathways between different nodes. Thus, knowledge or solutions of group member will be quickly disseminated inside groups. With sparse connectivity between groups, knowledge can still spread across the organization, but the semi-isolation shields local ideas from being extinguished through competition with globally dominant ideas, thus preserving global knowledge diversity and future chances of knowledge exploration.

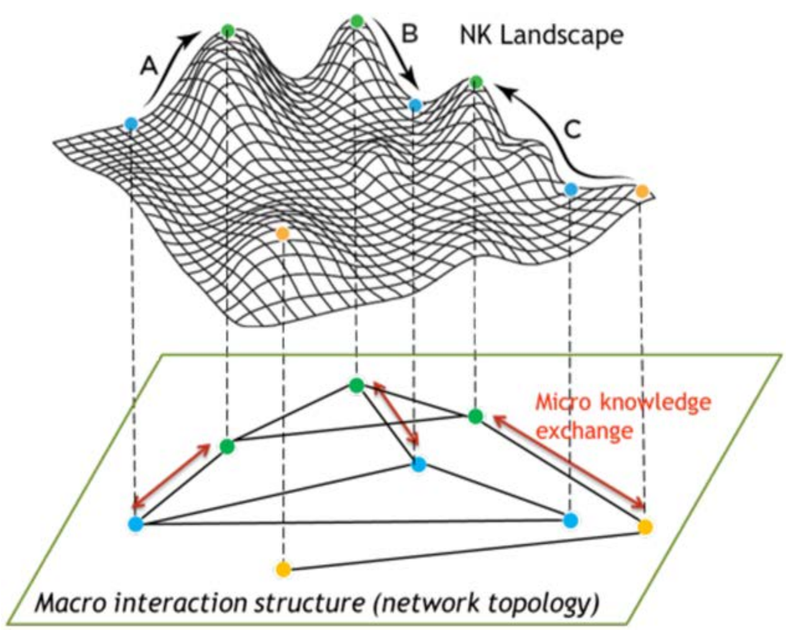


Figure 5. Individuals search on a NK landscape while interacting via a macro network

Empirical studies on social networks and social capital have compatible findings, mostly with regard to how two special network structures – closure (Coleman 1988) and brokerage (Burt 1992) – affect the characteristics of knowledge flows and the distribution of knowledge. Placed in the same network, closure and brokerage respectively refer to dense clusters and sparse areas between these clusters[8] (**Figure 6**). It has been found, using different research methods, that a network with loosely connected dense clusters are associated with higher organizational performance than a well-connected dense network or a fragmented sparse network (Balkundi & Harrison 2006; Cowan & Jonard 2004; Kwon et al. 2007; Mason et al. 2008; Oh et al. 2006; Reagans & McEvily 2008). Below are more specific findings on closure and brokerage structures' complementary effects on knowledge exchanges and knowledge creation.

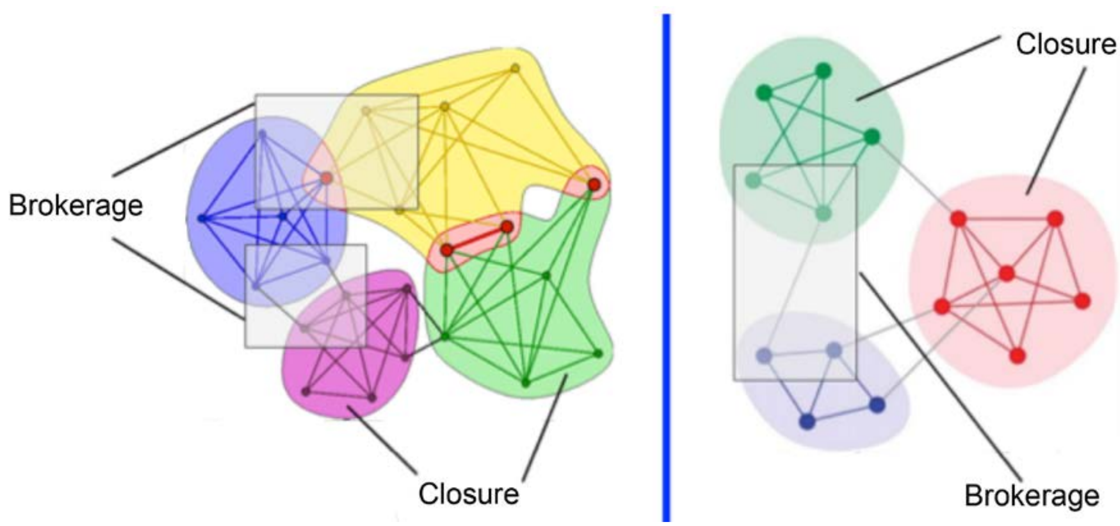


Figure 6. Examples of closure and brokerage

In a closure structure, the nodes (actors) are more connected with one another and thus have more opportunities to interact with one another than with the rest of the network. Frequent and

[8]These definitions emphasize the structural aspects rather than social meanings of closure and brokerage.

intense interactions strengthen social relations and strong ties facilitate knowledge transfer, especially when knowledge is tacit and complex or knowledge sharing behavior is unrewarded (Hansen 1999; Kmetz 1984; Orlikowski 1992; Reagans & McEvily 2003). Frequent and intense interactions also cultivate group cohesion featured by shared knowledge (e.g., jargons, norms, and routines), beliefs, identities, and high levels of trust, reciprocity, sanctioning, and predictability (Coleman 1988; 1990; Granovetter 1985; Krackhardt & Hanson 1993; Raub & Weesie 1990). The solidarity enables and motivates knowledge exchanges among group members, including previously unconnected ones (Bock et al. 2005; Borgatti & Cross 2003; Brown & Duguid 2001; Levin & Cross 2004). However, efficient knowledge dissemination and strong solidarity produce increasingly homogeneous and redundant knowledge within groups (Burt 1992). The strong norms and mutual identification also leads to over-embeddedness and inertia (Gargiulo & Benassi 1999; Gargiulo & Benassi 2000), which can block the inflow of new ideas by reducing the likelihood that people will look for external knowledge or by increasing the transfer costs of external knowledge (Hansen et al. 2005). Consequently, the creativity of group members is constrained because of their limited exposure to new information (Perry-Smith 2006; Perry-Smith & Shalley 2003).

As social interactions mostly happen within groups, there tend to be a scarcity of social interactions between groups, leaving “structural holes” in the network topology (Burt 1992). A brokerage structure comprises a structural hole and a small number of ties or nodes that span the hole and connect otherwise isolated groups (Burt 2000b). These ties and nodes are often referred to as bridges and brokers respectively, whereas intra-group ties are often referred to as bonds. Thus, brokerage structures always coexist with closure structures; bridges connect groups while indicating their boundaries. Since structural holes separate sources of diverse knowledge (i.e., actors from different groups), fresh non-redundant knowledge that can provide insights into unprecedented problems is more likely to flow through bridges than through bonds (Burt 1992; Granovetter 1973). However, knowledge flowing through a bridge is often low in quantity and quality, because the two ends of a bridge tend to have low commitment or liability and limited common understandings (Reagans & Zuckerman 2008). Two individuals connected by a weak bridge are less likely to proactively offer knowledge to each other. They are also less willing to devote time and effort to assist each other during knowledge exchange and creation. A network rich in structural holes facilitates the creation and the novel combination of ideas by exposing people to diverse knowledge domains. Being situated at the broker position, an individual may develop abilities in recognizing the value of new knowledge, transferring a broad range of knowledge to a broad audience, and persuading others to provide knowledge (Cohen & Levinthal 1990; Reagans & McEvily 2003; Rodan & Galunic 2004; Simon 1991). However, the knowledge exchanged between groups is not readily disseminated and exploited within groups (Burt 2004; Fleming et al. 2007). Knowledge exchange between distinct groups in the same organization can be inherently difficult, since involved parties tend to have different interests, perspectives, and languages (Hansen & Nohria 2004). Successful innovation or knowledge creation requires new knowledge to be pushed to all group members, which is better supported by a densely connected group (Obstfeld 2005).

From a social capital view, the above benefits and risks only come into being when individual actors are both capable and motivated to leverage the opportunities provided by corresponding social network structures. In other words, social network structure is a source but not the only source of social capital (Adler & Kwon 2002; Kwon & Adler 2014). Distinguishing between having and using social capital reflects the long-standing appeal for integrating structure and agency in social network studies (Emirbayer & Goodwin 1994; Ibarra et al. 2005; Kilduff & Krackhardt 1994; Kilduff et al. 2006; Stevenson & Greenberg 2000). This integrative view is against rational action theory, which assumes that all individuals are identically motivated by self-interests, and the strong version of formalistic network sociology, which posits motivation as the effect of network structure (Burt 1992). This view has inspired research on cognitive networks (Kilduff et al. 2008), potential and latent ties (Mariotti & Delbridge 2012), and individuals' psychological predispositions (Kalish & Robins 2006; Mehra et al. 2001; Oh & Kilduff 2008; Sasovova et al. 2010). By focusing on organizational members' leveraging of social capital, the current study also takes an integrative view that individual members engage in organizational problem solving as a result of both their intentional actions or agency behaviors and the macro social interaction network they are embedded in. Knowledge exchange is a social process that can be initiated by either side. The initiating side must first know and then determine whom she can acquire relevant knowledge from or whom her knowledge may help, and both sides must be willing or convinced to acquire or provide knowledge and assistance. The macro network provides channels that knowledge can but may not actually flow through. Individuals' motivation and ability also greatly impact the direction, content, quantity, and quality of actual knowledge flows (Aral & Alstyne 2011). As noted, closure and brokerage represent two major conceptualizations of social capital and their key micro-components are bonds and bridges respectively. Thus, we could generally refer to individuals' leveraging of the two types of social capital as bonding and bridging (Adler & Kwon 2002; Putnam 2000; Reagans & McEvily 2008).

2.4. Network Formation Models and the Formation of a Hybrid Network

While previous studies assumed a predefined static hybrid macro network that represents formal organizational structure, the current study is concerned with informal organizational structure that is better described as a dynamically formed macro network. There have been many research efforts on modeling network formation and dynamics (Jackson 2010; Toivonen et al. 2009). Based on their distinct perspectives, methods, and goals, existing network formation models can be generally divided into three types. One type of models contains an often stochastic process, in which one or two micro mechanisms drive tie dynamics. Inspired by the fact that many real-world networks (including organizational social networks) show common macro patterns such as fat-tailed degree distributions, high clustering, and low diameter (Albert & Barabási 2002; Newman & Park 2003; Newman 2001b; Watts 1999), these models are dedicated to account for common patterns by simple micro processes. Classic examples include the ER random network model[9] (Erdos & Renyi 1960), the WS small-world network model[10] (Watts & Strogatz 1998),

---

[9] Random tie creation is the basic micro process of the first network evolution model proposed in 1959 by Erdos and Renyi. The model described the process of growing a *random network*: n nodes connected by m edges randomly selected from all $n(n-1)/2$ possible edges with equal probability p. The generated network has a Poisson degree distribution. The other key feature is a sudden change of the network connectivity with the increase of p: when p is small, many clusters are small and isolated, but once p increases to be larger than a critical value ($1/n$), the network suddenly becomes very dense where almost all the nodes are linked to each other in a giant connected component.

[10] The small-world network originated from an experiment of Milgram (1967), in which selected persons were asked to deliver a letter to a target receiver by only passing the letter to their acquaintances. The average length of successful communication chains was short, around six steps. The phenomenon is well known as “small-world effect” or “six degrees of separation”. In a small-world network, the distance between a random pair of people is smaller than expected, implying that an individual is close to most of her friends but may also have a few distant friends. The WS model was designed to reproduce the small-world phenomenon by rewiring each link in a regular network with a probability p. When $p = 0$, the network is fully ordered; when $p = 1$, every edge is rewired so the generated network is a random network; when $0 < p < 1$, we obtain a small-world network with small average shortest path and high clustering coefficient.

and the BA scale-free network model[11] (Barabasi & Albert 1999). The second type of models are statistical models that try to predict the more-than-chance presence or absence of ties or network structures based on specific features in social network data sets. These models incorporate specific local structural patterns and characteristics of nodes and ties to work with empirical data. Two important classes of statistical models are Exponential Random Graph Models (ERGM)[12] (Snijders et al. 2006; Wasserman & Pattison 1996) and community detection models[13] (Abell et al. 2008). Both random and statistical models rely on tie dynamics (i.e., creation, change, and dissolution) and probability theory to handle the real-life complexity caused by individual heterogeneity. Random models are less powerful than statistical models with regard to describing and distinguishing various possible reasons for the same network dynamics, but the simplicity of random models makes them very powerful in explaining the process of network formation and evolution. The third type of models exist in the economics literature and see network formation as the outcome of agents' choices of relationships (Bala & Goyal 2000; Buskens & Van de Rijt 2008; Dutta & Jackson 2003; Galeotti et al. 2006; Goyal & Vega-Redondo 2005, 2007; Hummon

---

[11] A scale-free network has a power-law degree distribution, commonly seen in many real-world networks. Highly unbalanced degree distribution indicates that, in a large group of people, only a few are extremely popular and most others do not have too many contacts. The BA model was the first to generate a scale-free network with two simple mechanisms: continuously adding new nodes into the system ("growth") and connecting with other nodes with preference to the high-degree ones ("preferential attachment").

[12]ERGM is designed to statistically analyze how specific structural patterns and node/tie attributes interdependently affect the existence probability of a real social network or a particular tie in the network. As statistical models, ERGMs are very useful for identifying "more-than-chance" patterns and significant correlations, answering questions such as "are networks with a specific pattern more likely to appear than networks without this pattern?" ERGMs can correlate network dynamics with multiple confounding patterns/tendencies and estimate their different influential strength (via coefficients). However, they are deficient in disentangling causal relations. MCMC (Markov Chain Monte Carlo) techniques are usually used to estimate ERGMs.

[13] These models intend to detect natural underlying communities in a specific network and they usually follow certain definitions of communities. One class of definitions, represented by normalized cut (Shi & Malik 2000) and conductance (Kannan et al. 2004) rely on the normalized number of edges falling between communities to quantifying the profoundness of community separation in a network. The second class of definitions, represented by modularity (Newman & Girvan 2004) and surprise (Aldecoa & Marín 2013), quantify the extent to which a network displays community structure by comparing the network with a random network with similar properties. The third class are node similarity measures, with an underlying assumption is that communities are groups of nodes similar to each other.

2000; Jackson 2003; Jackson & Watts 2002; Jackson & Wolinsky 1996; Watts 2001). These models assume that each individual agent chooses relationships to maximize her payoffs or utility[14], which is a function of the emergent network topology. Among the three types of models, statistical models are mostly tie-based with some exception[15] (Snijders 1996, 2001; Snijders et al. 2010). Random models and economic models are both agent-based but have opposing assumptions on the rationality of agents[16]. It is possible to create agents with bounded rationality (Simon 1957) by integrating random and economic models. For example, in the current study's model, an individual agent's behavior is neither the result of rational choice nor caused by pure chance. Instead, both the agent and other agents' individual propensities, the agent's position in the macro interaction network[17], and randomness all have impacts.

The rest of this section will focus on random models whose micro processes reasonably mimic social processes and give rise to a combination of closure and brokerage structures, i.e., a hybrid network. Previous studies (Fang et al. 2010; Lazer & Friedman 2007) represented the hybrid network was by a small-world network, which is known for its high clustering and a few cluster-spanning bridges (Newman 2001a; Watts 1999). The first small-world network formation model randomly rewired the ties of a same-sized cluster network to build bridges (Watts 2003; Watts &

---

[14]These models often incorporates game theoretic techniques. Utility is different from payoffs in a game. An individual player can increase its utility by helping other players increase their payoffs while compromising its own payoff. In other words, individual utility may increase with collective rather than individual payoffs. However, every move an agent takes is to increase its own utility no matter how that is defined.

[15] This model is known as the stochastic actor-based model. It is a Markov-process based network formation model. Unlike ERGMs, the stochastic actor-based model represents changes in the network as the collective outcomes emerging from individual choices, which depends on individual agents' attributes and structural utilities.

[16] To reduce model complexity, modelers "either consider worlds composed of remarkably prescient and skilled agents or worlds populated by morons" (Miller and Page 2007). Economic models assume that individual agents are completely rational and act towards maximizing their gains, whereas random models assume that the behaviors of individual agents follow the same law and only differ by chance.

[17] For example, a well-connected agent are more likely to respond to rather than initiate knowledge exchange requests, as there tend to be many invitations from others.

Strogatz 1998). Later in small-world network formation models that has a social context, the random rewiring was replaced by two micro generative mechanisms – a pair of nodes meet randomly or via common contacts (Davidsen et al. 2002; Jin et al. 2001). If the two nodes are not already connected, the meeting will tie them up; otherwise, the meeting will reinforce their existing connection or do nothing.

Linking via common contacts is a typical micro mechanism in social networks known as triad closure: if two currently disconnected nodes in a social network have a common third-party contact, then they are likely to connect at some point in the future. This hypothesis was originally proposed by Simmel (1980) and confirmed by longitudinal studies on social networks of different types and sizes (Hammer 1980; Kossinets & Watts 2006; Leskovec et al. 2008). Moreover, the probability for two nodes to meet with each other increases with the number of their common contacts (Kwon & Adler 2014; Newman 2001b), as illustrated in **Figure 7**. Because of triad closure, social networks exhibit a unique feature of high clustering and dense areas tend to self-sustain or become even denser (Grabowicz et al. 2012; Newman & Park 2003).

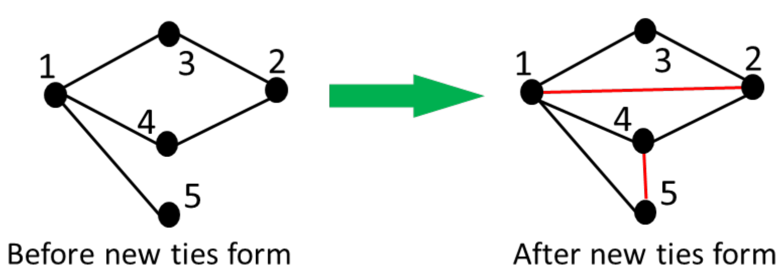


Figure 7. Examples of triad closure[18]

[18] If triad closure is the only force that impacts tie formation, then a new tie is more likely to form between 1 and 2 than between 4 and 5. Node 1 and 2 share two contacts (Node 3 and 4), so connecting 1 and 2 will close two triads (123 and 124). Node 4 and 5 only have one common contact Node 1, so connecting 4 and 5 will close one triad

As nodes link to each other via common contacts or by chance, the network gets denser and eventually each node may connect with every other node. However, real-life social networks usually reach a saturated state way before becoming fully connected, because human beings' finite cognitive and communication capacities limit the number of social ties we each can maintain active, with or without the help of today's advanced information technology (Dunbar 1992; Gonçalves et al. 2011; Leskovec et al. 2009; Miritello et al. 2013). Thus, another typical micro mechanism in random network formation models is tie removal. At each step of the process, with some probability, either a set of ties are randomly selected and deleted, or one node's incident ties are deleted or replaced by a single random tie. When triad closure coexists with (infrequent and conservative) tie removal, well-connected clusters in the network can self-sustain: if a tie in the cluster is accidentally deleted, the two involved nodes have so many mutual friends that they will soon meet again and reconnect. None of the preceding models distinguish between dynamics of the network and dynamics on the network (i.e., interactions), which may bring in more tie dynamics, such as tie availability and changes in tie strength.

2.5. Self-Organization and Complex Adaptive Systems (CAS)

Organizational researchers have been interested in self-organization for quite some time (Anderson 1999; Contractor 1994; 1999). Regarding the increasingly complex and turbulent business environment as well as the more and more autonomous and heterogeneous employees, an organization's ability to leverage individual members' self-organization can be the key to

(145). The probability that a new tie will form between 2 and 5 is even smaller, as the two nodes have no common contact.

success. However, traditional classical mechanics-based quantitative methods and tools cannot provide all the help organizational researchers need in order to develop or test self-organization related hypotheses. The main reason is that traditional methods and tools are essentially reductionist: they radically simplify research subjects for analytical tractability.

One typical way of simplification is to focus on representative or assume homogenous properties and behaviors. When the research subject is a system, system components are assumed to be infinitely many, infinitesimally small, or indistinguishable from one another. As noted earlier, organizations have a limited number of members. A primary mechanism of self-organization – variation – is premised on individual heterogeneity. Self-organized micro interactions also tend to produce macro quantities whose probability distributions have no representative value (Barabasi & Albert 1999). Another typical way of simplification is to divide and conquer: the research subject is dissected into smaller isolated parts, each of which is investigated separately. It requires negligible or linear interactions between dissected parts, so that partial findings can add up to explain the whole. Interactions that lead to self-organization are usually too complex[19] to meet such requirement. When the research object is a dynamic system, most traditional methods and tools focus on equilibrium states rather than the processes that lead to these states[20].

---

[19]Self-organization is one of the defining properties of complex systems. Different components of such systems interact with one another in a highly interdependent and often non-linear way; as a result, the whole is usually more than the sum of parts (due to extra complexity).

[20]Admittedly, some traditional analytical approaches are intended to deal with the transition between states. For example, the Markov process is a stochastic model defined by a finite number of states and transition probabilities for moving between these states. This process always converges to a unique distribution over states. Thus, what happens in the long run will not depend on where the process started or on what happened along the way; instead, it will be completely determined by the transition probabilities – the likelihoods of moving between the various states. In other word, the Markov process is often used to model a random system that changes states according to a transition rule that only depends on the current state. The Markov model are inadequate for understanding self-organization because they are premised on a fixed set of states and constant transition properties and assume away the effects of earlier states including the initial condition. Every time step especially the initial condition is important

However, the most interesting part of a self-organization phenomenon is the intermediate organizing process, i.e., how decentralized autonomous micro interactions give rise to adaptive macro structure without exogenous interventions (Kauffman 1993). Self-organizing processes often exhibit significant time lags, discontinuities, or thresholds that are challenges to traditional analysis methods/tools.

Organizational researchers have been applying theories of complex systems and methods to compensate for the deficiencies of traditional reductionist methods (Anderson 1999; Dooley 1997; Maguire et al. 2006). In recent years, a specific branch of complexity science named Complex Adaptive System (CAS) has been developed and used as a powerful analytical tool. A CAS is a multi-agent system characterized by decentralized adaptation of individual agents, interaction-based organization of individual adaptations, and emergent complexity of the self-organizing process. Self-organization refers to the emergence of structural or dynamical patterns at the macro level of the system, which represent system-level adaptations to the environment. The self-organizing process in a CAS entails three mechanisms – variation, interaction, and selection (Anderson 1999; Brownlee 2007; Dooley 1996; Gell-Mann 1994) [21] – and a fourth mechanism reproduction if the CAS is a social system (Fuchs 2003; Mingers 2004). These mechanisms are elaborated below.

First, micro-level variations serve as the source of and provide opportunities for adaptations. While earlier research on physical or chemical systems focuses on variations and ensuing

---

for understanding a self-organizing process. In addition, constant transition properties imply a static rather than a changing interaction structure.

[21] The terms may vary in different studies.

adaptations caused by exogenous factors (Prigogine & Stengers 1984); later studies on biological or social systems argue that the stochastic and creative nature of endogenous micro behavior[22] can naturally produce ongoing variations, which eventually lead to adaptations (Holland 1995, 1998; Kauffman 1993). Second, unpredictable heterogeneous micro mechanisms are selected by the present macro interaction structure, which act as established system order. Although each individual agent decides whom it will interact with, the options are constrained: the present macro interaction structure indicates whom in the population each agent is more likely to interact with than with the rest of the population (Strogatz 2001). In other words, the present interaction structure tends to support compatible micro interactions while inhibiting incompatible ones. This self-reinforcing tendency explains why the macro interaction structure (and existing interaction patterns) remains relatively stable and why the macro level does not change as frequently as the micro level. Sometimes the extant selection power can even stifle the adaptation and evolution of the entire system[23]. Third, since micro interactions continuously and collectively contribute to the macro interaction structure, new macro patterns can still emerge and sustain. Given that some micro mechanisms are more prevalent than others in the system, current interaction patterns reflect previous dominant micro mechanism(s). Because of the self-reinforcing tendency discussed above, the effect of any new micro mechanism that is not currently dominant (e.g., deviated or adapted individual behaviors) is likely to be dampened. In order for a new mechanism to take over, its macro-level effect must overcome the dampening effect, i.e., must change the present macro interaction structure. The updated macro structure will then reflect this

---

[22]The stochastic nature means that there is always a chance for individual behaviors to accidently deviate from normal patterns. The creative nature means that individuals can intentionally come up with new better options.

[23]This phenomenon is known as "complexity catastrophe" (Kauffman 1993). It means the external force of selection (via payoff or utility) that is supposed to drive system adaptation and evolution is suppressed by the internal order of a complex system that reflects statistically typical properties or central tendencies of the majority.

new mechanism and amplify it by influencing individual agents' future interaction decisions and behaviors. This micro-macro feedback loop keeps rotating as micro adaptations continuously generate new promising mechanisms. To understand the structural and dynamic complexity of this loop, a network representation of the macro structure has been suggested (Mingers 2006). Last, the self-organizing process in a social CAS is history-dependent (Deppa 2014). As mentioned earlier, current macro interaction structure grows out of old structures rather than start from scratch. In addition, individual decisions are unavoidably affected by their past experiences, and individual resistance to change is natural and common in social systems (Argyris & Schön 1974, 1978; Piderit 2000; Zander 1950). Thus, previous dominant micro-mechanisms may still exist in the system, even though their numbers and impacts are fewer than before. Individual behaviors are history-dependent, but they are not history-determined. As discussed earlier, unpredictable even random variations happen all the time. Since these variations are the sources of new powerful micro mechanisms, the latter are difficult to predict too. The current macro structure may be weakened by several new micro mechanisms, each representing a possible future direction of the system. Thus, when the system is lacking in order due to the dwindling macro structure, it is uncertain which micro mechanism will take over and which direction the system will head for.

A social CAS reproduces from an initial condition in a path-dependent yet unpredictable manner. This process is often visualized as a tree-like map of all available evolutionary paths (**Figure 8**). At each intersection point of the tree (also known as bifurcation point), there are two or more branches stretching out, each indicating a possible next state of the system. A set of subsequent branches constitutes a possible evolutionary path of the system (e.g., 1 → 2 → 4). If the system is at Branch 2, its next state is either Branch 4 or Branch 5 but will never be Branch 6 or Branch 3, indicating path dependence. When a system reaches a bifurcation point, it is like staying on a razor's edge. Even slight external disturbance can push it down any evolutionary path available at that moment. Since there are so many unknown factors (what disturbance there may be and how the system may respond), the exact next state of the system is usually unpredictable.

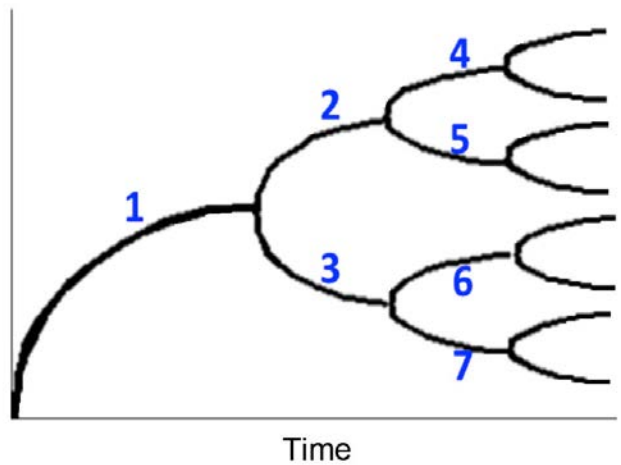


Figure 8. CAS evolutionary paths

In summary, a CAS is characterized by iterative micro-macro feedback loop that enables the system to self-organize and change as a whole. This signature feedback loop is supported by three mechanisms – variation, interaction, and selection. The system's macro structure impacts but not determines the behaviors of individual agents, so behavioral variations continuously occur at the micro level and spread in the system through the interactions of individual agents. Since the macro structure enables and constrains the interactions of agents, it "selects" different variations by affecting their dissemination. Variations that are widely spread collectively change the macro structure. Thereby, the feedback loop is iterative and the system keeps changing at the micro and macro levels. In a social CAS, the macro structure includes the system's history; so two subsequent states of a changing system have similar lower-level components, although they may be proportioned differently in the two states.

2.6. Agent-Based Modeling and Computer Simulation Theory Development

Originated from early studies in cellular automata (Wolfram 1994) and artificial life (Langton 1995), agent-based modeling (ABM) is now widely used in social science (Bonabeau 2002; Epstein 1999; Macy & Willer 2002; Smith & Conrey 2007) and organizational research (Burton & Obel 2011; Miller & Lin 2010). The ABM name itself implies the core of this method and its strong connection with CAS. ABM allows to investigate (often unexpected) phenomena of interest as emergent patterns of a dynamic system, which consists of multiple agents interacting with one another based on predefined rules in a predefined environment (Gilbert 2008).

The basic units of an ABM are agents, which nowadays are commonly implemented as objects in object-oriented programming (OOP) languages or by toolkits based on OOP languages. OOP distributes and encapsulates a system's data and operations in different objects as their attributes and behaviors. The state of an object is indicated by the values of its attributes, whose changes can trigger different behaviors of the object. An object can manipulate its own attributes and influence other objects' attributes by interacting with them; the latter means that an object can react to external stimuli from the environment including other agents. During the interaction, an object's behavior is visible to others, while its attributes are usually not. Thus, we can endow individual agents with the following characteristics and control their distributions in the agent population:

- Autonomy: agents choose their own actions based on internal rules and self-goals. There is no central or top-down control.
- Heterogeneity: agents have different identities. They can behave differently given different internal or external states with certain probabilities.
- Adaptability: agents have memory and can learn from the results of their past actions.

These characteristics are essential for a self-organized social CAS but often underrepresented or ignored in traditional analytical models [24]. ABM, however, can incorporate them in an individualized straightforward way.

Besides designing the attributes and behaviors of agents, building an agent-based model require specifying the environment (or context)[25]. Not only does the context provide a global view of the macro patterns emerging from agent behaviors and interactions, it enables the emergence in the first place. Usually each agent only interacts with its local context including other agents inside that area. Thus, when model a CAS, the global context defines both the interaction structure and the external environment of the CAS. The former impact the evolutionary path of the system endogenously by defining, among all agents, who interact with whom in what way; the latter impacts system dynamics exogenously by driving the adaptation of the system. Researchers can define the global context (from the simplest to the most complex) as a set of parameters, a generic data structure (such as vector or matrix), a complex topology (such as grid or network), a sub-model, or some kind of combination. Depending on the purpose of modeling, the global

[24] To reduce complexity traditional models either assume that individuals have a high level of homogeneity (typically for mathematical models) or aggregate/average individual characteristics (typically for statistical models).

[25] In order to distinguish the environment of an agent-based model with that of a CAS, I will refer to the former as "context" hereafter.

context can be static, altered by the adaptations and interactions of agents, or changed by external/temporal triggers[26]. Today almost every ABM platform allows researchers to observe the structural or behavioral characteristics of the global context graphically and to measure these characteristics quantitatively.

Finally, the agent-based models described here are discrete and stochastic. The term discrete refers to the treatment of time. In the real world events may happen continually and concurrently, but in agent-based models time is usually modeled as discrete steps and things happen at each step are executed by the computer one by one[27]. A stochastic model has inherent randomness and thus will not produce the same outputs when repeated with the same inputs. The stochasticity of agent-based models takes four major forms: (a) the probability distribution of some individual attributes or behaviors, (b) the order agents follow to take action at each time step, (c) random chance that impacts the interactions between agents, and (d) the decisions each agent makes at each time step[28]. In this regard, stochasticity is indispensable for modeling the complex individual heterogeneity in the real world (Gilbert 2008).

Computer modeling for theory development (Davis et al. 2007) is an iterative process consisting of model construction, validation and analysis. Various tasks of this process are highly intertwined (**Figure 9**) to serve the purpose and assure the rigor of modeling. The process starts

---

[26]For example, Miller et al. (2008) modeled the environment as a vector whose value randomly changes every 200 time steps.
[27]ABM simplifies time in this way not because computers can only execute one piece of code at a time. Computers that have multiple processors and thus can execute multiple threads are very common nowadays. The primary reason is that it is very difficult to understand the model results when many agents execute many actions simultaneously.
[28]as they are influenced by both random chance and the stochastic states of related agents (the focal agent and other agents interacting with the focal agent)

by constructing a conceptual model of a real-world social system. Basically, some features of the system are abstracted (usually including some independent and dependent variables and their relations) and translated into model elements and logics. To ensure model validity, it is suggested that researchers draw on existing theories and/or prior knowledge (Carroll & Harrison 1994). When prior theories and knowledge are incomplete, ambiguous, unquantifiable, or simply unavailable, speculation is inevitable and assumptions are often made wherever needed. When there are too many competing theories and contradictory empirical evidence, researchers need to decide what they would like to incorporate into the model. In general, the conceptual model should capture the essence of the research problem while keeping simplicity. When the purpose is to explore new theories or explanations (as opposed to make prediction or pursue analytical accuracy), the power of a model often comes from its simplification of reality (Harrison et al. 2007).

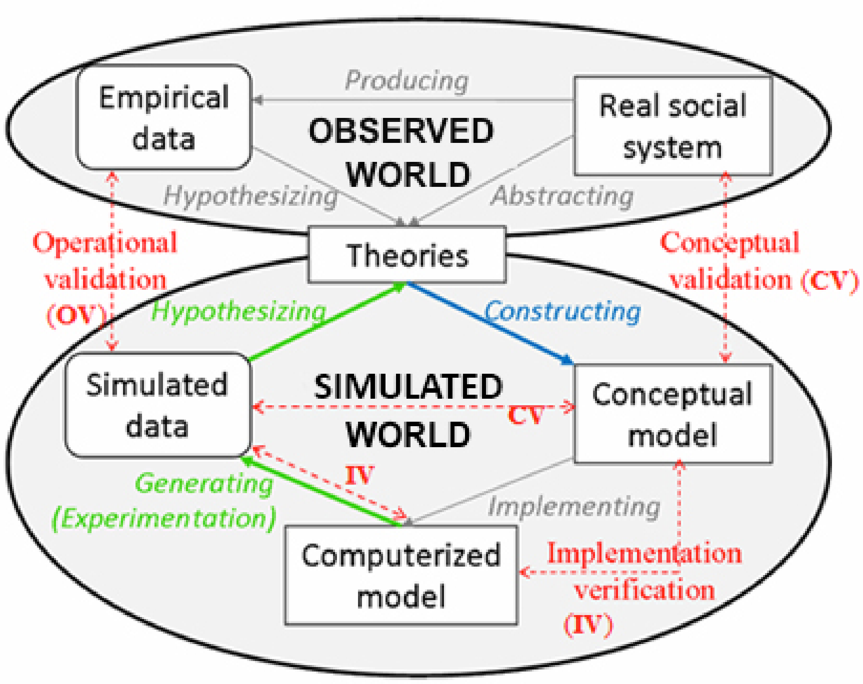


Figure 9. Computer modeling as a theory development process

(Modified from Sargent 2005)

The conceptual model is then implemented by a computer program (the computerized model), with explicitly defined parameters and outputs representing independent and dependent variables of the research question. Model parameters are primary but not the only model inputs. Other inputs include initial conditions of the model and artificial accessories added during implementation only to make the computerized model work or have some desirable feature such as efficiency (Galán et al. 2009). The effects of these non-primary inputs on model outputs may and often confound with the effects of primary model parameters. Implementation verification, a verification of the computerized model, thus has two purposes. One purpose is to make sure the conceptual design is correctly and completely coded with minimal artificial accessories. For this purpose, the extreme condition test is often conducted to assess whether the computer code works as expected under a variety of extreme conditions that may occur during simulations, and it often leads to the optimization of codes. The other purpose is to identify non-primary model inputs, whose effects need to be examined during model analysis. ABM researchers tend to expect interesting and surprising outcomes from their models, but chances are such outcomes are completely artificial and have nothing to do with the real social system being modeled. Thus, agent-based models need more rigorous implementation verification than other computational models.

The next step is model analysis, or model behavior exploration if little is known about the real-world phenomenon being modeled. Computer modeling is deemed an effective alternative to traditional statistical and inductive case methods when the real-world phenomenon involves multiple unknown or ambiguous variables and processes that interact with each other in a nonlinear and dynamic way. By modeling probable variables and processes and then running simulation experiments on the model, researchers try to identify significant relations underlying the real-world phenomenon by analyzing the transformation from primary model inputs (experimental conditions) to outputs (experimental results). Given a significant relation identified from the simulated data, how much can we say about the real-world relation? To answer this question, researchers rely on uncertainty analysis (UA), sensitivity analysis (SA), and/or robustness analysis (RA), which quantify the degree of confidence and the boundary conditions of their findings. Ideally the variations in model outputs (referred to as model uncertainty in UA and SA[29]) are supposed to be caused by only variations in primary model inputs, but model construction and implementation usually bring in unwanted variations or noises. In agent-based models some common sources of noise[30] include (but are not limited to) the initial condition of simulation, the scale of micro-to-macro aggregation, timing of observation, and stochastic elements[31]. To accurately evaluate the significance of primary model inputs, we need to reduce noises or take them into consideration while measuring model uncertainty.

Reducing noises usually takes two steps. First is to collect multiple instances of certain model outcomes given different values of a noise factor and the second is to use the representative value or central tendency (e.g., mean or median) of these instances for data analysis. The main challenge is to figure out how many instances should be considered to obtain the representative value: a larger number of instances improve accuracy but their collection needs more time and

---

[29]UA is concerned with the overall uncertainty in model outputs. SA focuses on the contributions of individual model inputs to the uncertainty shown in model outputs, with an attempt to evaluate the relative importance of different model inputs (primary or non-primary).
[30]referred to as noise factors hereafter
[31]The specific value of a stochastic element in each execution run of the model is randomly generated by a pseudo-random number generator based on the elements' probability distribution.

resources. A reduction approach is preferable when the noise factor meets three conditions: (a) it is known to affect the variability but not the central tendency of model outputs; (b) it only changes the effect size of any primary model inputs; (c) the noise factor value is easy to change. The variations of model outcomes caused by time and inherent stochasticity[32] are often reduced rather than explicitly analyzed: the former is reduced by collecting model outcomes when they are relatively stable; the latter is reduced by averaging the results of replicate runs. If any of the three conditions are not met, or if the relation between the noise factor and the model outcome is well formed[33], it is preferable to estimate the amount of noise. A common approach is to integrate the noise factor into the analysis model, so we can explicitly examine to what extent the relations between primary model parameters and model outputs are affected by the noise factor. This type of analyses, as they focus on how robust the effects of primary model inputs are against noise, is known as robustness analysis[34]. Fundamentally, robustness analysis tests a null hypothesis that there are no interactions between primary model inputs and noise factors.

The results of simulation experiments (including the results of UA/SA/RA) provide internally valid evidence that can verify/falsify existing theories/hypotheses or bolster new ones. Based on these results, researchers can refine prior theories or hypotheses by adding significant variables (e.g., previously unknown moderators or mediators) to or removing insignificant ones from the conceptual model. The computerized model can also be improved by changing initial conditions, distributions of stochastic elements, or the value ranges of some model inputs. Model refinement

---

[32]In stochastic models, the use of pseudo-random number generators can produce different simulation results despite the use of identical input values.
[33]We can formalize the relation as linear, monotonic, or quadratic.
[34]RA examines whether the relations between primary model inputs and model outputs remain significant despite changes in non-primary model inputs and implementation details.

is an iterative and continuing process until the model shows the following features: (a) the variations in primary model parameters explain most of the overall variations in model outputs (i.e., model uncertainty), (b) model outputs are robust to other variations such as those originated from initial conditions and stochastic elements. Only at this point can researchers be confident that the relations between specific model parameters and outputs somewhat reflect the relations between real-world independent and dependent variables.

Operational validation further regulates the simulated model by comparing the data obtained from simulation experiments with empirical data produced by the real social system. The major task of operational validation is to calibrate model inputs, initial conditions, and stochastic variations so that model outputs are as close as possible to existing empirical data, which can be either quantitative or qualitative (e.g., stylized facts, statistical regularities, and behavioral signature[35]). Conceptual validation and implementation verification are also about model validity, but they rely on existing theories (including their deduction) or previous versions of the model, neither of which is the real-world social system itself. Every theory in empirical science has boundary conditions beyond which it is inapplicable or simply wrong. In addition, the validity of a theory or a theorizing process has various meanings in social science (Campbell & Stanley 1963; Cook & Campbell 1979; Feldman & Arnold 1983). Thus, it is important to directly compare a simulated model and the reality (Edmonds & Moss 2005). But we should also avoid over-fitting a model to specific data for both theoretical and practical reasons (Fagiolo et al.

[35] In social sciences especially economics, a stylized fact is a simplified presentation of an empirical finding (typically on some macro-level phenomenon). It is a broad generalization which is essentially true despite inaccuracies in the detail (Fagiolo et al. 2007). An example of statistical regularities would be the heavy-tail shape of a power-law distribution, which has been widely observed as the distribution of many complex-system quantities. An example of behavioral signature would be discontinuous shifts (a.k.a. phase transitions or abrupt changes) in system behavior.

2007)[36]. After all, for complex systems, empirical data by itself is insufficient to understand system behavior (Bar-Yam 2013).

2.7. Perspective Remarks and Further Study Guide

Previous studies revealed a nonlinear relationship between an organization's long-run performance in solving an external problem and the speed with which organizational members' individual solutions are disseminated across the organization (by interpersonal knowledge exchange or by updating organizational knowledge). Generally, superior individual solutions coming up during the problem-solving process should be disseminated to the entire organization, but in a relatively "inefficient" manner (i.e., neither too fast nor too slowly), so that organizational members' exploitation of these solutions will not stifle their exploration of even better solutions. Recently researchers looked into the possibility of coordinating micro problem-solving behaviors through a macro social interaction network. The network topology enables some knowledge exchanges while constraining others. They found that a hybrid network topology with both closure and brokerage sub-structures can support an intermediate rate of knowledge (i.e., solution) dissemination and thus contribute to high organizational performance in the long run.

Despite previous findings and advances, a specific question, among others, may rise: how would

[36]For example, there are many potential influential factors for which data do not currently exist; some may not be amenable to quantitative measurement. Also, the quality of available empirical data is not guaranteed, in that the data collection process is guided and thus inherently biased by existing theories and that some complex phenomena rarely happen (probably only once). Lastly, when the phenomenon under study is sensitive to initial conditions and/or time, it is necessary (and very difficult) to determine the appropriate starting/ending points and sample segments in the (longitudinal) empirical data.

the overall organizational problem-solving performance be affected by organizational members' autonomous rather than regulated problem-solving behaviors, that is, knowledge exchange and creation by leveraging social capital? The first step towards answering this question would be modeling the coevolution of the macro interaction network and the micro interactions on the network. According to the CAS theories, this micro-macro coevolution hinges on an iterative feedback loop between micro behaviors and the macro structure. This loop can be created via three dynamic mechanisms – interactions of individual agents, variations in individual behaviors, and selections by the macro interaction structure. Given the previous findings on the effect of a hybrid network topology, the next step could be to investigate the conditions under which a hybrid interaction network resistant to micro variations can arise from organizational members' knowledge exchange interactions. In addition to the macro network topology, several other factors may also impact the speed with which superior individual solutions are created or disseminated, such as organizational size, problem complexity, the accuracy and frequency of interpersonal knowledge exchange and organizational members' independent knowledge creation abilities. The main and interactive effects of these factors should also be investigated in the further studies.

Chapter 3

MODEL

This chapter introduces the agent-based model developed in the current study. It starts with the conceptual model, which theorizes the organization as a social CAS system characterized by the coevolution of organizational members' autonomous problem-solving behaviors and a macro network emerging from interpersonal knowledge exchanges. Micro generative mechanisms of the model were designed around the generation and maintenance of a hybrid network. Next, a computational implementation of the conceptual model was specified, in which earlier identified micro generative mechanisms were formalized and associated with one or more model parameters. Thus, the main and interaction effects of different micro generative mechanisms on organizational problem-solving performances were tested through simulation experiments on these parameters, which are described in the next chapter.

3.1. Conceptual Model

The agent-based model of the current study was intentionally aligned with the models of previous studies in several aspects to make the results comparable. First of all, organizational problem solving was modeled as individual agents searching an NK landscape in parallel and exchanging knowledge with one another. Secondly, individual agents' selections of knowledge exchange partners were impacted by a macro interaction network. Thirdly, organizational problem-solving performance at a specific time was the average individual performance (i.e., the average value of all individual solutions) at that time. Despite these similarities, the model of the current study was essentially different from previous models in its emphasis on the coevolution of organizational members' autonomous problem-solving behaviors and a macro interaction network emerging from interpersonal knowledge exchanges. The micro-macro coevolution was modeled using complex adaptive system (CAS) theories, which suggest that three micro mechanisms – variation, selection, and interaction – can jointly produce an iterative micro-macro feedback loop and therefore drive the coevolution. Specifically, an organization was simulated as a social CAS and organizational members as individual agents who interact to exchange knowledge. The macro interaction network represented the macro structure. Individual agents' problem-solving behaviors (independent knowledge creation or knowledge exchange) collectively and gradually change the macro network via tie dynamics. Meanwhile, individual behaviors were influenced by individual propensities for social capital (i.e., whether or not use social capital and which type of social capital to use), individual agents' network surroundings, and randomness. This section describes primary micro generative mechanisms of the current study's model: how they were identified from the building blocks of a hybrid network, how they fit into the CAS theoretical framework, and how they support the micro-macro coevolution together.

To start with, the defining feature of a hybrid macro network, as noted earlier, is the concurrent presence of two structures – closure and brokerage. An examination of their lower-level building blocks[37] (**Figure 10**) reveals that there are only two types of ties: (a) bonds that exist within and

[37]That is, their building blocks at the nodal, dyadic, and triadic levels. A triad is a triple of connected or unconnected nodes. A dyad is a pair of connected or unconnected nodes. The disintegration ends up at the dyad rather than the

maintain a closure structure (e.g., DE); (b) bridges that cross over a structural hole and connect nodes from different closure structures (e.g., BC). Bonds and bridges represent intra-group and inter-group social relations respectively. Different combinations of bonds and bridges resulted in two types of triads. Bond triads, if fully connected (i.e., with three ties), had no bridge (e.g., DEF). Bridge triads, if fully connected, had one to two bridges (e.g., DEG and ABC)[38]. Bond triads exclusively constituted the closure structure, whereas the brokerage structure had at least one bridge triad.

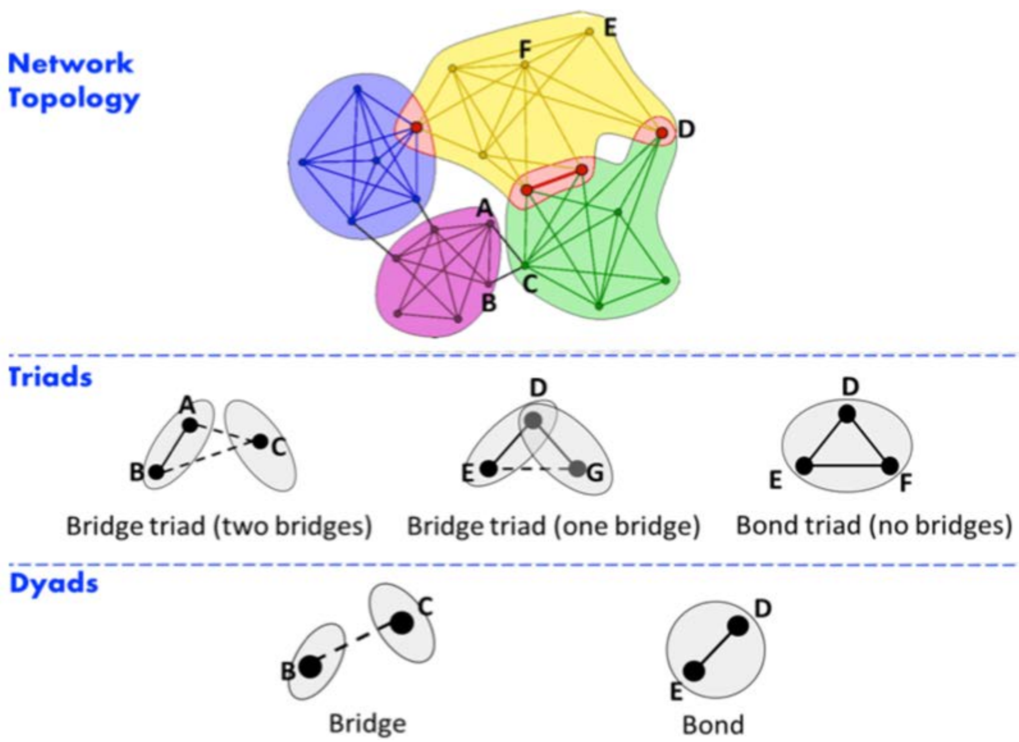


Figure 10. The characteristic microstructures of a hybrid network topology

---

nodal level because the current study does not consider nodal changes. Organizational members are assumed to have stable properties (e.g., ability of innovation, accuracy in knowledge exchange) and behavioral preferences (e.g., whether to exchange knowledge, exchange knowledge with whom) during the problem-solving process.

[38]Assume that a closure structure has at least two nodes.

The next step was to identify micro mechanisms that can generate or change these building blocks. Due to the dynamic complexity of the micro-macro feedback loop, simply pairing bonds/bridges with the changes of ties (e.g., creating a bond or reinforcing a bridge) may not work. First, a tie can change its role as a bond or bridge. Group merging turns all bridges into bonds, while group splitting turns some bonds into bridges. Second, the next role of a tie is path-dependent. A bridge in the middle of group merging will become a bond, but it will remain a bridge if the two groups are further splitting. Third, micro dynamics in a CAS are the results of individual actions. When it comes to modeling, defining bonds and bridges from an individual perspective is difficult; let alone that each agent has a different network position. Alternatively, we can take advantage of the fact that a tie in a hybrid network is more likely to be a bond than a bridge. There are more bonds than bridges in a hybrid network because the closure structure is denser than the brokerage structure. Thus, a hybrid network can be created by continuously maintaining (or increasing) the density of dense areas and occasionally building ties at sparse areas to ensure global connectivity. To this end, the current study's model incorporated two micro-mechanisms in terms of individual agents' selection of knowledge exchange partners. With certain probability, an agent will choose the knowledge exchange partner from within her current social circle, whom she connects to directly or indirectly via common contacts (i.e., an open triad). Otherwise the agent will just randomly select a partner from all other organizational members. The two behaviors are referred to hereafter as embedded and random knowledge exchanges respectively.

Random knowledge exchanges create bonds and bridges in equal chances at the early stage of network evolution, while embedded knowledge exchanges create only bonds. So, when random and embedded knowledge exchanges both exist (no matter in what proportion), eventually there will be more bonds than bridges, leading to the emergence of closure structures as dense areas in the network. After that, embedded knowledge exchanges tend to happen at dense areas (i.e., within closure structures), as there are more open triads in those areas. They create new bonds and make dense areas even denser. As the density of closure structures increases, embedded knowledge exchanges are more and more likely to happen on extant bonds instead of creating new bonds. In other words, over time embedded knowledge exchanges will stop creating new ties. The forgoing processes are illustrated in **Figure 11**.

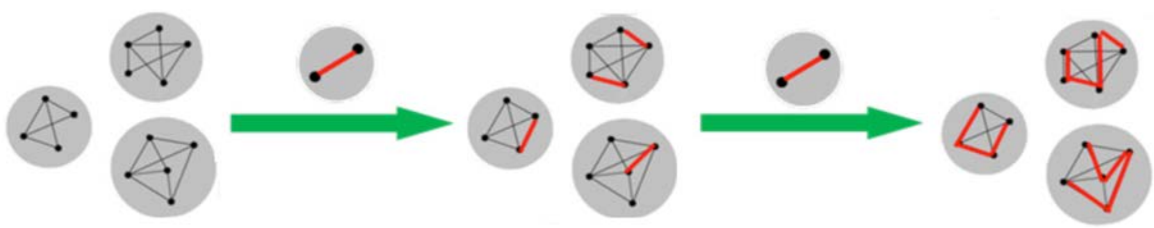


Figure 11. The creation and strengthening of bonds

Random knowledge exchanges are not confined to established closure structures, so there is always a chance for them to create new ties in sparse areas outside closure structures. By definition, these new ties are bridges. When closure structures are almost isolated, creating bridges reduces the closure extent of the entire network by increasing the number of open bridge triads. As shown in **Figure 12**, creating Bridge AB adds four more open triads ABC, ABD, ABE and ABF. But then more bridges will be created to close these triads through either embedded or random knowledge exchanges. More and more bridges gradually fill in structural holes, leading to the merge of adjacent closure structures and the elimination of brokerage structures in the middle. As a result, the closure level of the entire network goes up again and after the merge, bridges convert to bonds.

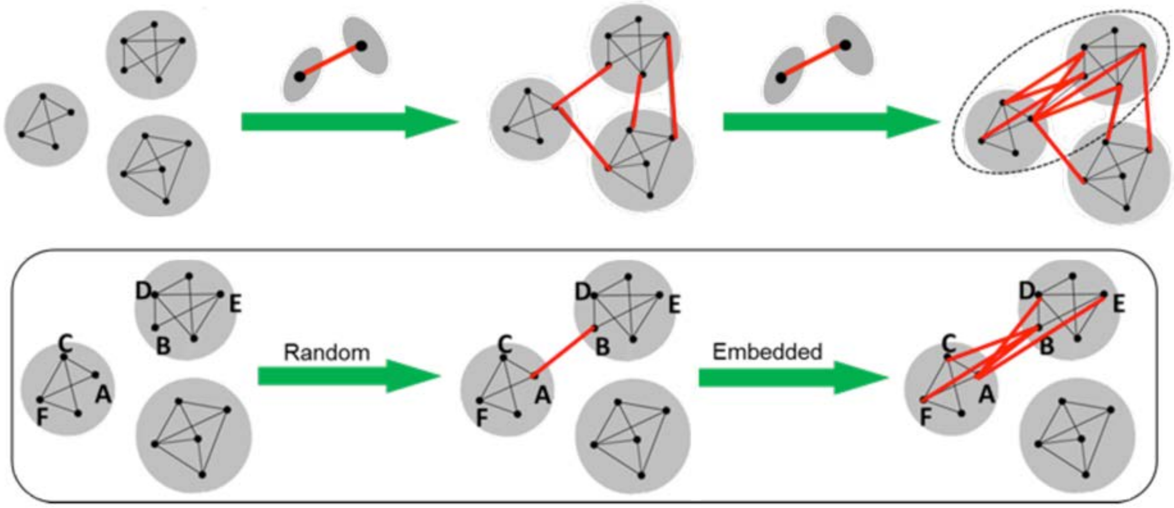


Figure 12. Adding bridges first reduce and then increase extent of network closure

The current study's model implemented a weighted network topology, in which the strength of a tie increases whenever there is a knowledge exchange on the tie or decreases whenever there is not[39]. Thus, tie strength contains information on the interaction history of two connected agents. When an agent considers a candidate partner for embedded knowledge exchange, she will take into account not only the number of their common contacts (see the triad closure mechanism in Section 2.4) but also her interaction histories with the candidate and their common contacts (if any). When past interactions largely strengthen ties, an agent tends to choose one of her closet social contacts over someone she interacted with only for a few times or a friends' friend she never met before. In other words, embedded knowledge exchanges are even more likely to close or strengthen bond triads than to close bridge triads (compare **Figure 12** with

[39]This process includes tie creation, growth, decay, and dissolution. Creating a tie means increasing tie strength from zero to a positive value. The decay of tie strength will eventually lead to the dissolution of a tie when its strength becomes zero. Admittedly, sudden tie removal can happen due to the turnover of organizational members: when an individual leaves an organization, all her connections in the organization are severed. However, in this study the network size was assumed to be fixed, so this situation (i.e., tie removal) was ruled out.

Figure 13). Thus, there tend to be a number of small dense closure structures with strong bonds (Kumpula et al. 2009) and these structures tend to emerge quickly and merge slowly. Moreover, two connected agents are increasingly likely to interact with each other (i.e., they interact more and more often), which simulated a well-observed real-life phenomenon known as relational inertia[40] (Briscoe & Tsai 2010; Gargiulo & Benassi 1999).

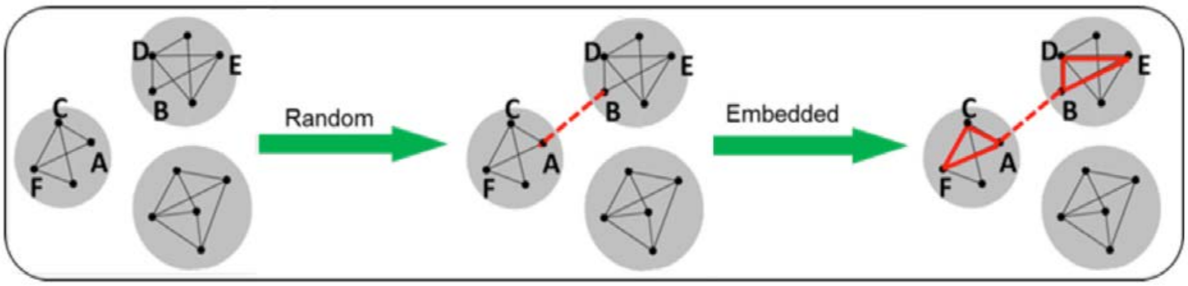


Figure 13. Embedded knowledge exchanges are confined within closure structures

Given different proportions of embedded and random knowledge exchanges and different rates of tie strength change, the network topology is expected to evolve through different paths (**Figure 14**). As modeled, the initial network topology is empty (State A) and no structural pattern has emerged yet at the early stage of network evolution (State B). When knowledge exchanges are completely random, no specific sub-structures would emerge (State C) (Erdos & Renyi 1960). When there are both random and embedded knowledge exchanges, a hybrid network containing brokerage and closure sub-structures tend to arise (State D) (Davidsen et al. 2002; Jin et al. 2001). When only embedded knowledge exchanges exist, bridges created at the early stage of network evolution will either dissolve because of disuse or turn into bonds after the merge of adjacent closure structures. Thus, no brokerage structures will sustain. The network will have one big closure structure (State E) or several isolated small closure structures (State G).

[40]a tendency to stick with established ties in lieu of initiating new ones

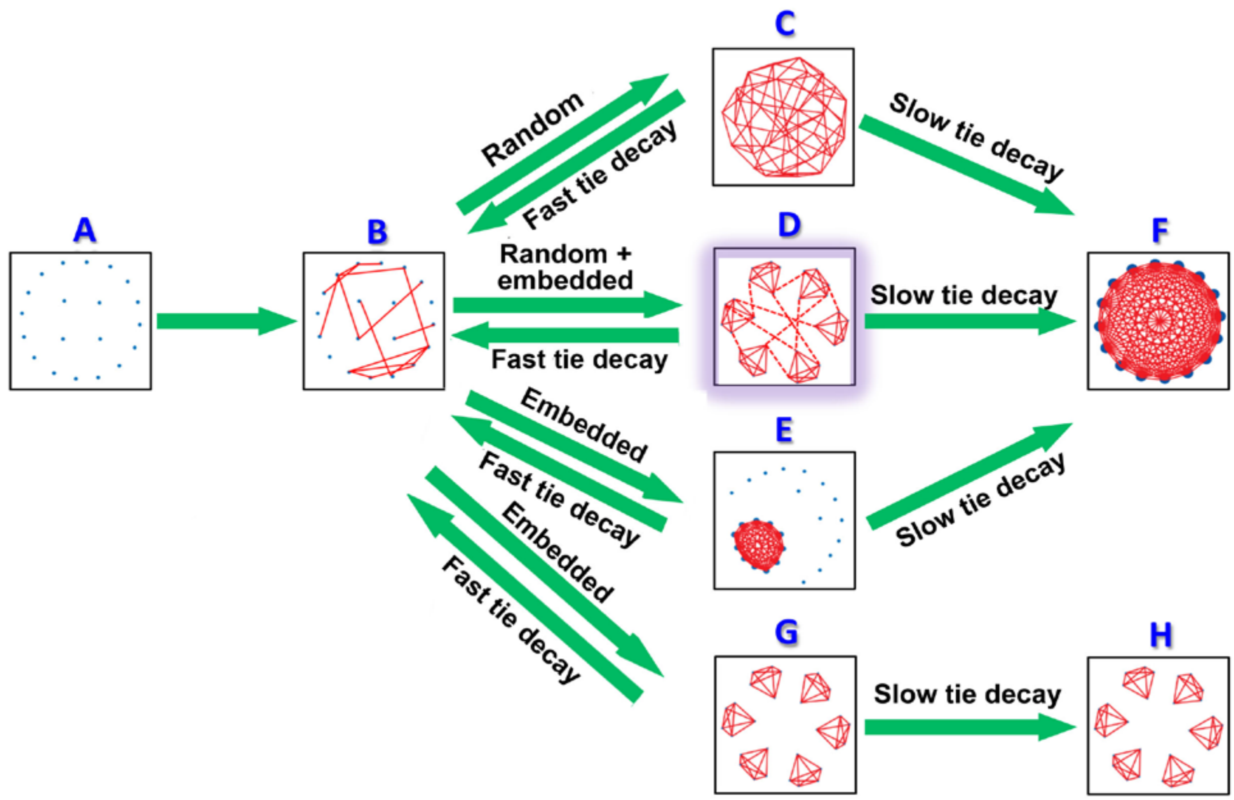


Figure 14. Evolutionary paths of the network topology in the model of this study
(Modified from Fronczak et al. 2007)

Given a high rate of tie decay, no specific structural patterns will emerge and the network topology will remain at State B. In contrast, if established ties never decay or tie decay is slower than tie formation, all closure structures in the network will become denser and denser until fully connected inside (State H). If the network is connected[41] at some point, then eventually every node will connect with all other nodes (State F). Due to the existence of random knowledge exchanges, both State C and State D will evolve into State F. As the network evolves from D to F, new ties mostly contribute to the merge of distinct closure structures. Since embedded knowledge exchanges tend to happen within existing closure structures, a higher proportion of

[41] That is, every two nodes in the network are connected by some path in the network.

embedded knowledge exchange will slow down the transformation from D to F. If network evolution is driven by embedded knowledge exchanges alone, once the network is disconnected at some point, it will remain that way, because embedded knowledge exchanges require that involved agents have at least one common contact. In this case, zero or slow tie decay will lead to multiple fully connected yet isolated closure structures (State H).

In sum, there are two types of micro interactions in the simulated CAS – random or embedded knowledge exchanges. The selecting power of the macro structure (i.e., the macro interaction network) originates from its extent of closure. The closer the network, the more likely micro interactions (random or embedded knowledge exchange) would create or strengthen bonds and bond triads, which maintain or reinforce the macro network's extent of closure. Micro interactions do not always follow the macro structure. Variations happen when random knowledge exchanges create bridges, thus introducing open triads and reducing the extent of network closure. These variations were then neutralized as open triads were closed by embedded knowledge exchanges. Finally, the evolution of the macro interaction network was path-dependent regarding the specific order it went through different states. The more individual behaviors conformed to and reinforced the macro structure (represented by the probability of embedded knowledge exchanges and the increase of tie strength per interaction), the sooner the network would be stabilized and the more resistant it would be to changes (i.e., it would have fewer state changes, such as State G and H in **Figure 14**). **Table 1** lists the primary elements of a social CAS and their counterparts in the current study's context.

Table 1. CAS elements and their counterparts in the current study

CAS element	Implementation
The system	An organization
Agents	Organizational members
Macro structure	A macro interaction network emerging from interpersonal knowledge exchanges (including no exchange)
Micro interactions	Random and embedded knowledge exchanges
Variation	New bridges and open triads created by random knowledge exchanges
Selection	Mutual reinforcement of closure structures and embedded knowledge exchanges

3.2. Computational Model

This section specifies the computer implementation of the above conceptual model, which was programmed in Java and was compiled and executed on Repast Simphony 2.1 (North et al. 2013). First, individual agents were characterized by a propensity for knowledge exchange P_{KE} ($0 \le P_{KE} \le 1$) and a propensity for random knowledge exchange P_{RM} ($0 \le P_{RM} \le 1$). While P_{RM} is the same for every agent[42], P_{KE} varies with agents. Each agent's P_{KE}, denoted by $P_{KE}(i)$ for Agent *i*, is assigned at model initialization from a power-law distribution[43] and then remains constant during a simulation run. Mathematically, the probability that P_{KE} has a value x ($0 \le x \le 1$) is $p(x) \sim x^{-\alpha}$. This probability and the expected number of knowledge exchangers vary with the

[42]Every agent has the same probability of choosing random knowledge exchange, but whether it actually chooses that is stochastic (i.e., nondeterministic).

[43]According to recent studies on human beings' telecommunication and online communication, the rate of a specific person interacting with other people is relatively stable and shows a power-law distribution within a community – only a few members are very active while the majority is not (Cattuto et al. 2010; Muchnik et al. 2013; Perra et al. 2012). In terms of this study's model, it means a knowledge exchanger or independent worker at the previous time step tends to remain a knowledge exchanger or independent worker at the current and the next time steps.

exponent parameter α. In the computer program, x is approximated by $y^{1/(\alpha+1)}$, where y is a uniformly distributed pseudorandom number on [0, 1] (Weisstein 2013). As shown in **Figure 15,** a smaller α means fewer knowledge exchangers in the population; when $\alpha = 0$, the probability distribution becomes a uniform distribution. At each time step Agent i participates in knowledge

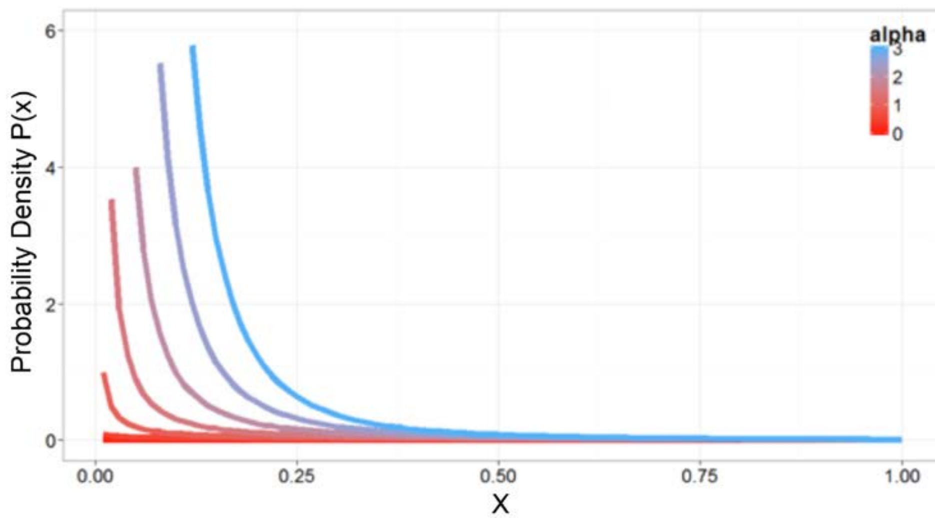


Figure 15. Probability density functions of $p(x) \sim x^{-\alpha}$ given different values of alpha

exchange with a probability $P_{KE}(i)$ and conducts independent knowledge creation otherwise. In the former case, Agent i conduct random knowledge exchange with a probability P_{RM} and embedded knowledge exchange otherwise. The computer program executes the above process by generating two pseudo-random numbers from a uniform distribution on [0, 1] and compared them with $P_{KE}(i)$ and P_{RM} respectively to determine the behaviors of Agent i. Besides the power-law distribution, three other probability distributions were also implemented for comparison. The four of them were distinguished by a categorical variable named *actType*.

In a model of the previous study (Lazer & Friedman 2007), individuals' propensity for knowledge exchange was modeled as the frequency of interpersonal knowledge exchange.

Basically, all agents share the same P_{KE}; at each time step, the computer program decides whether an agent will participate in knowledge exchange by generating a pseudorandom number (uniformly distributed) to compare with P_{KE}. Thus, every agent has the same frequency but stochastic timing of knowledge exchange. However, empirical evidence suggests that the number of knowledge exchangers per time step should follow a power-law distribution – only a few members are very active while the majority is not (Cattuto et al. 2010; Muchnik et al. 2013; Perra et al. 2012). One way to fix the preceding design is to generate power-law instead of uniformly distributed pseudo-random numbers, but it still cannot guarantee that every time the knowledge exchangers are mainly the same groups of agents. The current study's model, as described earlier, is more consistent with empirical observations.

In this model, agents were placed on an NK landscape to search for the highest peak(s) in parallel. Every point in the landscape was described using an N-dimensional vector of binary digits (0 or 1), which represents a problem solution involving N knowledge areas $a_i = \{0,1\}$, $i = 1,..., N$ [44]. The score of a solution (denoted as $A = \langle a_1, a_2,..., a_N \rangle$) was obtained by averaging the contribution of each area:

$$F(A) = \frac{1}{N}\sum_{i=1}^{N} f(a_i)$$

And the contribution of each area is affected by K other areas:

$$f(a_i) = f(a_i \mid a_1^i,...a_K^i)$$

a_j^i ($j = 1,..., K$ and $j \neq i$) is every other knowledge areas the contribution of a_i depends on.

[44]In this study, N = 20.

Creating a NK landscape includes predefining the performance of totally $2^{k+1}N$ solutions and K other dependent areas for each area. The former is randomly drawn from a uniform distribution over [0, 1], while the latter is randomly drawn from $\{1,..., N\} \setminus \{i\}$. Following prior studies, organizational performance $\overline{F}$ is measured by the (arithmetic) mean of all individual performances[45] at a sufficiently large time step. It indicates the system's ability to develop and propagate good solutions (i.e., to explore and exploit knowledge) in the long run.

Every execution of this study's model continues for a certain number of time steps until all individual solutions converge or the change of organizational performance clearly slows down and stabilizes. At each time step, three major processes are conducted in sequence: individual decision-making, individual solution improving and macro network update (see **Appendix A** for a pseudo-code[46] description of the main procedure). The first process is a propensity and priority based decision-making process that determines each agent's behavior at the current time step. Firstly, every agent chooses between independent knowledge creation and knowledge exchange. With a probability $P_{KE}(i)$, Agent i will improve its solution by independent knowledge creation and will not respond to other agents' knowledge exchange requests as well. Otherwise, Agent i will interact with another agent j to exchange knowledge. With a probability P_{RM}, Agent j will be randomly selected from the $(m – 1)$ agents (excluding Agent i) in the system. Otherwise, Agent j will be the one socially closest to Agent i. In this study's model, the social closeness of two

[45]Each individual score is normalized against the highest score (i.e., the global maximum) of an NK space and then monotonically enlarged by an exponential function $f(x) = x^8$. This transformation generates a performance distribution that has only a few very good solutions and a majority of quite bad ones, making it easy to detect changes in organizational performance (Lazer & Friedman 2007). Otherwise, the distribution of normalized individual performances is similar to a normal curve with most solutions having moderate performance scores, and the variance of these scores further decreases when problem complexity (K) increases.
[46]An informal mixture of natural language and programming conventions that makes the structure and flow of a program clear without requiring the reader to be familiar with any particular programming language.

agents i and j is measured by Burt's (1992) *local network constraint* (*LNC*) that integrates the number of mutual friends and the strength of ties:

$$LNC_{ij} = (\tilde{w}_{ij} + \sum_q \tilde{w}_{iq}\tilde{w}_{qj})^2 \text{ , } i \neq q \neq j$$

where w_{ij} is the weight of the tie connecting i and j; $w_{ij} = 0$ if i and j are not connected. w_{ij} (as well as w_{iq} and w_{ig}) is normalized with the strength of i by $\tilde{w}_{ij} = w_{ij} \big/ \sum_g w_{ig}$ ($i \neq g$) to make sure that LNC_{ij} falls between 0 and 1 (inclusively). $\sum_q \tilde{w}_{iq}\tilde{w}_{qj}$ is positively related with the number of i and j's mutual friends and the strength of their respective connections with these friends. LNC_{ij} thus has two components: (a) the amount of relational investment Agent i devotes to Agent j in proportion to the first Agent i's total relational investment, and (b) the extent of triad closure regarding the two agents (i and j). If i has no connection in the network, it will interact with a random agent regardless of its P_{RM}.

In the models of previous studies (Fang et al. 2010; Lazer & Friedman 2007), an agent obtained knowledge from multiple sources at one time. The current study's model applied a finer time scale focusing on one-to-one knowledge exchanges, each of which then became a building block of the dynamic network topology. To avoid potential conflicts, an agent will accept another agent's knowledge exchange request only if the former agent has decided not to self-learn at this step and is not involved in any knowledge exchange. After accepting a request, the receiver will not respond to other requests or send out a request itself. Each agent is allowed to send zero or one request per step (independent workers have no requests). If its request is rejected (i.e., the receiver decides to self-learn or has been involved in another knowledge exchange), the agent will remain active until it is asked to participate in a knowledge exchange by someone else.

Agents with a denser ego network or more local constraints have a better chance of being asked. As in the real life, such individuals tend to spend more time on knowledge exchange than knowledge creation (McFadyen & Cannella 2004). If no one asks, the agent will waste an opportunity (as a time step) to improve its performance.

The second process handles individual agents' independent knowledge creation or knowledge exchanges and the ensuing improvement of individual solutions. Since individual solutions are formalized as binary vectors in a NK model, agents who work independently improve their solutions by randomly selecting ω different knowledge areas in its solution and flip their values (from 0 to 1 or vice versa). The resultant new solution will be adopted only if it is strictly[47] better than the old one. As described in Section 2.2, ω distinguishes the adaptive walking ($\omega = 1$) and random long jumping ($\omega > 1$) of individual agents. As for knowledge exchange, the maximal percent of knowledge areas that can be exchanged between two agents depends on the strength of the network tie that connects them. Specifically, the maximal percent is equal to the ratio of the tie's current weight and maximally possible weight, which is calculated by assuming that the tie was continuously strengthened up till the current time step. Unlike the models in previous studies (Fang et al. 2010; Lazer & Friedman 2007) where a knowledge exchange only benefits the worse side of two interacting agents, i.e., only one individual solution is improved, the current study's model gives either side an opportunity to improve its performance by creating a new solution based on the original solutions of both sides.

Between two interacting agents, the one with better original solution (and better individual

[47] I use the "strictly better" criterion to avoid unnecessary fluctuation (Fang et al. 2010; Lazer & Friedman 2007).

performance) was named as Agent *i* and the other one as Agent *j*. Agent *i* acts first. It identifies some different areas from all exchangeable knowledge areas (see above) and creates a new solution by flipping the values of these areas in its own original solution. As mentioned, the probability for an agent to miss a different knowledge area in the other agent's solution is defined as the error rate of inter-agent solution imitation (ε). Given totally z different exchangeable knowledge areas, the probability that Agent *i* identifies all of them is $(1 - \varepsilon)^z$. The probability is higher given smaller z, mimicking the real-life positive relation between the ease of knowledge exchanges and the amount of common knowledge already shared by two participants (Hansen 1999; Reagans & McEvily 2003)[48]. Because of incomplete knowledge exchange (limited by tie strength) and inaccurate discrepancy detection (caused by imitation error rate), the new solution is almost always a recombination of two original solutions. Agent *i* will adopt the new solution if it is strictly better than Agent *i*'s original solution and will discard the new solution otherwise. Agent *j* acts next and follows almost the same procedure as described, except that the new solution Agent *j* creates is based on her original solution and Agent *i*'s latest solution, which may be different from Agent *i*'s original solution.

Based on the micro behaviors occurring in the second process, the third process updates the existing macro interaction network by changing tie strength. In the current study's model, tie

[48]In addition, this model allows more knowledge areas to be exchanged through stronger ties per interaction, no matter whether they are different or not. If the error rate ε was applied to every exchangeable knowledge area rather than just different ones, the overall accuracy $(1–\varepsilon)^z$ decreases with the number of exchanged knowledge areas, leading to an unrealistic implication that stronger ties were less reliable than weak ties in terms of knowledge exchange.

strength resembles memory in terms of development and impacts[49], so the change of tie strength was modeled based on the ACT (Adaptive Control of Thought) theory on learning and forgetting of memory contents (Anderson 1992). According to ACT, activating a memory trace[50] requires accessing the corresponding memory item. Each access reinforces the activity of the memory trace, but the effect decays over time. Thus, the activity of a memory trace is increasingly smaller than the total number of accesses. Without constant access, active memory trace will be deactivated. In the current study's model, each tie is a memory trace and tie strength amounts to activity of the memory trace; each knowledge exchange interaction on the tie is an access. More specifically, each knowledge exchange between two agents either creates a new tie of an initial weight δ or increases the weight of an existing tie by δ. Every existing tie not strengthened at the current time step will decay. The weight of a tie after the nth and before the $(n + 1)$th reinforcement is given by

$$S = \delta \sum_{j=1}^{n} t_j^{-d}$$

where t_j is the current time step or tick. It represents the elapsed time since the jth increment. The exponent d ($0 < d \leq 1$) controls how fast the tie decays and therefore models the speed of memory loss. **Figure 16** shows the strength of a tie (S) after 100 times of uninterrupted reinforcements. It is clear that S has a near-linear positive relation with δ and monotonically (not linearly) decreases with d.

---

[49]A tie goes through the stages of formation, growth, decay, and dissolution with continuous strength change. Tie strength is positively related with the duration and the freshness of the tie. The strength of a tie represents the interaction history of two agents and affects their future interactions with each other and with other agents.
[50]It is defined as the hypothetical structural alteration in brain cells following learning.

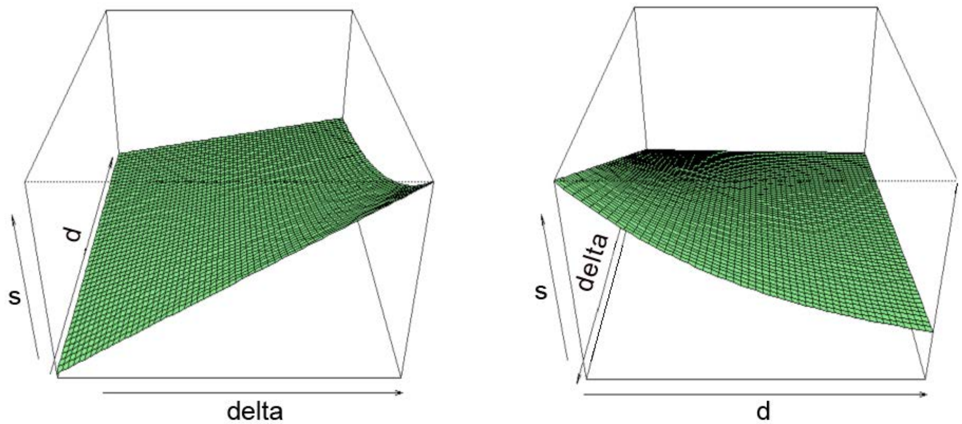


Figure 16. 3D plots of the tie strength after t = 100 uninterrupted reinforcements given the changing values of *d* and *δ*

Table 2 summarizes the model parameters and outcomes described in this section, with the last column displaying the name by which each variable will be referred to hereafter. Among all the model parameters, *randLink* (the relative probability of random and embedded knowledge exchanges in the system), *decayRate* (the decay rate of network ties not used for interactions), *wtGain* (the increment of tie weight or strength per interaction), *actType* (indicating the probability that an agent prefers knowledge exchange over independent knowledge creation at each time step[51]) and *actDist* (exponent of the power-law distribution) would directly affect the generation and maintenance of a hybrid network topology and then *avgScore* (organizational performance). The impacts of these variables on *avgScore* were examined through simulation experiments on the computational model, as described in the next chapter.

Another influential factor, according to the literature, is the accuracy of interpersonal knowledge exchange. This factor (designated as *learnErr*) was operationalized in the current study the same way as in the previous study (Lazer & Friedman 2007). Its effect on *avgScore* was also explored

[51]Whether the interaction actually happens also depends on whether the other agent is idle at that moment. If the other agent has been occupied, the interaction will not happen.

through simulation experiments. In addition, the effects of *tick* (time), *space_k* (complexity of external problems), and *innoRange* (organizational members' innovation ability during independent knowledge creation) on *avgScore* were examined to ensure that the NK model of parallel problem solving was successfully integrated into the current study's agent-based model. Finally, Model robustness was tested regarding three noise factors: the scale of aggregation measured by *orgSize* (the number of agents), the initial condition of model execution identified by *nkSpace*, and the probability distribution of individual agents' knowledge exchange intention differentiated by *actType* (limited to the difference between a power-law and a normal distribution).

Table 2. Parameters and outcomes of the current study's model

Parameter/ outcome	Description	Related model component	Alias
γ	Exponent of a power-law distribution that determines the probability that an individual will choose knowledge exchange over independent knowledge creation (P_{KE}), if given a chance to choose	Individual behavior	*actDist*
-	A categorical variable distinguishing different probability distributions of knowledge exchange intentions of all the agents	Individual behavior	*actType*
P_{RM}	If given a chance to choose, the probability that an individuals will choose random rather than embedded knowledge exchange	Individual behavior	*randLink*
δ	The initial weight of a tie; the increment of tie weight per interaction	Macro interaction structure	*wtGain*
d	The exponent in the tie decay function that indicates the rate of tie decay	Macro interaction structure	*decayRate*
K	The number of interdependent knowledge areas	NK space	*space_k*
ω	The number of knowledge areas an individual can explore per independent knowledge creation	Individual attribute	*innoRange*
ε	The probability of missing one different knowledge area during solution imitation	Individual attribute	*learnErr*
m	The number of individual agents; network size; organizational size	Macro interaction structure; NK space	*orgSize*
-	A categorical variable identifying the initial NK space and the set of starting points on the space	NK space	*nkSpace*
t	Time step/tick	Model dynamics	*tick*
$\overline{F}_t$	Organizational performance at a specific time	Model outcome	*avgScore*

Chapter 4

SIMULATION EXPERIMENTS AND RESULTS

Next, simulation experiments were conducted on the above computational model. This chapter describes in detail the experimental design, statistical analysis and results. In addition to the main experiments, extreme condition tests were also conducted and the results verified the integrity of the computational model (**Appendix I**). In organizational research, no common procedures or methods are currently available for exploring and testing agent-based models (Burton & Obel 1995; Carley 1996). Thus, the methods and techniques applied in the current study, as introduced below, mostly came from other research areas (Alden et al. 2013; Chalom & de Prado 2012; Fagiolo et al. 2007; Marino et al. 2008; Marks 2007; Richiardi et al. 2006). Both experimental design and result analysis were conducted in the R statistical computing environment (R Core Team 2014).

4.1. Experimental Design

Mathematically, the transformation from model inputs to outputs can be represented by a function $Y = F(X)$, where $Y = [y_1, y_2,..., y_{nY}]$ is a vector of model outputs (or responses) and $X = [x_1, x_2,..., x_{nX}]$ is a vector of imprecisely known model inputs (or factors). Function F is often referred to as a response surface (in the sense that particular responses form a surface over a multi-dimensional value space of X) or a meta-model (in the sense that the purpose of data analysis is to create a statistical model of the original model to help understand the latter). Every sample is a point on the surface representing a specific model input ($X_k = [x_{1k}, x_{2k},..., x_{nX,k}]$) and the resultant output ($Y_k = [y_{1k}, y_{2k},..., y_{nY,k}]$). While single sample points only provide partial information on model behavior, a well selected sample set (X_k , Y_k), k = 1, 2, …, nS show a bigger and better picture of the input-output transformation.

Since samples are collected through computer simulation experiments, a specific sample set corresponds to a specific experimental design that plan different experimental conditions (also known as design points)[52] and the number of replicate runs[53]. An experimental design is typically described by a matrix: Row *i* represents the *i*th experimental run, Column *j* represents the *j*th independent variable (or factors), and the element at the intersection of Row *i* and Column *j* represents the value of the *j*th factor in the *i*th run. Since simulation experiments can be time consuming and resource demanding, a good experimental design should be able to get as much information of a response surface as possible with a relatively small number of high-quality[54] design points (i.e., samples).

It is contended that global space-filling designs should be used for analyzing complex agent-based models (Collins et al. 2013; Marino et al. 2008; Timmis et al. 2011). Such models are featured by multiple interacting mechanisms each associated with more than one model parameter. Exploring model behaviors thus requires considering and quantifying the interactive

[52]An experimental condition is a specific value setting of independent variables (or factors) represented by model parameters.
[53]A run is defined as a single replication experiment of a design point. In computer simulation experiments, each design point typically has multiple runs.
[54]High-quality design points satisfy multiple criteria developed for physical (Box et al. 2005; Myers 2009) or simulated experiments (Kleijnen et al. 2005). For instance, one criterion states that the design points should meet the assumptions of different statistical models/tests.

effects of multiple parameters or even the overall effects of all model inputs. The common one-at-a-time (OAT) design[55] is inadequate for this purpose, as it focuses on only one model parameter while ignoring significant interactive effects among different model parameters (Manson 2003; Saltelli & Annoni 2010). In contrast, so-called global designs (or global sampling techniques) simultaneously perturb the values of all model parameters based on their probability distributions on respective ranges of values. Global designs thus provide a more representative indication of the relative influence of different model parameters on model behavior than local designs[56].

When researchers have little prior knowledge about a response surface or expect it to have a complex shape, space-filling designs are preferable to classical ones such as central composite design and factorial design (Sacks et al. 1989; Simpson et al. 2001). Classical designs sample at the center and extreme points of the model parameter space and takes multiple samples (replicates) at each point, as shown in **Figure 17a**. In contrast, space-filling designs sample through the model parameter space uniformly and seldom take replicates (**Figure 17b**). Both

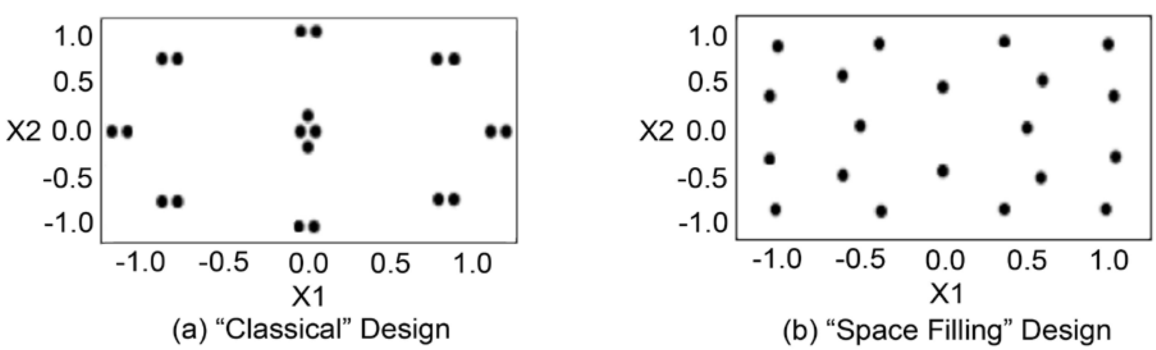


Figure 17. Classical versus space-filling designs (Booker 1998)

[55]It samples the value range of one model parameter while holding all other parameters fixed.
[56]The OAT method is often considered as a local design or local sampling technique.

classical and space-filling designs can identify low-complexity features (e.g., second-order changes) of the response space by sampling not only at extremes but also somewhere in between. However, since classical designs only sample at the corners and the center, it cannot identify high-complexity features such as the presence of thresholds (i.e., sudden changes), while space-filling designs can do so.

A specific experimental sampling technique used in this study is Latin Hypercube Sampling (LHS)[57]. LHS is a type of stratified Monte-Carlo sampling without replacement technique. It divides the distribution of each model parameter into equal probability intervals and then samples each interval of each parameter exactly once, so that each model parameter has its whole value range well scanned and represented according to its probability distribution on the range. More specifically, assume that the response surface we want to estimate has the form of $Y = f(X)$, and that we decide to get $S = 5$ LHS samples by varying $d = 3$ model parameters a, b, and c. The first two parameters are uniformly distributed: $a \sim Unif\,(a_{min}, a_{max})$ and $b \sim Unif\,(b_{min}, b_{max})$. The third one follows a normal distribution $c \sim Normal\,(\mu_c, \sigma_c)$. Firstly, the range of each parameter is partitioned into $S = 5$ equally probable intervals and independent samples of parameter values are drawn from each interval. Then an $S \times d$ LHS design matrix is built by assembling (without replacement) the acquired value samples. Each row of the matrix is a unique combination of parameter values, known as a design point (a prescribed experimental condition). The next step is to run the model using these design points. Corresponding model outputs are collected and stored in Y. **Figure 18** illustrates the above process.

[57]Also called Latin Hypercube Design (LHD)

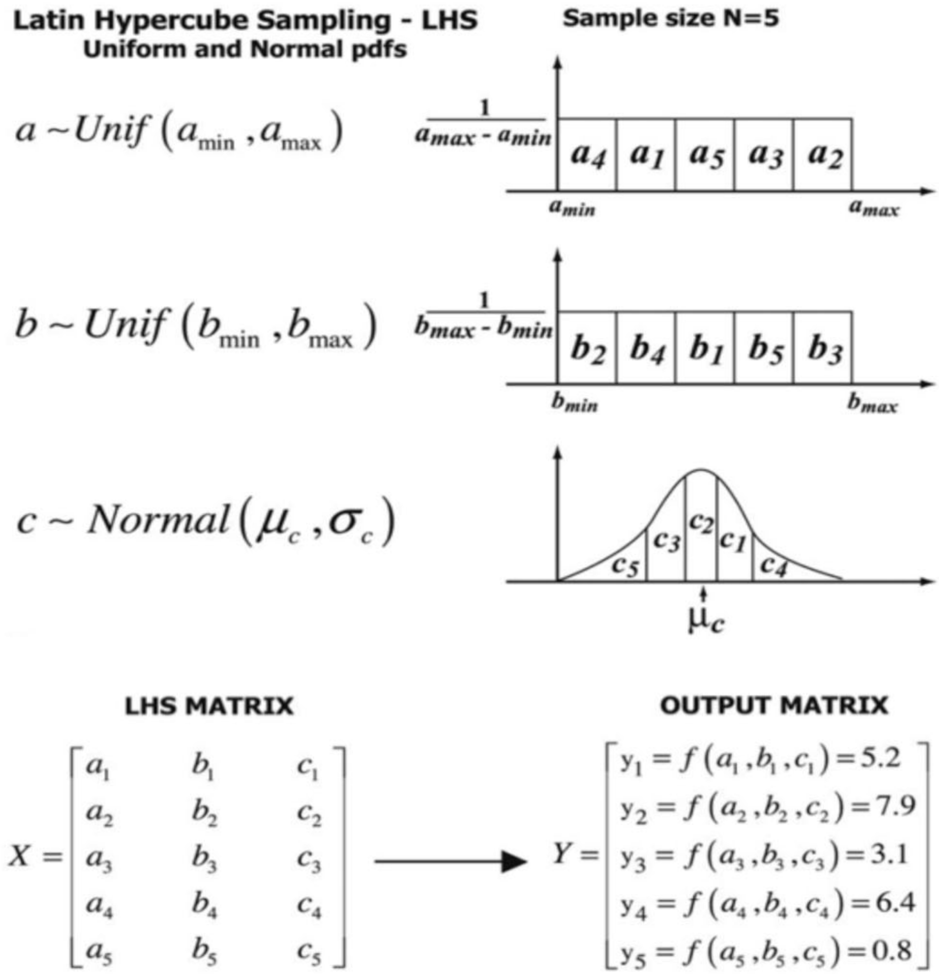


Figure 18. An exemplary LHS process
(modified from Marino et al. 2008)

LHS can cover the model parameter space adequately while minimizing the correlations between the values of different parameters. It also requires fewer samples than simple random sampling to achieve the same accuracy (Helton & Davis 2003; McKay et al. 1979). There is no a priori rule for determining the sample size for LHS. The minimum value is $nX + 1$, where nX is the number of model parameters being varied, but researchers tend to take much more samples to ensure accuracy (Marino et al. 2008). The LHS sample set used in this study had $S = 300$

samples[58] by varying d = 6 primary model parameters, including *actDist* (indicating the probability that an agent prefers knowledge exchange over independent knowledge creation), *randLink* (an agent's probability of random knowledge exchange), *wtGain* (tie strength increment per interaction), *decayRate* (indicating the rate of tie decay), *learnErr* (an agent's probability of missing one different area when exchanging knowledge), and *innoRange* (the number of knowledge areas an agent explores every time when it is independent knowledge creation). These parameters were all assumed to follow a uniform distribution on their respective value ranges. The sample set was generated by the R package "pse" (Chalom & de Prado 2013) and was orthogonal: the input vectors of any two parameters are uncorrelated (Spearman's *rho* = 0)[59]. This study requires an orthogonal (or near-orthogonal) design because a major statistical analysis technique of this study is sensitive to strong collinearity among independent variables (De Veaux et al. 1993)[60]. **Appendix B** includes the entire LHS sample set, a summary of the sampled values of each parameter, and the correlation matrix of all these parameters.

When there are many variables whose main and interaction effects need to be explored, even an efficient experimental design like LHS requires a large number of samples. Collecting and adjusting[61] these samples take a lot of time and computing resources. To further improve the efficiency of data collection, researchers often employ a crossed design (Kleijnen et al. 2005), in

---

[58]The sample size x is calculated by $1-0.99^{x} > 0.95$. According to Helton et al (2000), a LHS sample of size ≥ 298 can generate an adequate number of complementary cumulative distribution functions (CCDFs), so that the maximum CCDF generated exceeds the 99th percentile of the population of CCDFs with at least a 0.95 probability.

[59]Orthogonality means no multi-collinearity, which is the basic assumption of many statistical tests and makes testing results easy to interpret. The "pse" package uses a method that changes the order of design points to force the correlation matrix into a prescribed value.

[60]The current study applied MARS (Multivariate Adaptive Regression Spline), which automatically drops one of two highly correlated independent variables during analysis.

[61]for example, changing the order of samples to minimize correlation

which model parameters are categorized into several groups based on their anticipated impacts and each group is sampled differently (i.e., using different experimental designs). Generally, the less important or ambiguous a group of parameters are, the less accuracy is required in estimating their effects on model outcomes, so the fewer samples need to be collected for this group of parameters[62]. For example, primary model parameters and noise factors are usually treated as different groups in a cross design and sampled separately. In the end all subgroup designs are crossed, with every unique combination of sub design points corresponds to one and only one crossed design point.

The non-primary model parameters include *space_k* and *orgSize*. They were not the primary interests of this study: the impact of problem complexity (*space_k*) on organizational performances has been well studied before; organizational size is just a noise factor. As problem complexity and organizational performance has a monotonic relationship, two design points for *space_k* would be sufficient. Thus, the value of *space_k* alternated between 1 and 5 in this study's simulation experiments, indicating a low and a high level of problem complexity (i.e., a smooth or rugged NK landscape) respectively. The variable *orgSize* has three sample points (50, 100, and 200) to capture the curvilinear[63] relationship between the performance and the size of an organization (Keller 1986; McGrath 1984; Shaw et al. 1981).

This study applied two crossed designs for simulation experiments. In the first design, the LHS samples of six primary model parameters, *actDist, randLink, wtGain, decayRate, learnErr,*

[62]In terms of experimental design, it means fewer experimental conditions or design points.
[63]Up to a point, group performance tends to increase with size, owing to the added knowledge of the additional members. Once past the optimal size, however, performance often decreases because subgroups develop and coordination involves costs.

innoRange, were crossed with *space_k* (= 1, 5), *orgSize* (= 50, 100, 200), and *actType* (= A, B, C, D) that indicates different probability distributions of knowledge exchange intent. When *actType* = A, there is no knowledge exchange. Thus, of six primary parameters, only *innoRange* (the number of knowledge areas an agent explores during independent knowledge creation) is ineffective. When *actType* = B, the probability an agent is willing to exchange knowledge follows a power-law distribution on [0, 1], so all six primary parameters are effective. When *actType* = C, all agents are willing to exchange knowledge all the time. When *actType* = D, the probability an individual agent is willing to exchange knowledge follows a normal distribution on [0, 1] centered on 0.5. In the last two cases, 5 out of 6 primary model parameters (except for *actDist*) is effective[64]. In the second crossed design, the LHS samples were crossed with *space_k* (= 1, 5) and *nkSpace* (= 0, 1, 2, 3, or 4) that distinguishes 5 initial conditions; *actType* = B and *orgSize* = 100. **Table 3** lists all the aforementioned variables regarding their types and value ranges in the experimental design. All design matrices are partially presented in **Appendix C**.

The next step is to decide how many replicated runs are needed for each experimental condition (i.e., design point). Researchers often rely on replicate runs to reduce variations in experimental results. Since computer programs are executed step by step exactly, replicate runs only make sense when the model being experimented has stochastic elements, as these elements lead to different model outputs per run even though model inputs are the same. Thus, the purpose of replicate runs in simulation experiments is to estimate model uncertainty arising from stochastic elements[65]. As many agent-based models, the current study's model is stochastic. The number of

---

[64]I keep ineffective variables in the experimental design in order to get a balanced sample set.
[65]There are two types of uncertainty. One of them is called aleatory uncertainty. It is caused by the stochastic elements in a model. The other one is called epistemic uncertainty. It results from a lack of knowledge about (a) the

replicate runs was determined through a procedure originally developed by Read et al (2012) and later implemented by the R package "Spartan" (Alden et al. 2013). According to the results (see **Appendix D**), each experimental condition in this study should have 300 replicate runs (each run used a different random seed).

Table 3. Type and value range of variables used in experimental design

Variable	Description	Type	Range
actDist	Exponent of a power-law distribution that determines the probability that an individual will choose knowledge exchange over independent knowledge creation, if given a chance to choose	Continuous	[0,3]
randLink	If given a chance to choose, the probability that an individuals will choose random rather than embedded knowledge exchange	Continuous	[0,1]
wtGain	The initial weight of a tie; the increment of tie weight per interaction	Continuous	[0.1,3]
decayRate	The exponent in the tie decay function that indicates the rate of tie decay	Continuous	[0.5,1]
innoRange	The number of knowledge areas an individual can explore per independent knowledge creation	Discrete	{1,2,3,4,5}
learnErr	The probability of missing one different knowledge area during solution imitation	Continuous	[0,1]
space_k	The number of interdependent knowledge areas	Discrete	{1,5}
orgSize	The number of individual agents; network size; organizational size	Discrete	{50, 100, 200}
actType	A categorical variable distinguishing the probability distribution of knowledge exchange intentions of all agents	Discrete	{A, B, C, D}
nkSpace	A categorical variable identifying the initial NK space and the set of starting points on the space	Discrete	{1,2,3,4,5}

Finally, the duration of a simulation run (i.e., the number of time steps) is an important decision regarding ABM in general and this study in particular. As mentioned, ABM is mainly used to study the emergence or self-organization phenomenon in a complex system – how macro patterns arise from individual properties and micro mechanisms. A simulation run thus should

specific values that model inputs should be assigned, and (b) everything in the real domain that is absent in the model. Epistemic uncertainty is more of a problem when we try to determine whether an observed relation is an artifact of model implementation or actually exists in the real life.

last long enough for the dynamic process to completely unfold. A premature simulation risks missing important complex changes that tend to happen suddenly and shortly[66]. To determine the end time, ABM researchers typically monitor model behaviors through the values of one or more aggregated or emergent measures[67] and terminate a simulation run when the value reaches a conceptual threshold or stabilizes for a while. The results of this study were directly impacted by the duration of a simulation run, as organizational performance (measured by *avgScore*) is expected to improve over time (see **Figure 23** in Section 4.3). Thus, one option for the end time is the moment when *avgScore* reaches its maximum value 100%. However, preliminary experimental results showed that in some conditions *avgScore* could not reach this level in an acceptable period of time; for example, sometimes after as many as 5,000 steps *avgScore* still lingered around 60%. Thus, in this study a simulation run was terminated at tick 1000 or 1200 if *avgScore* did not reach 100% before that.

4.2. Statistical Analysis

The data derived from simulation experiments are sample points (X_k, Y_k), $k = 1, 2, \ldots, nS$ from the response surface of the computational model. They describe an input-output transformation $Y_k = F_k(X_k)$, where $X_k = [x_{1k}, x_{2k}, \ldots, x_{nX,k}]$ and $Y_k = [y_{1k}, y_{2k}, \ldots, y_{nY,k}]$. In this study, x_{ik} is the value of a primary model parameter or noise factor under a certain experimental condition; y_{jk} is the value of *avgScore* at a specific time step averaged across all replicate runs, with j indicating the time. F_k approximates F, the real input-output transformation (i.e., the response surface) of

[66]Abrupt changes are very common in complex systems due to so-called threshold effect.
[67]A macro measure is aggregated if it can be defined on the basis of averaging or aggregating micro features or dynamics. An emergent measure cannot be reduced to micro features or dynamics.

the model. Response surface F contains all the relations between individual input variables x_i, i = 1, 2, …, nX and individual output variables y_j, j = 1, 2, …, nY. Without the loss of generality, each of these relations is denoted as f_{ij}, i.e., $y_j = f_{ij}(x_i)$. By analyzing the results of simulation experiments, researchers intend to estimate the form and significance of f_{ij} and to assess the robustness of the estimated form and significance against noise factors. The analysis is guided by hypotheses and/or driven by data, depending on researchers' prior knowledge of f_{ij}.

Specific statistical analysis was determined by the stage of analysis, the expected form of f_{ij}, and the characteristics of model inputs and outputs. More than one technique was used to collect complementary information. Data analysis of the current study started with graphic visualization, as plots can reveal complex (nonlinear or non-monotonic) relations, thresholds, and interactions, helping to understand model behavior and plan for more sophisticated analysis. Moreover, plotting can be particularly enlightening when LHS stratifies over the full range of each input variable. However, plots become unclear and incomprehensible when an output variable is simultaneously affected by more than three input variables interacting with one another[68]. Also, graphics cannot tell us whether a relation $y_j = f_{ij}(x_i)$ actually exists or how important it is compared with other relations. Imagine that the samples (x_{ik}, y_{jk}), k = 1, 2, …, nS form a certain shape on the response surface captured by abovementioned plots. Is the shape an actual non-random pattern conditional on the marginal distributions of x_i and y_j? Answering this question needs inferential statistical analysis.

In the current study, different values of each discrete model parameter were analyzed using the

[68] Unless only one or two of the inputs have dominant effects; then the dominant effects will stand out in the plot.

Kruskal-Wallis (KW) test and the Mann-Whitney (MW) test. The KW test examines the overall difference. It is the non-parametric alternative to ANOVA and does not require equal group size. The MW test (with Bonferroni correction) was conducted for post-hoc pairwise comparison. It provides not only the significance level but also the effect size of a difference: anything greater than 0.5 is large, 0.5-0.3 is moderate, 0.3-0.1 is small, and anything smaller than 0.1 is trivial (Cohen 1988). In simulation experiments, it is relatively easy to collect a large number of samples (i.e., paired model inputs and outputs). Given a large enough sample, a statistical comparison will always show a significant result unless the effect size is exactly zero. Thus, regarding computer experiment results, the effect size or the impact of an input-output relationship is more meaningful than the relationship's statistical significance, which indicates the relationship's magnitude against chance(Troitzsch 2014; White et al. 2014).

The preferable inferential analysis used for continuous model parameters is regression analysis, given that the relation $y_j = f_{ij}(x_i)$ has a "shaped" form[69] (Storlie & Helton 2008). Regression analysis deals with relations between dependent variables and independent variables (model outputs and inputs in this study). It estimates separate and joint impact of independent variables on the dependent variable (the form of f_{ij}) and therefore the relative importance of each independent variable (the significance of f_{ij})[70]. Conventional parametric regression is unsuitable

---

[69]From a sample-based sensitivity analysis perspective, a statistical significant relation between model inputs and outputs can be classified, in terms of increasing complexity, as a linear relation, a monotonic relation, a trend in central location, and a trend in variability (spread) (Chalom & de Prado 2012; Kleijnen & Helton 1999). Here a "shaped" relation should be at least a trend in central location. A trend in variability is usually detected by a variation decomposition approach known as eFAST, which relies on a special heavy sampling technique different from LHS (Alden et al. 2013; Marino et al. 2008).

[70]usually based on (a) whether x_i can be selected into the regression model through an automated stepwise process and if so in what order, (b) the proportion of variance in y_j that can be accounted for as x_i enters the regression model, or (c) the standardized regression coefficient of x_i in the fitted regression model.

for this study. Owing to the structural and dynamic complexity of CAS, the relations investigated in this study are mostly non-linear and non-monotonic, so neither linear nor rank regressions can help. Moreover, there is not enough knowledge to predetermine the mathematical functions of these relations, so polynomial and nonlinear regressions are out of the picture as well[71]. Also, parametric regressions assume data normality, yet the data collected in the current study were not normally distributed. Thus, nonparametric regression techniques were used for the current study. Instead of imposing rigid assumptions on the underlying data distribution, these techniques derive the regression function directly from the data, which makes them useful for data exploration as well.

Nonparametric regressions operate locally rather than globally (Kleijnen & Helton 1999). They divide the response surface into small regions according to the data and summarize how dependent variables respond to independent variables in each region separately. Sometimes separate summaries need to be combined and that requires smoothing at region boundaries. The specific nonparametric regression technique employed in this study is Multivariate Adaptive Regression Splines (MARS®). It estimates nonlinear relations by automatically partitioning the hyperspace of independent variables into separate regions and then fitting a piecewise linear regression within each region (Friedman 1991).

Mathematically speaking, a spline is defined by a set of polynomial functions and has sufficient smoothness at transition points of the polynomial pieces. At each transition point (referred to as

[71]When using either of these two regressions, researchers need to guess a function (including all its nonlinear transition points), check its goodness of fit, and refine it iteratively.

knot), the slope of the regression line is allowed to change from one region to another. Each knot thus marks the end of one local region and the beginning of another; together these knots indicate piecewise linear, continuous behavioral changes of the regression model (i.e., predicted response surface). **Figure 19** shows a spline with two knots at *x*1 and *x*2. Knots are mathematically represented by so-called basis functions. A basis function is defined as *max* $(0, x - c)$ or *max* $(0, c - x)$, where x is an independent variable[72] and c is a constant value within the range of x indicating the location of a knot. The function *max* $(0, x - c)$ transforms the original variable x into a new variable x^* that equals to 0 for all values of x up to c and equals to $(x - c)$ for all values of x greater than c. This transformation reduces the value range of x to a subset, zeroing out everything outside that subset.

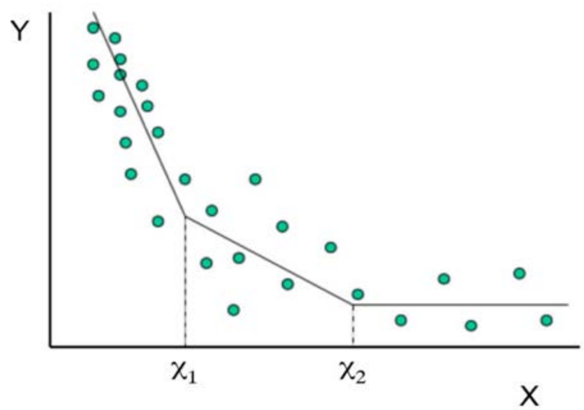


Figure 19. Example of a two-knot spline
(Adapted from Briand et al. 2004)

Figure 20 shows two basis functions *max* $(0, x - 3.1)$ and *max* $(0, 3.1 - x)$ that hinge on $x = 3.1$. More than one knot (i.e., more than one pair of basis functions) can be specified for an independent variable, allowing complex nonlinear relations to be formalized. A key property of basis functions is their ability to locally specify the main effects of corresponding variables, as they are nonzero only at part of the range of the variables. MARS describes the interaction effect

[72]The variable can be categorical or continuous. If it is categorical, the basis function is written as max(0, *xc* - 0).

of multiple variables using the product of their basis functions, so the interaction effect is also specified locally: it is confined to a sub-region described by the nonzero parts of every relevant variable's basis function, rather than across the full range of these variables. Thus, MARS produces a parsimonious regression model that contains local nonzero components only applied to where they are needed. Mathematically, the regression model has the form

$$f(X) = \beta_0 + \sum \beta_k b_k(X)$$

Each term $b_k(X)$ is a basis function or a product of two or more basis functions indicating the interaction effects of corresponding variables. The coefficient β_k of each basis function or their products defines the slope of the non-zero section.

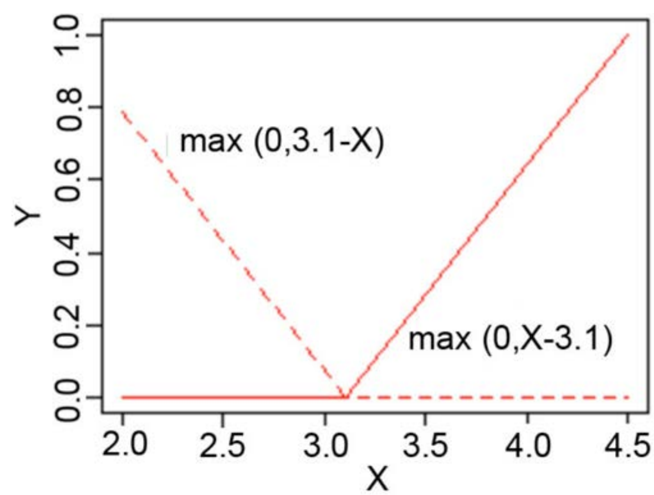


Figure 20. A pair of reflective basis functions
(knot c = 3.1)

The major task of spline regression is to determine the number of knots and their locations, which is a big challenge when the data is of high dimension (i.e., multivariate). Unlike traditional spline regression in which the knot positions are predetermined and often evenly spaced, MARS automatically searches for optimal knot locations based on the data. The search process adds one knot at a time and steadily increases the number, as illustrated in **Figure 21**. Given multiple candidate variables, each of which has one or more knots that can be placed at any positions within the value range, the algorithm selects basis functions whose entry to the regression function will produce the largest decrease in the residual sum of squares (RSS). The selection process continues until a predetermined maximum model size (i.e., the number of basis functions in the MARS model) is reached. Then the algorithm switches to a backward pruning procedure because the forward stepwise addition procedure runs the risk of overfitting, i.e., capturing not only the main features but also random fluctuations in the data. The pruning procedure remove one at a time any non-constant basis functions that no longer make sufficient contribution to the model. At this stage, a variable can be dropped from the MARS model, if none of its basis functions remain. The backward procedure produces a sequence of models, among which the one with the lowest value of generalized cross-validation (GCV) (Craven & Wahba 1978) will be chosen by the algorithm.

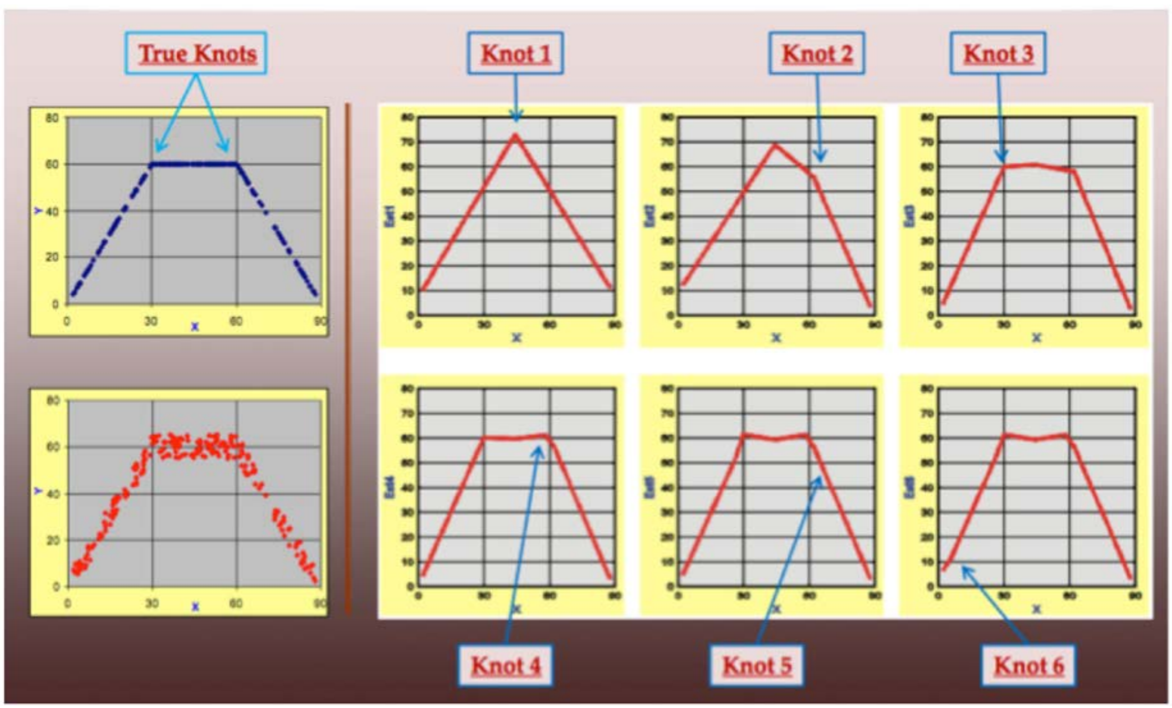


Figure 21. A stepwise knot selection procedure of the MARS algorithm (Steinberg 2013)

The goodness of a MARS model was evaluated using four measures. The first two measures –

the residual sum of squares (RSS) and the coefficient of determination (RSq = 1 – RSS/total sum of square) – are borrowed from traditional regression analysis. The third measure GCV is RSS penalized for model complexity. Specifically,

$$GCV = RSS / n\,[1 - c(m)/n]^2$$

where n is the number of observations in the data set and $c(m)$ is the cost complexity measure of a model containing m basis functions[73]. The purpose of using GCV is to avoid overfitting the data or creating overly big MARS models. The fourth measure GCVSq normalizes GCV the same way as RSq normalizes RSS. In general, a good MARS model should show large RSS and GCV and have RSq and GRSq close to 1.

By flexibly representing the response surface and by using a two-stage predictor selection procedure, MARS is capable of reliably tracking complex (nonlinear) patterns hidden in high-dimensional data. MARS is similar to the well-known CART (Classification and Regression Tree) model as they both rely on automatic recursive partitioning to select predictor variables. But MARS can better handle numeric data because basis functions (smooth curves) are more appropriate for continuous variables than the step functions (constant segmentation) used by CART. Moreover, MARS allows for more explicit representation and easier interpretation of interaction effects. Figure 22 compares the estimated response surface of the same data set using CART and MARS. The R package "earth" (version 3.2-7) [74] (Milborrow 2014) was applied to estimate MARS models from experimental results. The algorithm was tuned to detect up to four-way interactions, i.e., products of four different basis functions.

[73] For MARS, $c(m)$ = m + Penalty * (m – 1) / 2, where Penalty is 2 or 3. (m – 1) / 2 is the number of hinge-function knots, so the formula penalizes the addition of knots.
[74]MARS® is one of the main algorithms used in Salford Systems commercial predictive analytics software. To avoid trademark infringements, some open-source implementations of MARS are called "Earth".

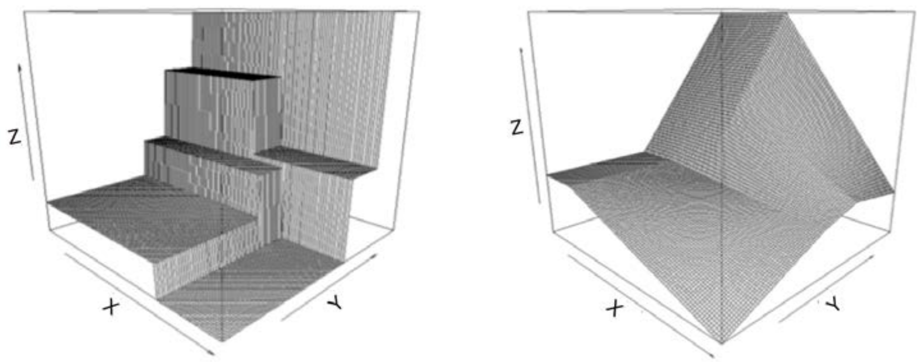


Figure 22. The response surface of the same data set estimated by CART (left) versus MARS (right) (Deppa 2014)

4.3. Results of Simulation Experiments

Simulation experiments were conducted on the computational model using the experimental designs mentioned above. The first and second crossed designs yielded 7,200 (300 × 2 × 3 × 4) and 3,000 (300 × 2 × 5) design points respectively. Every design point has 300 replicate simulation runs. The initial setting of each simulation run – the NK space and the starting position of every agent on the space – was created using Friedman's source code (Friedman 2007) with predefined values of N and K. The five different initial settings required by the second crossed design were created using the same N and K yet different random seeds. Each simulation run lasted for 1,000 or 1,200 time steps or ticks. The measure of organizational problem-solving performances (*avgScore*) were collected at each time step for each run and the medians[75] across 300 replicated runs were calculated as stepwise outcomes. This section describes the results of simulation experiments (**Table 4**), which confirmed most of the previous findings while providing some new insights.

[75]I used median to avoid making assumptions on underlying distributions of the data. In this regard, median is a non-parametric counterpart of mean.

Table 4. Major results of experiment simulations

	Related variables	Result	Interpretation
1	*tick*	The time variable (*tick*) had a decreasingly positive impact on organizational performance (*avgScore*).	The organizational performance improved over time but the improvement was smaller and smaller. In some conditions it stopped improving before achieving the maximally possible value.
2	*randLink*	Organizational members' propensity for embedded or random knowledge exchange (*randLink*) has an inverted-U relationship with *avgScore*.	The organizational performance was better when there was a certain combination of embedded and random knowledge exchanges than when all knowledge exchanges were embedded or random.
3	*actType*	The value of *avgScore* is significantly higher when organizational members' propensity for knowledge exchanges somewhere between 0 and 1 (*actType* = B or D) than exactly 0 (*actType* = A) or 1 (*actType* = C).	The organizational performance was better when there was a certain combination of independent workers and knowledge exchangers than when all members tried to solve the problem alone or by learning from one another.
4	*actType*	When organizational members' propensity for knowledge exchanges follows a normal distribution (*actType* = D) or a power-law distribution (*actType* = B), the difference in *avgScore* has statistical significance but trivial effect size.	There was no substantial change in organizational performance whether or not there is a group of individuals who consistently prefer knowledge exchange to independent knowledge creation.
5	*decayRate*	Except for extremely large values, the decay rate of established network ties (*decayRate*) positively affects *avgScore*.	The organizational performance was better when past knowledge exchanges less efficiently improved the quality and increased the number of knowledge exchange channels.
6	*learnErr*	The accuracy of interpersonal knowledge exchange (*learnErr*) positively affects *avgScore*.	The organizational performance was better when the exchange of different knowledge was more accurate.

	Related variables	Result	Interpretation
7	*innoRange*	Organizational members' independent knowledge creating abilities (*innoRange*) has an inverted-U relationship with *avgScore*	The organizational performance was better when independent workers explored new solutions neither too conservatively nor too progressively.
8	*space_k*	Low interdependence among knowledge areas (*space_k*) corresponds to high *avgScore*	The organizational performance was better when the external problem is less complex.
9	*orgSize*	The value of *avgScore* increases with organization size (orgSize) with a declining rate.	When organization size is sufficiently big, further enlarging the organization will not improve organizational performances.
10	*randLink, tick*	Given high problem complexity and complete knowledge exchanges, the positive effect of *randLink* on *avgScore* was more apparent when *tick* is large. Given low problem complexity and partial knowledge exchanges, the same effect was more apparent when *tick* is small.	When problem complexity was low and there was a small group of frequent knowledge exchangers in the organization, the positive effect of random knowledge exchanges was more apparent in the short run. The same effect was more apparent in the long run when problem complexity was high and all individual members were willing to exchange knowledge.
11	*randLink, decayRate*	Given small *randLink*, the positive effect of *decayRate* on *avgScore* was stronger.	Fast tie decay improve organizational performances even more when organizational members prefer embedded knowledge exchanges.

The analysis first examined whether the NK space based parallel problem solving has been appropriately integrated into this study's model. To this end, how organizational performances (*avgScore*) changed with time (tick), problem complexity (*space_k*), or individuals' independent knowledge creating ability (*innoRange*) was examined. **Figure 23** shows the temporal changes of *avgScore* under different experimental conditions[76]. The trend lines all follow the same pattern – going up and then entering a relatively stable state[77], but not all of them reach the maximal possible value (*avgScore* =1) in the end (*tick* = 1,200). The results indicate that organizational performances were improved smoothly over time with a declining rate. Different experimental conditions impacted the improving rate and/or the stable value of organizational performances but never reversed the progress. Generally, the trend lines in the first three columns (*space_k* = 1) have a higher stable value than the trend lines in the last three columns (*space_k* = 5). It means long-run organizational performances (stabilized *avgScore*) were lower when the problem was more complex. This negative effect of *space_k* ($p < 0.001$, effect size $|r| = 0.7$) was further confirmed by KW and MW test results (Appendix E). Finally, if the NK model was successfully implemented, there should be an inverted-U relationship between individual independent knowledge creating ability (*innoRange*) and *avgScore* and the relationship would be more apparent over time or when the problem is more complex (see Section 2.2 **Figure 4** and in-text description where the pace of individual agents' random jumping amounts to *innoRange*). These expectations were confirmed by plotting partial results of the first cross design (*actType* = A) where *innoRange* is the only effective primary model parameter (**Figure 24**). The data points form a curve, which, as expected, is clearer at higher ticks and when *space_k* = 1.

[76]Based on the results obtained from the first crossed design
[77]The value of *avgScore* either completely or almost levels off.

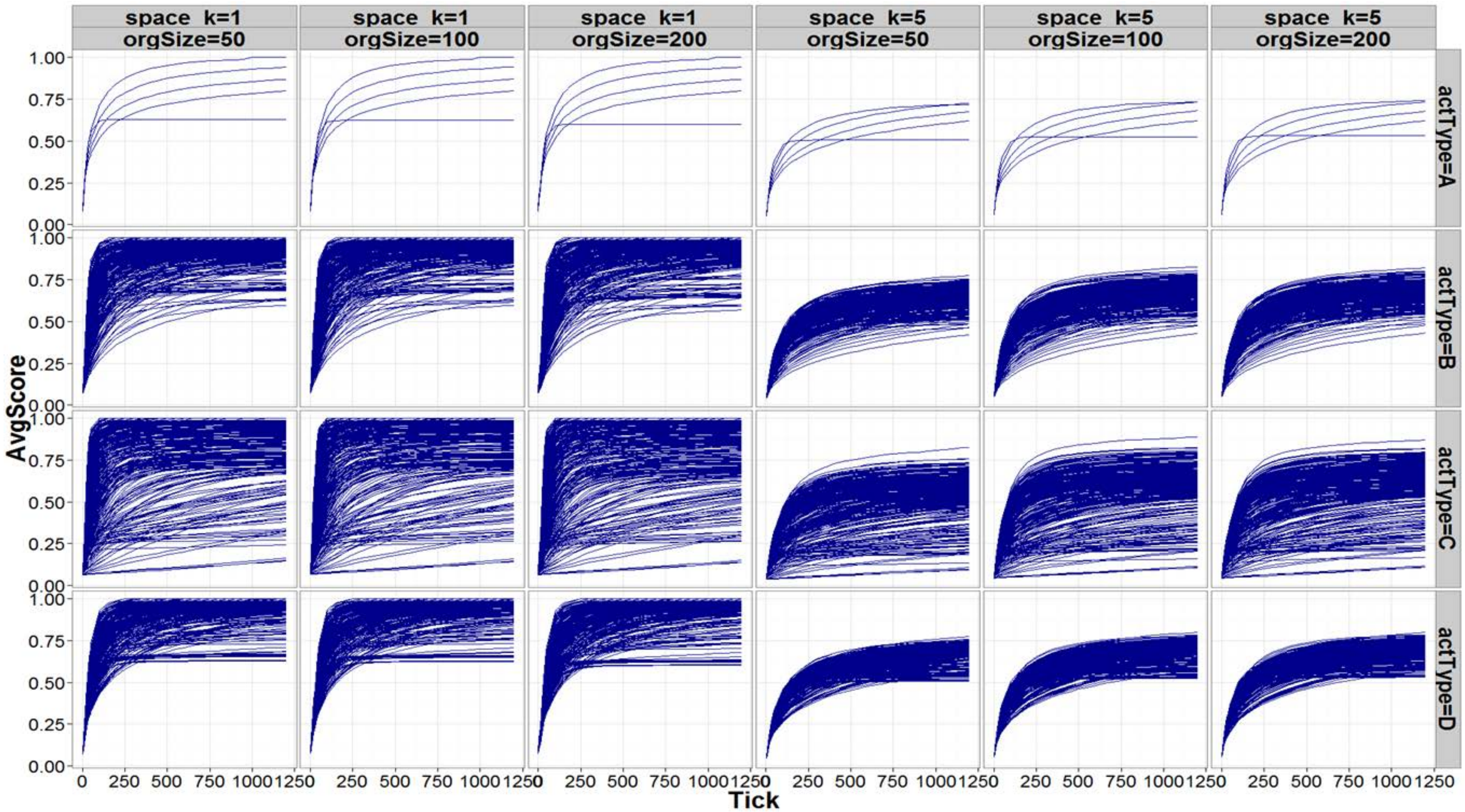


Figure 23. Changes in *avgScore* over time under different experimental conditions[78]

[78] The x-axis indicates elapsed time (*tick*), and the y-axis indicates organizational performance (*avgScore*). The columns represent different combinations of problem complexity (*space_k* = 1 or 5) and organization size (50, 100, or 200). The rows represent different probability distributions of individual agents' willingness to exchange knowledge: *actType* = A, B, C, or D indicates that the probability is all zero, power-law distributed, all one, or normally distributed. When *actType* = B, C, or D, each trend line corresponds to 1 out of 300 LHS design points. When *actType* = A, none of the primary model parameters except for *innoRange* is effective, so there are only five trend lines corresponding to the five different values of *innoRange* (1, 2, 3, 4, or 5).

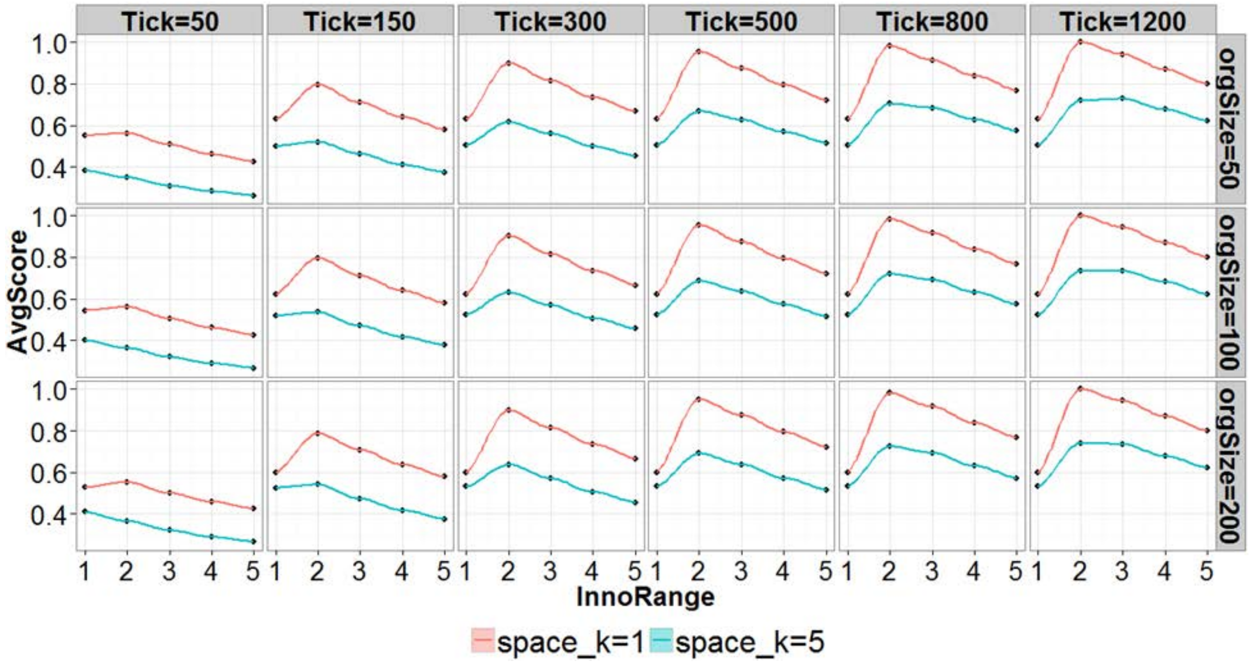


Figure 24. The inverted-U effect of *innoRange* on *avgScore*

To better observe the distributions of *avgScore* in different experimental conditions, the density functions of *avgScore* (across 300 LHS design points) at *tick* = 1,200 is plotted (**Figure 25**)[79]. In majority, *actType* = B or D (i.e., some agents prefer knowledge exchange) outperformed *actType* = C (i.e., all agents prefer knowledge exchange) and *actType* =D (individual propensity for knowledge exchange follows a normal distribution around 0.5) slightly outperformed *actType* =B (individual propensity for knowledge exchange follows a power-law distribution[80]). If agents were good at independent knowledge creation (e.g., *innoRange* = 5), *actType* = A (i.e., no agent prefers knowledge exchange) outperformed *actType* = B or D, especially when the problem is complex (*space_k* = 5). The KW and MW tests on *actType* (**Appendix E**) confirmed that *avgScore* was significantly high ($p < 0.001$, $|r| > 0.1$) when some but not all members are willing to exchange knowledge (*actType* = B or D vs. *actType* = A or C). The difference between

[79] **Appendix F** presents a few figures for other ticks.
[80] There are a few agents who constantly exchange knowledge while most others only occasionally do so.

actType = B and *actType* = D was only significant ($p < 0.01$) in the long run and the effect size was small ($|r| < 0.1$).

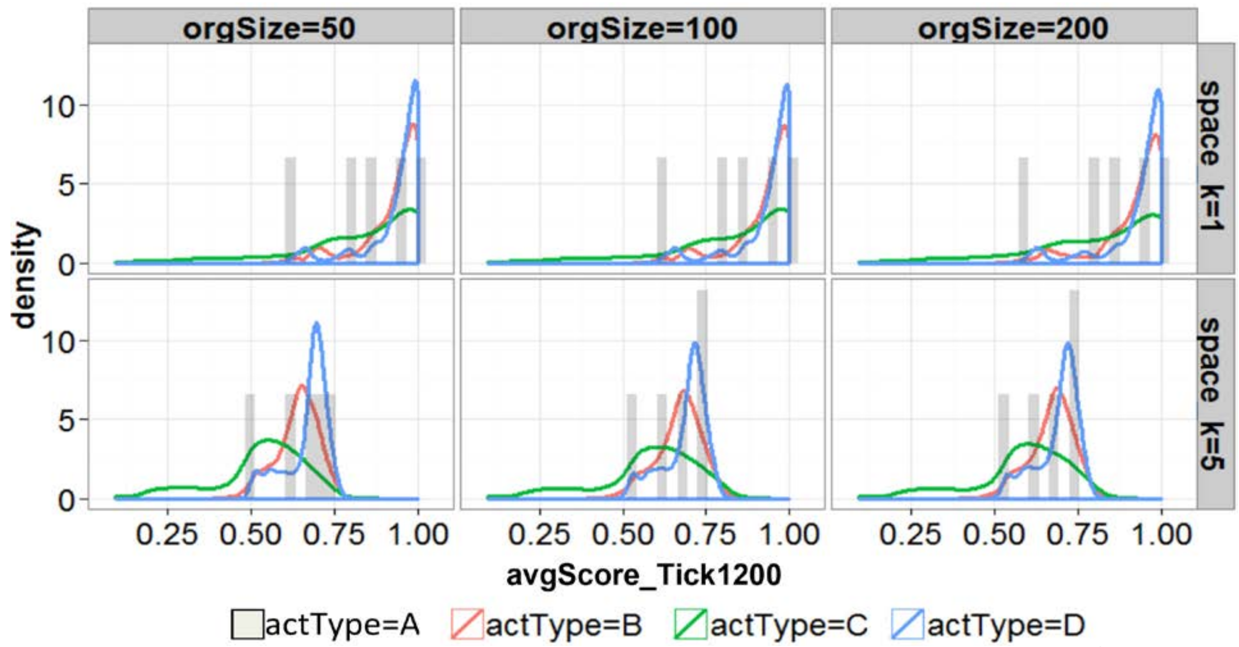


Figure 25. The probability density functions of *avgScore* (*tick* = 1200)[81]

The effects of continuous model parameters were examined using MARS[82] as implemented by R package "earth". In the MARS models presented hereafter, the dependent variable was always *avgScore*, whereas independent variables, knots, and basis functions varied with the underlying data set. For simplicity and the ease of result interpretation, the algorithm was tuned to detect up to three-way interactions between basis functions (i.e., the product of no more than three different basis functions). The first MARS model (**MARS-I**) was built on partial results of the first crossed design – at 32 out of 1,200 time steps and without the subset of *actType* = A. The model explained 93.2% of the variations in *avgScore*, the dependent variable. The regression

---

[81]The results of *actType* = A were shown in histograms because in this condition only one of the six LHS variables (*innoRange*) affects model outcomes and it is discrete (*innoRange* = 1, 2, 3, 4, or 5). When *space_k* = 1, the density plots are truncated at their right side because *avgScore* cannot increase outside its defining range ([0, 1])
[82]MARS can handle both continuous and discrete variables.

function was a linear combination of 17 basis functions containing 7 uncorrelated independent variables (*tick*, *space_k*, *actType*, *innoRange*, *randLink*, *decayRate*, *learnErr*) automatically selected from the ten parameters of this study's agent-based model (**Table 3**). All basis functions were statistically significant. The other three model parameters (*orgSize*, *actDist*, and *wtGain*) were considered as having no impacts on *avgScore*. Details on **MARS-I** are presented in **Table 5**.

Table 5. MARS-I and the corresponding experimental conditions

Experimental Condition			
tick		1, 6, 11, 16, 21, 26, 31, 36, 50, 100, 150, … 1,200	
actType		B, C, D	
Number of observations[83]		172,800	
Number of independent variables		7	
Regression model			
Component	**Description**	**Coefficient**	***P* value**
Intercept		0.818	< 0.0001
BF1	max (0, Tick – 150)	0.001	< 0.0001
BF2	max (0, 150 – Tick)	–0.004	< 0.0001
BF3	space_k = 5	–0.285	< 0.0001
BF4	max (0, randLink – 0.382)	–0.066	< 0.0001
BF5	max (0, 0.382 – randLink)	–0.102	< 0.0001
BF6	max (0, decayRate – 0.978)	1.340	< 0.0001
BF7	max (0, 0.978 – decayRate)	–0.155	< 0.0001
BF8	max (0, learnErr – 0.855)	–0.821	< 0.0001
BF9	max (0, 0.855 – learnErr)	0.194	< 0.0001
BF10	BF2 * BF3	0.002	< 0.0001
BF11	BF1 * (actType = C)	–0.001	< 0.0001
BF12	BF1 * max (0, innoRange – 2)	–0.001	< 0.0001
BF13	BF1 * max (0, 2 – innoRange)	–0.001	< 0.0001
BF14	(actType = C)* BF5	–0.607	< 0.0001
BF15	(actType = C)* BF8	–1.700	< 0.0001
BF16	BF2 * BF3 * max (0, learnErr – 0.868)	0.008	< 0.0001
BF17	BF2 * BF3 * max (0, 0.868 – learnErr)	–0.001	< 0.0001
Goodness-of-fit			
GCV (Generalized Cross Validation)		0.00555	
GCV-Squared		0.932	
RSS (Residual Sum of Squares)		959	
R-Squared		0.932	

[83]The result set has 172,800 items: 3 (*orgSize* = 50, 100, or 200) × 3 (*actType* = B, C, or D) × 2 (*space_k* = 1 or 5) × 300 (LHS design points) × 32 (*tick* = 1, 6, 11, 16, 21, 26, 31, 36, 50, 100, 150, …, 1,200).

Interpreting MARS results requires combining related basis functions. For example, the general effect of *decayRate* is a combination of BF6 and BF7 in **Table 5**: 1.34 × max (0, *decayRate* – 0.978) – 0.155 × max (0, 0.978 – *decayRate*). When *decayRate* > 0.978, the effective coefficient of *decayRate* is 1.34; when *decayRate* < 0.978, this coefficient is 0.155. Overall, it means that *avgScore* is positively related with *decayRate*. Interaction effects (if any) should be considered as well. For example, the general effect of *randLink* is – 0.066 × max (0, *randLink*–0.382) – 0.102 × max (0, 0.382 – *randLink*) – 0.607 × (*actType* = C) × max (0, 0.382 – *randLink*). When *actType* = C, the equation is simplified to – 0.066 × max (0, *randLink* – 0.382) – 0.709 × max (0, 0.382 – *randLink*); when *actType* ≠ C, the equation is simplified to – 0.066 × max (0, *randLink* – 0.382) – 0.102 × max (0, 0.382 – *randLink*). It suggests a curvilinear relation between *randLink* and *avgScore*: when *randLink* > 0.382, its increment slows down the growth of *avgScore*; otherwise (*randLink* < 0.382), *avgScore* grows faster as *randLink* increases. In addition, when *randLink* < 0.382, the same increment of *randLink* accelerates the growth of *avgScore* even more when *actType* = C.

MARS-I showed some earlier results obtained from data visualization and/or KW and MW tests, validating the application of MARS on experimental results. First, organizational performance *avgScore* increased with *tick* but the rate significantly decreased after *tick* >150 (BF1 & 2). Second, *avgScore* was bigger when *space_k* = 5 (BF3). Third, *avgScore* increased more slowly in the long run when *actType* = C (BF11). Fourth, *avgScore* grew at an increasing rate as *innoRange* went up to a threshold (*innoRange* = 2); after that, as *innoRange* increased, *avgScore* grew more slowly (BF12 & 13). **MARS-I** also detected new effects. First, *avgScore* grew at an increasing rate as individual agents' propensity for random knowledge exchange *randLink* went up to a threshold (*randLink* = 0.382); once passing the threshold, the increment of *randLink* would slow down the growth of *avgScore* (BF4 & 5). This finding was consistent with the postulation of the earlier conceptual model (**Figure 13**). Among all states of the macro interaction network shown in this figure, only State D is a hybrid topology with both closure and brokerage sub-structures. The evolutionary paths crossing State D suggest that for the network to enter and stay at this state, there must be both random and embedded knowledge exchanges in the system.

Second, the error rate of solution imitation *learnErr* had a negative effect on the growth of *avgScore* (BF8 & 9), which is intuitively expected because inaccurate knowledge exchanges would be a waste of time. The same effect was studied before with the same measure[84] yet limited to a small error rate (20%) and based on a static interaction network (Lazer & Friedman 2007)[85]. It was found that comparing with zero error rate, a small error rate led to better long-run organizational performances by preserving individual diversity for a longer time. This effect dwindled over time as individual solutions gradually assimilated and it was more apparent when the problem was more complex[86]. If the current study's agent-based model also implemented these effects, the prevailing negative effect of *learnErr* should become relatively weak in the conditions when the previous study detected significant positive effects – at the early stage of problem solving or given high problem complexity. These interaction effects were confirmed by

---

[84] In both the previous and the current studies, the accuracy of knowledge exchanges was defined as the probability that an agent misidentifies a knowledge area when she imitates the solution of another agent.

[85] In contrast, the current study investigated a wider range of error rate (from 0 to 1, inclusive) based on a dynamic network that evolved with knowledge exchanges on top of it. In this network, every individual could interact with different others and strangers during the problem solving process rather than repeatedly with the same others each time she decides to exchange knowledge. Thus, individual solutions converge at a slower rate and thus individual diversity is better preserved in the current study than in the previous study.

[86] Recall that keeping diversity was important for overcoming semi-optimal peaks on a rugged NK landscape.

MARS-I (BF16 & 17), suggesting that the seemingly conflicting results of the current and the previous studies were actually compatible.

Third, the variable that controlled the speed of tie decay, *decayRate*, had a positive effect on the improvement of *avgScore* (BF5 & 6). There are two interrelated explanations for this result. On the one hand, in this study's agent-based model, tie strength serves as the "bandwidth" of knowledge exchanges – the maximal amount of knowledge exchangeable between two agents per interaction increases with the strength of their tie. Fast tie decay would help maintain a relatively low level of bandwidth, therefore slowing down knowledge dissemination and preserving knowledge diversity. On the other hand, infrequently used ties kept decaying and eventually dissolved, so fast tie decay would lead to low global connectivity, which also slowed down knowledge dissemination and preserved knowledge diversity.

Finally, the effects of *randLink* and *learnErr* on *avgScore* was stronger when *actType* = C (BF14 & 15), as both variables were influenced by the frequency of knowledge exchanges, which achieved the highest level when agents all choose knowledge exchange over independent knowledge creation. This interaction between knowledge exchange frequency and interaction-related model parameters probably explains the dispersed trend lines when *actType* = C in **Figure 23**. **Figure 26** visualizes the main effects described above using partial dependence plots, one plot for an independent variable. In each plot, the x axis indicates values of the focal independent variable, and the y axis indicates *avgScore*. The blue line is the loess smoothing[87] of

[87] The name "loess" are derived from the term "locally weighted scatter plot smooth," as this method uses locally weighted linear regression to smooth data. The smoothing process is local – each smoothed value is determined by

actual *avgScore* in the data set used to create the **MARS-I** model. The black line is the MARS model prediction of *avgScore* by varying the focal independent variable while fixing other independent variables at their median values. The predicted curve and the actual curve (after smoothing) of *avgScore* have similar shapes in these plots, but the gap between two curves suggests an overestimation of **MARS-I** on the extent to which each independent variable impacts *avgScore*.

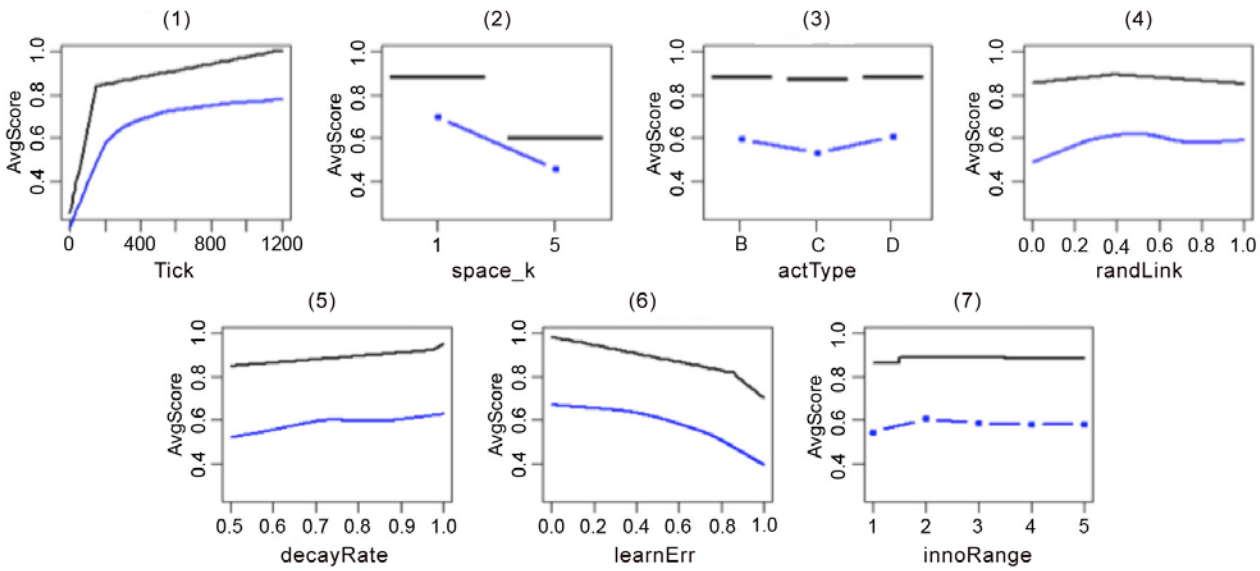


Figure 26: Visualization of the main effects in MARS-I

MARS-I was built on a data set aimed to cover the response surface of the agent-based model as wide and equally as possible. Since the MARS algorithm selects variables and knots based on how much they reduce RSS (or GCV), independent variables with strong global effects (i.e., significant across an extensive area of the response surface) will first enter the regression model, possibly shadowing locally significant variables (i.e., variables whose effects are significant only at some constrained areas of the response surface). As a result, the effects of globally significant

---

neighboring data points falling in a predefined span. The process is weighted because a regression weight function is defined for the data points contained within the span. Finally, the process used a quadratic polynomial model in the regression.

variables may be overestimated. As shown in **Table 6**, *avgScore* was mostly affected by time (*tick*) affects, and then to decreasing extent by problem complexity (*space_k*), the accuracy of interpersonal knowledge exchange (*learnErr*), the probability distribution of knowledge exchange intentions (*actType*), organizational members' propensity for random rather than embedded knowledge exchanges (*randLink*), the decay rate of network ties not used for interactions (*decayRate*), organizational members' independent knowledge creating ability (*innoRange*). The most influential variables are globally significant ones such as *tick*, *space_k*, and *learnErr*. The variables of the most interests (*actType*, *randLink*, *decayRate*, and *wtGain*), that is, variables directly influencing the generation and maintenance of a hybrid network topology, either are not so important or do not exist in the regression function at all.

Table 6. Descending importance of MARS-I variables in terms of predicting *avgScore*[88]

Variable	Importance (-gcv)	Importance (-rss)
tick	100.0	100.0
space_k	60.3	59.9
learnErr	42.1	41.4
actType	34.6	33.7
randLink	34.6	33.7
decayRate	15.2	15.6
innoRange	12.6	13.0

To more accurately estimate the effects of locally significant independent variables, additional MARS analysis were conducted on four subsets of the experimental results obtained from the first crossed design. These subsets were identified using **MARS-I**. In each subset, the values of *space_k* and *actType* were fixed (*space_k* = 1 or 5, *actType* = B or C) and the value ranges of *tick*

[88] The number shown in both "Importance" columns indicates the relative importance (percentage) of all variables as compared to the most important one.

and *learnErr* were curtailed (*tick* = 100 to 300 with an interval of 50, *learnErr* ≤ 0.15). The fitted MARS regression models (**MARS-II**, **MARS-III**, **MARS-IV**, and **MARS-V**) and the importance ranking of each model's independent variables are shown in **Appendix G**. Some earlier detected local effects was more apparent in new MARS models where global effects were better controlled. For example, the curvilinear relationships between *randLink* and *avgScore* is clearer in **Figure 27** than in **Figure 26** (D).

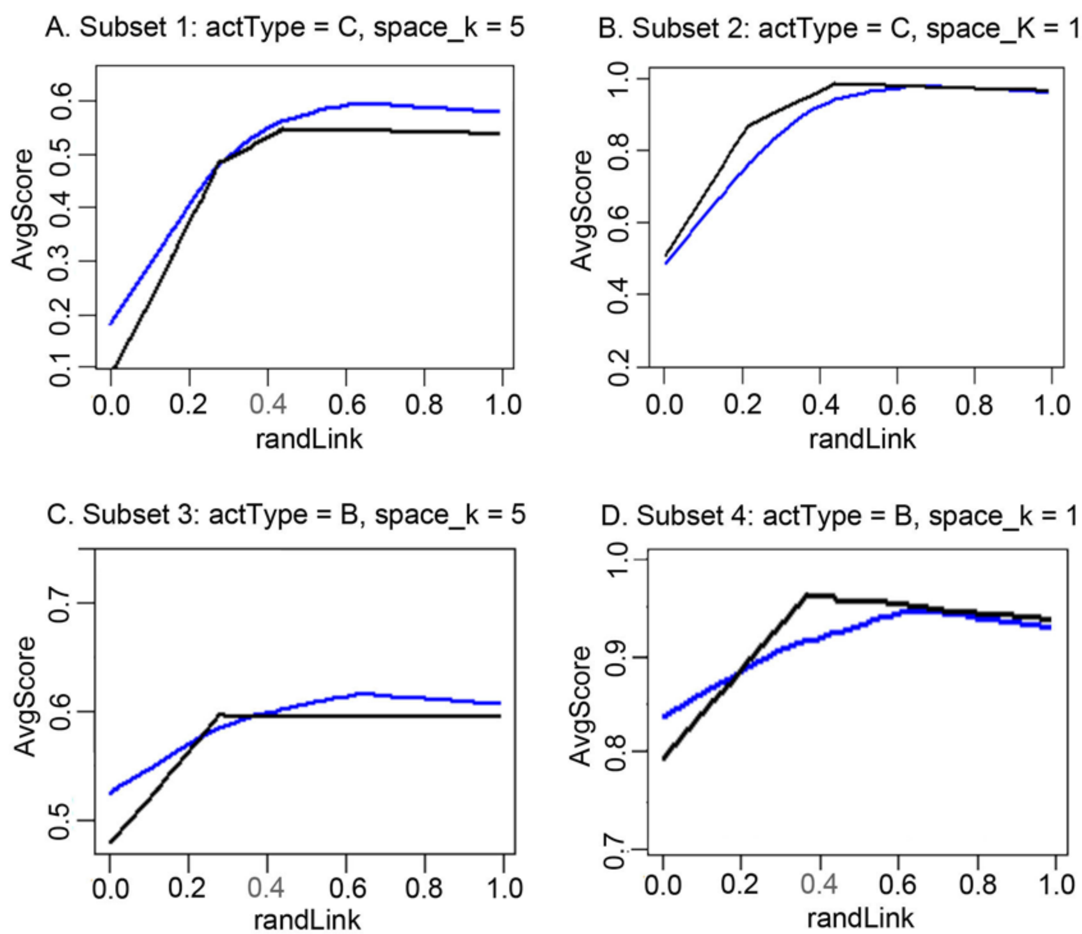


Figure 27: The effect of *randLink* in MARS-II to MARS-V (partial dependence plots)

In addition, new influential factors and new effects stood out. First, while **MARS-I** suggested a monotonically positive relationship between *decayRate* and *avgScore*, further analysis revealed that a negative effect when *decayRate* is very large (e.g., **MARS II** BF8 & 9, **MARS III** BF3 &

4). This finding was consistent with the postulation of the conceptual model (**Figure 13**) that overly fast tie decay would undermine organizational performances by preventing the emergence of any sub-structures in the macro network, including closure and brokerage structures. Moreover, *randLink* and *decayRate* had significant interactive effects. Small *randLink* reinforced the positive effect and weakened the negative effect of *decayRate* on *avgScore* (e.g., **MARS II** BF14-16, **MARS III** BF10, 11, 15, & 16). Larger *decayRate* weakened the positive effect of *randLink* (e.g., **MARS IV** BF13 & 14), as fast tie decay reduced the chance for random knowledge exchanges to create bridges. To better understand how the impact of *decayRate* on *avgScore* were affected by *randLink*, a series of simulation experiments were further conducted, focusing on structural features of the emergent macro network (see **Appendix H**). Notably, when *randLink* was small, increasing *decayRate* (yet the absolute value remains moderate) counterintuitively promoted the creation of new ties by reducing the tendency for knowledge exchanges to happen on existing ties. In other words, when knowledge exchanges tended to be embedded, fast tie decay improves *avgScore* by keeping the macro network from evolving into a "lock-in" state where different knowledge clusters remain isolated (State G in **Figure 13**).

In addition, increasing *wtGain* reinforced the (positive) effect of small *randLink* but had no impact on the (negative) effect of big *randLink* (**MARS-II** BF12 & 13, **MARS-III** BF8 & 9). This result reflected the model design that *wtGain* acted like memories guiding (only) embedded knowledge exchanges. The impact of *wtGain* was stronger when its value was bigger and when there were more frequent interactions on the network (*actType* = C). The variable *actDist* controlled the number of active knowledge exchangers and it was only effective when *actType* = B. General, *actDist* showed a positive effect on organizational performance and that effect interacted with *randLink*. Finally, given high problem complexity (*space_k* = 5) and frequent knowledge exchanges (*actType* = C), the positive effect of *randLink* on *avgScore* was more apparent in the long run (*tick* >150). Given low problem complexity (*space_k* = 1) and infrequent knowledge exchanges (*actType* = B), however, the effect was more apparent in the short run (*tick* < 200).

The last step is to examine the robustness of earlier results against certain noise factors. The first noise factor is organizational size (*orgSize*). In **Figure 23** and **Figure 25**, there is no visual difference in the growth of *avgScore* as *orgSize* varies, regardless of the value of *space_k*. The KW and MW tests (**Appendix E**) showed that *avgScore* was significantly lower ($p < 0.02$) when *orgSize* = 50 than when *orgSize* = 100 or 200, but the effect size was trivial ($|r| \leq 0.06$). **MARS-I** did not select *orgSize* into the regression function. **MARS-II** and **MARS-IV** (*space_k* = 5) included *orgSize* = 100 and *orgSize* = 200 as positive terms of the regression function, but neither of them interacted with other independent variables (**Appendix G**, **Table A8** and **Table A11**). Moreover, both terms had similar coefficients, indicating similar effect sizes. Together, these results suggest that *avgScore* increased with *orgSize* in a declining rate (**Figure 28**), as documented in the literature (Keller 1986; McGrath 1984; Shaw et al. 1981), and that the effects of other important variables (especially network topology related variables) on *avgScore* were robust to changes in *orgSize*, as also found in a previous study (Fang et al. 2010).

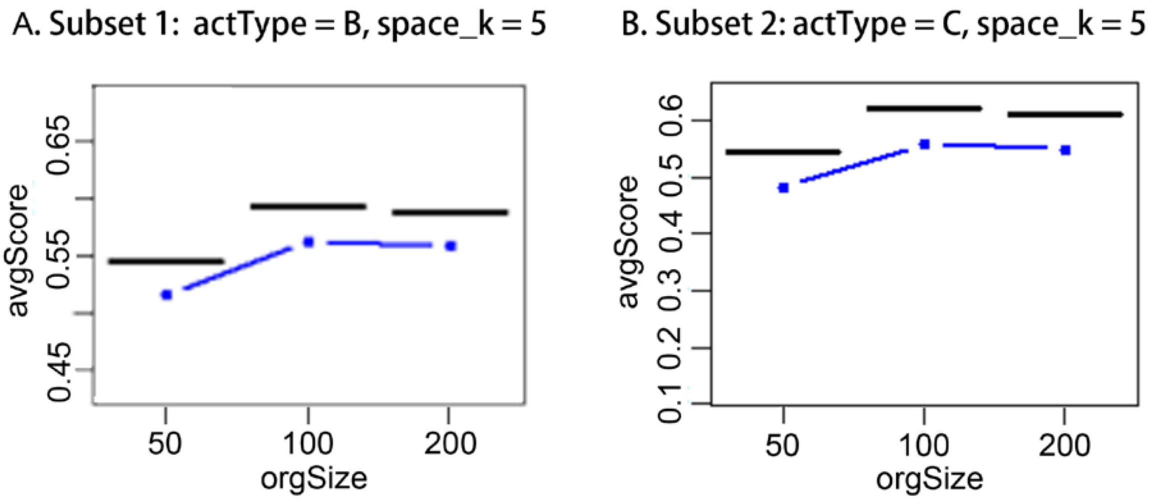


Figure 28: The effect of *orgSize* in MARS-II and MARS-IV (partial dependence plots)

Another noise factor is the NK space that serves as the initial condition of a simulation run (identified by *nkSpace*). Some modeling studies (Fang et al. 2010; Lazer & Friedman 2007) preferred to reduce the variations in simulation results caused by this factor rather than test the robustness of simulation results to this factor. Since the variations are expectedly large, reducing them requires running the model on a large number of different NK spaces[89] and then averaging all the results. The model of the current study is computational intense, so running it on many NK spaces would be very time-consuming. Thus, I decided to evaluate the model's robustness to different NK spaces. To this end, MARS analysis was conducted on the experimental results of the second crossed design. The fitted model (**MARS-VI**) is presented in **Table A15 (Appendix G**). The regression function included *nkSpace* as an important independent variable, as some NK spaces (*nkSpace* = 4 and 5) produced lower *avgScore* than others and some *nkSpace* had an interaction effect with *space_k*, a defining property of NK spaces. However, there was no interaction between *nkSpace* and independent variables that are irrelevant to NK spaces.

[89] These NK spaces are defined by the same *N* and *K*. Lazer and Friedman used 10,000 NK spaces, whereas Fang and colleagues used 200 NK spaces for each experimental condition.

The third noise factor was the probability distribution of individual agents' tendency to exchange knowledge. Although *avgScore* was generally higher when *actType* = A or C than when *actType* = B or D, no significant difference was expected between *actType* = B (a power-law probability distribution) and D (a normal probability distributions). Graphic visualization (**Figure 25**) and the MW test result (**Appendix E**) suggest that a normal distribution[90] outperformed a power-law distribution[91] ($p < 0.001$ when *tick* = 800 and 1200) in improving *avgScore*, but the difference is trivial ($|r| \leq 0.08$). Also, **MARS-I** did not select *actType* = B or D into the regression function.

[90]Fifty percent agents are willing to exchange knowledge.
[91]a few agents constantly exchange knowledge while most others only occasionally would like to participate.

Chapter 5

DISCUSSION AND CONCLUSION

This chapter discusses the contributions and limitations of the current study. The major findings have theoretical and practical implications on several areas such as organizational ambidexterity, organizational social capital, and organizational networks. The extended network topology of this study's model is a modeling innovation that can benefit future research. The limitations of the current study lie in its simulation essence and potential extensions in both modeling and analysis.

5.1. Major Findings and Implications

The current study was motivated by a recent discovery that a hybrid macro network (with both closure and brokerage structures) is a promising structural approach to balancing micro knowledge exploitation and exploration activities and achieving high collective performance in the long run (Fang et al. 2010; Lazer & Friedman 2007). However, this discovery was made based on a predefined and static network; thus, the question of how to generate and maintain such a network remains. To tackle this problem, I looked into micro mechanisms that would strengthen or weaken closure and brokerage structures and mapped out possible evolutionary paths of a hybrid macro network (**Figure 13**). The overarching hypothesis was that the longer a hybrid network persisted, the better the collective performance would be in the long run. The hypothesis was tested on an agent-based model in which the preceding micro mechanisms and the macro network were co-evolving, so that the network could be emergent and dynamic.

The simulation results confirmed the existence of a trade-off between (a) rapidly disseminating knowledge to improve organizational problem-solving performances in the short run and (b) enduringly maintaining knowledge diversity to guarantee organizational problem-solving performances in the long run. This trade-off reflects the fundamental need for an organization to resolve the tension between exploitation and exploration (March 1991). The results also echoed previous findings on the moderating effects of exogenous factors such as organization size, problem complexity, organizational members' independent problem-solving attributes, and the configuration of intra-organizational interaction structure. As for new findings, the results revealed that long-run organizational problem-solving performances were affected by the collaboration propensities of regular organizational members and the residual influence of organizational members' past collaboration.

In organizational research, the specific research topic and modeling scenario of the current study – how individual members' social-capital-based knowledge exchanging and creating behaviors affect organizational problem-solving performances – is related to more general research on how the structural dimension of social capital affects knowledge creation or exchange in organizations (Inkpen & Tsang 2005; Maurer et al. 2011; McFadyen & Cannella 2004; Nahapiet & Ghoshal 1998; Tsai & Ghoshal 1998; Wei et al. 2011), a topic following the knowledge-based view of organizations[92]. The scenario is also related with the study of contextual ambidexterity (Gibson & Birkinshaw 2004; McCarthy & Gordon 2011), which relies on behavioral and social means to balance two organizational learning approaches, exploitation and exploration, at the organizational or unit level. Both lines of research are concerned with outcomes, i.e., individual or organizational performances. They overlap when social capital serves as the dominant resource (Turner et al. 2013) for balancing exploitation and exploration. In this regard, social capital refers to "the sum of actual and potential resources embedded within, available through, and derived from…the network possessed by an individual (or a collectivity), and the assets…mobilized through that network" (Nahapiet & Ghoshal 1998). The structural dimension of social capital is the overall configuration of connections between actors, which constitute information channels that reduce the amount of time and investment required to access knowledge. From a network structure perspective, there are two major types of social capital rooted in closure and brokerage structures respectively and having complementary effects (Adler & Kwon 2002; Burt 2000b; Oh 2004; Oh et al. 2006; Putnam 2000; Reagans & McEvily 2008). Thus, a hybrid network consisting of both structures has better outcomes than a pure closure or brokerage structure with regard to balancing knowledge exploration and exploitation and improving organizational performances, as shown by empirical evidence (Tiwana 2008) and simulation results (Fang et al. 2010; Lazer & Friedman 2007).

The major contributions of the current study relate to three research areas – organizational ambidexterity, organizational social capital, and organizational social networks (**Table 7**). First

[92] The knowledge-based view of organizations extends the traditional resource-based view of organizations by identifying knowledge as the primary resource for creating new values and gaining competitive advantages (Grant 1996; Kogut & Zander 1992; Nonaka & Takeuchi 1995; Spender 1996; Zander & Kogut 1995)

and foremost, the current study contributes to organizational ambidexterity studies by linking organizational capability in balancing exploitation and exploration to regular organizational members' characteristics that affect their autonomous knowledge exploitation and exploration activities. Organizational ambidexterity studies and the knowledge-based research paradigm of organizations traditionally focus on the higher level, attributing organizational performances to individual-independent factors such as structures, routines, and capabilities (Brown & Duguid 1991; Eisenhardt & Martin 2000; Kogut 2000; Kogut & Zander 1992; Nahapiet & Ghoshal 1998; Spender 1996; Tsoukas 1996; Zander & Kogut 1995; Zollo & Winter 2002). These studies implicitly assume homogeneity at the individual level, neglecting how organizational members might influence organizational balance between exploration and exploitation. The importance of individual characteristics as antecedents to organizational performances has been supported by previous studies that regard individuals as the loci of knowledge[93] (Felin & Hesterly 2007; Grant 1996; Nonaka 1994; Simon 1991) and reinforced by recent arguments that a micro-foundation view is necessary for advancing research on knowledge-based value creation and organizational capabilities (Abell et al. 2008; Felin et al. 2012; Foss 2011). Besides their explanatory power, individual characteristics also have high managerial relevance. As some researchers have pointed out, "what is missing is a clear articulation of those specific managerial actions that facilitate the simultaneous pursuit of exploitation and exploration … what is needed is greater insight into the specific micro-mechanisms …" (O'Reilly III & Tushman 2011). Thus, there have been calls for cross-level research that links organizational ambidexterity to individual attributes or behaviors (Raisch & Birkinshaw 2008; Raisch et al. 2009; Turner et al. 2013). Most studies responding to

[93] "The locus problem may be described as that of selecting the ultimate subject-matter for inquiry in behavioral science, the attribute space for its description, and the conceptual structure within which hypotheses about it are to be formulated" (Kaplan 1964).

these calls looked into the role of senior or middle managers rather than regular organizational members in meeting the contradictory demands of exploitation and exploration (Alexiev et al. 2010; Carmeli & Halevi 2009; Huy 2002; Jansen et al. 2009; Jansen et al. 2008; O'Reilly III & Tushman 2011), because “such decisions cannot be left to the discretion of lower level employee but, at some point, required senior managers to provide the resources and legitimacy…”(O'Reilly & Tushman 2013). This statement neglects the possibility that regular organizational members’ knowledge exploitation and exploration can be balanced without managerial intervention, which is the concern of the current study.

Table 7. Major contributions to related research areas

Research area	Specific issue	Contribution of the current study
Organizational ambidexterity	*Ignorance of regular organizational members* • Lack of cross-level research • Lack of research on the underlying micro-mechanisms of contextual ambidexterity	*The collective power of regular organizational members investigated* • Link organizational performances to regular organizational members’ characteristics that impact independent and collaborative problem solving • Provide a micro-level and informal structure-based demonstration of contextual ambidexterity
Organizational social capital	*Lack of an appropriate synthesis of various social capital sources* • Overemphasis on network positions • Assume network positions are antecedents to motivations and abilities	*Multiple sources of social capital addressed* • Jointly consider individual members’ opportunities, motivations, and abilities to utilize social capital • Separate individuals’ motivations and abilities from their network positions
Organizational social networks	*Lack of an appropriate combination of agency and network structure* • Predominance of structure • A local perspective on agency • Insufficient research on the genesis and dynamics of networks	*Structuration theory faithfully modeled* • Implement the iterative mutual impacts between agency and the global network • Model an emergent and dynamic network whose evolution is pushed by endogenous and exogenous (random) factors

The current study investigated the connection between organizational performances and organizational members' propensities for three collaborative problem-solving behaviors, that is, independent knowledge creation, knowledge exchanges embedded in one's social circle, and knowledge exchanges with random others. Embedded knowledge exchanges create and maintain clusters or closure structures. Random knowledge exchanges occasionally create brokerage structures by bridging the structural holes concomitant with clusters[94]. Instead of creating social capital, independent knowledge creation avoids social liability conveyed by the same social structure that produces social capital (Gabbay & Leenders 2002)[95]. Each of the three behaviors could improve organizational performances over time (**Result 1** in **Table 4** & **Figure 23**), but the (simulated) organization performed the best when all three behaviors existed with certain probabilities (**Result 2 and 3** in **Table 4**). That means these behaviors complemented or reinforced one another's effect. Thus, to improve organizational problem-solving performances in the long run, managers are advised to gather information on how frequently and between whom knowledge exchanges tend to happen inside their organizations and allocate resources accordingly to support all three behaviors[96]. Merely encouraging knowledge exchanges among individual members, as many early studies have suggested, is not enough.

Second, the current study also contributes to organizational social capital studies. While the knowledge-based view of organizations and organizational ambidexterity studies tend to ignore

---

[94] Organizational members are organized around and differentiated by clusters representing various domains of organizational knowledge base. Frequent intra-cluster interactions maintain efficient knowledge dissemination inside clusters while separating clusters from each other. Thus, the emergence of multiple clusters also gives rise to inter-cluster gaps or structural holes.

[95] Staying (temporarily) isolated from the rest of the network, independent workers can independently and efficiently create novel knowledge without being blinded or slowed down (Levine & Prietula 2011)

[96] For example, libraries for independent knowledge creation, chat rooms for embedded knowledge exchanges, social events for random knowledge exchanges

the individual level, organizational social capital studies started with an individual-level focus (Gabbay & Leenders 2002). However, cross-level research is still rare. Even when multiple levels are involved, characteristics of the collective level (e.g., teams, units, or organizations) often serve as moderators of the relationship between individual-level social capital and its individual-level outcomes (Wei et al. 2011). Moreover, the independent variables in previous organizational social capital studies were predominantly about individual members' network positions, that is, certain structural characteristics of an individual member's local network (e.g., degree centrality)[97]. However, an individual's network position only provides opportunities. Whether and how well these opportunities can be seized also depend on the individual's motivation and ability (Burt et al. 1998; Kalish & Robins 2006; Klein et al. 2004; Mehra et al. 2001; Oh & Kilduff 2008; Sasovova et al. 2010). Opportunity, motivation, and capability are three often intertwined sources of social capital (Adler & Kwon 2002). Also, motivation and ability-related individual characteristics (e.g., personality, beliefs, and skills) are more stable, reliable and easier for managers to capture than opportunity-related individual characteristics (i.e., individuals' network positions), especially when the network is informal (i.e., interactions and relationship building are autonomous) and subjects to change (Hallinan & Kubitschek 1988; Kalish & Robins 2006; Snyder & Gangestad 1982), as assumed in the current study. To evaluate individual network positions, managers need to map out relevant intra-organizational networks (partially or entirely), which has been reported as a real challenge for the management (Casciaro 1998; Krackhardt 1990).

[97] This is probably owing to the influence of formalistic sociologists such as Ronald Burt, who advocated focusing on network analysis to better cumulate research on social capital (Burt 2000b). Thus, organizational social capital studies and social network studies have been closely related.

As mentioned earlier, the current study was a cross-level research on the collective outcome of individual-level social capital. Moreover, this study not only accounted for the dynamic nature of opportunity, but also connected it with motivation and ability. Regarding the context of this study, the three sources need to be jointly considered, as knowledge potentially available through a network tie can only reach an individual when both the source and the recipient are motivated and able to exchange knowledge. Specifically, the current study linked organizational problem-solving performances to individual members' network positions (opportunity), knowledge exchange-related individual propensities (motivation), and independent knowledge creation and knowledge-exchange skills (ability), thus providing a holistic and dynamic representation of organizational practice. The impacts of individual motivations on organizational performances have been discussed earlier (**Result 2 and 3** in **Table 4**). The results also showed that organizational performances were positively related with individual members' abilities to accurately exchange knowledge (**Result 6** in **Table 4**). Along with the result on individuals' independent knowledge creation abilities (**Result 7** in **Table 4**), it is suggested that organizations tend to have high problem-solving performances, if organizational members have superior knowledge exchange skills and average independent knowledge creation skills. Another implication on hiring comes from the result on organization size. There seems to be a threshold in organization size; once the threshold is met, hiring more people would have diminishing returns or even undermine organizational performances in parallel problem solving (**Result 9** in **Table 4**). Future research could look into optimal organization size regarding problem characteristics.

Third, the current study contributes to organizational social networks studies. Overemphasizing individuals' network positions at the expense of more fundamental individual characteristics such as motivations and abilities mirrors a problem long existing in organizational social networks studies[98], that is, how to combine agency with network structure (Emirbayer & Goodwin 1994; Ibarra et al. 2005; Kilduff & Krackhardt 1994; Kilduff et al. 2006; Srivastava & Gulati 2014; Stevenson & Greenberg 2000). The lack of appropriate theoretical and technical approaches might be the primary reason why this problem has not yet been thoroughly solved. The current study contributed in both aspects. Theoretically this study followed Gidden's structuration theory (1984), which states that structure shapes agency and the resultant actions while is subsequently affected by those actions. Accordingly, the theoretical framework of this study highlighted iterative mutual impacts between the macro interaction network and individual members' problem-solving behaviors. Recently Srivastava and Gulati (2014) proposed an theoretical framework based on the same logic, but they took the local perspective of a single constrained actor who exercises agency. The global network, which includes multiple actors as well as their relations and interactions, often has characteristics that cannot be reduced to individual actors. The current study was about the global network and it applied Complex Adaptive System (CAS) theories, which attribute irreducible system characteristics to the micro interactions of system components. Thus, the simulated organization was constructed as a complex adaptive social system and organizational members' knowledge exchanges as micro interactions inside the system. Besides interaction, two other CAS mechanisms – selection and variation – are also critical for theorizing the preceding mutual influences. They were

[98] The strong version of formalistic network sociology posits motivation as the effect of network structure (e.g., Burt 1992: 32-34).

implemented as embedded and random knowledge exchanges respectively in the model. For simplicity purpose, the model attributed individual heterogeneity to pure chance (Kossinets & Watts 2006), but it can easily incorporate more complex and substantial sources of micro variations, such as intentional or strategic adaptations of individual actors, as will be discussed in Section 5.3.

Technically, studying the micro-macro mutual influences needs finding the connections between emergent macro patterns and micro interaction. Since the influences are iterative and evolving, we also need to take time into account. Agent-based modeling as a research method was created to analyze the dynamic process of emergence. This study's agent-based model implemented a stochastic network generating process usually found in network formation models (Jackson 2010; Toivonen et al. 2009). During this process, one or more predefined micro mechanisms are executed with certain probabilities at each time step. The resultant micro structural changes accumulate at the macro level to "generate" the network topology. The process lasts until the generated network topology exhibits the structural patterns of interest. According to the current study's earlier results, organizational performances can be improved by accommodating all three collaborative behaviors. The stochastic characteristic of this study's model thus suggests that synchrony is not required for such a mechanism; the three behaviors do not have to coexist all the time in the problem-solving process. Neither is behavioral consistency required. As long as there is a mix of independent workers and knowledge exchangers in the organization, organizational performances had no substantial difference whether or not there was a group of individuals who consistently chose knowledge exchange over independent knowledge creation (**Result 4** in **Table 4**). Thus, this study provides a micro-level, informal structure-based demonstration of the contextual approach to organizational ambidexterity (Gibson & Birkinshaw 2004): the macro interaction network emerging from organizational members' problem-solving activities can enable and encourage individuals to divide their time between exploitative and exploratory activities in a way that leads to an organizational balance. The result also shed some light on the timing of switching between exploitation and exploration (**Result 10** in **Table 4**), which is known to be important yet difficult to decide. More specifically, when problem complexity was low and there was a small group of frequent knowledge exchangers in the organization, the positive effect of encouraging random knowledge exchanges would be more observable in the short run than in the long run. The same effect would be more observable in the long run when problem complexity was high and all individual members were enthusiastic about exchanging knowledge. Previous studies on contextual ambidexterity are mostly about the role of formal structures and processes and have not been able to specify the underlying mechanisms or to provide concrete managerial advice (O'Reilly & Tushman 2013).

The techniques used in the current study also allowed us to study endogeneity: given stable individual propensities, how does the macro structure's impact on organizational performances change over time? It has been reported that the same form of individual social capital may have stronger, weaker, or even opposite effects over time (Baum et al. 2010; Gargiulo & Benassi 1999; Soda et al. 2004), but little is known about the endogenous forces that drive these changes. The current study investigated endogenous forces by making established network ties decay at a certain rate whenever they were not being used for knowledge exchange. It turned out that there was an inverted-U relationship between tie decay rate and organizational performance (**Result 5** in **Table 4**), indicating that the emergent macro network both facilitates and constrains organizational problem solving over time. This dual impact is consistent with the aforementioned role of "inefficient" knowledge dissemination in balancing exploitation and exploration. Over time the knowledge exchanges of organizational members improve the quantity or the quality of interpersonal social relations as knowledge channels, generating a macro interaction network increasingly efficient in knowledge dissemination. The decay of established network ties thus can slow down this process. As knowledge dissemination needs deceleration but not elimination, an intermediate rate of tie decay leads to the best organizational performances in the long run. Since the emergent macro structure was created and maintained by past knowledge exchanges, it records the history of organization learning. In this regard, the preceding result demonstrated the importance of organizational unlearning as paving the way for further organizational learning (Holan & Nelson 2004; Tsang & Zahra 2008). Given the current study's context, unlearning can be seen as the devaluation of established social relations that are efficient knowledge exchange channels but stand in the way of obtaining new knowledge or recombining existing knowledge and therefore impede future improvement in organizational performances. Again, there is a question about what to unlearn and abandoning everything that has been learned is definitely not the answer.

In practice, trying to adjust the decay rate of intra-organizational social relations may not be an effective management strategy. Leaving aside the feasibility issue, doing so will influence the macro structure's negative and positive effects indistinguishably. It would be better is to identify and prohibit just the endogenous force responsible for the negative impact. Because the same rate of tie decay led to higher organizational performances when there were more embedded knowledge exchanges (**Result 11** in **Table 4**), embedded knowledge exchanges are the micro-level origin of the emergent macro structure's negative impact on organizational performances. When most organizational members preferred embedded knowledge exchange, the macro network displayed a strong tendency towards strengthening existing ties or forming ties inside existing clusters. Repeated knowledge exchanges with existing partners and the formation of transitivity circles made the knowledge flows in the network increasingly redundant. As a result, the problem-solving process would be hindered or even stopped. The accumulation of past embedded knowledge exchanges gradually "locked" the organization into a state when knowledge exchanges no longer contributed to problem solving. Since embedded knowledge exchanges were guided by tie strength (i.e., follow existing strong ties), tie decay would logically slow down the "locked-in" process and pave the way for knowledge recombination. Thus, there is no need to worry unless most organizational members are very fond of embedded knowledge exchanges. But if they are, how should managers intervene? Promoting random knowledge exchange might help, but there is no promise since individuals have a strong preference for embedded knowledge exchange. A possible solution is to remove or deactivate existing strong ties by, for example, reorganization or rotation (Scholl 2014), so that individual members can be somewhat freed from the knowledge clusters they are currently in. Purposive removal or deactivation of social relations has also been documented as inter-organizational strategies (Davis 2010; Mariotti & Delbridge 2012). This solution challenges the traditional view that organizations should preserve informal social connections especially the fast-decaying bridges (Burt 2002)[99] by illustrating the idea of "creative destruction"[100] – new things emerge from the demise of whatever existed before.

---

[99] The fundamental proposition of social capital theory is that established social relations provide access to valuable resources such as information and knowledge. These resources are important for supporting action but costly to

5.2. Methodology-related Innovation: an Extended Network Topology

The cross-level and dynamic view of the current study was implemented through an extended network topology, which differs from the network topology used in the previous studies in several noteworthy aspects (**Table 8)**. First, the extended network was dynamic rather than static. The underlying assumption of a static network is that interpersonal knowledge exchanges rarely deviate from established network ties[101] and even if they do, there are only temporary impacts on network topology. However, contemporary organizations provide abundant opportunities[102] for their members to create new social relations and to continue exchanging knowledge with new contacts. Moreover, given today's ever changing business environment, balancing organizational exploitation and exploration has become a continuous adaptive process. A potential avenue to studying this process is the evolution of the macro interaction network, which results from organizational members' knowledge exploitation and exploring activities during problem solving.

Table 8. Contributions of the current study in modeling an extended network topology

Network topology in previous studies	Network topology in the current study
Static	Dynamic
Predefined	Emergent

---

gather. Social relations, often established for other purposes, constitute channels that reduce the amount of time and investment required to gather these resources.
[100] This term was coined by Schumpeter (1942) to denote a "process of industrial mutation that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one."
[101]In other words, once two organizational members exchange knowledge, they will always exchange knowledge with each other instead of with someone else.
[102]such as project-based organizational structure, communities of practice, and Enterprise 2.0

Exogenous	Endogenous
Random	Non-random & random
Binary	Weighted
Interactions on the network all result from pure chance	Some interactions are more frequent than others

Second, the extended network is endogenous rather than exogenous. The macro informal structure is both the cause and the effect of organization members' problem-solving activities, as indicated in **Figure 1** in *Chapter 1*. Thus, it is important to model the endogenous forces that drive and coordinate micro activities, which collectively generate and change the macro network. In the real life, the endogenous forces can be random or non-random but in previous studies they were mostly simplified as random. For example, the inter-group ties in a hybrid network can be created either randomly, when strangers bumped into each other in a random place or event (Kossinets & Watts 2006), or non-randomly, when individuals from different groups meet via common contacts[103] (Davis 1970; Holland & Leinhardt 1971, 1972; Rapoport 1953, 1957) or reach out to each other (Burt 1992; Granovetter 1973). However, in previous studies the hybrid network was defined as a randomly generated small-world network[104] (Newman 2001a; Watts 1999; Watts & Strogatz 1998) or its variant, which would have been of oversimplification regarding the context of this study. The endogenous mechanisms of the current study's model has a "locking-in" effect, as shown in the results. Thus, the simulated problem-solving process was not a Markov process: an agent's behavioral outcomes at any moment are not only affected by probability-based individual propensity but also depend on (the effective part of) a macro

[103]Structurally it turns a two-tie triad into a completely connected triad, a process known as triad closure.

[104]In a small-world network, multiple dense groups of node are connected by a few inter-group ties that maintain global connectivity of the network. Therefore, the distance between any two nodes remains short; that is, any two nodes can reach each other through a small number of intermediaries. Previous studies generated the small-world network using or modifying Watts and Strogatz's model (Watts 1999; Watts & Strogatz 1998) that creates inter-group ties by randomly rewiring the ties of a same-size regular network.

interaction network gradually shaped by long-term interactions on the network (see Footnote 12).

Third, the extended network topology was weighted, while in most of previous studies it was binary – the social interactions were treated as “presence” or “absence” without considering their strength. A binary network makes sense only when the effect of tie strength is negligible or can be substituted by other network structural features. However, in fact, tie strength plays a unique and irreplaceable role in interpersonal knowledge exchanges (Cross & Sproull 2004; Gabbay & Leenders 2002; Grabowicz et al. 2012; Hansen 1999; Kang & Kim 2010; Kossinets & Watts 2006; Levin & Cross 2004; Reagans & McEvily 2003; Wei et al. 2011) and cannot be overlooked. Up to date, there is no standard way to model tie strength, as it is still an elusive concept with many theoretical dimensions and even more operational definitions (Gilbert & Karahalios 2009). In the extended network, tie strength increased with the number of interactions on the tie and decreased at a certain rate once there was no interaction on the tie. Thereby, ties change gradually with growth and decay stages rather than abruptly appear or disappear. More recent interactions contribute more to tie strength, so that the sustainability of a tie is positively related with the duration and the freshness of the tie, as reported by multiple empirical studies (Baum et al. 2010; Burt 2000a, 2002; McEvily et al. 2012; Raeder et al. 2011). As for the effect of tie strength, empirical evidence has shown that the upper bound of knowledge exchangeable between two agents per interaction increases with tie strength (Aral & Alstyne 2011) in a decreasing rate (Reagans & McEvily 2003). It has also been observed that weak ties tend to provide access to novel knowledge (Granovetter 1973; Hansen 1999) whereas strong ties facilitate the exchange of complex knowledge (Aral & Alstyne 2011; Reagans & McEvily 2003).

The current study's model integrated all these effects[105].

Fourth, in previous studies, the heterogeneity of interactions on the network in terms of frequency and partner of knowledge exchange was usually ignored. It was assumed that each agent exchanges knowledge in the same frequency with all partners being equally available. However, in the real life, only a few individuals are active in social interactions while the majority are not (Eckmann et al. 2004; Krackhardt 1990; Moran & Ghoshal 1996; Rybski et al. 2010); an individual interacts with some of her contacts more frequently than with others (Barrat & Cattuto 2013; Gabbay & Leenders 2002; Rybski et al. 2010; Song et al. 2012; Srivastava & Gulati 2014). Thus, human interactions, initiated by a specific individual or between two specific individuals, tend to occur in bursts within short periods of time followed by long periods of inactivity. Borrowing ideas from some earlier work (Jo et al. 2011; Karsai et al. 2013), the burstiness of interactions were integrated into the model as a result of individual agents' decision-making processes[106] (Karsai et al. 2012; Krackhardt 1990; Min et al. 2009; Moran & Ghoshal 1996; Oliveira & Vazquez 2009; Stehlé et al. 2010; Wu et al. 2010; Zhao & Bianconi 2011), which is more realistic than those in previous studies.

---

[105] For parsimony purpose, the strong and weak tie effects were implemented based on the fact that tie strength and knowledge exchange outcomes are both affected by the frequency of interactions on the tie: strong ties facilitate knowledge exchange because frequent interactions produce similar knowledge bases, while weak ties bring in new knowledge because two connected agents interact only occasionally.

[106]For each individual, interacting with someone or no one is described as a Type-A or B task respectively. Each Type-A task has a priority value indicating the individual's propensity for interacting with a specific other, which modelers can tune to reflect their research purposes. At each time step, an individual first chooses between Type-A and B tasks. If it is the former, the individual will either pick a Type-A task with the highest priority or a random Type-A task regardless of its priority. This process filters high priority tasks rapidly through the priority list, while forcing long inactive periods on low priority tasks.

5.3. Limitations and Future Work

Despite aforementioned advantages, the current study has some limitations that may lead to future extensions. The major limitation is that the findings came from computer simulation rather than real organizations. Regarding the research topic of this study, two discrepancies between a real organization and the simulated organizational environment used in this study are noteworthy. For one thing, this study focused on the micro interactions for knowledge exchange purpose only, while interpersonal interactions in a real organization have various ends and they all impact the emergence and dynamics of informal organizational structure. For another, this study was concerned with a self-organizing process devoid of central control or any kind of top-down interventions, while complete self-organization of organizational members does not exist in real organizations. That being said, computer modeling and simulation are not only suitable but also crucial for the current study. On the one hand, this study attempts to synthesize multiple vaguely related theories, some of which are underdeveloped[107], was well supported by the inherent precision and flexibility of modeling and simulation[108]. On the other hand, simulated data are more precise than empirical data on the dynamics and evolution of organizational social

---

[107] Underdeveloped theories are characterized by a few weakly conceptualized constructs, vaguely understood processes, modest empirical grounding, and/or rough theoretical logic (Davis et al. 2007).

[108] For a computational model to run, every aspect of the model – inputs, outputs, assumptions, and processes in which outputs are derived from inputs under the regulation of assumptions – must be well specified. Thus, translating an underdeveloped theory into (part of) a computational model enhances the accuracy and internal validity of that theory (Davis et al. 2007; Taber & Timpone 1996). On the other hand, computational models are flexible in their ability to represent a variety of behaviors and contextual settings. Using computational models, we can construct new problems, collate isolated processes, and modify prior assumptions. More importantly, all these efforts can be formalized as testable hypotheses that incorporate existing theories with the modeler's ideas (Hummon & Fararo 1995).

networks[109]. The influential factors and testable hypotheses revealed by simulated experiments then shed light on and set up directions for future empirical studies.

To better understand model behavior and capture essential aspects of the research problem, the complexity of this study's model was intentionally kept to the minimum level, thus leaving substantial room for extension. For starter, it would be useful to incorporate other influential factors, such as personnel turnover and environmental turbulence. The current model assumed that the same group of organizational members conducted parallel problem solving. To introduce personnel turnover, we can have each individual agent leave the organization at each time step with certain probability and have vacant positions filled by new agents with randomly assigned knowledge. The current model also assumed that the organization faced a single problem, which represents a static organizational environment. To introduce environmental turbulence, we can reassign the performance score of each individual solution with certain probability every certain time steps[110]. Another potential extension is to advance the current design of agent behaviors to make them more realistic (and more complex). For example, in the current model individual propensities for whether and with whom to exchange knowledge are fixed. A more realistic

---

[109] There are multiple ways to collect longitudinal data of real organizational social networks, primary methods including survey (questionnaire or interview), observation, and experiments. When using observation, researchers often need to compromise on the number of observations (Corten & Buskens 2010), the number of time points (Lazer 2001), or the depth of observation (Moody et al. 2005). In the end there may not be enough information to make significant inference. Computational simulation allows researchers to track the variation of any key measures for any time duration (theoretically). Self-reported social-network data (i.e., data collected from surveys) may systematically deviate from the reality due to cognitive limitations, illusions, and irrationality of human beings (Bernard et al. 1984; Krackhardt 1987), especially when the social network is informal (Marsden 1990). Simulation data do not have this problem. Real network experiments more or less "manipulate" their human subjects (e.g., planting rumor), who may react by altering their behaviors in various unpredictable ways, making experimental results incomparable. In simulated experiments, however, researchers can simply erase the memory of artificial agents and run them anew. Besides primary network data sources, researchers also use secondary sources, such as key informants and proxies of interactions – email records, membership lists, and so on (Aral & Alstyne 2011). Unfortunately the lack of precision is still a problem: the former source suffers from the same data distortion as network survey. The latter may not indicate real social relations and there tends to be a large amount of noise.

[110] The period of time should be long enough to allow sufficient adaptation of the informal interaction network.

design is for agents to change their propensities based on the outcomes of their behaviors (e.g., how well their solutions are improved). Finally, by integrating general modeling frameworks such as the NK space and the network topology, the current study's model can be applied (with modifications) to other research problems that involve self-organization.

The analysis on this study's model is also limited and expandable. An immediate extension is to bring in network analysis. Since the model can easily produce detailed network data at every time step, we can visualize and analyze the extended network topology and its evolution during the problem-solving process. Previous studies measured the hybrid network topology using the characteristic statistics of a small-world structure[111] (Watts & Strogatz 1998). We can observe the temporal changes of these statistics under different experimental conditions using plots similar to **Figure 23**. We can also investigate these statistics' relationships with organizational problem-solving performance and the primary parameters of the agent-based model. Moreover, recent progresses in capturing the structural and temporal characteristics of community structures in a network (Abell et al. 2008) provide advanced analysis techniques for studying the evolution of a hybrid network topology.

5.4. Conclusion

---

[111] There are two primary statistics. One statistic is the average local clustering coefficient that measures the extent of clustering in a network – the probability that two nodes are connected increases with the number of their common neighbors. The other statistic is the average shortest path length (also known as characteristic path length or average geodesic distance) that measures the extent of separation in a network. It is the mean of shortest-path lengths over all node pairs.

In the face of highly diversified and discontinuous problems in their external environments, more and more organizations try to improve their problem-solving performances by harnessing individual members' autonomous independent or collaborative problem-solving activities, such as independent knowledge creation and knowledge exchange. These individual behaviors can generally be featured as exploitative or exploratory. An organizational-level balance of exploitation and exploration has been considered crucial for organizational performances (March 1991). In terms of solving external problems, this balance means a trade-off between (a) rapidly disseminating knowledge to improve organizational problem-solving performances in the short run and (b) enduringly maintaining knowledge diversity to guarantee organizational problem-solving performances in the long run (Lazer & Friedman 2007). Temporally good solutions obtained in the problem-solving process should be disseminated to the entire organization, yet in an "inefficient" manner, so that the exploitation of these solutions will not stifle the exploration of new ideas that may lead to better solutions in the future.

Since the rate of organization-wide knowledge dissemination is the collective outcome of interpersonal knowledge exchanges, which are fundamentally social interactions, a properly formed macro interaction structure may support "inefficient" knowledge dissemination and benefit organizational problem-solving performances. Indeed, researchers have found that a hybrid macro network (with both closure and brokerage structures) shows the preceding effect (Fang et al. 2010; Lazer & Friedman 2007). Notably, this finding was obtained by assuming that the network was predefined and static[112] during problem solving, and that organizational members followed the network precisely to exchange knowledge. These rigid assumptions are

[112] The network topology is unchanged.

appropriate only when the macro interaction structure is formal structure and the interpersonal knowledge exchanges are formally regulated. The current study focused on organizational members' autonomous problem-solving behaviors such as choosing whether and with whom to exchange knowledge. These behaviors are less affected by formal structure than by informal interaction structure that is emergent, subject to change during problem solving, and thus better represented by a dynamic network. Moreover, since the macro structure is both the medium and the outcome of micro behaviors (Giddens 1984), the network is supposed to coevolve with organizational members' problem-solving behaviors. Given the impact of such a network, how would individuals' problem-solving behaviors collectively impact organizational problem-solving performances? Under what conditions would the organization achieve the best long-run performance, that is, organizational members' autonomous problem-solving behaviors self-organize into a collective balance?

The current study tackled these questions using computer simulation and in particular agent-based modeling (ABM). ABM has been widely used to study the often unexpected collective outcomes of individual behaviors. These outcomes are difficult to predict because they usually contain additional complexity resulting from the interaction of individual behaviors, that is, they are usually more than the aggregation of individual behaviors. ABM allows to investigate the collective phenomena of interest as emergent patterns of a dynamic system that has multiple agents interacting with one another based on predefined rules in a predefined environment (Gilbert 2008). It thus provides a way to connect the collective outcomes with the attributes and/or behavioral rules of individual agents. In the current study, an organization was modeled as a dynamic system and organizational members as agents in the system. Individual agents solved the same problem in parallel and interacted with one another via a macro network to exchange solutions from time to time during the process. The organizational performance at a specific time was measured by the average score of all individual solutions at that time, which was determined using an NK model (Kauffman 1993).

The specific design of this model drew on existing theories, modeling results, and empirical evidence. The overarching hypothesis was that the longer the emergent network maintains a hybrid topology during the problem solving process, the better the organizational performance would be in the long run. Accordingly, certain micro mechanisms that would impact the genesis and maintenance of a hybrid network were identified and integrated into the model, as well as contingent factors such as problem complexity. At each time step, an individual agent has a tendency to deal with the problem by independently creating knowledge or exchanging knowledge with familiar or random others. A knowledge exchange may be self-initiated or requested by another agent and the result is influenced by how much knowledge both parties have already shared and how well they learn from the difference. Thus, individual performances were jointly affected by individual motivations and abilities (stochastically) as well as endogenous and exogenous opportunities. Exogenous opportunities were presented by chance. Endogenous opportunities were associated with individual agents' positions in the macro network, which changed over time as a result of individual agents' problem-solving activities on the network and the decay of established but unused ties. Drawing on complex adaptive system (CAS) theories, an iterative micro-macro feedback loop was implemented to support the coevolution of the macro interaction network and individual agents' problem-solving behaviors.

In the model, each micro mechanism of interest was associated with one or more model parameters, whose main and interaction effects on model outputs (i.e., temporal organizational problem-solving performance) were examined through computer simulation experiments. To handle the nonlinear relationships between model inputs and outputs, the experiments followed Latin Hypercube design and the data were analyzed using Multivariate Adaptive Regression Splines (MARS). The results not only confirmed previously identified factors to organizational problem-solving performances, such as time, problem complexity, and individual members' independent knowledge creating skills, but also revealed new factors only visible from a multi-level and dynamic perspective, such as individual members' collaborative problem solving related characteristics and the impacts of past knowledge exchange interactions. According to the major findings, an organization has the best problem-solving performances in the long run when all three individual problem-solving behaviors – self learning, embedded knowledge exchange, and random knowledge exchange – are present in proportion to one another during the problem-solving process. That said, they do not have to coexist each time or alternate but can be decided by organizational members independently. Because of the gradual "lock-in" effect of embedded knowledge exchange, sufficiently but not extremely fast tie decay contributes to organizational performances by reducing the influences of past knowledge exchanges but not jeopardizing the existence of the emergent macro structure. Organizational performances are positively related with organizational members' collaborative problem-solving abilities, but has an invert-U relationship with their independent problem-solving abilities. The number of organizational members has a diminishing return on organizational performances.

Thus, the current study extended previous findings on “inefficient” knowledge dissemination and the balancing power of a hybrid network to a dynamic and self-organized context and therefore provided a richer picture for researchers and practitioners. Theoretically, the current study adds to the literature of organizational ambidexterity, organizational social capital, and organizational social networks. Methodological, the agent-based model developed in this study is more realistic than previous models and can benefit future studies. Despite these contributions, future work is still needed to better understand and extend the model and to empirically test the findings in a real organizational environment.

APPENDICE

Appendix A: Pseudo–Code Description of the Agent-based Model

1. Main Procedure

INITIALIZATION

READ model parameter values from an external file

GET the nk space and initial individual solutions from external files

WHILE the number of agents < ***orgSize***

    Randomly SELECT an initial solution without replacement

    CALL *anOrgMember.initialize* with ***actDist*** and the selected initial condition

ENDWHILE

CREATE an empty network and randomly place each agent on a network node

EXECUTION PER TIME STEP

REPEAT

    Randomly SELECT *anOrgMember* without replacement

    IF *anOrgMember.Idle* is TRUE THEN

        DETERMINE *anOrgMember.ActType* with *anOrgMember.ActRate* and ***randLink***

        CALL *anOrgMember.step*

    ENDIF

UNTIL all agents have been called

CALCULATE and OUTPUT ***avgScore***

REPEAT

    Randomly SELECT anOrgMember without replacement

        SET *anOrgMember.Idle* to FALSE

UNTIL all agents have been processed

UPDATE the network with ***decayRate*** and ***wtGain***

2. Definition of an agent: the *OrgMember* class

VARIABLES / ATTRIBUTES

- *ActRate*: a continuous value with a power distribution on [0,1]
- *CurrentSln*: a binary vector of 20 dimensions representing problem solution
- NewSln: same as above (for knowledge exchange)
- ActType: a categorical value from {selfLearn, randomExchange, closeExchange}

- Idle: a Boolean variable indicating whether the agent is currently occupied

METHODS / BEHAVIORS

FUNCTION *initialize*(***actDist***, *anInitialSln*)

SET *this.CurrentSln* to *anInitialSln*

SET *this.NewSln* to *this.CurrentSln*

DETERMINE *this.ActType* with ***actDist***

ENDFUNCTION

FUNCTION *step*

IF *this.Idle* is FALSE THEN

RETURN

ENDIF

CASE *this.ActType* OF

“selfLearn”:

CALL *this.selfLearn*

BREAK OUT

“closeExchange”:

IF current agent has at least one neighbor in the network THEN

CALL *this.closeExchange*

BREAK OUT

ELSE

CONTINUE TO “randomExchange”

ENDIF

“randomExchange”:

CALL *this.randomExchange*

BREAK OUT

ENDCASE

ENDFUNCTION

FUNCTION *selfLearn*

FOR *i* = 1 to ***innoRange***

Randomly SELECT a solution dimension *d*

SET *this.NewSln*[*d*] to its opposite value (0 to 1 or vice versa)

ENDFOR

CALL *this.updateSolution*

ENDFUNCTION

```
FUNCTION closeExchange
    SET neighborList → GET the current agent's direct and indirect neighbors in the network
    CALCULATE the current agent's local constraint with each agent in neighborList
    SET partner → the agent with whom the current agent has the highest local constraint
    IF partner.Idle is TRUE AND partner.ActType is NOT "selfLearn" THEN
        SET scoreA → CALCULATE the score of this.CurrentSln
        SET scoreB → CALCULATE the score of partner.CurrenSln
        IF scoreA ≥ scoreB THEN
            this.exchangeKnowledge (this, partner)
        ELSE
            this.exchangeKnowledge (parnter, this)
        ENDIF
    ENDIF
ENDFUNCTION

FUNCTION randomExchange
    SET partner →GET another random agent
    IF partner.Idle is TRUE AND partner.ActType is NOT "selfLearn" THEN
        SET scoreA → CALCULATE the score of this.CurrentSln
        SET scoreB → CALCULATE the score of partner.CurrenSln
        IF scoreA ≥ scoreB THEN
            this.exchangeKnowledge (this, partner)
        ELSE
            this.exchangeKnowledge (parnter, this)
        ENDIF
    ENDIF
ENDFUNCTION

FUNCTION exchangeKnowledge (firstOrgMember, secondOrgMember)
    SET firstOrgMember.Idle to FALSE
    SET secondOrgMember.Idle to FALSE
    IF a network tie exists between firstOrgMember and secondOrgMember THEN
        INCREASE tie strength based on wtGain and decayRate
    ELSE
        CREATE a tie between firstOrgMember and secondOrgMember with wtGain
    ENDIF
```

```
    CALCULATE bandwidth based on current strength of the above tie
    DETERMINE the max number of exchangeable solution dimensions with bandwidth
    SET exchangeDimList → Randomly SELECT that number of solution dimensions
    FOR each dimension d in exchangeDimList
        IF firstOrgMember.CurrentSln[d] ≠ secondOrgMember.CurrentSln[d] THEN
            SET firstOrgMember.NewSln[d] to secondOrgMember.CurrentSln[d] with learnErr
        ENDIF
    ENDFOR
    CALL firstOrgMember.updateSolution
    FOR each dimension d in exchangeDimList
        IF firstOrgMember.CurrentSln[d] ≠ secondOrgMember.CurrentSln[d] THEN
            SET secondOrgMember.NewSln[d] to firstOrgMember.CurrentSln[d] with learnErr
        ENDIF
    ENDFOR
    CALL secondOrgMember.updateSolution
ENDFUNCTION

FUNCTION updateSolution
    SET newScore → CALCULATE the score of this.NewSln
    SET currentScore → CALCULATE the score of this.CurrentSln
    IF newScore > currentScore THEN
        SET this.CurrentSln to this.NewSln
    ENDIF
ENDFUNCTION
```

Appendix B: Latin hypercube sampling (LHS) and sample set used in this study

This study used *S* = 300 LHS samples by varying *d* = 6 model parameters. **Table A1** shows the entire LHS sample set. The sampled values of each parameter are summarized in **Table A2** and their correlation matrix is given in **Figure A2**.

Table A9. The complete LHS sample set

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
1	1.585	0.178	0.1628	0.5958	0.418	3
2	1.825	0.602	1.3712	0.5642	0.875	2
3	2.785	0.022	0.4432	0.9508	0.842	4
4	1.975	0.042	2.6955	0.7792	0.275	5
5	0.035	0.458	0.7235	0.6892	0.845	5
6	0.415	0.815	0.5495	0.7558	0.385	4
7	2.975	0.215	1.5742	0.9225	0.048	5
8	0.405	0.525	1.1875	0.5992	0.572	2
9	0.905	0.395	0.1822	0.5808	0.095	5
10	1.705	0.835	1.3905	0.6692	0.542	1
11	1.215	0.338	0.1338	0.8692	0.732	1
12	1.445	0.898	2.2798	0.6058	0.295	2
13	2.285	0.662	0.5978	0.6292	0.885	4
14	2.505	0.138	2.5698	0.7392	0.565	4
15	2.155	0.038	1.2165	0.6258	0.785	1
16	1.015	0.085	0.1532	0.5208	0.718	1
17	2.425	0.378	0.3948	0.5058	0.608	3
18	2.855	0.168	2.5312	0.7592	0.398	4
19	2.275	0.765	1.4582	0.6375	0.352	2
20	0.325	0.782	0.4045	0.9292	0.858	1
21	2.595	0.345	0.6462	0.8275	0.865	5
22	2.935	0.792	2.4635	0.8675	0.838	5
23	1.575	0.518	2.1252	0.6608	0.025	2
24	2.875	0.735	1.1972	0.8258	0.722	3
25	0.625	0.548	2.9468	0.6492	0.145	5
26	1.905	0.985	1.3132	0.5425	0.702	4
27	2.095	0.788	0.7138	0.6208	0.768	4
28	1.155	0.728	1.2745	0.6225	0.882	3
29	2.465	0.948	1.0425	0.9775	0.685	1
30	0.085	0.978	1.8062	0.5008	0.938	1
31	2.105	0.945	1.3325	0.9692	0.432	4
32	0.265	0.725	0.2305	0.8592	0.825	4
33	0.935	0.425	2.2605	0.7708	0.112	3

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
34	0.475	0.435	2.7148	0.7525	0.098	5
35	2.685	0.428	0.2015	0.6808	0.435	3
36	1.925	0.295	2.1832	0.8392	0.312	5
37	1.775	0.568	0.2208	0.9675	0.085	1
38	2.915	0.405	2.9275	0.7908	0.518	1
39	2.535	0.615	1.7385	0.8408	0.255	1
40	2.385	0.132	1.0232	0.7208	0.212	4
41	2.985	0.055	2.1155	0.6458	0.968	3
42	1.095	0.445	2.3475	0.9825	0.358	5
43	0.155	0.415	0.4238	0.7225	0.575	2
44	0.075	0.398	2.4442	0.8808	0.218	2
45	0.605	0.262	2.2122	0.5942	0.952	5
46	2.765	0.352	2.1058	0.9842	0.725	2
47	2.025	0.832	2.8018	0.6592	0.728	4
48	1.865	0.198	1.9608	0.9342	0.392	3
49	2.835	0.382	2.0575	0.8942	0.655	4
50	0.855	0.182	0.2788	0.5758	0.562	3
51	0.465	0.422	1.7675	0.8358	0.848	3
52	1.285	0.855	0.3368	0.5792	0.052	1
53	1.735	0.515	0.3658	0.9425	0.755	3
54	0.275	0.088	0.6268	0.9858	0.332	1
55	0.365	0.608	1.6225	0.6008	0.238	2
56	1.625	0.252	0.8588	0.9992	0.665	4
57	2.705	0.875	0.6945	0.8025	0.368	4
58	2.215	0.148	2.7632	0.5608	0.832	4
59	2.295	0.332	2.7922	0.5708	0.552	3
60	1.335	0.502	0.9845	0.9542	0.568	2
61	1.195	0.638	2.6278	0.9458	0.402	3
62	1.715	0.015	1.4678	0.7875	0.168	2
63	0.545	0.885	1.0715	0.5092	0.915	4
64	1.535	0.048	2.4152	0.5875	0.638	2
65	0.985	0.122	0.5302	0.5775	0.038	3
66	2.775	0.335	2.6665	0.7292	0.378	4
67	1.245	0.675	0.4335	0.9442	0.812	2
68	1.655	0.008	2.0768	0.6125	0.815	1
69	0.025	0.472	1.4195	0.8658	0.652	4
70	1.005	0.202	2.5795	0.6642	0.302	3
71	0.165	0.118	2.2025	0.9275	0.165	3
72	2.745	0.305	0.1242	0.6175	0.452	2
73	1.795	0.532	2.8405	0.5725	0.458	2
74	0.755	0.045	2.1445	0.9708	0.918	5
75	1.665	0.802	1.5452	0.9092	0.888	1

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
76	1.305	0.688	1.0908	0.5525	0.282	3
77	2.735	0.625	2.1928	0.8008	0.508	5
78	2.905	0.068	0.1048	0.5375	0.522	5
79	2.695	0.715	2.0188	0.7075	0.372	5
80	2.755	0.292	2.9952	0.8542	0.395	5
81	0.045	0.545	1.2455	0.6342	0.792	1
82	1.225	0.595	0.9362	0.5658	0.408	5
83	1.505	0.528	1.9512	0.5242	0.088	3
84	0.815	0.125	1.4098	0.6508	0.628	4
85	2.715	0.868	2.8985	0.6625	0.005	1
86	2.625	0.658	0.6172	0.6992	0.325	4
87	0.335	0.825	2.6858	0.7892	0.472	1
88	2.605	0.818	0.7042	0.5675	0.598	5
89	0.745	0.592	0.3078	0.6158	0.925	2
90	0.615	0.222	1.1682	0.6042	0.548	1
91	2.085	0.925	0.7912	0.5892	0.878	2
92	2.645	0.718	0.5205	0.8425	0.258	1
93	0.575	0.628	1.0618	0.7508	0.208	2
94	1.135	0.535	1.6032	0.5625	0.248	5
95	1.435	0.538	2.7052	0.7158	0.278	5
96	0.345	0.418	2.2702	0.7008	0.625	2
97	0.635	0.738	2.7245	0.5742	0.582	3
98	0.955	0.695	2.5602	0.5258	0.805	5
99	0.975	0.812	1.6515	0.9058	0.955	5
100	0.395	0.115	0.9748	0.6575	0.182	4
101	1.415	0.002	2.8695	0.8058	0.032	4
102	0.515	0.018	2.4248	0.5592	0.782	2
103	0.995	0.212	2.2218	0.9558	0.855	2
104	0.065	0.655	2.0865	0.8242	0.132	5
105	2.455	0.965	0.7428	0.9892	0.712	5
106	0.145	0.475	0.8202	0.9192	0.945	2
107	0.425	0.448	2.3088	0.6025	0.288	2
108	0.485	0.152	0.8685	0.8758	0.742	4
109	1.265	0.932	1.6128	0.7658	0.532	4
110	1.755	0.365	2.2508	0.6925	0.068	3
111	0.535	0.968	2.7342	0.5308	0.775	5
112	1.315	0.432	2.0382	0.5925	0.105	3
113	2.005	0.705	1.6708	0.5458	0.992	1
114	2.415	0.318	0.9168	0.6108	0.682	4
115	0.595	0.742	1.7578	0.5358	0.758	1
116	0.845	0.888	1.1585	0.7675	0.065	3
117	1.815	0.408	1.7095	0.5692	0.405	5

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
118	2.805	0.902	0.9652	0.9008	0.215	4
119	1.165	0.228	0.7622	0.6825	0.012	5
120	2.165	0.952	0.7332	0.5542	0.695	1
121	1.545	0.342	1.5355	0.9608	0.648	1
122	0.725	0.635	1.5935	0.7175	0.892	3
123	1.185	0.158	0.4528	0.6708	0.082	5
124	0.175	0.972	1.0522	0.5108	0.375	5
125	1.175	0.328	0.6075	0.8292	0.185	1
126	2.635	0.098	1.2842	0.9942	0.428	5
127	2.375	0.752	2.4732	0.7542	0.502	2
128	1.565	0.808	1.5838	0.6192	0.958	1
129	1.345	0.078	2.6762	0.8825	0.262	1
130	1.295	0.915	1.5162	0.9875	0.078	2
131	1.145	0.988	2.9855	0.8875	0.055	4
132	2.335	0.275	1.8158	0.9575	0.998	4
133	2.525	0.912	0.6365	0.6942	0.138	5
134	0.205	0.698	1.8642	0.6775	0.222	3
135	2.055	0.165	1.5645	0.6975	0.135	3
136	0.685	0.865	0.7525	0.9392	0.412	4
137	0.005	0.358	0.3272	0.9108	0.448	5
138	2.185	0.488	0.1435	0.9042	0.355	1
139	1.035	0.385	2.4345	0.8608	0.492	3
140	1.785	0.065	1.2648	0.6875	0.692	1
141	0.315	0.955	2.6568	0.9958	0.178	1
142	0.785	0.908	2.3572	0.6275	0.975	1
143	2.925	0.702	1.1005	0.7692	0.528	5
144	1.745	0.375	0.8492	0.5975	0.018	2
145	0.735	0.918	1.8835	0.7342	0.008	1
146	2.485	0.402	1.4002	0.8925	0.808	3
147	2.015	0.682	1.4775	0.7958	0.612	1
148	2.895	0.778	1.6322	0.8092	0.868	4
149	0.795	0.218	1.7965	0.7425	0.765	4
150	0.245	0.452	2.9565	0.8792	0.545	3
151	1.835	0.562	1.8932	0.5175	0.252	5
152	1.845	0.092	2.4055	0.6908	0.525	2
153	0.945	0.012	1.3035	0.8575	0.142	2
154	1.045	0.225	2.8502	0.7042	0.912	3
155	2.325	0.482	0.6752	0.5042	0.382	5
156	0.185	0.355	1.1392	0.7275	0.045	4
157	2.815	0.748	1.2552	0.6858	0.798	1
158	0.375	0.192	1.7868	0.6842	0.162	2
159	0.115	0.315	0.4818	0.5842	0.748	5

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
160	1.765	0.285	1.9318	0.6092	0.175	3
161	0.285	0.755	0.9555	0.7842	0.942	3
162	0.305	0.882	0.8395	0.7108	0.198	3
163	1.605	0.258	1.1778	0.6142	0.485	3
164	2.135	0.575	2.6472	0.8342	0.988	4
165	2.725	0.998	1.5258	0.8108	0.272	1
166	2.345	0.145	2.7728	0.9142	0.298	4
167	1.875	0.845	2.3378	0.8142	0.995	3
168	2.245	0.062	2.9662	0.7442	0.228	2
169	2.065	0.272	0.5592	0.8725	0.795	2
170	1.635	0.322	1.1102	0.7058	0.555	2
171	2.145	0.245	2.2315	0.7192	0.698	2
172	0.655	0.598	0.2692	0.9808	0.595	5
173	1.885	0.142	2.8115	0.6725	0.195	4
174	1.105	0.412	2.0962	0.7408	0.362	2
175	2.575	0.268	0.8878	0.7375	0.148	1
176	2.615	0.828	2.0092	0.7325	0.778	5
177	1.555	0.288	0.3175	0.6325	0.125	3
178	1.955	0.618	1.0812	0.5408	0.365	4
179	0.665	0.878	1.6998	0.9642	0.488	1
180	1.615	0.282	1.1295	0.8908	0.622	3
181	0.525	0.578	2.8888	0.8708	0.735	1
182	2.495	0.872	2.7535	0.5158	0.422	4
183	0.455	0.442	1.2262	0.7242	0.498	4
184	2.655	0.645	2.0478	0.9125	0.118	2
185	1.205	0.498	1.6805	0.7775	0.445	3
186	0.645	0.905	2.5215	0.7092	0.232	2
187	1.805	0.858	0.8298	0.5392	0.738	1
188	1.595	0.938	2.6375	0.9758	0.982	4
189	2.885	0.235	1.6902	0.6442	0.172	5
190	0.925	0.058	2.3862	0.9175	0.705	2
191	2.965	0.892	2.9178	0.7758	0.002	2
192	0.355	0.588	2.5988	0.8175	0.895	4
193	2.585	0.958	2.4925	0.5125	0.108	4
194	0.135	0.392	0.5688	0.9658	0.658	5
195	2.075	0.072	1.4485	0.6675	0.592	5
196	2.675	0.082	2.8598	0.7125	0.902	1
197	2.435	0.648	0.2595	0.7942	0.538	1
198	1.425	0.775	0.3755	0.8525	0.328	4
199	0.055	0.678	2.6182	0.7808	0.965	3
200	1.675	0.188	2.1348	0.5142	0.662	4
201	1.465	0.995	0.7718	0.5075	0.242	3

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
202	0.565	0.685	2.1542	0.8775	0.822	3
203	0.295	0.848	0.9458	0.9358	0.102	5
204	1.995	0.632	0.2498	0.6075	0.668	4
205	1.355	0.665	2.5022	0.9475	0.772	4
206	0.385	0.468	2.3668	0.7742	0.338	2
207	1.125	0.362	2.4538	0.5558	0.802	2
208	1.115	0.302	2.2895	0.5292	0.632	3
209	2.315	0.112	1.1198	0.9792	0.152	2
210	2.845	0.478	1.7772	0.8892	0.908	1
211	1.645	0.232	0.9072	0.5275	0.058	5
212	1.985	0.672	0.4142	0.7625	0.115	4
213	1.945	0.298	0.5012	0.6658	0.335	1
214	2.795	0.758	0.3562	0.7642	0.932	3
215	1.375	0.348	1.4968	0.5025	0.922	1
216	0.125	0.485	2.7825	0.5225	0.898	5
217	2.565	0.585	0.2402	0.8992	0.455	4
218	1.965	0.668	0.8105	0.6392	0.465	5
219	1.365	0.768	2.8308	0.6558	0.495	4
220	2.555	0.558	1.8545	0.5342	0.305	3
221	2.865	0.522	1.2068	0.9975	0.935	2
222	0.775	0.942	2.3765	0.9625	0.515	4
223	2.445	0.462	0.9942	0.9908	0.415	4
224	2.235	0.238	1.9028	0.6542	0.075	1
225	0.865	0.935	0.6558	0.6475	0.292	3
226	1.255	0.388	1.8255	0.7308	0.308	2
227	1.455	0.722	0.6655	0.7975	0.615	2
228	1.325	0.692	0.1918	0.8375	0.862	2
229	1.475	0.032	1.8738	0.8158	0.588	4
230	0.435	0.572	0.4625	0.5492	0.558	5
231	1.055	0.555	1.3808	0.5575	0.752	2
232	0.915	0.508	2.8212	0.6742	0.342	1
233	0.675	0.795	1.9705	0.5858	0.708	1
234	2.665	0.928	2.8792	0.8442	0.852	5
235	1.085	0.308	1.9898	0.5442	0.972	4
236	0.255	0.208	1.8352	0.9025	0.225	2
237	2.195	0.265	0.5882	0.8842	0.318	1
238	0.555	0.005	0.6848	0.9742	0.642	3
239	0.875	0.732	1.4872	0.7925	0.188	2
240	0.445	0.095	1.9802	0.9242	0.245	2
241	0.495	0.135	1.2938	0.7825	0.235	3
242	2.255	0.822	1.3615	0.9075	0.905	3
243	1.495	0.785	1.4388	0.6792	0.438	5

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
244	0.805	0.465	0.7815	0.9375	0.388	4
245	0.715	0.582	1.3422	0.7142	0.345	1
246	1.895	0.195	2.3282	0.6408	0.482	2
247	1.395	0.762	0.8782	0.8625	0.872	1
248	2.475	0.035	0.8975	0.8458	0.602	3
249	0.965	0.622	1.3228	0.6958	0.062	1
250	2.175	0.495	1.6418	0.9408	0.835	5
251	1.915	0.798	0.3852	0.7575	0.122	2
252	2.825	0.922	2.1638	0.9308	0.035	3
253	1.275	0.155	1.4292	0.8958	0.962	5
254	1.855	0.185	0.5398	0.7475	0.675	2
255	1.235	0.325	1.8448	0.8642	0.605	1
256	0.195	0.438	1.0038	0.7492	0.128	2
257	2.265	0.105	1.0328	0.7258	0.578	5
258	1.515	0.605	1.9995	0.9492	0.072	3
259	0.765	0.372	2.9082	0.8858	0.468	1
260	1.025	0.895	1.1488	0.8325	0.688	3
261	0.095	0.992	2.4828	0.6758	0.462	5
262	0.835	0.028	2.5408	0.5825	0.788	2
263	2.405	0.712	2.0285	0.7025	0.015	4
264	0.885	0.492	1.7192	0.9208	0.268	4
265	0.215	0.708	1.7288	0.8558	0.678	4
266	2.955	0.838	2.5505	0.5508	0.022	1
267	2.125	0.242	1.2358	0.9325	0.985	3
268	1.685	0.075	2.7438	0.6525	0.672	2
269	1.725	0.455	0.3465	0.8975	0.155	4
270	0.895	0.642	1.5548	0.8225	0.205	2
271	2.945	0.745	0.9265	0.5908	0.475	3
272	2.995	0.512	1.6612	0.8508	0.645	4
273	0.015	0.278	2.0672	0.9725	0.042	2
274	0.825	0.975	2.1735	0.9925	0.158	5
275	1.385	0.962	2.5118	0.8042	0.512	5
276	2.035	0.842	2.2992	0.6358	0.478	3
277	0.585	0.128	1.7482	0.8492	0.818	3
278	2.355	0.542	1.9125	0.9258	0.192	3
279	2.395	0.102	0.5108	0.7608	0.978	2
280	2.225	0.025	1.9222	0.7992	0.762	1
281	2.515	0.862	2.2412	0.7858	0.715	3
282	1.525	0.505	1.9415	0.9592	0.322	5
283	2.205	0.172	0.4915	0.8075	0.028	4
284	0.695	0.852	0.4722	0.8192	0.285	1
285	1.075	0.052	0.8008	0.8475	0.425	2

designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
286	1.485	0.552	0.2885	0.6242	0.442	3
287	2.115	0.175	0.5785	0.5192	0.202	1
288	0.705	0.368	1.5065	0.8125	0.618	4
289	1.405	0.612	1.3518	0.6425	0.635	3
290	1.695	0.772	2.3185	0.8308	0.092	2
291	2.045	0.982	0.2982	0.8208	0.315	1
292	2.545	0.312	2.6085	0.5475	0.585	1
293	0.235	0.162	0.1145	0.5325	0.828	5
294	0.505	0.565	2.5892	0.8742	0.348	1
295	0.225	0.652	0.2112	0.9525	0.745	3
296	0.105	0.108	0.1725	0.7358	0.265	5
297	1.935	0.205	2.9372	0.7725	0.505	5
298	2.365	0.255	1.0135	0.7458	0.535	5
299	2.305	0.805	2.9758	0.9158	0.948	5
300	1.065	0.248	2.3958	0.6308	0.928	1

Table A10. Summary of model parameters varied in the LHS sample set

Variable	actDist	randLink	wtGain	decayRate	learnErr	innoRange
Minimum	0.005	0.001667	0.1048	0.5008	0.002	1
1st Quartile	0.7525	0.250833	0.8274	0.6254	0.250833	2
Median	1.5	0.5	1.55	0.75	0.5	3
Mean	1.5	0.5	1.55	0.75	0.5	3
3rd Quartile	2.2475	0.749167	2.2726	0.8746	0.749167	4
Maximum	2.995	0.998333	2.9952	0.9992	0.998333	5

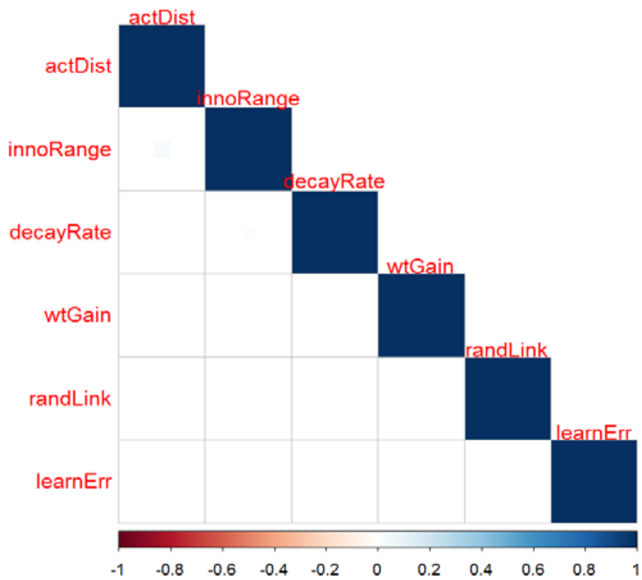


Figure A29. The correlation matrix of model parameters varied in the LHS sample

Appendix C: Design Matrices for two crossed design

Note that in the first crossed design, Group 2 sub-design is crossed with Group 1 LHS design differently when *actType* = A than when *actType* = B/C/D; they are presented in the following two separate tables.

Table A11. A partial crossed design when actType = A

Crossed design point	Group 1 (LHS sample)							Group 2			
	LHS design point	*actDist*	*randLink*	*wtGain*	*decayRate*	*learnErr*	*innoRange*	Design point	*orgSize*	*space_k*	*actType*
1	10	1.705	0.835	1.3905	0.669167	0.54167	1	1	50	1	A
2	...	...	...	...	...	...	1	2	100	1	A
3	...	...	...	...	...	...	1	3	200	1	A
4	...	...	...	...	...	...	1	4	50	5	A
5	...	...	...	...	...	...	1	5	100	5	A
6	...	...	...	...	...	...	1	6	200	5	A
7	11	1.215	0.338333	0.133833	0.869167	0.731667	1	1	50	1	A
8	...	...	...	...	...	...	1	2	100	1	A
9	...	...	...	...	...	...	1	3	200	1	A
10	...	...	...	...	...	...	1	4	50	5	A
11	...	...	...	...	...	...	1	5	100	5	A
12	...	...	...	...	...	...	1	6	200	5	A
...	...	...	...	...	...	...	...	...	...	...	...
355	300	1.065	0.248333	2.395833	0.630833	0.928333	1	1	50	1	A
356	...	...	...	...	...	...	1	2	100	1	A
357	...	...	...	...	...	...	1	3	200	1	A
358	...	...	...	...	...	...	1	4	50	5	A
359	...	...	...	...	...	...	1	5	100	5	A
360	...	...	...	...	...	...	1	6	200	5	A
361	2	1.825	0.601667	1.371167	0.564167	0.875	2	1	50	1	A
362	...	...	...	...	...	...	2	2	100	1	A
363	...	...	...	...	...	...	2	3	200	1	A
364	...	...	...	...	...	...	2	4	50	5	A
365	...	...	...	...	...	...	2	5	100	5	A
366	...	...	...	...	...	...	2	6	200	5	A
...	...	...	...	...	...	...	...	...	...	...	...
1440	4	1.975	0.041667	2.6955	0.779167	0.275	5	1	50	1	A
1441	...	...	...	...	...	...	5	2	100	1	A
1442	...	...	...	...	...	...	5	3	200	1	A
1443	...	...	...	...	...	...	5	4	50	5	A
1444	...	...	...	...	...	...	5	5	100	5	A
1445	...	...	...	...	...	...	5	6	200	5	A
...	...	...	...	...	...	...	...	...	...	...	...
1795	299	2.305	0.805	2.975833	0.915833	0.948333	5	1	50	1	A
1796	...	...	...	...	...	...	5	2	100	1	A
1797	...	...	...	...	...	...	5	3	200	1	A
1798	...	...	...	...	...	...	5	4	50	5	A
1799	...	...	...	...	...	...	5	5	100	5	A
1800	...	...	...	...	...	...	5	6	200	5	A

Table A12. A partial crossed design when actType = B, C, and D

Crossed design point	Group 1 (LHS sample)							Group 2			
	LHS design point	*actDist*	*randLink*	*wtGain*	*decayRate*	*learnErr*	*innoRange*	Design point	*orgSize*	*space_k*	*actType*
1	1	1.585	0.178333	0.162833	0.595833	0.418333	3	1	50	1	B
2	1	...	...	...	...	...	...	2	100	1	B
3	1	...	...	...	...	...	...	3	200	1	B
4	1	...	...	...	...	...	...	4	50	5	B
5	1	...	...	...	...	...	...	5	100	5	B
6	1	...	...	...	...	...	...	6	200	5	B
7	1	...	...	...	...	...	...	7	50	1	C
8	1	...	...	...	...	...	...	8	100	1	C
9	1	...	...	...	...	...	...	9	200	1	C
10	1	...	...	...	...	...	...	10	50	5	C
11	1	...	...	...	...	...	...	11	100	5	C
12	1	...	...	...	...	...	...	12	200	5	C
13	1	...	...	...	...	...	...	13	50	1	D
14	1	...	...	...	...	...	...	14	100	1	D
15	1	...	...	...	...	...	...	15	200	1	D
16	1	...	...	...	...	...	...	16	50	5	D
17	1	...	...	...	...	...	...	17	100	5	D
18	1	...	...	...	...	...	...	18	200	5	D
19	2	1.825	0.601667	1.371167	0.564167	0.875	2	1	50	1	B
20	2	...	...	...	...	...	...	2	100	1	B
21	2	...	...	...	...	...	...	3	200	1	B
22	2	...	...	...	...	...	...	4	50	5	B
23	2	...	...	...	...	...	...	5	100	5	B
24	2	...	...	...	...	...	...	6	200	5	B
...	...	...	...	...	...	...	...	...	...	...	...
5395	300	1.065	0.248333	2.395833	0.630833	0.928333	1	1	50	1	D
5396	300	...	...	...	...	...	...	2	100	1	D
5397	300	...	...	...	...	...	...	3	200	1	D
5398	300	...	...	...	...	...	...	4	50	5	D
5399	300	...	...	...	...	...	...	5	100	5	D
5400	300	...	...	...	...	...	...	6	200	5	D

Table A13. A crossed design incorporating different initial conditions

Crossed design point	Group 1 (LHS sample)							Group 2		
	LHS design point	*actDist*	*randLink*	*wtGain*	*decayRate*	*learnErr*	*innoRange*	Design point	*nkSpace*	*space_k*
1	1	1.585	0.178333	0.162833	0.595833	0.418333	3	1	1	1
2	1	...	...	...	...	...	...	2	2	1
3	1	...	...	...	...	...	...	3	3	1
4	1	...	...	...	...	...	...	4	4	1
5	1	...	...	...	...	...	...	5	5	1
6	1	...	...	...	...	...	...	6	1	5
7	1	...	...	...	...	...	...	7	2	5
8	1	...	...	...	...	...	...	8	3	5
9	1	...	...	...	...	...	...	9	4	5
10	1	...	...	...	...	...	...	10	5	5
11	2							1	1	1
12	2	...	...	...	...	...	...	2	2	1
13	2	...	...	...	...	...	...	3	3	1
14	2	...	...	...	...	...	...	4	4	1
15	2	...	...	...	...	...	...	5	5	1
...	...	...	...	...	...	...	...	...	...	...
2991	300	1.065	0.248333	2.395833	0.630833	0.928333	1	1	1	1
2992	300	...	...	...	...	...	...	2	2	1
2993	300	...	...	...	...	...	...	3	3	1
2994	300	...	...	...	...	...	...	4	4	1
2995	300	...	...	...	...	...	...	5	5	1
2996	300	...	...	...	...	...	...	6	1	5
2997	300	...	...	...	...	...	...	7	2	5
2998	300	...	...	...	...	...	...	8	3	5
2999	300	...	...	...	...	...	...	9	4	5
3000	300	...	...	...	...	...	...	10	5	5

Appendix D: A procedure for determining the number of replicate runs

Provided the same model inputs, stochastic model still produces different outputs per simulation run. To obtain representative model outputs, a common approach is to perform replicate runs and identify central tendency from the results (e.g., mean or median). But it is not always clear how many replicate runs are necessary for this purpose. In addition, while the number of replications should be large enough to ensure desired accuracy, we cannot afford too many replications due to often limited time and computing resources. This study drew on a special procedure for determining the number of replicate runs for simulation experiments (Read et al. 2012). Basically, the procedure quantifies the extent to which the statistical consistency of model-output distributions changes with the number of samples comprising the distributions. Since each sample is obtained from a replicate run, the result sheds light on the appropriate number of replicate runs.

The procedure unfolds as follows, illustrated by its application in the current study. As we know, more replicate runs (larger sample size) mitigate the effects of inherent stochasticity on the outputs of stochastic models and produce increasingly identical distributions of model outputs. Thus, the procedure starts by picking a sequence of increasing sample sizes, one of which will become the final choice. We selected 9 candidates – from 50 to 450 (inclusive) with an interval of 50. Then multiple sets of simulation results are collected for each possible sample size[113]; each set contains that number of results. As recommended by the original article (Read et al. 2012), we collected 20 sets of results for each sample size: 20 sets each containing the results of 50 replicate runs, another 20 sets each containing the results of 100 replicate runs, right through to 20 sets each containing the results of 450 replicate runs. Then each sample size is analyzed by (i) generating the median of the target model output for each of the 20 sets, and (ii) comparing the medians obtained from Set 2 to Set 20 with the median obtained from Set 1 using the Vargha-Delaney A-Test (Vargha & Delaney 2000), a non-parametric effect size test indicating the statistical consistency of two distributions. Finally, a plot is created summarizing the A-Test scores of all

[113] The same set of model inputs is used for all simulation runs. It is assumed that changing these values has no impact on the effects of stochastic elements.

sample sizes. It helps researchers identify the smallest sample size that produces an acceptably low level of variations (i.e., aleatory uncertainty) in simulation results[114]. The number of replicate runs should be no smaller than that sample size.

The core of Vargha-Delaney A-Test is a measure of stochastic superiority (denoted by *A*). The stochastic superiority of population 1 over population 2 is defined as

$$A_{12} = P(X_1 > X_2) + 0.5P(X_1 > X_2)$$

where *X* is a variable of at least ordinal scale defined for population 1 and 2 (denoted by X_1 and X_2 respectively). A_{12} represents the probability that a randomly chosen sample taken from population 1 is larger than a randomly chosen sample from population 2. $A_{12} = 1 - A_{21}$, both in the range of [0, 1]. If *X* has the same distribution in the two populations, $A_{12} = A_{21} = 0.5$. The two populations are said to be stochastically equal to each other with respect to *X*. A value of *A* above 0.71 or below 0.29 (i.e., 1 – 0.71) indicates a large difference between two populations. Since *A* measures the magnitude of effect rather than statistical significance, the detected difference is conceptually significant. **Table A4** lists some *A*-measure values and corresponding effect size.

Table A14. Effect size indicated by the *A* measure (Vargha & Delaney 2000)

Effect size	Large	Medium	Small	None	Small	Medium	Large
A measure	0.29	0.36	0.44	0.50	0.56	0.64	0.71

The Vargha-Delaney A-Test has several features that make it desirable for this study. First of all, it is an effect size test, which tends to provide more meaningful results than a statistical significance test in the context of computer simulation experiments[115]. Secondly, it is agnostic to the underlying distribution and the variance heterogeneity of the test data, so no strict assumptions on the experimental results (e.g.,

---

[114]The level of variation is measured by the maximal or median A-Test score between the 1st set of results and the remaining 19 sets.
[115]Computer simulation experiments often generate hundreds and thousands of samples; with so many samples, even a conceptually meaningless difference can turn out to be statistically significant.

normal and independent distribution) are needed. Lastly, the *A* measure is easy to interpret, efficient to compute, and applicable for both discrete and continuous values.

Following the procedure, we got two plots (**Figure A2** and **Figure A3**) that summarize the maximal and median A-Test scores for all potential sample sizes (50 to 450 with a constant interval of 50) at multiple time steps (Tick 2, 10, 50, 100, 150, and 200) based on model output *avgScore*. In both plots, the horizontal dash lines represent the boundaries for levels of differences. They are drawn at the values of 0.56 (small difference), 0.64 (medium difference), and 0.71 (large difference). As shown **Figure A2** and **Figure A3**, a sample size of 300 (i.e., 300 replicate runs) is required to reduce the amount of aleatory uncertainty to an effect size less than “small” for model outputs at various time steps.

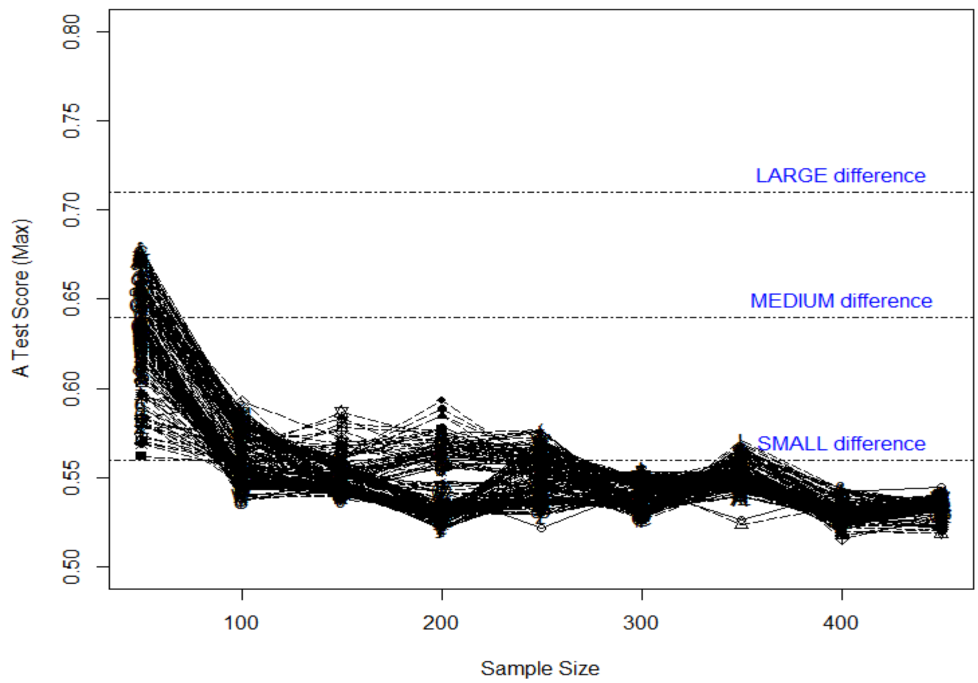


Figure A30. Maximal A-Test Scores for each Sample Size at Multiple Time Steps

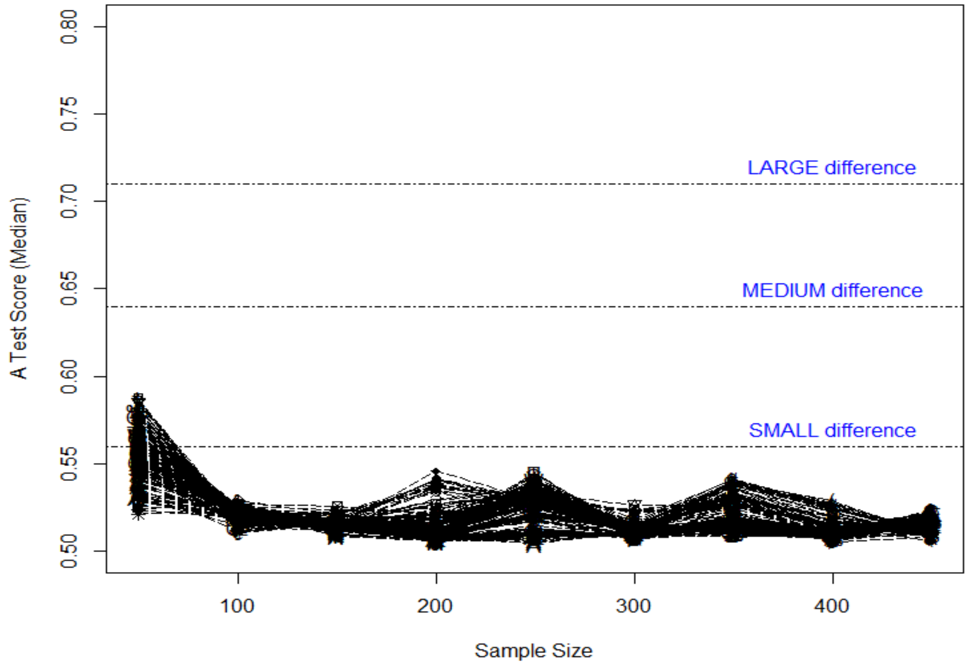


Figure A31. Median A-Test Scores for each Sample Size at Multiple Time Steps

Appendix E. Descriptive and inferential analysis for discrete model parameters

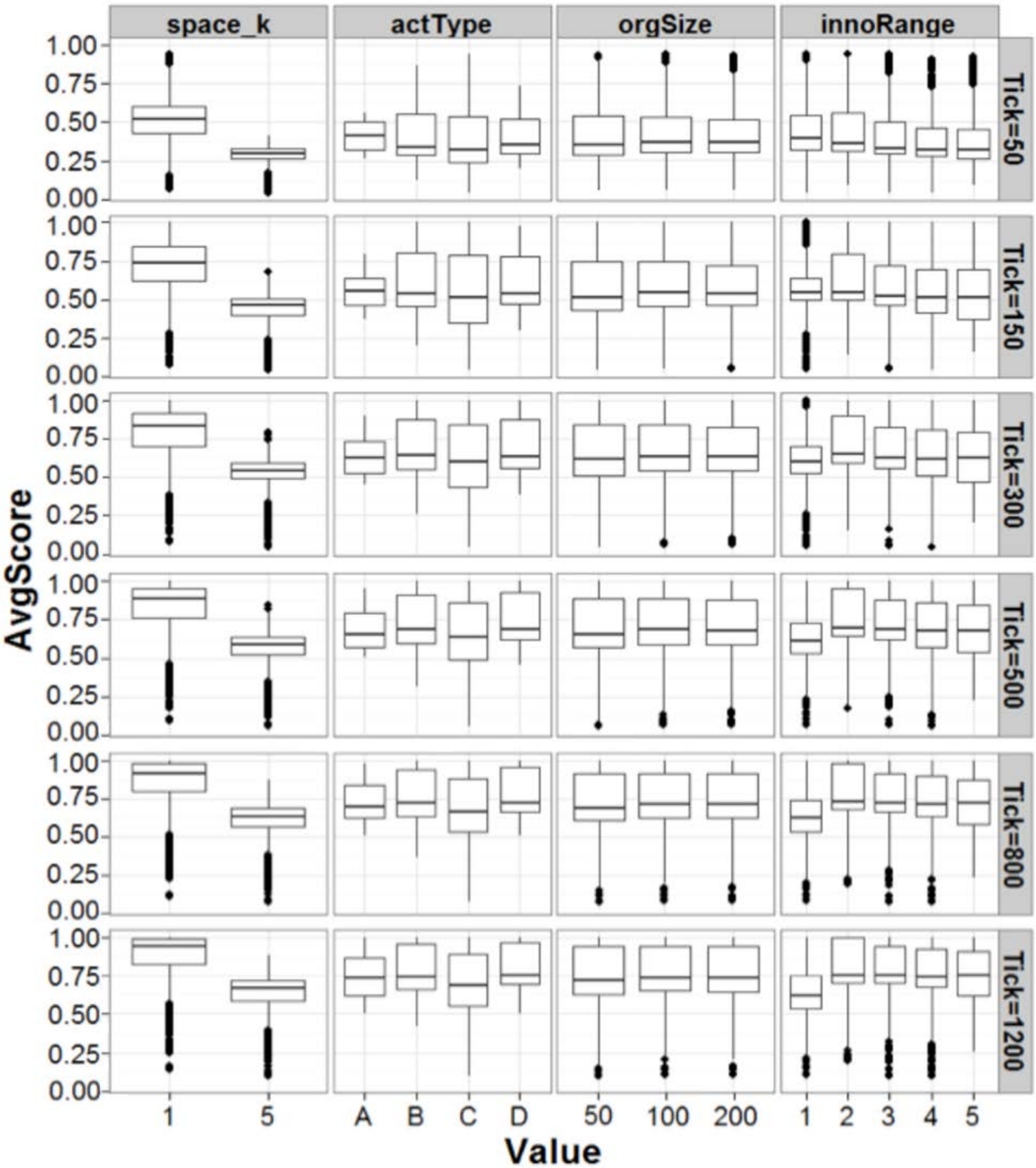


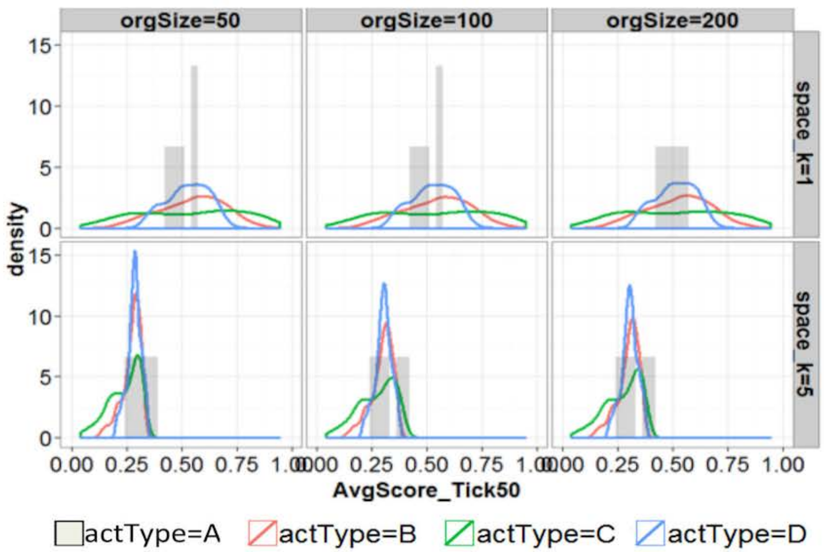
Figure A32: Relations between discrete model parameters and model outcomes

Table A15. Results of inferential analysis for discrete model parameters[116]

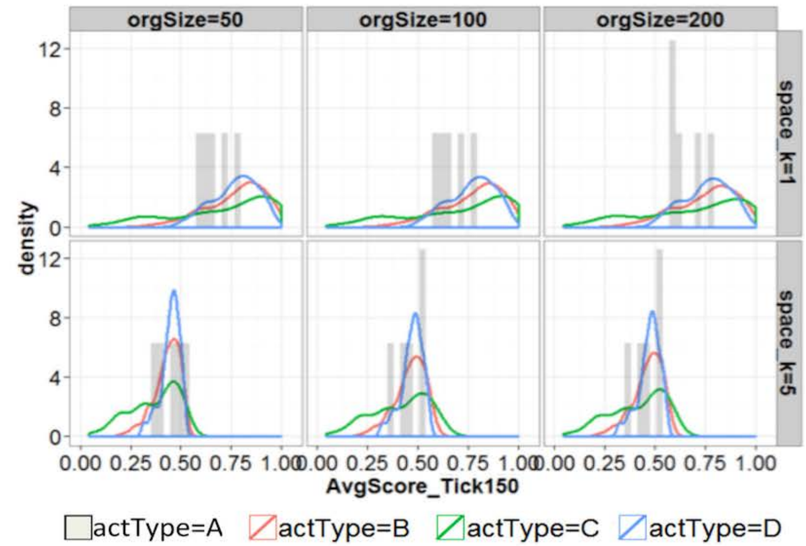
Tick	space_k					actType					orgSize					innoRange				
	KW (H)	MW				KW (H)	MW				KW (H)	MW				KW (H)	MW			
		v1	v2	p	|r|		v1	v2	p	|r|		v1	v2	p	|r|		v1	v2	p	|r|
50	2755.5*	1	5	*	**0.7***	127.6*	A	B	*	0.06*	16.4*	50	100	0.002	0.06*	193.0*	1	2	1	
							A	C	*	**0.19***		50	200	0.02	0.05(0.002)		1	3	*	0.12*
							A	D	0.12	0.04(0.019)		100	200	1			1	4	*	**0.15***
							B	C	*	0.11*							1	5	*	**0.18***
							B	D	1								2	3	*	0.12*
							C	D	*	0.13*							2	4	*	**0.17***
																	2	5	*	**0.19***
																	3	4	*	0.09*
																	3	5	*	0.11*
																	4	5	0.04	0.05(0.004)
150	2859.0*	1	5	*	**0.7***	108.1*	A	B	*	0.08*	15.4*	50	100	*	0.06*	149.6*	1	2	*	0.10*
							A	C	*	0.08*		50	200	0.01	0.05(0.002)		1	3	1	
							A	D	*	0.11*		100	200	1			1	4	0.002	0.07*
							B	C	*	0.13*							1	5	*	0.12*
							B	D	1								2	3	*	0.13*
							C	D	*	0.14*							2	4	*	**0.17***
																	2	5	*	**0.19***
																	3	4	*	0.08*
																	3	5	*	0.10*
																	4	5	0.06	0.05(0.006)
300	2969.2*	1	5	*	**0.7***	183.7*	A	B	*	0.13*	16.0*	50	100	*	0.006*	241.9*	1	2	*	**0.26***
							A	C	0.01	0.05(0.002)		50	200	0.007	0.05(0.002)		1	3	*	0.15*
							A	D	*	**0.18***		100	200	1			1	4	0.01	0.06*
							B	C	*	**0.15***							1	5	1	
							B	D	0.27								2	3	*	0.14*
							C	D	*	**0.17***							2	4	*	**0.18***
																	2	5	*	**0.20***
																	3	4	*	0.08*
																	3	5	*	0.11*
																	4	5	0.04	0.05(0.004)
500	2986.3*	1	5	*	**0.7***	246.7*	A	B	*	0.15*	14.6*	50	100	0.001	0.06*	454.6*	1	2	*	**0.36***
							A	C	0.001	0.06*		50	200	0.008	0.05(0.002)		1	3	*	**0.28***
							A	D	*	**0.20***		100	200	1			1	4	*	**0.19***
							B	C	*	**0.17***							1	5	*	0.13*
							B	D	0.01	0.05(0.002)							2	3	*	**0.15***
							C	D	*	**0.20***							2	4	*	**0.19***
																	2	5	*	**0.21***
																	3	4	*	0.09*
																	3	5	*	0.13*
																	4	5	0.02	0.06(0.002)
800	2954.7*	1	5	*	**0.7***	293.7*	A	B	*	0.14*	13.4	50	100	0.002	0.06*	678.4*	1	2	*	**0.42***
							A	C	*	0.09*	(0.001)	50	200	0.019	0.05(0.004)		1	3	*	**0.36***
							A	D	*	**0.20***		100	200	1			1	4	*	**0.30***
							B	C	*	**0.19***							1	5	*	**0.24***
							B	D	*	0.07*							2	3	*	0.14*
							C	D	*	**0.24***							2	4	*	**0.19***
																	2	5	*	**0.22***
																	3	4	*	0.09
																	3	5	*	0.14*
																	4	5	0.004	0.07*
1200	2916.1*	1	5	*	**0.7***	328.6*	A	B	*	0.12*	13.1	50	100	0.002	0.05*	821.2*	1	2	*	**0.46***
							A	C	*	0.12*	(0.001)	50	200	0.02	0.05(0.006)		1	3	*	**0.40***
							A	D	*	**0.18***		100	200	1			1	4	*	**0.35***
							B	C	*	**0.21***							1	5	*	**0.30***
							B	D	*	0.08*							2	3	*	0.11*
							C	D	*	**0.26***							2	4	*	**0.19***
																	2	5	*	**0.22***
																	3	4	*	0.10*
																	3	5	*	0.14*
																	4	5	0.001	0.07*

[116] *H* is the statistic of the KW test; v1 and v2 indicate a variable's two value levels that are being compared; |*r*| is the absolute value of effect size; p is the p value; the symbol * indicates a p value < 0.001, while larger p values are directly provided (sometimes in parentheses).

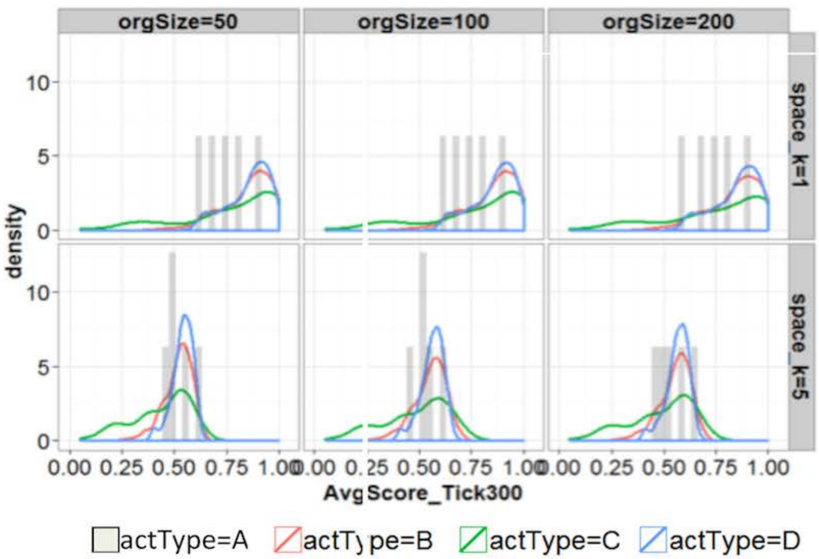
Appendix F: The probability density functions of *avgScore* over time



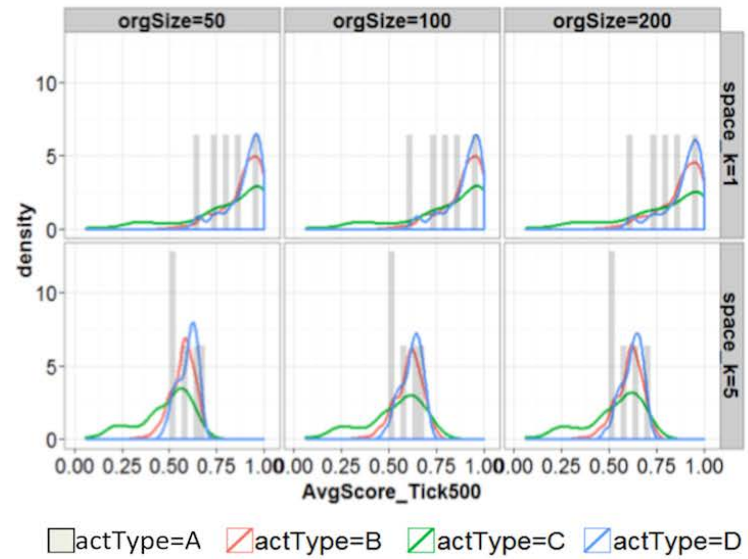
A. The probability density functions of avgScore (tick = 50)



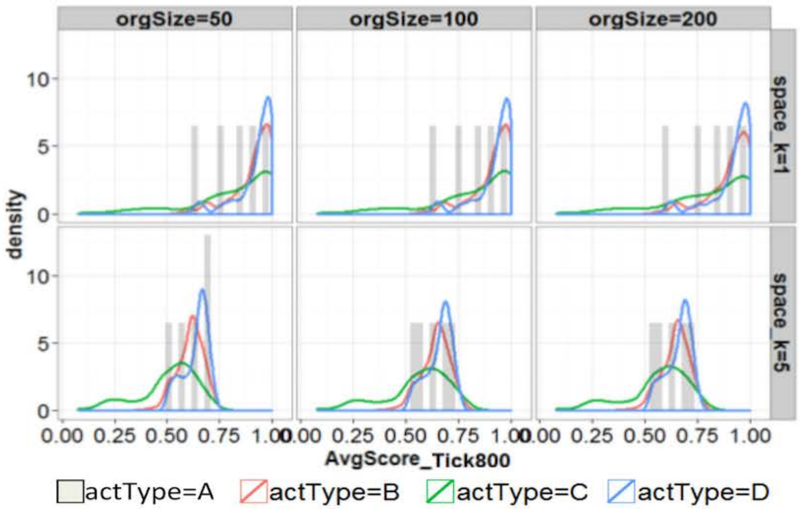
B. The probability density functions of avgScore (tick = 150)



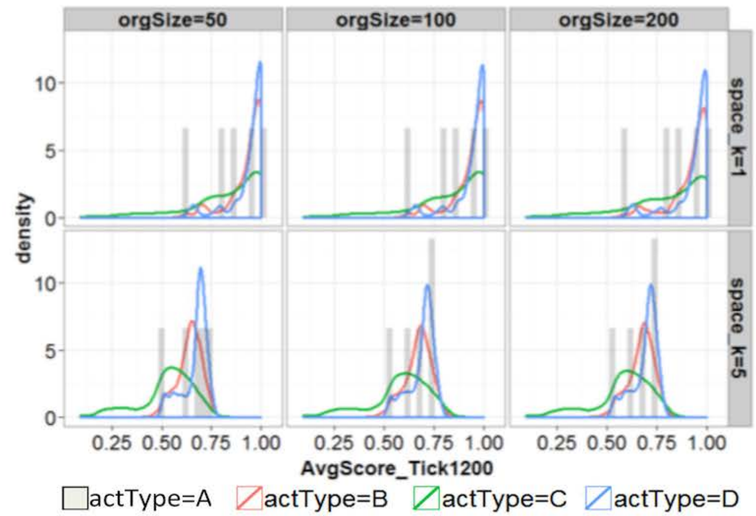
C. The probability density functions of avgScore (tick = 300)



D. The probability density functions of avgScore (tick = 500)



E. The probability density functions of avgScore (tick = 800)



F. The probability density functions of avgScore (tick = 1200)

Appendix G. MARS regression models on different subsets of the experimental results

Table A16. MARS model II and the corresponding experimental conditions

Experimental condition			
tick		100, 150, 200, 250, 300	
learnErr		<=0.15	
space_k		5	
actType		C	
Number of observations		4,500	
Number of independent variables		5	
Regression model			
Component	***Description***	***Coefficient***	***P value***
Intercept		0.483	< 0.0001
BF1	max (0, Tick – 150)	0.001	< 0.0001
BF2	max (0, 150 – Tick)	–0.002	< 0.0001
BF3	orgSize = 100	0.076	< 0.0001
BF4	orgSize = 200	0.066	< 0.0001
BF5	max (0, randLink – 0.278)	0.398	< 0.0001
BF6	max (0, 0.278 – randLink)	–1.354	< 0.0001
BF7	max (0, randLink – 0.438)	–0.416	< 0.0001
BF8	max (0, decayRate – 0.858)	–1.805	< 0.0001
BF9	max (0, 0.858 – decayRate)	–0.224	< 0.0001
BF10	BF1 * BF6	–0.001	< 0.0001
BF11	BF2 * BF6	0.005	< 0.0001
BF12	BF6 * max (0, wtGain – 0.907)	–0.037	0.0020
BF13	BF6 * max (0, 0.907 – wtGain)	0.669	< 0.0001
BF14	BF7 * BF8	4.173	< 0.0001
BF15	BF8 * max (0, randLink – 0.922)	–5.675	0.0070
BF16	BF8 * max (0, 0.922 – randLink)	4.496	< 0.0001
Goodness–of–fit			
GCV (Generalized Cross Validation)		0.000244	
GCV–Squared		0.985	
RSS (Residual Sum of Squares)		0.145	
R–Squared		0.987	

Table A17. Descriptive analysis of the underlying subset of MARS-II and MARS-III[117]

Variable	randLink	wtGain	decayRate	learnErr
Minimum	0.001667	0.1048	0.5008	0.001667
1st Quartile	0.230833	0.8274	0.6254	0.250833
Median	0.500000	1.5500	0.7500	0.500000
Mean	0.500000	1.5500	0.7500	0.500000
3rd Quartile	0.749167	2.2726	0.8746	0.749167
Maximum	0.998333	2.9952	0.9992	0.998333

[117]The descriptive analysis results of the two result subsets were identical

Table A18. MARS model III and the corresponding experimental conditions

Experimental condition			
tick		100, 150, 200, 250, 300	
learnErr		<=0.15	
space_k		1	.
actType		C	
Number of observations		4,500	
Number of independent variables		5	
Regression model			
Component	***Description***	***Coefficient***	***P value***
Intercept		0.900	< 0.0001
BF1	max (0, randLink – 0.215)	0.372	< 0.0001
BF2	max (0, 0.215 – randLink)	–1.150	< 0.0001
BF3	max (0, decayRate – 0.771)	–0.092	< 0.0001
BF4	max (0, 0.771 – decayRate)	–0.244	< 0.0001
BF5	max (0, Tick – 150)	0.001	< 0.0001
BF6	max (0, 150 – Tick)	–0.001	< 0.0001
BF7	max (0, randLink – 0.518)	–0.471	< 0.0001
BF8	BF2 * max (0, wtGain – 1.3)	–1.375	< 0.0001
BF9	BF2 * max (0, 1.3 – wtGain)	–0.704	0.0010
BF10	BF4 * max (0, randLink – 0.672)	–0.357	0.0040
BF11	BF4 * max (0, 0.672 – randLink)	–0.618	< 0.0001
BF12	BF2 * max (0, Tick – 200) * max(0, 1.3 – wtGain)	0.007	< 0.0001
BF13	BF2 * max (0, 200 – Tick) * max(0, 1.3 – wtGain)	–0.012	< 0.0001
BF14	BF4 * max (0, 200 – Tick) * max(0, 0.672 – randLink)	–0.008	< 0.0001
BF15	BF3 * max (0, randLink – 0.622)	0.729	< 0.0001
BF16	BF3 * max (0, 0.622 – randLink)	1.349	< 0.0001
BF17	BF2 * max (0, learnErr – 0.132)	–221.4	< 0.0001
Goodness–of–fit			
GCV (Generalized Cross Validation)		0.000273	
GCV–Squared		0.990	
RSS (Residual Sum of Squares)		0.161	
R–Squared		0.991	

Table A19. MARS model IV and the corresponding experimental conditions

Experimental condition			
tick		100, 150, 200, 250, 300	
learnErr		<=0.15	
space_k		5	
actType		B	
Number of observations		4,500	
Number of independent variables		6	
Regression model			
Component	***Description***	***Coefficient***	***P value***
Intercept		0.582	< 0.0001
BF1	max (0, Tick – 200)	0.001	< 0.0001
BF2	max (0, 200 – Tick)	–0.001	< 0.0001
BF3	max (0, Tick – 150)	–0.001	< 0.0001
BF4	max (0, randLink – 0.278)	–0.012	< 0.0001
BF5	max (0, 0.278 – randLink)	–0.467	< 0.0001
BF6	max (0, decayRate – 0.824)	0.175	< 0.0001
BF7	max (0, 0.824 – decayRate)	–0.064	< 0.0001
BF8	orgSize = 100	0.047	< 0.0001
BF9	orgSize = 200	0.043	< 0.0001
BF10	BF4 * max (0, actDist – 0.735)	0.013	< 0.0001
BF11	BF4 * max (0, 0.735 – actDist)	–0.188	< 0.0001
BF12	BF7 * max (0, innoRange – 2)	–0.065	< 0.0001
BF13	BF5 * max (0, decayRate – 0.831)	8.217	< 0.0001
BF14	BF5 * max (0, 0.831 – decayRate)	0.461	< 0.0001
BF15	BF3 * BF6 * max (0, 2 – innoRange)	–0.002	< 0.0001
BF16	BF7 * max (0, 150 – Tick) * max(0, 2 – innoRange)	0.001	< 0.0001
Goodness–of–fit			
GCV (Generalized Cross Validation)		0.000118	
GCV–Squared		0.977	
RSS (Residual Sum of Squares)		0.070	
R–Squared		0.980	

Table A20. Descriptive analysis of the underlying subset of MARS IV and MARS V[118]

Variable	actDist	randLink	wtGain	decayRate	learnErr	innoRange
Minimum	0.0050	0.001667	0.1048	0.5008	0.001667	1
1st Quartile	0.7525	0.250833	0.8274	0.6254	0.250833	2
Median	1.5000	0.500000	1.5500	0.7500	0.500000	3
Mean	1.5000	0.500000	1.5500	0.7500	0.500000	3
3rd Quartile	2.2475	0.749167	2.2726	0.8746	0.749167	4
Maximum	2.9950	0.998333	2.9952	0.9992	0.998333	5

[118] The descriptive analysis results of the two result subsets were identical

Table A21. MARS model V and the corresponding experimental conditions

Experimental condition			
tick		100, 150, 200, 250, 300	
learnErr		<=0.15	
space_k		1	
actType		B	
Number of observations		4,500	
Number of independent variables		6	
Regression model			
Component	***Description***	***Coefficient***	***P value***
Intercept		0.988	< 0.0001
BF1	max (0, decayRate – 0.858)	0.147	< 0.0001
BF2	max (0, 0.858 – decayRate)	–0.300	< 0.0001
BF3	max (0, Tick – 150)	0.001	< 0.0001
BF4	max (0, 150 – Tick)	–0.001	< 0.0001
BF5	max (0, randLink – 0.365)	–0.042	< 0.0001
BF6	max (0, 0.365 – randLink)	–0.496	< 0.0001
BF7	max (0, actDist – 0.945)	0.011	< 0.0001
BF8	max (0, 0.945 – actDist)	–0.052	< 0.0001
BF9	max (0, innoRange – 2)	–0.003	< 0.0001
BF10	max (0, 2 – innoRange)	–0.035	< 0.0001
BF11	BF6 * max (0, actDist – 1.42)	0.254	< 0.0001
BF12	BF6 * max (0, 1.42 – actDist)	0.410	< 0.0001
BF13	BF6 * max (0, Tick – 200)	0.001	< 0.0001
BF14	BF6 * max (0, 200 – Tick)	–0.002	< 0.0001
BF15	BF3 * BF7	–0.001	< 0.0001
BF16	BF4 * BF7	0.001	< 0.0001
BF17	BF2 * max (0, 0.0383 – learnErr)	4.082	< 0.0001
BF18	BF2 * max (0, 200 – Tick)	–0.001	< 0.0001
Goodness–of–fit			
GCV (Generalized Cross Validation)		0.000134	
GCV–Squared		0.966	
RSS (Residual Sum of Squares)		0.078	
R–Squared		0.970	

Table A22. Relative importance of the variables in terms of predicting organizational performance

Experimental condition	Variable	Importance (–gcv)	Importance (–rss)
space_k = 5	randLink	100.0	100.0
actType = C	Tick	58.7	58.6
	decayRate	46.7	46.5
	orgSize	32.5	32.2
	wtGain	32.5	32.2
space_k = 1	randLink	100.0	100.0
actType = C	Tick	36.5	36.5
	decayRate	36.5	36.5
	wtGain	16.1	16.0
	learnErr	15.4	15.3
space_k = 5	Tick	100.0	100.0
actType = B	randLink	62.8	62.6
	decayRate	54.8	54.6
	orgSize	43.4	42.9
	innoRange	43.4	42.9
	actDist	31.0	30.6
space_k = 1	decayRate	100.0	100.0
actType = B	randLink	83.3	83.0
	Tick	66.1	65.8
	actDist	43.2	42.9
	innoRange	25.0	24.6
	learnErr	15.8	15.5

Table A23. MARS model VI and the corresponding experimental conditions

Experimental Condition			
tick		1, 6, 11, 16, 21, 26, 31, 36, 41, 46, 50, 100, 150, …, 1,000	
actType		B	
Number of observations[119]		90,000	
Number of independent variables		6	
Regression model			
Component	**Description**	**Coefficient**	***P* value**
Intercept		0.852	< 0.0001
BF1	max (0, Tick – 150)	0.001	< 0.0001
BF2	max (0, 150 – Tick)	–0.004	< 0.0001
BF3	space_k = 5	–0.370	< 0.0001
BF4	nkSpace = 4	–0.045	< 0.0001
BF5	nkSpace = 5	–0.042	< 0.0001
BF6	max (0, decayRate – 0.814)	0.149	< 0.0001
BF7	max (0, 0.814 – decayRate)	–0.097	< 0.0001
BF8	max (0, learnErr – 0.822)	–0.849	< 0.0001
BF9	max (0, 0.822 – learnErr)	0.234	< 0.0001
BF10	max (0, Tick – 200)*BF3	0.001	< 0.0001
BF11	max (0, 200 – Tick)*BF3	0.001	< 0.0001
BF12	BF1 * max (0, learnErr – 0.978)	–0.009	< 0.0001
BF13	BF1 * max (0, 0.978 – learnErr)	–0.001	< 0.0001
BF14	BF1 * max (0, innoRange – 2)	–0.001	< 0.0001
BF15	BF1 * max (0, 2 – innoRange)	–0.001	< 0.0001
BF16	BF3 * (nkSpace = 2)	–0.077	< 0.0001
BF17	BF3 * (nkSpace = 4)	–0.066	< 0.0001
BF18	BF3 * max (0, 200–Tick)*max (0, learnErr–0.842)	0.003	< 0.0001
BF19	BF3 * max (0, 200–Tick)*max (0, 0.842–learnErr)	–0.001	< 0.0001
Goodness-of-fit			
GCV (Generalized Cross Validation)		0.00481	
GCV-Squared		0.941	
RSS (Residual Sum of Squares)		433	
R-Squared		0.941	

[119] The result set has 9,000 items: 1 (*orgSize* = 100) × 1 (*actType* = B) × 2 (*space_k* = 1 or 5) × 5 (*nkSpace* = 0, 1, 2, 3, 4) × 300 (LHS design points) × 30 (*tick* = 1, 6, 11, 16, 21, 26, 31, 36, 50, 100, 150, …, 1,000).

Appendix H. Interactive effects of *randLink* and *decayRate* on the emergent network

To better understand how *decayRate* affected *avgScore* independently and together with *randLink*, we conducted a new set of simulation experiments focusing on the emergent macro network. In the experimental design, the value ranges of *decayRate* and *randLink* were both from 0 to 1 with an increment of 0.25. Other model parameters were fixed: *orgSize* = 200, *actType* = C, *space_k* = 5, *innoRange* = 1, *learnErr* = 0.15, and *wtGain* = 1. Every experiment lasted for 800 ticks. We collected data from *tick* = 50 to *tick* = 800 with an interval of 10 ticks. **Figure A5** to **Figure A7** show the effects of *decayRate* on the macro network structure given different values of *randLink*. In these figures, every trend line correspond to a specific tick. The relations between *decayRate* and specific network statistics were consistent over different ticks in most cases. Accelerating the decay of infrequently used ties would eventually lead to a sparse network with weak ties and few clusters. However, increasing *decayRate* (yet remaining moderate) counterintuitively increased network density and this effect was more apparent when *randLink* was small, i.e., when knowledge exchanges tend to be embedded (**Figure A5**). Meanwhile, global clustering coefficient kept declining as *decayRate* increased (**Figure A6**), indicating that new ties did not contribute to cluster formation. Together these results suggest that increasing *decayRate* reduced the tendency for knowledge exchanges to happen between already connected individuals, a tendency that is stronger when *randLink* is smaller. Finally, the relationships between *decayRate* and *avgScore* revealed by this set of experiments (**Figure A8**) was generally consistent with earlier MARS results. On the one hand, moderate *decayRate* had a positive effect on *avgScore* while large *decayRate* negatively impacted *avgScore*. On the other hand, small *randLink* reinforced the positive effect and weakened the negative effect of *decayRate*. When *randLink* was smaller, the value range of *decayRate* that showed negative effects seemed to be narrower.

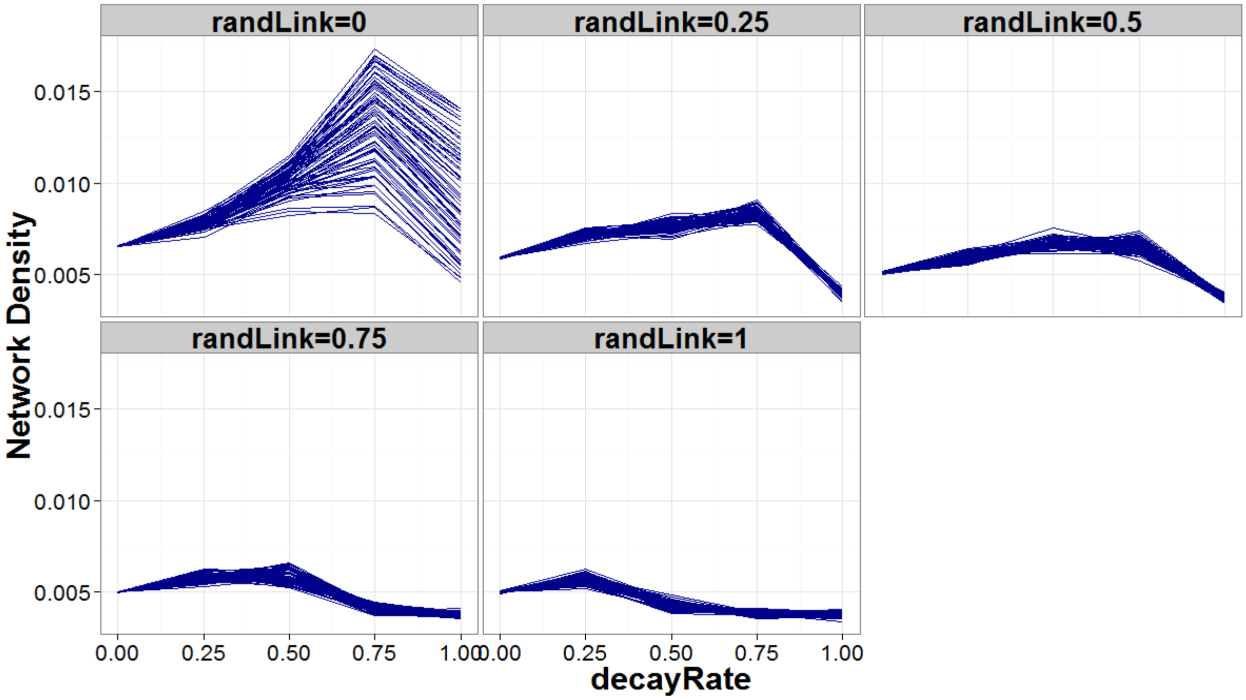


Figure A33: Relations between *decayRate* and network density given different *randLink*

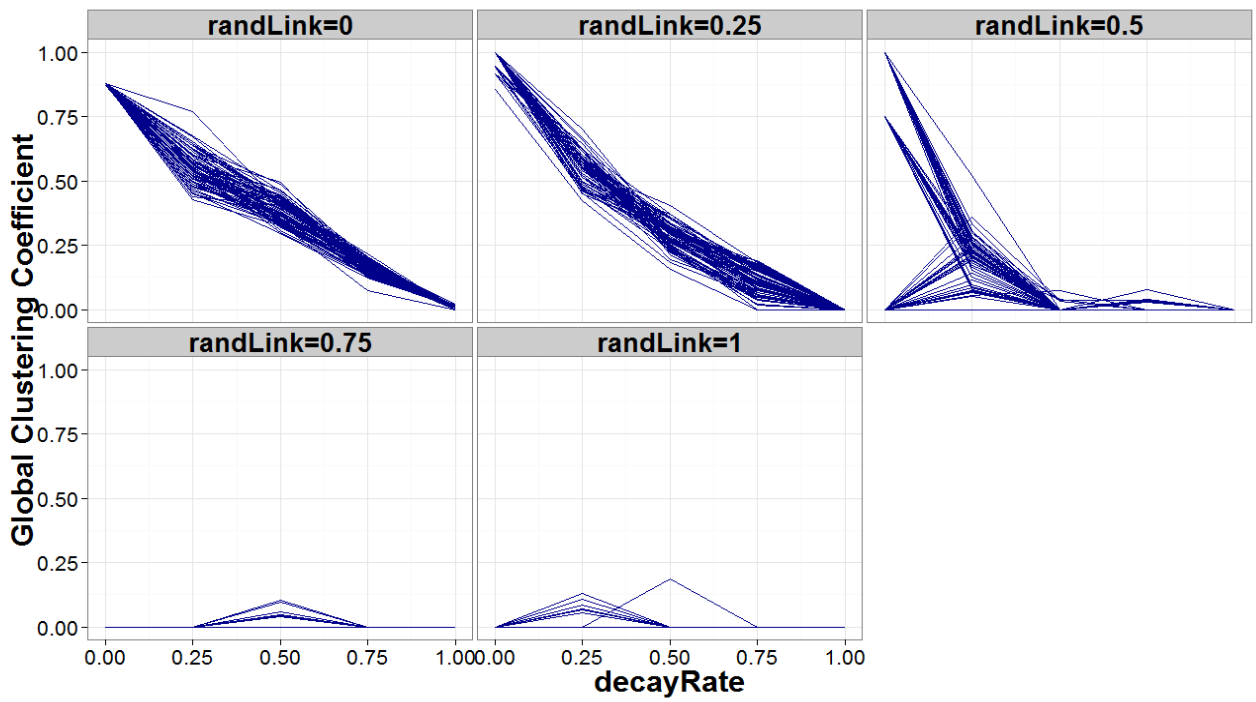


Figure A34: Relations between *decayRate* and global clustering coefficient[120]

[120] Global clustering coefficient is a network statistic measuring the clustering extent of the entire network. It is defined as 3 * number of triangles in the network / number of paths of length 2.

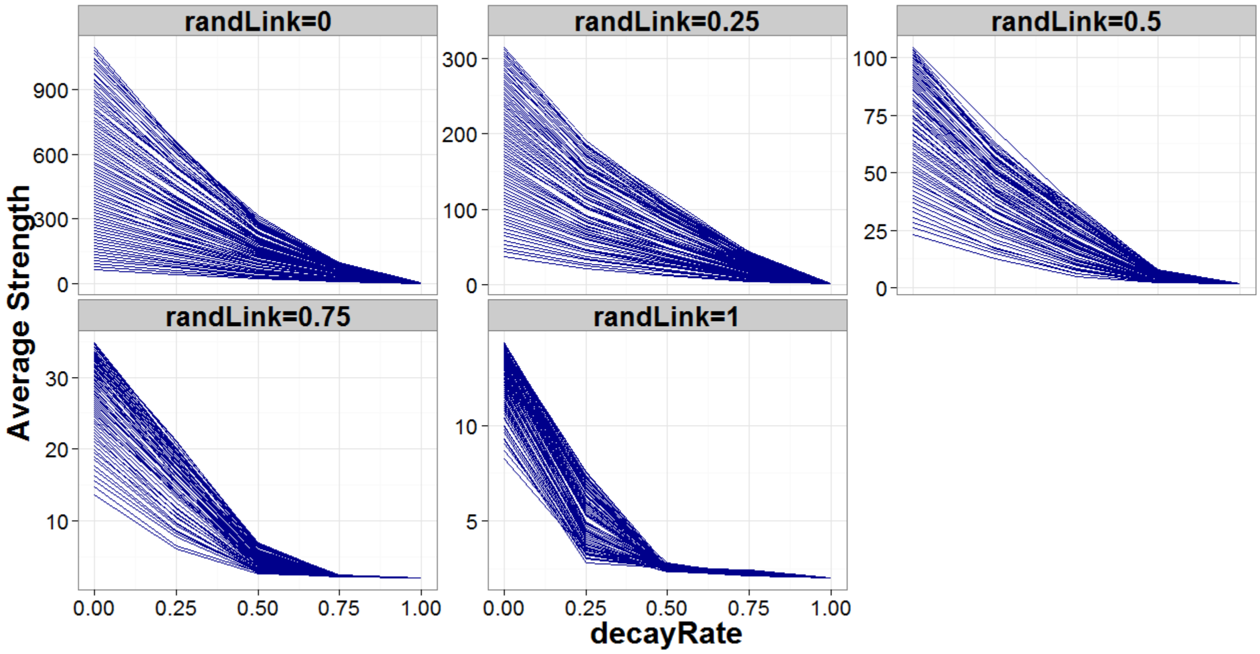


Figure A35: Relations between *decayRate* and average tie strength

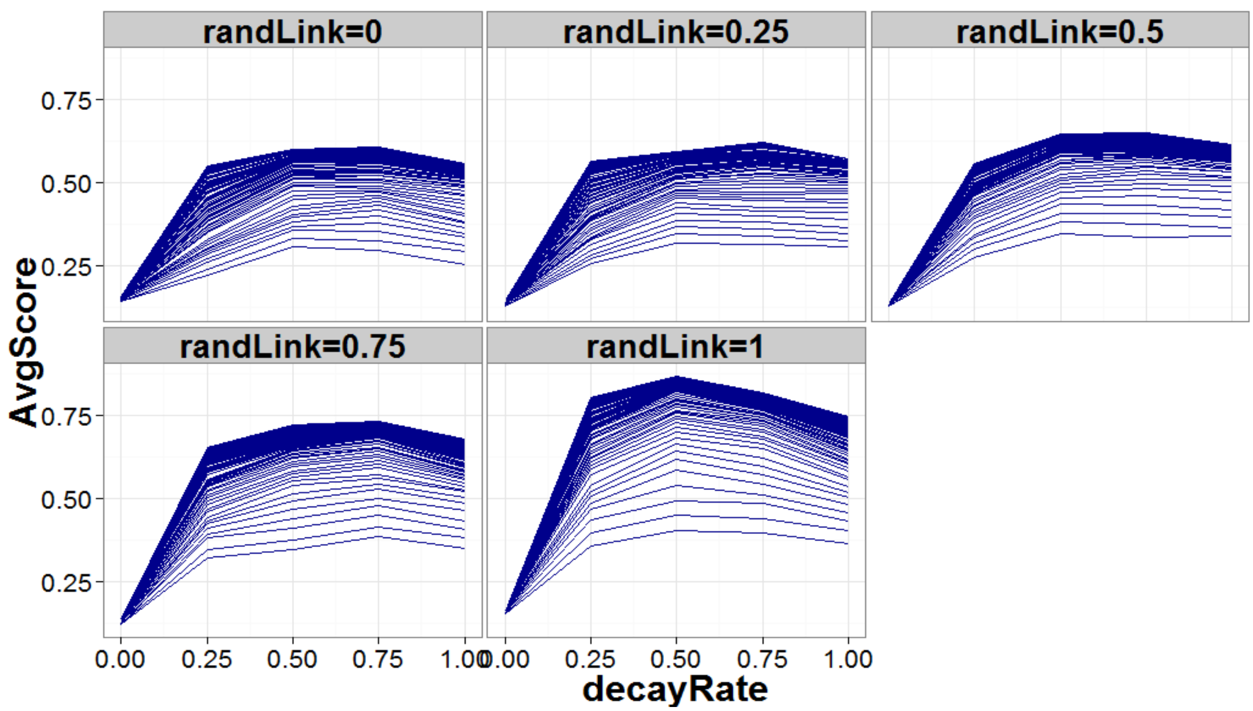


Figure A36: Relations between *decayRate* and *avgScore* given *randLink*

Appendix I. Extreme condition tests and results

To further verify the integrity of the computational model, the computer program was tested in several extreme conditions (**Table A16**). The results were presented in **Figure A9** and **Figure A10**. Organizational performance *avgScore* [121] increased smoothly over time in each condition. The relationships between *avgScore* and primary model parameters, as reflected in these extreme conditions, were mostly consistent with the results of more sophisticatedly designed simulation experiments (Section 4.1 in the main text). The performances were always better when the problem was less complex (*space_k* = 1). The performances were generally better when individual members solved problems by independent knowledge creation and knowledge exchange (*actType* = B) than by knowledge exchange only (*actType* = C). When *actType* = C, the organizational performance stopped improving while it was still low in half of the extreme conditions. The first condition (*designPt* = 1) represented the "lock-in" effect (State H in **Figure 13**): when (a) individuals kept exchanging knowledge with the same others (*randLink* = 0), (b) network tie decayed slowly (*decayRate* = 0.01), and (c) there were few mutations (*learnErr* = 0), the knowledge (or individual solutions) being exchanged soon became almost identical. As a result, no new solutions were created and organizational improvement stopped. While everything else remained the same but the mutation rate changed to extremely high (*designPt* = 2), organizational performances were low in the short run but high in the long run. This phenomenon has been observed in a previous static network based model (Lazer & Friedman 2007). The negative effect of high *learnErr* on short-term performances was also observed in several other extreme conditions (*designPt* = 4, 6, 8, 10, 14), comparing to corresponding conditions that only differ in *learnErr* (*designPt* = 3, 5, 7, 9, 13). When individual members all preferred random knowledge exchange (*randLink* = 1), the network soon became well connected with no closure structure (State C in **Figure 13**). On this type of networks, knowledge is efficiently disseminated and assimilated. Thus, the organizational performance leveled off at a low value (*designPt* = 9 – 15). However, there was an exception (*designPt* = 16) when used and unused ties were respectively strengthened and weakened in a high rate (*wtGain* = 3, *decayRate* = 1) and the mutation rate

[121] The values presented in Figure A9 and Figure A10 are averaged across 50 replicate runs.

was high as well (*learnErr*). In this condition, the positive effect of *learnErr* on organizational performance (by increasing knowledge diversity) was reinforced by high *wtGain*, which made used ties stronger and therefore allowed for more different knowledge areas to be exchanged and then mutated. This interactive effect also explained why *designPt* = 61 (*actType* = B) had better performances than *designPt* = 63 while *designPt* = 53 had worse performances than *designPt* = 55. All in all, it can be concluded that the behaviors of the current study's model continue to make sense at boundary or extreme conditions.

Table A24. Extreme conditions[122]

	designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
actType = B,	1	0.01	0	0.01	0.01	0	1
space_k = 1	2	0.01	0	0.01	0.01	0	5
and 5	3	0.01	0	0.01	0.01	1	1
	4	0.01	0	0.01	0.01	1	5
	5	0.01	0	0.01	1	0	1
	6	0.01	0	0.01	1	0	5
	7	0.01	0	0.01	1	1	1
	8	0.01	0	0.01	1	1	5
	9	0.01	0	3	0.01	0	1
	10	0.01	0	3	0.01	0	5
	11	0.01	0	3	0.01	1	1
	12	0.01	0	3	0.01	1	5
	13	0.01	0	3	1	0	1
	14	0.01	0	3	1	0	5
	15	0.01	0	3	1	1	1
	16	0.01	0	3	1	1	5
	17	0.01	1	0.01	0.01	0	1
	18	0.01	1	0.01	0.01	0	5
	19	0.01	1	0.01	0.01	1	1
	20	0.01	1	0.01	0.01	1	5
	21	0.01	1	0.01	1	0	1
	22	0.01	1	0.01	1	0	5
	23	0.01	1	0.01	1	1	1
	24	0.01	1	0.01	1	1	5
	25	0.01	1	3	0.01	0	1
	26	0.01	1	3	0.01	0	5

[122] When *actType* = C, all individual agents are willing to exchange knowledge and none of them will do independent knowledge creation. Thus, *actDist* and *innoRange* are no longer effective.

	designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
actType = B,	27	0.01	1	3	0.01	1	1
space_k = 1	28	0.01	1	3	0.01	1	5
and 5	29	0.01	1	3	1	0	1
	30	0.01	1	3	1	0	5
	31	0.01	1	3	1	1	1
	32	0.01	1	3	1	1	5
	33	3	0	0.01	0.01	0	1
	34	3	0	0.01	0.01	0	5
	35	3	0	0.01	0.01	1	1
	36	3	0	0.01	0.01	1	5
	37	3	0	0.01	1	0	1
	38	3	0	0.01	1	0	5
	39	3	0	0.01	1	1	1
	40	3	0	0.01	1	1	5
	41	3	0	3	0.01	0	1
	42	3	0	3	0.01	0	5
	43	3	0	3	0.01	1	1
	44	3	0	3	0.01	1	5
	45	3	0	3	1	0	1
	46	3	0	3	1	0	5
	47	3	0	3	1	1	1
	48	3	0	3	1	1	5
	49	3	1	0.01	0.01	0	1
	50	3	1	0.01	0.01	0	5
	51	3	1	0.01	0.01	1	1
	52	3	1	0.01	0.01	1	5
	53	3	1	0.01	1	0	1
	54	3	1	0.01	1	0	5
	55	3	1	0.01	1	1	1
	56	3	1	0.01	1	1	5
	57	3	1	3	0.01	0	1
	58	3	1	3	0.01	0	5
	59	3	1	3	0.01	1	1
	60	3	1	3	0.01	1	5
	61	3	1	3	1	0	1
	62	3	1	3	1	0	5
	63	3	1	3	1	1	1
	64	3	1	3	1	1	5
actType = C,	1	NA	0	0.01	0.01	0	NA
space_k = 1	2	NA	0	0.01	0.01	0.9	NA
and 5	3	NA	0	0.01	1	0	NA
	4	NA	0	0.01	1	0.9	NA

	designPt	actDist	randLink	wtGain	decayRate	learnErr	innoRange
actType = C,	5	NA	0	3	0.01	0	NA
space_k = 1	6	NA	0	3	0.01	0.9	NA
and 5	7	NA	0	3	1	0	NA
	8	NA	0	3	1	0.9	NA
	9	NA	1	0.01	0.01	0	NA
	10	NA	1	0.01	0.01	0.9	NA
	11	NA	1	0.01	1	0	NA
	12	NA	1	0.01	1	0.9	NA
	13	NA	1	3	0.01	0	NA
	14	NA	1	3	0.01	0.9	NA
	15	NA	1	3	1	0	NA
	16	NA	1	3	1	0.9	NA

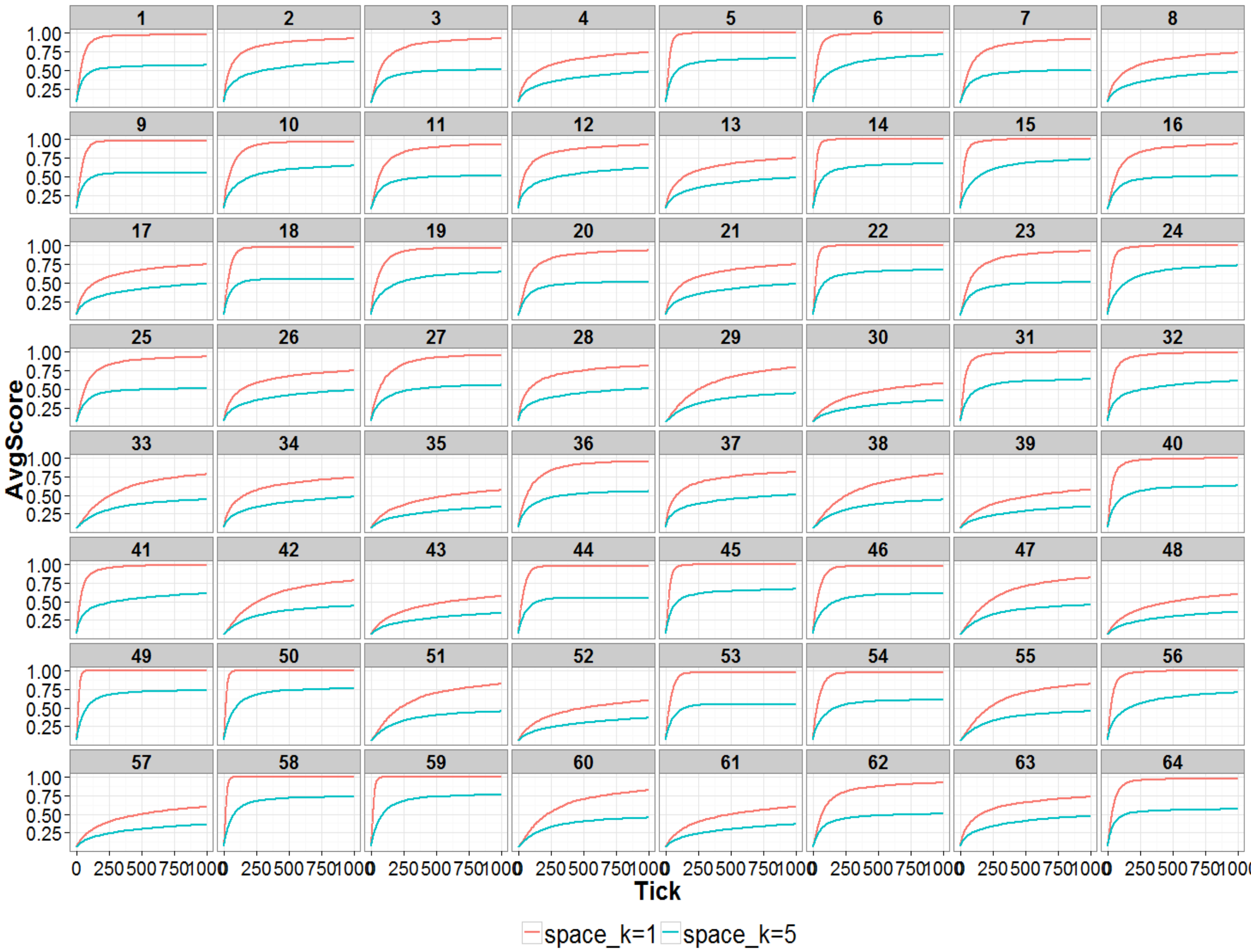


Figure A37: Changes in *avgScore* over time in different extreme conditions (*actType* = *B*)

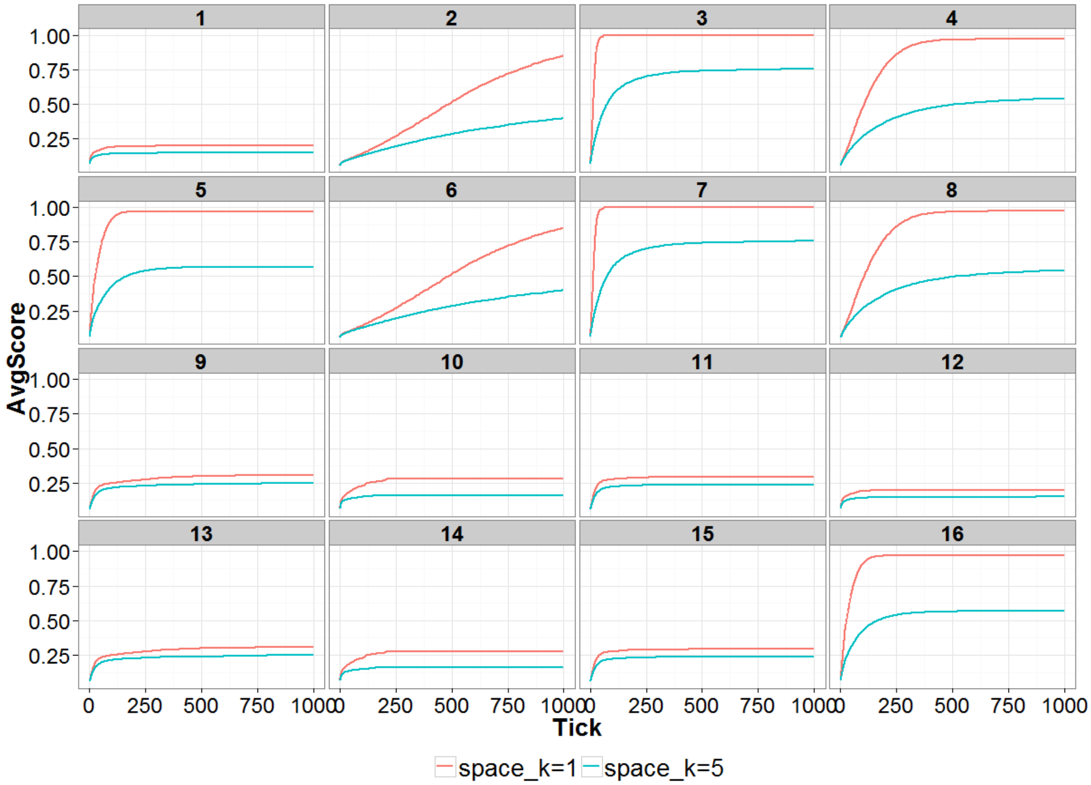


Figure A38: Changes in *avgScore* over time in different extreme conditions (*actType* = *C*)

REFERENCES

Abell, P., Felin, T., & Foss, N. (2008). Building micro-foundations for the routines, capabilities, and performance links. *Managerial and Decision Economics, 29*(6), 489-502. doi: 10.1002/mde.1413

Adler, P. S., Goldoftas, B., & Levine, D. I. (1999). Flexibility versus Efficiency? A Case Study of Model Changeovers in the Toyota Production System. *Organization Science, 10*(1), 43-68.

Adler, P. S., & Kwon, S.-W. (2002). Social Capital: Prospects for a New Concept. *The Academy of Management Review, 27*(1), 17-40.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly, 25*(1), 107-136.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47.

Aldecoa, R., & Marín, I. (2013). Surprise maximization reveals the community structure of complex networks. *Sci Rep, 3*.

Alden, K., Read, M., Timmis, J., Andrews, P. S., Veiga-Fernandes, H., & Coles, M. (2013). Spartan: A Comprehensive Tool for Understanding Uncertainty in Simulations of Biological Systems. *PLoS Comput Biol, 9*(2), e1002916. doi: 10.1371/journal.pcbi.1002916

Alexiev, A. S., Jansen, J. J., Van den Bosch, F. A., & Volberda, H. W. (2010). Top management team advice seeking and exploratory innovation: the moderating role of TMT heterogeneity. *Journal of Management Studies, 47*(7), 1343-1364.

Anderson, J. R. (1992). Automaticity and the ACT theory. *The American journal of psychology*, 165-180.

Anderson, P. (1999). Complexity Theory and Organization Science. *Organization Science, 10*(3), 216-232.

Aral, S., & Alstyne, M. V. (2011). The Diversity-Bandwidth Trade-off. *American Journal of Sociology, 117*(1), 90-171.

Argote, L., McEvily, B., & Reagans, R. (2003). Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes. *Management Science, 49*(4), 571-582.

Argyris, C., & Schön, D. A. (1974). *Theory in practice: increasing professional effectiveness*: Jossey-Bass Publishers.

Argyris, C., & Schön, D. A. (1978). *Organizational learning: a theory of action perspective*: Addison-Wesley Pub. Co.

Bala, V., & Goyal, S. (2000). A Noncooperative Model of Network Formation. *Econometrica, 68*(5), 1181-1229.

Balkundi, P., & Harrison, D. A. (2006). Ties, Leaders, and Time in Teams: Strong Inference about Network Structure's Effects on Team Viability and Performance. [Article]. *Academy of Management Journal, 49*(1), 49-68.

Bar-Yam, Y. (2013). The Limits of Phenomenology: From Behaviorism to Drug Testing and Engineering Design. *arXiv preprint arXiv:1308.3094*.

Barabasi, A., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science, 286*(5439), 509-512. doi: 10.1126/science.286.5439.509

Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America, 101*(11), 3747-3752. doi: 10.1073/pnas.0400087101

Barrat, A., & Cattuto, C. (2013). Temporal networks of face-to-face human interactions *Temporal Networks* (pp. 191-216): Springer.

Baum, J. A. C., McEvily, B., & Rowley, T. J. (2010). Better with Age? Tie Longevity and the Performance Implications of Bridging and Closure. *Organization Science*. doi: 10.1287/orsc.1100.0566

Bernard, H. R., Killworth, P., Kronenfeld, D., & Sailer, L. (1984). The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology, 13*(ArticleType: research-article / Full publication date: 1984 / Copyright © 1984 Annual Reviews), 495-517.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(4–5), 175-308. doi: http://dx.doi.org/10.1016/j.physrep.2005.10.009

Bock, G.-W., Zmud, R. W., Kim, Y.-G., & Lee, J.-N. (2005). Behavioral intention formation in knowledge sharing: examining the roles of extrinsic motivators, social-psychological forces, and organizational climate (1). *MIS Quarterly, 29*(1), 87(25).

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America, 99*(Suppl 3), 7280-7287.

Booker, A. J. (1998). Design and analysis of computer experiments. *AIAA paper*, 98-4757.

Borgatti, S. P., & Cross, R. (2003). A Relational View of Information Seeking and Learning in Social Networks. *Management Science, 49*(4), 432-445.

Box, G. E. P., Hunter, J. S., & Hunter, W. G. (2005). *Statistics for experimenters : design, innovation, and discovery*. Hoboken, N.J.: Wiley-Interscience.

Briscoe, F., & Tsai, W. (2010). *Relational Inertia and Client Sharing: How Organizational Members Respond to Acquisition Events in a Law Firm.*

Brown, J. S., & Duguid, P. (1991). Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. *Organization Science, 2*(1), 40-57.

Brown, J. S., & Duguid, P. (2001). Knowledge and Organization: A Social-Practice Perspective. *Organization Science, 12*(2), 198-213.

Brown, S. L., & Eisenhardt, K. M. (1997). The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations. *Administrative Science Quarterly, 42*(1), 1-34.

Brownlee, J. (2007). Complex Adaptive Systems. Melbourne, Australia: Complex Intelligent Systems Laboratory, Centre for Information Technology Research, Swinburne University of Technology.

Burgelman, R. A. (2002). Strategy as Vector and the Inertia of Coevolutionary Lock-in. *Administrative Science Quarterly, 47*(2), 325-357.

Burt, R. S. (1992). *Structural holes : the social structure of competition*. Cambridge, Mass.: Harvard University Press.

Burt, R. S. (2000a). Decay functions. *Social Networks, 22*(1), 1-28.

Burt, R. S. (2000b). The network structure of social capital. *Research in Organizational Behavior, 22*(2), 345.

Burt, R. S. (2002). Bridge decay. *Social Networks, 24*(4), 333-363.

Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology, 110*(2), 349-399. doi: doi:10.1086/421787

Burt, R. S., Jannotta, J. E., & Mahoney, J. T. (1998). Personality correlates of structural holes. *Social Networks, 20*(1), 63-87. doi: 10.1016/s0378-8733(97)00005-1

Burton, R. M., & Obel, B. (1995). The validity of computational models in organization science: From model realism to purpose of the model. *Computational & Mathematical Organization Theory, 1*(1), 57-71.

Burton, R. M., & Obel, B. (2011). Computational Modeling for What-Is, What-Might-Be, and What-Should-Be Studies—And Triangulation. *Organization Science, 22*(5), 1195-1202. doi: 10.1287/orsc.1100.0635

Buskens, V., & Van de Rijt, A. (2008). Dynamics of networks if everyone strives for structural holes. *American Journal of Sociology, 114*(2), 371-407.

Butts, C. T. (2009). Revisiting the Foundations of Network Analysis. *Science, 325*(5939), 414-416. doi: 10.1126/science.1171022

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-experimental Designs for Research*. Chicago, IL: Rand McNally.

Carley, K. (1996). *Validating Computational Models*. Carnegie Mellon University. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.87.9019

Carmeli, A., & Halevi, M. Y. (2009). How top management team behavioral integration and behavioral complexity enable organizational ambidexterity: The moderating role of contextual ambidexterity. *The Leadership Quarterly, 20*(2), 207-218.

Carroll, G. R., & Harrison, J. R. (1994). On the Historical Efficiency of Competition Between Organizational Populations. *The American Journal of Sociology, 100*(3), 720-749.

Casciaro, T. (1998). Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Social Networks, 20*(4), 331-351.

Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.-F., & Vespignani, A. (2010). Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLoS ONE, 5*(7), e11596. doi: 10.1371/journal.pone.0011596

Chalom, A., & de Prado, P. I. d. K. (2012). Parameter space exploration of ecological models. *arXiv preprint arXiv:1210.6278*.

Chalom, A., & de Prado, P. I. K. L. (2013). pse: Parameter space exploration with Latin Hypercubes.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly, 35*(1), 128-152.

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *The American Journal of Sociology, 94*, S95-S120.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Collins, A. J., Seiler, M. J., Gangel, M., & Croll, M. (2013). Applying Latin hypercube sampling to agent-based models: Understanding foreclosure contagion effects. *International Journal of Housing Markets and Analysis*, *6*(4), 422-437.

Contractor, N. S. (1994). Self-organizing systems perspective in the study of organizational communication. In B. Kovacic (Ed.), *New approaches to organizational communication* (pp. 39-65). Albany, NY: SUNY Press.

Contractor, N. S. (1999). Self-Organizing Systems Research in the Social Sciences. *Management Communication Quarterly, 13*(1), 154-166. doi: 10.1177/0893318999131009

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: design & analysis issues for field settings*. Boston, MA: Houghton Mifflin.

Corten, R., & Buskens, V. (2010). Co-evolution of conventions and networks: An experimental study. *Social Networks, 32*(1), 4-15. doi: DOI: 10.1016/j.socnet.2009.04.002

Costa, L. d. F., Rodrigues, F. A., Travieso, G., & Boas, P. R. V. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics, 56*(1), 167 - 242.

Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control, 28*(8), 1557-1575.

Craven, P., & Wahba, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik, 31*(4), 377-403.

Cross, R. L., Borgatti, S. P., & Parker, A. (2002). Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration. *California Management Review, 44*(2), 25-46.

Cross, R. L., Rice, R. E., & Parker, A. (2001). Information Seeking in Social Context: Structural Influences and Receipt of Information Benefits. *IEEE transactions on systems, man and cybernetics. Part C, Applications and reviews, 31*(4), 438-448.

Cross, R. L., & Sproull, L. (2004). More Than an Answer: Information Relationships for Actionable Knowledge. *Organization Science, 15*(4), 446-462. doi: 10.1287/orsc.1040.0075

Crozier, M. (1972). The Relationship Between Micro and Macrosociology: A Study of Organizational Systems as an Empirical Approach to the Problems of Macrosociology. *Human Relations, 25*(3), 239-251. doi: 10.1177/001872677202500304

Davenport, T. H., & Prusak, L. (1998). *Working knowledge : how organizations manage what they know*. Boston, Mass: Harvard Business School Press.

Davidsen, J., Ebel, H., & Bornholdt, S. (2002). Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks. *Physical Review Letters, 88*(12), 128701.

Davis, J. A. (1970). Clustering and Hierarchy in Interpersonal Relations: Testing Two Graph Theoretical Models on 742 Sociomatrices. *American Sociological Review, 35*(5), 843-851.

Davis, J. P. (2010). Agency and knowledge problems in network dynamics: brokers and bridges in innovative interorganizational relationships.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing Theory Through Simulation Methods. *Academy of Management Review, 32*(2), 480-499.

De Veaux, R. D., Psichogios, D. C., & Ungar, L. H. (1993). A comparison of two nonparametric estimation schemes: MARS and neural networks. *Computers & Chemical Engineering, 17*(8), 819-837. doi: http://dx.doi.org/10.1016/0098-1354(93)80066-V

Deppa, B. (2014). Tutorials and Handouts: Multivariate Adaptive Regression Splines (MARS) Retrieved 10/15/2014, from http://course1.winona.edu/bdeppa/Stat%20425/Tutorials/tutmain.htm

Dooley, K. J. (1996). A Nominal Definition of Complex Adaptive Systems. *The Chaos Network, 8*(1), 2-3.

Dooley, K. J. (1997). A Complex Adaptive Systems Model of Organization Change. *Nonlinear Dynamics, Psychology, and Life Sciences, 1*(1), 69-97. doi: 10.1023/a:1022375910940

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution, 22*(6), 469-493. doi: http://dx.doi.org/10.1016/0047-2484(92)90081-J

Dutta, B., & Jackson, M. O. (2003). *Networks and groups : models of strategic formation*. New York: Springer.

Eckmann, J.-P., Moses, E., & Sergi, D. (2004). Entropy of dialogues creates coherent structures in e-mail traffic. *Proceedings of the National Academy of Sciences of the United States of America, 101*(40), 14333-14337. doi: 10.1073/pnas.0405728101

Edmonds, B., & Moss, S. (2005). From KISS to KIDS - an 'anti-simplistic' modelling approach. In P. Davidsson, B. Logan & K. Takadama (Eds.), *Multi Agent Based Simulation 2004. Lecture Notes in Artificial Intelligence* (Vol. 3415, pp. 130-144): Springer.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic management journal, 21*(10-11), 1105-1121.

Emirbayer, M., & Goodwin, J. (1994). Network Analysis, Culture, and the Problem of Agency. *American Journal of Sociology, 99*(6), 1411-1454. doi: doi:10.1086/230450

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity, 4*(5), 41-60.

Epstein, J. M. (2006). *Generative social science : studies in agent-based computational modeling*. Princeton: Princeton University Press.

Erdos, P., & Renyi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5*, 17-61.

Fagiolo, G., Moneta, A., & Windrum, P. (2007). A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems. *Computational Economics, 30*(3), 195-226.

Fang, C., Lee, J., & Schilling, M. A. (2010). Balancing Exploration and Exploitation Through Structural Design: The Isolation of Subgroups and Organizational Learning. *Organization Science, 21*(3), 625-642. doi: 10.1287/orsc.1090.0468

Feldman, D. C., & Arnold, H. J. (1983). *Managing Indiviudal and Group Behavior in Organizations*. New York, NY: McGraw-Hill.

Felin, T., Foss, N. J., Heimeriks, K. H., & Madsen, T. L. (2012). Microfoundations of routines and capabilities: Individuals, processes, and structure. *Journal of Management Studies, 49*(8), 1351-1374.

Felin, T., & Hesterly, W. S. (2007). THE KNOWLEDGE-BASED VIEW, NESTED HETEROGENEITY, AND NEW VALUE CREATION: PHILOSOPHICAL CONSIDERATIONS ON THE LOCUS OF KNOWLEDGE. [Article]. *Academy of Management Review, 32*(1), 195-218. doi: 10.5465/AMR.2007.23464020

Fleming, L., Mingo, S., & Chen, D. (2007). Collaborative Brokerage, Generative Creativity, and Creative Success. *Administrative Science Quarterly, 52*(3), 443-475. doi: 10.2189/asqu.52.3.443

Foss, N. J. (2011). Invited editorial: Why micro-foundations for resource-based theory are needed and what they may look like. *Journal of Management, 37*(5), 1413-1428.

Friedman, A. (2007). Code and discussion of an ongoing project on parallel problem solving and network dynamics Retrieved from https://sites.google.com/site/parallelproblemsolving/

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 1-67.

Fronczak, P., Fronczak, A., & Hołyst, J. A. (2007). Phase transitions in social networks. *The European Physical Journal B-Condensed Matter and Complex Systems, 59*(1), 133-139.

Fuchs, C. (2003). Structuration Theory and Self-Organization. *Systemic Practice and Action Research, 16*(2), 133-167. doi: 10.1023/a:1022889627100

Gabbay, S. M., & Leenders, R. T. A. (2002). *Social capital of organizations: From social structure to the management of corporate social capital*: University of Groningen.

Galán, J. M., Izquierdo, L. R., Izquierdo, S. S., Santos, J. I., del Olmo, R., López-Paredes, A., & Edmonds, B. (2009). Errors and Artefacts in Agent-Based Modelling. *Journal of Artificial Societies and Social Simulation, 12*(1).

Galeotti, A., Goyal, S., & Kamphorst, J. (2006). Network formation with heterogeneous players. *Games and Economic Behavior, 54*(2), 353-372.

Gargiulo, M., & Benassi, M. (1999). The Dark Side of Social Capital. In R. T. Leenders & S. M. Gabbay (Eds.), *Corporate social capital and liability* (pp. 298-322). Boston, MA: Kluwer.

Gargiulo, M., & Benassi, M. (2000). Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital. *Organization Science, 11*(2), 183-196. doi: 10.1287/orsc.11.2.183.12514

Gell-Mann, M. (1994). complex adaptive systems. In G. A. Cowan, D. Pines & D. Meltzer (Eds.), *Complexity: Metaphors, Modes, and Reality* (pp. 17-46): Addison-Wesley.

Gibson, C. B., & Birkinshaw, J. (2004). The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity. *The Academy of Management Journal, 47*(2), 209-226.

Giddens, A. (1984). *The constitution of society : outline of the theory of structuration*. Cambridge [Cambridgeshire]: Polity Press.

Gilbert, E., & Karahalios, K. (2009). *Predicting tie strength with social media.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Gilbert, G. N. (2008). *Agent-based models*. Los Angeles: Sage Publications.

Goldstein, S. G. (1985). Organizational Dualism and Quality Circles. *The Academy of Management Review, 10*(3), 504-517.

Gonçalves, B., Perra, N., & Vespignani, A. (2011). Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number. *PLoS ONE, 6*(8), e22656. doi: 10.1371/journal.pone.0022656

Goyal, S., & Vega-Redondo, F. (2005). Network formation and social coordination. *Games and Economic Behavior, 50*(2), 178-207.

Goyal, S., & Vega-Redondo, F. (2007). Structural holes in social networks. *Journal of Economic Theory, 137*(1), 460-492.

Grabowicz, P. A., Ramasco, J. J., Moro, E., Pujol, J. M., & Eguiluz, V. M. (2012). Social Features of Online Networks: The Strength of Intermediary Ties in Online Social Media. *PLoS ONE, 7*(1), e29358. doi: 10.1371/journal.pone.0029358

Granovetter, M. S. (1973). The Strength of Weak Ties. *The American Journal of Sociology, 78*(6), 1360-1380.

Granovetter, M. S. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *The American Journal of Sociology, 91*(3), 481-510.

Grant, R. M. (1996). Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal, 17*, 109-122.

Gulati, R., & Puranam, P. (2009). Renewal Through Reorganization: The Value of Inconsistencies Between Formal and Informal Organization. *Organization Science, 20*(2), 422-440. doi: 10.1287/orsc.1090.0421

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The Interplay between Exploration and Exploitation. *The Academy of Management Journal, 49*(4), 693-706.

Haas, M. R., & Hansen, M. T. (2007). Different knowledge, different benefits: toward a productivity perspective on knowledge sharing in organizations. *Strategic Management Journal, 28*(11), 1133-1153. doi: 10.1002/smj.631

Hallinan, M. T., & Kubitschek, W. N. (1988). The Effects of Individual and Structural Characteristics on Intransitivity in Social Networks. *Social Psychology Quarterly, 51*(2), 81-92.

Hammer, M. (1980). Predictability of social connections over time. *Social Networks, 2*(2), 165-180.

Hansen, M. T. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly, 44*(1), 82-111.

Hansen, M. T. (2009). When Internal Collaboration Is Bad for Your Company. [Article]. *Harvard Business Review, 87*(4), 82-88.

Hansen, M. T., Mors, M. L., & Lovas, B. (2005). Knowledge Sharing in Organizations: Multiple Networks, Multiple Phases. *Academy of Management journal., 48*(5), 776-793.

Hansen, M. T., & Nohria, N. (2004). How To Build Collaborative Advantage. (cover story). [Article]. *MIT Sloan Management Review, 46*(1), 22-30.

Hansen, M. T., Nohria, N., & Tierney, T. (1999). What's your strategy for managing knowledge? *Harvard Business Review, 77*(2), 106-116.

Harrison, J. R., Zhiang, L. I. N., Carroll, G. R., & Carley, K. M. (2007). Simulation Modeling in Organizational and Management Research. *Academy of Management Review, 32*(4), 1229-1245.

Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety, 81*(1), 23-69.

Helton, J. C., Martell, M. A., & Tierney, M. S. (2000). Characterization of subjective uncertainty in the 1996 performance assessment for the Waste Isolation Pilot Plant. *Reliability Engineering & System Safety, 69*(1–3), 191-204. doi: http://dx.doi.org/10.1016/S0951-8320(00)00032-6

Holan, P. M. d., & Nelson, P. (2004). Remembrance of Things Past? The Dynamics of Organizational Forgetting. *Management Science, 50*(11), 1603-1613.

Holland, J. H. (1995). *Hidden order : how adaptation builds complexity*. Reading, Mass.: Addison-Wesley.

Holland, J. H. (1998). *Emergence : from chaos to order*. Reading, Mass.: Addison-Wesley.

Holland, P. W., & Leinhardt, S. (1971). Transitivity in Structural Models of Small Groups. *Small Group Research, 2*(2), 107-124. doi: 10.1177/104649647100200201

Holland, P. W., & Leinhardt, S. (1972). Holland and Leinhardt Reply: Some Evidence on the Transitivity of Positive Interpersonal Sentiment. *The American Journal of Sociology, 77*(6), 1205-1209.

Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports, 519*(3), 97-125.

Hummon, N. P. (2000). Utility and dynamic social networks. *Social Networks, 22*(3), 221-249.

Hummon, N. P., & Fararo, T. J. (1995). The emergence of computational sociology. *The Journal of Mathematical Sociology, 20*(2-3), 79-87. doi: citeulike-article-id:4100436

Huy, Q. N. (2002). Emotional Balancing of Organizational Continuity and Radical Change: The Contribution of Middle Managers. *Administrative Science Quarterly, 47*(1), 31-69.

Ibarra, H., Kilduff, M., & Tsai, W. (2005). Zooming In and Out: Connecting Individuals and Collectivities at the Frontiers of Organizational Network Research. *Organization Science, 16*(4), 359-371. doi: 10.1287/orsc.1050.0129

Inkpen, A. C., & Tsang, E. W. K. (2005). Social capital, networks, and knowledge transfer. *Academy of Management Review, 30*(1), 146-165.

Jackson, M. O. (2003). A survey of models of network formation: stability and efficiency. In G. Demange & M. Wooders (Eds.), *Group Formation in Economics: Networks, Clubs, and Coalitions*: Cambridge University Press.

Jackson, M. O. (2010). An overview of social networks and economic applications. *The handbook of social economics*, 511-585.

Jackson, M. O., & Watts, A. (2002). The Evolution of Social and Economic Networks. *Journal of Economic Theory, 106*(2), 265-295.

Jackson, M. O., & Wolinsky, A. (1996). A Strategic Model of Social and Economic Networks. *Journal of Economic Theory, 71*(1), 44-74.

Jansen, J. J., Vera, D., & Crossan, M. (2009). Strategic leadership for exploration and exploitation: The moderating role of environmental dynamism. *The Leadership Quarterly, 20*(1), 5-18.

Jansen, J. J. P., George, G., Van den Bosch, F. A. J., & Volberda, H. W. (2008). Senior Team Attributes and Organizational Ambidexterity: The Moderating Role of Transformational Leadership. *Journal of Management Studies, 45*(5), 982-1007. doi: 10.1111/j.1467-6486.2008.00775.x

Jin, E. M., Girvan, M., & Newman, M. E. J. (2001). Structure of growing social networks. *Physical Review E, 64*(4), 046132.

Jo, H.-H., Pan, R. K., & Kaski, K. (2011). Emergence of Bursts and Communities in Evolving Weighted Networks. *PLoS ONE, 6*(8), e22687. doi: 10.1371/journal.pone.0022687

Kalish, Y., & Robins, G. (2006). Psychological predispositions and network structure: The relationship between individual predispositions, structural holes and network closure. *Social Networks, 28*(1), 56-84. doi: DOI: 10.1016/j.socnet.2005.04.004

Kane, G. C., & Alavi, M. (2005). *Information Technology and Organizational Learning: An Investigation of Exploitation and Exploration Processes.* Paper presented at the International Conference on Information Systems (ICIS).

Kang, M., & Kim, Y.-G. (2010). A multilevel view on interpersonal knowledge transfer. *Journal of the American Society for Information Science and Technology, 61*(3), 483-494. doi: 10.1002/asi.21267

Kannan, R., Vempala, S., & Vetta, A. (2004). On clusterings: Good, bad and spectral. *Journal of the ACM (JACM), 51*(3), 497-515.

Kaplan, A. (1964). *The conduct of inquiry; methodology for behavioral science*. San Francisco: Chandler Pub. Co.

Karsai, M., Kaski, K., Barabási, A.-L., & Kertész, J. (2012). Universal features of correlated bursty behaviour. *Sci Rep, 2*.

Karsai, M., Perra, N., & Vespignani, A. (2013). The emergence and role of strong ties in time-varying communication networks. *arXiv preprint arXiv:1303.5966*.

Kauffman, S. A. (1993). *The origins of order : self-organization and selection in evolution*. New York: Oxford University Press.

Keller, R. T. (1986). Predictors of the Performance of Project Groups in R & D Organizations. *The Academy of Management Journal, 29*(4), 715-726. doi: 10.2307/255941

Kilduff, M., Crossland, C., Tsai, W., & Krackhardt, D. (2008). Organizational network perceptions versus reality: A small world after all? *Organizational Behavior and Human Decision Processes, 107*(1), 15-28.

Kilduff, M., & Krackhardt, D. (1994). Bringing the Individual Back in: A Structural Analysis of the Internal Market for Reputation in Organizations. *The Academy of Management Journal, 37*(1), 87-108.

Kilduff, M., Tsai, W., & Hanke, R. (2006). A paradigm too far? A dynamic stability reconsideration of the social network research program. *Academy of Management Review, 31*(4), 1031-1048.

Kleijnen, J. P. C., & Helton, J. C. (1999). Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1: Review and comparison of techniques. *Reliability Engineering & System Safety, 65*(2), 147-185. doi: http://dx.doi.org/10.1016/S0951-8320(98)00091-X

Kleijnen, J. P. C., Sanchez, S. M., Lucas, T. W., & Cioppa, T. M. (2005). A User's Guide to the Brave New World of Designing Simulation Experiments. *INFORMS Journal on Computing, 17*(3), 263-289.

Klein, K. J., Lim, B.-C., Saltz, J. L., & Mayer, D. M. (2004). How Do They Get There? An Examination of the Antecedents of Centrality in Team Networks. *The Academy of Management Journal, 47*(6), 952-963.

Kmetz, J. L. (1984). An Information-Processing Study of a Complex Workflow in Aircraft Electronics Repair. *Administrative Science Quarterly, 29*(2), 255-280.

Kogut, B. (2000). The Network as Knowledge: Generative Rules and the Emergence of Structure. *Strategic Management Journal, 21*(3), 405-425.

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

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science, 311*(5757), 88-90.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks, 9*(2), 109-134.

Krackhardt, D. (1990). Assessing the political landscape: Structure, cognition, and power in organizations. *Administrative Science Quarterly*, 342-369.

Krackhardt, D., & Hanson, J. R. (1993). Informal networks: The company behind the chart. *Harvard Business Review, 71*(4), 104-111.

Kumpula, J. M., Onnela, J. P., Saramäki, J., Kertész, J., & Kaski, K. (2009). Model of community emergence in weighted social networks. *Computer Physics Communications, 180*(4), 517-522.

Kwon, D., Oh, W., & Jeon, S. (2007). Broken Ties: The Impact of Organizational Restructuring on the Stability of Information-Processing Networks. *J. Manage. Inf. Syst., 24*(1), 201-231. doi: 10.2753/mis0742-1222240106

Kwon, S.-W., & Adler, P. S. (2014). Social Capital: Maturation of a Field of Research. *Academy of Management Review, 39*(4), 412-422. doi: 10.5465/amr.2014.0210

Langton, C. G. (1995). *Artificial life : an overview*. Cambridge, Mass.: MIT Press.

Law, A. M., & Kelton, W. D. (1991). Simulation modeling and analysis, from http://catalog.hathitrust.org/api/volumes/oclc/22111327.html

Lazer, D. (2001). The Co-evolution of individual and network. *Journal of Mathematical Sociology, 25*(1), 69-108.

Lazer, D., & Friedman, A. (2007). The Network Structure of Exploration and Exploitation. *Administrative Science Quarterly, 52*(4), 667-694.

Leana, C. R., & Barry, B. (2000). Stability and Change as Simultaneous Experiences in Organizational Life. *The Academy of Management Review, 25*(4), 753-759.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal, 13*(S1), 111-125. doi: 10.1002/smj.4250131009

Leskovec, J., Backstrom, L., Kumar, R., & Tomkins, A. (2008). *Microscopic evolution of social networks*. Paper presented at the Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining.

Leskovec, J., Lang, K. J., Dasgupta, A., & Mahoney, M. W. (2009). Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics, 6*(1), 29-123.

Levin, D. Z., & Cross, R. (2004). The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer. *Management Science, 50*(11), 1477-1490.

Levine, S. S., & Prietula, M. J. (2011). How Knowledge Transfer Impacts Performance: A Multilevel Model of Benefits and Liabilities. *Organization Science*. doi: 10.1287/orsc.1110.0697

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal, 14*(S2), 95-112. doi: 10.1002/smj.4250141009

Levinthal, D. A., & Warglien, M. (1999). Landscape Design: Designing for Local Action in Complex Worlds. *Organization Science, 10*(3), 342-357.

Lewis, M. W. (2000). Exploring Paradox: Toward a More Comprehensive Guide. *The Academy of Management Review, 25*(4), 760-776.

Macy, M. W., & Willer, R. (2002). From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology, 28*(1), 143-166.

Maguire, S., McKelvey, B., Mirabeau, L., & Oztas, N. (2006). Complex Science and Organization Studies. In S. R. Clegg, C. Hardy, T. B. Lawrence & W. R. Nord (Eds.), *The SAGE handbook of Organization Studies* (2nd ed., pp. 165-214). London: SAGE Publications.

Manson, S. M. (2003). Validation and Verification of Multi-Agent Models for ecosystem management. In M. Janssen (Ed.), *Complexity and Ecosystem Management: The Theory and Practice of Multi-Agent Approaches* (pp. 63-74). Northampton, Massachusetts Edward Elgar Publishers

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science, 2*(1), 71-87.

Marino, S., Hogue, I. B., Ray, C. J., & Kirschner, D. E. (2008). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology, 254*(1), 178-196.

Mariotti, F., & Delbridge, R. (2012). Overcoming Network Overload and Redundancy in Interorganizational Networks: The Roles of Potential and Latent Ties. *Organization Science, 23*(2), 511-528. doi: 10.1287/orsc.1100.0634

Marks, R. E. (2007). Validating simulation models: a general framework and four applied examples. *Computational Economics, 30*(3), 265-290.

Marsden, P. V. (1990). Network Data and Measurement. *Annual Review of Sociology, 16*, 435-463.

Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences, 109*(3), 764-769.

Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General, 137*(3), 422.

Maurer, I., Bartsch, V., & Ebers, M. (2011). The value of intra-organizational social capital: How it fosters knowledge transfer, innovation performance, and growth. *Organization Studies, 32*(2), 157-185.

McCarthy, I. P., & Gordon, B. R. (2011). Achieving contextual ambidexterity in R&D organizations: a management control system approach. *R&D Management, 41*(3), 240-258.

McDermott, R., & Archibald, D. (2010). Harnessing your staffs informal networks. *Harvard Business Review, 88*(3), 82-89.

McDonough, E. F., & Leifer, R. (1983). Using Simultaneous Structures to Cope with Uncertainty. *Academy of Management Journal, 26*(4), 727-735.

McEvily, B., Jaffee, J., & Tortoriello, M. (2012). Not All Bridging Ties Are Equal: Network Imprinting and Firm Growth in the Nashville Legal Industry, 1933–1978. *Organization Science, 23*(2), 547-563.

McFadyen, M. A., & Cannella, A. A. (2004). Social Capital and Knowledge Creation: Diminishing Returns of the Number and Strength of Exchange. *The Academy of Management Journal, 47*(5), 735-746.

McFadyen, M. A., Semadeni, M., & Cannella Jr, A. A. (2009). Value of strong ties to disconnected others: Examining knowledge creation in biomedicine. *Organization Science, 20*(3), 552-564.

McGrath, J. E. (1984). *Groups: Interaction and performance* (Vol. 14): Prentice-Hall Englewood Cliffs, NJ.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*(2), 239-245.

Mehra, A., Kilduff, M., & Brass, D. J. (2001). The Social Networks of High and Low Self-Monitors: Implications for Workplace Performance. *Administrative Science Quarterly, 46*(1), 121-146.

Milborrow, S. (2014). earth: Multivariate Adaptive Regression Spline Models. Retrieved from http://CRAN.R-project.org/package=earth

Miller, K. D., & Lin, S.-J. (2008). Different Truths in Different Worlds. *Organization Science*, orsc.1080.0409. doi: 10.1287/orsc.1080.0409

Miller, K. D., & Lin, S.-J. (2010). Different Truths in Different Worlds. *Organization Science, 21*(1), 97-114. doi: 10.1287/orsc.1080.0409

Miller, K. D., Zhao, M., & Calantone, R. J. (2006). Adding Interpersonal Learning and Tacit Knowledge to March's Exploration-Exploitation Model. [Article]. *Academy of Management Journal, 49*(4), 709-722.

Min, B., Goh, K.-I., & Kim, I.-M. (2009). Waiting time dynamics of priority-queue networks. *Physical Review E, 79*(5), 056110.

Mingers, J. (2004). Can Social Systems be Autopoietic? Bhaskar's and Giddens' Social Theories. *Journal for the Theory of Social Behaviour, 34*(4), 403-427. doi: 10.1111/j.1468-5914.2004.00256.x

Mingers, J. (2006). *Realising systems thinking : knowledge and action in management science*. New York, NY: Springer.

Miritello, G., Lara, R., Cebrian, M., & Moro, E. (2013). Limited communication capacity unveils strategies for human interaction. *Scientific reports, 3*, 1950.

Moody, J., McFarland, D., & Bender-deMoll, S. (2005). Dynamic Network Visualization. *American Journal of Sociology, 110*(4), 1206-1241. doi: doi:10.1086/421509

Moran, P., & Ghoshal, S. (1996). *VALUE CREATION BY FIRMS.* Paper presented at the Academy of Management Proceedings.

Muchnik, L., Pei, S., Parra, L. C., Reis, S. D. S., Andrade Jr, J. S., Havlin, S., & Makse, H. A. (2013). Origins of power-law degree distribution in the heterogeneity of human activity in social networks. [Article]. *Sci. Rep., 3*. doi: 10.1038/srep01783

Myers, R. H. M. D. C. A.-C. C. M. (2009). *Response surface methodology : process and product optimization using designed experiments*. Hoboken, N.J.: Wiley.

Nahapiet, J., & Ghoshal, S. (1998). Social Capital, Intellectual Capital, and the Organizational Advantage. *The Academy of Management Review, 23*(2), 242-266.

Newman, M. E. (2001a). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences, 98*(2), 404-409.

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69*(2), 026113.

Newman, M. E., & Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E, 68*(3), 036122.

Newman, M. E. J. (2001b). Clustering and preferential attachment in growing networks. *Physical Review E, 64*(2), 025102.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45*(2), 167-256.

Nickerson, J. A., & Zenger, T. R. (2002). Being Efficiently Fickle: A Dynamic Theory of Organizational Choice. *Organization Science, 13*(5), 547-566.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science, 5*(1), 14-37.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company : how Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

North, M., Collier, N., Ozik, J., Tatara, E., Macal, C., Bragen, M., & Sydelko, P. (2013). Complex adaptive systems modeling with Repast Simphony. *Complex Adaptive Systems Modeling, 1*(1), 3.

O'Dell, C., & Grayson, C. J. (1999). Knowledge transfer: discover your value proposition. *Strategy Leadership Strategy & Leadership, 27*(2), 10-15.

O'Reilly, C., & Tushman, M. (2013). Organizational ambidexterity: Past, present and future. *The Academy of Management Perspectives*, amp. 2013.0025.

O'Reilly III, C. A., & Tushman, M. L. (2011). Organizational ambidexterity in action: How managers explore and exploit. *California Management Review, 53*(4), 5-22.

O'Reilly III, C. A., & Tushman, M. (2011). Organizational ambidexterity in action: How managers explore and exploit. *Calif. Manage. Rev. California Management Review, 53*(4), 5-22.

O'Reilly III, C. A., & Tushman, M. L. (2008). Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. *Research in Organizational Behavior, 28*(0), 185-206. doi: 10.1016/j.riob.2008.06.002

Obstfeld, D. (2005). Social Networks, the Tertius Iungens Orientation, and Involvement in Innovation. *Administrative Science Quarterly, 50*(1), 100-130.

Oh, H. (2004). Group social capital and group effectiveness: The role of informal socializing ties. *Academy of Management Review, 47*(6), 860-865.

Oh, H., & Kilduff, M. (2008). The Ripple Effect of Personality on Social Structure: Self-Monitoring Origins of Network Brokerage. *Journal of Applied Psychology, 93*(5), 1155-1164. doi: 10.1037/0021-9010.93.5.1155

Oh, H., Labianca, G., & Chung, M.-H. (2006). A Multilevel Model of Group Social Capital. *Academy of Management Review, 31*(3), 569-582.

Oliveira, J., & Vazquez, A. (2009). Impact of interactions on human dynamics. *Physica A: Statistical Mechanics and its Applications, 388*(2), 187-192.

Orlikowski, W. J. (1992). *Learning from Notes: organizational issues in groupware implementation*. Paper presented at the Proceedings of the 1992 ACM conference on Computer-supported cooperative work, Toronto, Ontario, Canada.

Perlow, L. A. (1999). The Time Famine: Toward a Sociology of Work Time. *Administrative Science Quarterly, 44*(1), 57-81.

Perra, N., Gonçalves, B., Pastor-Satorras, R., & Vespignani, A. (2012). Activity driven modeling of time varying networks. *Sci. Rep., 2*.

Perry-Smith, J. E. (2006). Social Yet Creative: The Role of Social Relationships in Facilitating Individual Creativity. *The Academy of Management Journal, 49*(1), 85-101.

Perry-Smith, J. E., & Shalley, C. E. (2003). The Social Side of Creativity: A Static and Dynamic Social Network Perspective. *The Academy of Management Review, 28*(1), 89-106.

Piderit, S. K. (2000). Rethinking Resistance and Recognizing Ambivalence: A Multidimensional View of Attitudes Toward an Organizational Change. *Academy of Management Review, 25*(4), 783-794. doi: 10.5465/amr.2000.3707722

Prigogine, I., & Stengers, I. (1984). *Order out of chaos : man's new dialogue with nature*. Toronto; New York, N.Y.: Bantam Books.

Putnam, R. D. (2000). *Bowling alone : the collapse and revival of American community*. New York: Simon & Schuster.

R Core Team. (2014). R: A Language and Environment for Statistical Computing. Vienna, Austria.

Raeder, T., Lizardo, O., Hachen, D., & Chawla, N. V. (2011). Predictors of short-term decay of cell phone contacts in a large scale communication network. *Social Networks, 33*(4), 245-257. doi: http://dx.doi.org/10.1016/j.socnet.2011.07.002

Raisch, S., & Birkinshaw, J. (2008). Organizational Ambidexterity: Antecedents, Outcomes, and Moderators. *Journal of Management, 34*(3), 375-409. doi: 10.1177/0149206308316058

Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. (2009). Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance. *Organization Science, 20*(4), 685-695. doi: 10.1287/orsc.1090.0428

Rapoport, A. (1953). Spread of information through a population with socio-structural bias: I. Assumption of transitivity. *Bulletin of Mathematical Biology, 15*(4), 523-533.

Rapoport, A. (1957). Contribution to the theory of random and biased nets. *Bulletin of Mathematical Biology, 19*(4), 257-277.

Raub, W., & Weesie, J. (1990). Reputation and Efficiency in Social Interactions: An Example of Network Effects. *The American Journal of Sociology, 96*(3), 626-654.

Read, M., Andrews, P. S., Timmis, J., & Kumar, V. (2012). Techniques for grounding agent-based simulations in the real domain: a case study in experimental autoimmune encephalomyelitis. *Mathematical and Computer Modelling of Dynamical Systems, 18*(1), 67-86.

Reagans, R., & McEvily, B. (2003). Network Structure and Knowledge Transfer: The Effects of Cohesion and Range. *Administrative Science Quarterly, 48*(2), 240-267.

Reagans, R., & McEvily, B. (2008). Contradictory or Compatible? Reconsidering the "Trade-off" between Brokerage and Closure on Knowledge Sharing. In J. A. C. Baum & T. J. Rowley (Eds.), *Network strategy (Advances in Strategic Management)* (Vol. 25, pp. 275-313). Bingley, UK: Emerald JAI.

Reagans, R. E., & Zuckerman, E. W. (2008). Why knowledge does not equal power: the network redundancy trade-off. *Industrial and Corporate Change, 17*(5), 903-944. doi: 10.1093/icc/dtn036

Richiardi, M., Leombruni, R., Saam, N., & Sonnessa, M. (2006). A Common Protocol for Agent-Based Social Simulation. *Journal of Artificial Societies & Social Simulation, 9*(1).

Rivkin, J. W., & Siggelkow, N. (2003). Balancing Search and Stability: Interdependencies among Elements Organizational Design. *Management Science, 49*(3), 290-311.

Rodan, S., & Galunic, C. (2004). More than network structure: how knowledge heterogeneity influences managerial performance and innovativeness. *Strategic Management Journal, 25*(6), 541-562.

Rybski, D., Buldyrev, S. V., Havlin, S., Liljeros, F., & Makse, H. A. (2010). Communication activity: temporal correlations, clustering, and growth. *arXiv preprint arXiv:1002.0216*.

Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical science, 4*(4), 409-423.

Saltelli, A., & Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software, 25*(12), 1508-1517. doi: http://dx.doi.org/10.1016/j.envsoft.2010.04.012

Sasovova, Z., Mehra, A., Borgatti, S. P., & Schippers, M. C. (2010). Network Churn: The Effects of Self-Monitoring Personality on Brokerage Dynamics. *Administrative Science Quarterly, 55*(4), 639-670. doi: 10.2189/asqu.2010.55.4.639

Sastry, M. A. (1997). Problems and Paradoxes in a Model of Punctuated Organizational Change. *Administrative Science Quarterly, 42*(2), 237-275.

Scholl, H. J. (2014). Mobile ICTs in Government Field Operations: A Socio-Technical Innovation Project. *International Journal of Electronic Government Research (IJEGR), 10*(2), 60-81.

Schumpeter, J. (1942). *Capitalism, Socialism, and Democracy*. New York: Harper & Bros.

Shaw, M. E., Robbin, R., & Belser, J. R. (1981). *Group dynamics: the psychology of small group behavior*: McGraw-Hill.

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22*(8), 888-905.

Siggelkow, N., & Levinthal, D. A. (2003). Temporarily Divide to Conquer: Centralized, Decentralized, and Reintegrated Organizational Approaches to Exploration and Adaptation. *Organization Science, 14*(6), 650-669. doi: 10.1287/orsc.14.6.650.24840

Simon, H. A. (1957). A Behavioral Model of Rational Choice *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley.

Simon, H. A. (1991). Bounded Rationality and Organizational Learning. *Organization Science, 2*(1), 125-134.

Simpson, T. W., Poplinski, J., Koch, P. N., & Allen, J. K. (2001). Metamodels for computer-based engineering design: survey and recommendations. *Engineering with computers, 17*(2), 129-150.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory-building in social psychology. *Personality and social psychology review, 11*(1), 87-104.

Smith, W. K., & Lewis, M. W. (2011). Toward a Theory of Paradox: A Dynamic Equiibrium Model of Organizing. *Academy of Management Review, 36*(2), 381-403.

Smith, W. K., & Tushman, M. L. (2005). Managing Strategic Contradictions: A Top Management Model for Managing Innovation Streams. *Organization Science, 16*(5), 522-536.

Snijders, T. A. B. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology, 21*(1-2), 149-172.

Snijders, T. A. B. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology, 31*, 361-395.

Snijders, T. A. B., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New Specifications for Exponential Random Graph Models. *Sociological Methodology, 36*, 99-153.

Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks, 32*(1), 44-60. doi: 10.1016/j.socnet.2009.02.004

Snyder, M., & Gangestad, S. (1982). Choosing social situations: Two investigations of self-monitoring processes. *Journal of Personality and Social Psychology*, *43*(1), 123-135. doi: 10.1037/0022-3514.43.1.123

Soda, G., Usai, A., & Zaheer, A. (2004). Network Memory: The Influence of Past and Current Networks on Performance. *The Academy of Management Journal, 47*(6), 893-906.

Song, C., Wang, D., & Barabasi, A.-L. (2012). Joint Scaling Theory of Human Dynamics and Network Science.

Sosa, M. E. (2011). Where do creative interactions come from? The role of tie content and social networks. *Organization Science, 22*(1), 1-21.

Spender, J. C. (1996). Making Knowledge the Basis of a Dynamic Theory of the Firm. *Strategic Management Journal*, *17*, 45-62.

Srivastava, S. B., & Gulati, R. (2014). Bringing Agency Back into Network Research: Constrained Agency and Network Action *Contemporary Perspectives on Organizational Social Networks* (pp. 73-93).

Stehlé, J., Barrat, A., & Bianconi, G. (2010). Dynamical and bursty interactions in social networks. *Physical Review E, 81*(3), 035101.

Steinberg, D. (2013). Reuse MARS Regression Spline Basis Functions in a New Dataset Retrieved 10/15/2014, from http://1.salford-systems.com/blog/bid/283558/Reuse-MARS-Regression-Spline-Basis-Functions-in-a-New-Dataset

Stevenson, W. B., & Greenberg, D. (2000). Agency and Social Networks: Strategies of Action in a Social Structure of Position, Opposition, and Opportunity. *Administrative Science Quarterly, 45*(4), 651-678.

Storlie, C. B., & Helton, J. C. (2008). Multiple predictor smoothing methods for sensitivity analysis: Description of techniques. *Reliability Engineering & System Safety, 93*(1), 28-54.

Strogatz, S. H. (2001). Exploring complex networks. [10.1038/35065725]. *Nature, 410*(6825), 268-276.

Taber, C. S., & Timpone, R. J. (1996). *Computational modeling*. Thousand Oaks, CA: Sage Publications.

Timmis, J., Read, M. N., Andrews, P. S., & Kumar, V. (2011). Techniques for Grounding Agent-Based Simulations in the Real Domain. *a case study in Experimental Autoimmune Encephalomyelisis, 18*(Article), 296-302. doi: 10.1080/13873954.2011.601419

Tiwana, A. (2008). Do bridging ties complement strong ties? An empirical examination of alliance ambidexterity. *Strategic Management Journal, 29*(3), 251-272.

Toivonen, R., Kovanen, L., Kivelä, M., Onnela, J.-P., Saramäki, J., & Kaski, K. (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks, 31*(4), 240-254. doi: http://dx.doi.org/10.1016/j.socnet.2009.06.004

Troitzsch, K. G. (2014). Simulation Experiments and Significance Tests *Artificial Economics and Self Organization* (pp. 17-28): Springer.

Tsai, W., & Ghoshal, S. (1998). Social Capital and Value Creation: The Role of Intrafirm Networks. *The Academy of Management Journal, 41*(4), 464-476.

Tsang, E. W. K., & Zahra, S. A. (2008). Organizational unlearning. *Human Relations, 61*(10), 1435-1462. doi: 10.1177/0018726708095710

Tsoukas, H. (1996). The firm as a distributed knowledge system: a constructionist approach. *Strategic management journal, 17*(S2), 11-25.

Turner, N., Swart, J., & Maylor, H. (2013). Mechanisms for managing ambidexterity: a review and research agenda. *International Journal of Management Reviews, 15*(3), 317-332.

Tushman, M., & O'Reilly III, C. A. (1996). Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review, 38*(4), 8-30.

Vargha, A., & Delaney, H. D. (2000). A critique and improvement of the CL common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics, 25*(2), 101-132.

Wasserman, S., & Faust, K. (1994). *Social network analysis methods and applications*. Cambridge; New York: Cambridge University Press.

Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs andp. *Psychometrika, 61*(3), 401-425.

Watts, A. (2001). A Dynamic Model of Network Formation. *Games and Economic Behavior, 34*(2), 331-341.

Watts, D. J. (1999). Networks, Dynamics, and the Small-World Phenomenon. *The American Journal of Sociology, 105*(2), 493-527.

Watts, D. J. (2003). *Six degrees : the science of a connected age*. New York: Norton.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature, 393*(6684), 440-442.

Wei, J., Zheng, W., & Zhang, M. (2011). Social capital and knowledge transfer: A multi-level analysis. *Human Relations*. doi: 10.1177/0018726711417025

Weisstein, E. W. (Ed.) (2013). MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/RandomNumber.html.

White, J. W., Rassweiler, A., Samhouri, J. F., Stier, A. C., & White, C. (2014). Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos, 123*(4), 385-388.

Wolfram, S. (1994). *Cellular automata and complexity: collected papers* (Vol. 1): Addison-Wesley Reading.

Wu, Y., Zhou, C., Xiao, J., Kurths, J., & Schellnhuber, H. J. (2010). Evidence for a bimodal distribution in human communication. *Proceedings of the National Academy of Sciences, 107*(44), 18803-18808.

Zander, A. (1950). Resistance to change—its analysis and prevention. *Advanced Management Journal*.

Zander, U., & Kogut, B. (1995). Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities: An Empirical Test. *Organization Science, 6*(1), 76-92.

Zhao, K., & Bianconi, G. (2011). Social interactions model and adaptability of human behavior. *Frontiers in physiology, 2*.

Zollo, M., & Winter, S. G. (2002). Deliberate learning and the evolution of dynamic capabilities. *Organization science, 13*(3), 339-351.