ABSTRACT

Information systems development (ISD) is essentially a search process by which the ISD team seeks to find an optimal system configuration among numerous feasible configurations that produces the greatest business value. As information systems are embodiments of application domain knowledge and technical knowledge, ISD requires both in order to be effective. The business unit is ultimately responsible for making business design choices whereas the IS unit is largely responsible for making technical design choices. Complexity in ISD arises when these design choices are interdependent so that the contribution of one design choice to business value may depend on other design choices. We argue that knowledge overlaps between business and IS play an important role in the ISD process. However, little research has examined the complex and nuanced effects of knowledge overlaps on ISD performance under different conditions of design choice interdependencies. Using an NK fitness landscapes model of ISD, this research investigates how knowledge overlaps influence ISD performance (1) when the level of interdependencies among design choices varies, (2) for different distributions of within-unit and between-unit interdependencies, and (3) when between-unit interdependencies are balanced or skewed. We report the results of a simulation study and discuss their implications and insights. We present propositions drawn from the simulation results.

Keywords: information systems development (ISD), knowledge overlap, knowledge interdependence, complexity, complex adaptive systems, NK fitness landscape model, simulation
KNOWLEDGE OVERLAP IN INFORMATION SYSTEMS DEVELOPMENT: AN NK FITNESS LANDSCAPES MODEL

INTRODUCTION

It is no secret that information systems development (ISD) has not demonstrated a high success rate to date (Standish Group 2009). While both application domain knowledge and technical knowledge play important roles in ensuring successful ISD (Khatri et al. 2006), one of the important factors contributing to the low success rate is the knowledge gap between business users and systems development personnel (Lyytinen and Robey 1999, Nelson and Cooprider 1996). The IS unit possesses technical knowledge but tends to have relatively less application domain knowledge, and the business unit possesses application domain knowledge but tends to possess relatively less technical knowledge (Bassellier and Benbasat 2001, 2004). Although both technical knowledge and application domain knowledge are required and need to be integrated for effective ISD (Faraj and Sproull 2000, Khatri et al. 2006, Rus and Lindvall 2002), the distribution of knowledge across the organizational units involved in the ISD process makes the integration of necessary knowledge quite difficult (Tiwana 2004).

Notwithstanding, the division of knowledge to specialized units is not always so clear-cut (Takeishi 2002). It has become increasingly common to find business units to possess some technical knowledge and IS units to have ample knowledge of application domains, circumstances referred to as knowledge overlaps (Tiwana 2004). Knowledge overlap is important for organizational work because individuals need to have a certain level of common ground in their individual knowledge bases to effectively communicate and coordinate collective action (Alavi and Leidner 2001, Cramton 2001, Espinosa et al. 2007b). Knowledge overlap between the business and IS units has been found to be an important factor in the ISD process.
Prior research has argued that knowledge overlap between the business unit and the IS unit affects various aspects of ISD performance (Bassellier and Benbasat 2004, Bassellier et al. 2003, Nelson and Cooprider 1996, Reich and Benbasat 2000, Tiwana 2009). For example, knowledge overlap was found to strengthen business-IT alignment (Reich and Benbasat 2000), solidify business-IT partnerships (Nelson and Cooprider 1996), increase IT use (Boynton et al. 1994), increase outsourcing performance (Tiwana 2004) and enable more effective ISD project decision control and management (Tiwana 2009). The application domain knowledge of IT professionals affects the computer program comprehension process (Shaft and Vessey 1995, 1998) and the performance of understanding ISD models such as entity-relationship diagrams (Burton-Jones and Weber 1999) and facilitates team learning during ISD (Walz et al. 1993). It enables them to participate in important organizational decision-making processes (Feeny and Willcocks 1998) since knowledge of IT-business integration enables them to identify synergies between IT and business activities and to understand how various business and technical parts fit together (Bassellier and Benbasat 2004). Conversely, the lack of business knowledge within a software team leads to errors and mistakes in ISD (Curtis et al. 1988). The technical knowledge of business managers is positively associated with their propensity to innovate using IT (Nambisan et al. 1999) and also with their intentions to champion IT (Bassellier et al. 2003). Line managers with IT education and training are more likely to assume a leadership role in IT work (Rockart et al. 1996).

While the general benefits of knowledge overlap in ISD seem intuitively obvious, extant research has yet to fully uncover the complex and nuanced effects of the different degrees, distributions, and configurations of knowledge overlap on ISD performance. For example, how
much knowledge overlap is necessary or sufficient for effective ISD? When is knowledge overlap more or less beneficial to ISD performance? When is its impact the strongest? How do degrees and patterns of knowledge interdependence interplay with knowledge overlap in affecting ISD performance? Do projects require an equal amount of knowledge overlap across the business and IS units to be effective? These are some of the important questions for which the current literature does not provide clear answers. The main contribution of this research is to address these questions and fill the gap in the literature, thus deepening our understanding of the role of knowledge overlap in ISD. Since knowledge overlap should contribute to ISD performance through improved decisions about the design of the IS (Alavi and Leidner 2001, Bassellier and Benbasat 2004), decision structure and complexity are important factors that may influence the impact of knowledge overlap. Accordingly, we focus our analyses on how knowledge overlap influences ISD performance under different degrees and patterns of knowledge interdependence, which are important determinants of problem structure and complexity (Wood 1986, Xia and Lee 2005).

We adopt a complexity theory perspective and employ a complex adaptive systems approach using agent-based simulations to address the research questions. While complexity theory has traditionally been within the purview of the physical and biological sciences, it has recently gained much attention from management and organization scholars. Complex systems arise from the bottom-up emergent behavior of a population of interacting and co-evolving agents that act on limited and local interactions (Holland 1995). Such emergent behaviors often exhibit non-linearity and are as a result very difficult to predict and manage (Anderson 1999,)

---

1 See recent special issues of Organization Science (Anderson et al. 1999) and Management Science (Amaral and Uzzi 2007) on the application of complexity theory and complex systems approaches to management research.
ISD represents one such complex organizational activity (Benbya and McKelvey 2006). ISD requires the integration of knowledge and expertise from distinct disciplinary areas (i.e., application and technology domains), which are typically distributed across different groups of agents (i.e., business users / managers vs. IT professionals). Furthermore, the ISD process unfolds in an environment of complex and dynamic user and organizational requirements (Xia and Lee 2004). Such complexities inherent in the ISD process make complexity theory a fitting perspective (Benbya and McKelvey 2006).²

One of the key contributions of complexity theory to management has been to show that analytical frameworks and tools from the complexity sciences (e.g., complex adaptive systems, fitness landscapes, computational modeling, self-organization / stigmergic systems etc.) can be brought to bear on a diverse set of complex organizational problems (Anderson et al. 1999). Complexity theory-based research methods allow researchers to examine complex theoretical relationships that analytical models and/or field studies cannot fully uncover (Miller and Page 2007). Analytical models rely on drastically simplified representations of organizations for analytical tractability and, as a result, cannot faithfully represent the richness of actual organizations. Field studies of actual organizations enjoy the luxury of realism, but it is oftentimes difficult, if not impossible, to observe, measure or manipulate the theoretical constructs of interest to fully explore the complex theoretical relationships (Amaral and Uzzi 2007). The complex adaptive systems approach using agent-based simulations enables us to incorporate a greater number of interdependent elements in a formal model than is possible with a closed-form analytical approach (Davis et al. 2007). In addition, the simulation methodology

² However, despite the apparent potential benefits from a complexity theory perspective in the study of ISD, the IS discipline has yet to fully embrace the theoretical perspectives and analytical tools of complexity theory (Merali and McKelvey 2006).
also allows the researcher to freely manipulate the theoretical constructs of interest to help acquire non-intuitive and/or nuanced theoretical insights through various combinations of experimental conditions, which is not possible with field studies (Amaral and Uzzi 2007, Davis et al. 2007). Among the various simulation approaches (e.g., systems dynamics, genetic algorithms, cellular automata, etc.) for studying complex adaptive systems, the NK fitness landscapes model (Kauffman 1993) is most appropriate when organizational adaptation can be framed as problem solving or search (Davis et al. 2007). Since ISD is essentially a problem solving process whereby the ISD team seeks to derive an information systems design that produces gains in organizational effectiveness and/or efficiency (Cerveny et al. 1990), the NK fitness landscape model is chosen for this research. Our study extends the basic NK fitness landscape model to fit the ISD context by incorporating the notion of incremental and iterative search in a setting where knowledge is partitioned and possibly overlapping across decision-making units.

A MODEL OF INFORMATION SYSTEMS DEVELOPMENT AS SEARCH

The NK fitness landscapes model, originally developed by Kauffman (1993), provides a simple, yet powerful analytical framework to study organizational problem solving from the perspective of adaptive search. While the model itself was developed for the study of evolutionary biology, it has been extensively applied to management research to study such topics as organizational adaptation (Gavetti and Levinthal 2000, Levinthal 1997), organizational structure and design (Rivkin and Siggelkow 2003, Siggelkow and Rivkin 2005), and organizational innovation (Almirall and Casadesus-Masanell 2010, Ethiraj and Levinthal 2004, Ethiraj et al. 2008). Although, the NK modeling framework has recently gained broad
acceptance within the strategic management and organization sciences literatures, it is still relatively unknown within the IS literature. Accordingly, we briefly review the fundamental concepts of the NK fitness landscapes model before proceeding with the details of our model, which extends the basic NK landscapes model to the ISD context. The interested reader is directed to Davis et al. (2007) for a brief overview of the NK modeling approach. More technical details can be found in Kauffman (1989, 1993).

The NK Fitness Landscapes Model

In the NK fitness landscapes model, a complex adaptive system (e.g., an organization, and ISD team) is conceptualized as operating in an environment (e.g., a strategy, product, ISD etc.) involving \( N \) decision variables and \( K \) epistatic interactions (or interdependencies) among these decision variables. Each configuration of the set of decision variables is associated with a fitness value, which can be interpreted as performance if that particular configuration is implemented. The system uses search strategies (e.g., incremental/local hill-climbing, trial and error search, imitation of peers, long jumps, etc.) to navigate within the fitness landscape to find positions of greatest fitness (i.e., best strategy, best product design, best IS design etc.) These search strategies are heuristics/routines the system uses to configure and reconfigure the values of the \( N \) decision variables. The two parameters, \( N \) and \( K \), of the NK model allow the modeler to create “tunable” fitness landscapes of the decision environment of varying degrees of complexity on which to test the efficacy of various search strategies.

Formally, a configuration is represented by an \( N \)-element vector \( \mathbf{d} = <d_1, \ldots, d_N> \), where \( d_i \) can take the value of 0 or 1.\(^3\) Each decision \( d_i \) makes a contribution to the overall fitness of \( \mathbf{d} \),

\(^3\) The NK model has been shown to be robust to this binary simplification (Kauffman 1989). The model can be extended to an arbitrary finite number of possible values without altering the qualitative properties of the model.
The value contribution $c_i$ of each decision depends not only on the choice made concerning that decision ($d_i = 0$ or $1$) but also on choices made regarding $K$ other decisions that interact with the focal decision: $c_i = c_i(d_i|K$ other $d_j$’s). The interdependencies among decision variables are captured in an $N \times N$ influence matrix ($\text{INF}$), where $\text{INF}_{ij} = 1$ if (column) decision variable $j$ influences the value contribution of (row) decision variable $i$, or 0 otherwise. With $K$ interdependencies among decision variables, there will be $K+1$ 1’s in each row of the influence matrix – 1 for the value contribution of the focal decision variable $(d_i)$ itself and $K$ interdependencies with other decision variables $(d_j$’s). Figure 1 shows an example of an influence matrix for $N = 10$ and $K = 4$.

With the patterns of interdependencies set, we may stochastically assign fitness values to each of the $2^N$ possible configurations (Kauffman 1993). For each possible realization of $d_i$ and the $K$ relevant other $d_j$’s, $c_i$ is drawn randomly from a uniform $U[0, 1]$ distribution and the overall fitness value for contribution configuration $d$ is the average over $N$ decision level contributions.

$$F(d) = \frac{1}{N} \sum_{i=1}^{N} c_i(d_i|K$ other $d_j$’s)$$

This procedure for stochastically assigning fitness values allows the modeler to “tune” landscapes to varying degrees of ruggedness by defining the patterns of decision interdependencies based on the parameter $K$. When there is little interaction among decision variables (i.e., low $K$), the resulting fitness landscape is “smooth” (see Figure 2a). Conversely, when the decision variables become highly interdependent (i.e., high $K$), the resulting fitness landscape is “rugged” and multi-peaked (see Figure 2b). The ruggedness of the fitness landscape determines the efficacy of the various search strategies. For example, in an environment of low complexity (i.e., within a smooth fitness landscape, or low $K$), incremental hill-climbing
strategies are very effective and the system will eventually find the position with the globally optimal fitness level. However, in a highly complex environment (i.e., within a rugged fitness landscape, or high $K$), search for a high position is profoundly more difficult as incremental local search strategies are prone to lead the system to get stuck in a basin of attraction (also called a “sticking point”) – a local peak/optimum, but not necessarily the global optimum (Kauffman 1989) – from which it is difficult for the system to escape since changes to any single configuration choice will result in a fitness level that is lower than the status quo.

The $NK$ fitness landscapes model is a theory of the statistical properties of complex adaptive systems (Kauffman 1993). The primary application of the model is to analyze statistically typical behaviors of adaptive organizations. Once the structure of the interdependencies among decision variables have been set, fitness landscapes can be stochastically created onto which agents can be seeded. The agents use adaptive behaviors to seek configurations of higher fitness. They don’t have a bird’s-eye view of the entire fitness landscape and thus resort to local and experiential search. By simulating many different agents’ behaviors, one can analyze the statistical properties of the adaptation (e.g., average fitness level at equilibrium, average time to reach equilibrium, etc.).

While early management research that used the $NK$ modeling framework focused on analyzing and comparing the efficacy of various search strategies, recent efforts have shifted the focus of attention to issues related to organizational design. For example, Siggelkow and Rivkin (2005) analyze the effects of departmental information processing power, incentive structures, cross-departmental coordination and information flow on the speed and diversity of organizational search. The $NK$ model of ISD, which we will present next, also extends this line
of research by investigating the impact of various degrees and patterns of knowledge overlap between the business unit and the IS unit, and the effect of the extent and patterns of interdependencies between the decisions of the business unit and those of the IS unit on ISD performance.

**Information Systems Development as Search**

As discussed earlier, ISD is essentially a problem solving process whereby the ISD team searches for an IS configuration that delivers the greatest value to the organization. ISD requires the combination and integration of both technical knowledge and application domain knowledge (Benbya and McKelvey 2006, Kim and Lerch 1997, Robillard 1999, Rus and Lindvall 2002, Tiwana and McLean 2005). As a result, ISD must involve not only technical specialists from the IS unit but also domain experts from business units (Tiwana 2009). In the ISD process, a group of project team members are responsible for determining user requirements whereas another group of project team members are responsible for technical design and implementation. Task partitioning arises because the two groups tend to possess different sets of knowledge and competencies (von Hippel 1990) – the former group consists mainly of users and business analysts who have application domain knowledge, whereas the latter group consists mainly of system analysts and developers who possess technical expertise.

However, the need for combining and integrating business knowledge and technical knowledge (Faraj and Sproull 2000) complicates the ISD process because knowledge is not evenly distributed across the organization (Tiwana 2004). The IS unit tends to have relatively less application domain knowledge, and the business unit tends to possess relatively less technical knowledge (Bassellier and Benbasat 2001, 2004). Such knowledge partitioning can be problematic because it creates knowledge barriers that inhibit transfer of knowledge across
organizational units (DeSouza 2003, Ko et al. 2005, Slaughter and Kirsch 2006, Szulanski 1996). To cope with such knowledge difficulties, systems development methodologies prescribe an incremental and iterative process whereby the integration and combination of knowledge across business and technical domains occurs seamlessly across increments and iterations (Beck and Andres 2005, Lee and Xia 2010). Conventionally, business configuration choices are made during the user requirements determination phase while technology configuration choices are made during the systems design and development phases (Tiwana 2003).

Difficulties in ISD also arise due to the complex and ill-structured nature of the design task (Boland 1978, Curtis et al. 1988, Goel and Pirolli 1992). The complexity of a task such as ISD increases when multiple components exist, when these components are interdependent, and when they change over time (Ribbers and Schoo 2002, Simon 1962, 1973, Speier et al. 2003, Wood 1986, Xia and Lee 2005). The problems in building complex systems often arise in the interfaces between hardware, software, and human components (Curtis et al. 1988, Leveson 1997). Interdependencies among these elements make it difficult to predict the project’s process and outcome as the causality between inputs and outputs becomes ambiguous (Campbell 1988, Wood 1986). Knowledge interdependence in ISD affects how the outcome of a business or technology configuration choice is influenced by other configuration choices.

The complexity of ISD is further exacerbated by the interaction of the above two factors – the division of ISD knowledge across the business and IS units and the interdependencies that may exist among ISD design elements. In the ISD process, business requirements and technological choices often interact with one another and these interdependencies may exhibit different patterns and distributions depending on the type of ISD project and the business application problems the project is proposed to solve. For example, although for some types of
information systems, interdependencies may exist mostly within business or technology domains, most complex information systems may also exhibit a great level of interdependencies across knowledge domains. Even though the partitioned organizational structure may be effective for less complex ISD projects where little interdependencies exist across knowledge domains, it is not clear how limiting the organizational structure may be for more complex ISD projects where interdependencies across knowledge domains are prevalent as organizational units must make configuration choices without fully appreciating the influence of the design choices of the other decision making unit. However, the division of knowledge to specialized units is not always so clear-cut (Takeishi 2002), and in reality, knowledge overlap across business and IS units exists for many real-world ISD projects (Tiwana 2004). Knowledge overlap may help ISD teams overcome the aforementioned difficulties arising from knowledge partitioning and interdependence. The focus of this paper is to understand how knowledge overlap influences the ISD process and outcomes.

In the next section, we develop an NK fitness landscapes model that incorporates the following elements to fit the ISD context. First, the overall problem space is partitioned into separate knowledge domains (i.e., business vs. technical knowledge). Second, the ISD team is comprised of two distinct decision making units (i.e., business unit vs. IS unit) with specialized knowledge. ISD is an incremental and iterative process through which each decision-making unit chooses a configuration for the design elements within its knowledge domain that leads to increased business value. Third, the complexity of the ISD project is defined by the amount and patterns of interdependencies among design elements. Finally, we allow for the possibility of knowledge overlaps such that each unit may have knowledge of the other domain.
An NK Fitness Landscapes Model of Information Systems Development

To capture the essence of the information systems development (ISD) process in its realistic yet parsimonious form, we model an ISD project as consisting of \( N \) systems design choices (i.e., features or decision variables), where each design choice can take one of two values – 0 or 1. A systems design choice could represent an application domain related decision (e.g., make or buy intermediate products) or a technology related decision (e.g., distribute or centralize the database). Systems design choices can interact in the sense that the value contribution of one variable may depend on the configuration of other design choices. Such interdependencies can exist within the business or IS domains (e.g., the value contribution of the *business* decision to make or buy intermediate products may depend on a related *business* decision to acquire or develop complementary skills) and/or across domains (e.g., the value contribution of the *technical* decision to centralize or decentralize the enterprise data may depend on a related *business* decision to centralize management authority at the headquarters or empower regional branch managers by providing them with P&L responsibilities).

With this setup, the ISD project can be represented as an \( N \) element vector of design configurations: \( \mathbf{d} = <d_1, d_2, \ldots, d_N> \), which as a result can take on \( 2^N \) possible configurations. Each ISD project configuration is associated with a fitness value, which can be interpreted as performance if that particular configuration is implemented. The objective of the ISD process is to find the highest value point in the fitness landscape of ISD project configurations.

**Organizational Structure: Allocation of Decisions**

As discussed previously, in the ISD project some design choices determine how business processes and business models are configured and embodied in the system, whereas others determine how hardware, software, and network technologies are configured and implemented.
In ISD projects, these design choices are made by different units due to specialized knowledge (von Hippel 1990). To incorporate the allocation of decisions to different organizational units, we model the organization as composed of two units – a business unit and an IS unit. Since the ISD project is characterized as consisting of $N$ design choices, each unit is responsible for half of the design choices.\(^4\) Formally, assuming that $N$ is even, the ISD project configuration vector $\mathbf{d}$ is partitioned into two subsets $\langle \mathbf{d}_{\text{bus}}, \mathbf{d}_{\text{IS}} \rangle$, where $\mathbf{d}_{\text{bus}} = \langle d_1, d_2, \ldots, d_{N/2} \rangle$ and $\mathbf{d}_{\text{IS}} = \langle d_{N/2+1}, d_{N/2+2}, \ldots, d_N \rangle$.

**Knowledge Interdependence and ISD Project Complexity**

As noted earlier, the extent of interdependencies among design variables (i.e., parameter $K$) determines the overall ruggedness (i.e., complexity) of the fitness landscape, and as a result the difficulty of searching for the design configuration that yields the greatest performance. In the context of ISD, the notion of design choice interdependence becomes more complex as design choices are related to multiple knowledge domains. For example, interdependence exists within domain if the business value of a business (technology) design choice depends on other business (technology) design choices. Conversely, interdependence exists across domains if the business value of a business (technology) design choice depends on other technology (business) design choices.

As discussed above, the patterns of interdependencies among design variables can be defined and represented in an influence matrix. The influence matrix allows us to model the

---

\(^4\) For simplicity, we assume that the number of decision variables is equally distributed among the two units – $N/2$ design choices for each of the two units. In reality, some projects may involve a greater number of design choices for either of the two units. Analyzing the impact of the distribution of decisions is beyond the scope of this paper.
complex nature of the within-unit and between-unit interdependencies. Figure 3 shows some examples of influence matrices for various patterns of interdependencies.⁵

[INSERT FIGURE 3 HERE]

**Specialized Knowledge**

Each of the two units involved in the ISD project has specialized knowledge. The IS unit has technical knowledge, and the business unit has application domain knowledge. First, let’s consider the case of fully specialized knowledge where the business unit does not know anything in the technology domain, and the IS unit does not know anything about the application domain. Later, we will relax this assumption to allow knowledge overlap across units.

Although the ISD project involves setting the configurations for N design choices, since each unit only has knowledge of N/2 of the design choice elements, each unit can only assess the performance implications of the design choice elements for which they are knowledgeable. In other words, although each unit can assess the performance implications of their local design choices (i.e., which configuration of business design choices have the highest fitness value given some unknown, not fully understood, configuration of the technology design choices, and vice versa), its assessment of the full implications of their design choices will be imperfect. We model the imperfect assessment due to knowledge specialization using Gavetti and Levinthal’s (2000) model of simplified cognitive representations. The intuition underlying the model of simplified cognitive representations is that cognition is based on an actor’s representation of the decision context. Actors are boundedly rational, and as a result, although their representations of

---

⁵ The patterns of design choice interdependencies influence the complexity of the ISD project not only by determining the ruggedness of the ISD project fitness landscape but also by impacting the accuracy of fitness estimates given the allocation of decisions between units and distribution of specialized knowledge across units (Almirall and Casadesus-Masanell 2010, Gavetti and Levinthal 2000, McKelvey 1999, Rivkin and Siggelkow 2007).
the decision context are generally consistent with the actual landscape, they are only simplified and crude abstractions of the actual landscape. In the NK fitness landscapes framework, this can be achieved by reducing the dimensionality of the landscape, which is used as the basis for the actor’s simplified cognitive representation. Formally, although the structure of the landscape is defined by $N$ design choices (or decision variables), an actor can be conceptualized as having a boundedly rational representation consisting of $N'$ dimensions, where $N' < N$. In other words, the actor is assumed to only be mindful of a subset of the design choices that govern the fitness structure of the problem domain. The fitness contribution of a design choice based on an actor’s simplified cognitive representation can be estimated as the average of the fitness contributions of the known decision variables across the configurations of the unknown ones. Ultimately, simplified cognitive representations result in a reduced perceived degree of interdependencies among the decision variables (i.e., a reduction in the apparent $K$ value) as interactions among the $N - N'$ variables are not reflected in the local representation of the landscape (Gavetti and Levinthal 2000).

In our ISD context with fully specialized units (i.e., no knowledge overlap), each unit is mindful of only $N/2$ design choices that impact the fitness of the $N$-element ISD configuration (i.e., $N' = N/2$). As a result, each unit will have a local assessment of the fitness contribution of the ISD configuration, which in turn will guide the design choices that they enact. To illustrate, let’s reconsider the influence matrix in Figure 3c and suppose the ISD project configuration is specified by the ten-dimensional array $d = (1, 0, 0, 1, 1, 1, 0, 1, 0, 0)$. The fitness contribution of the value of the first element ($d_1 = 1$) of this array depends on the values of three other business design choices ($d_2$, $d_4$, and $d_5$) and two other technology design choices ($d_7$ and $d_{10}$). So even if the

---

6 The average will be an unbiased estimate of the fitness associated with the actual landscape.
fitness contribution of \( d_i \) (given \( d_2, d_4, \) and \( d \)) depends on the current values of \( d_i = 0 \) and \( d_7 = 0 \), the business unit’s assessment of \( c_i \), based on the unit’s simplified five-dimensional cognitive representation of the fitness landscapes, will be the average of the fitness contributions given the four possible combinations of \( d_7 \) and \( d_{10} \) values (i.e., \( \langle d_7, d_{10} \rangle = \{(0, 0), (0, 1), (1, 0), (1, 1)\} \)). See Table 1 for a numerical example.

[INSERT TABLE 1 HERE]

So although the fitness contribution of the first design variable \( d_i = 1 \) may be quite large (e.g., \( c_i(d_i \mid d_2, d_4, d_5, d_7 = 0, d_{10} = 1) = 0.954 \)) it can also be quite poor for some configurations of the unknown design variables (e.g., \( c_i(d_i \mid d_2, d_4, d_5, d_7 = 1, d_{10} = 0) = 0.127 \)). Nonetheless, the business unit who is assessing the value of \( d_i = 1 \), estimates that, on average, the fitness contribution of setting the first design choice \( d_i \) to 1 will be 0.674. We can easily infer from this numerical example that each unit’s assessment of the fitness contributions of each design choice will be less accurate when there are a greater number of knowledge interdependencies across knowledge domains.

**Knowledge Overlap**

We model knowledge overlap with parameter \( o \), which represents the number of elements in the other unit’s knowledge domain, for which the focal unit has knowledge. At the extremes, if \( o = 0 \), then there is no overlap; if \( o = N/2 \), then there is full overlap.\(^7\)

With knowledge overlap, the units’ cognitive representations are now more refined as they are mindful of \( N/2 + o \) design choices. To illustrate, let’s suppose again that the ISD project

---

\(^7\) Since knowledge overlap need not necessarily be evenly distributed across units (i.e., the business unit need not have the same amount of technical knowledge as the amount of business knowledge of the IS unit), we use subscripts “bus” and “IS” to denote the extent of technical knowledge of the business unit \((o_{bus})\) and the extent of application domain knowledge of the IS unit \((o_{IS})\), respectively.
configuration is specified by the array $d = (1, 0, 0, 1, 1, 1, 0, 1, 0, 0)$, that $o = 1$ and that it was determined that the business unit also has knowledge about $d_7$. In this case, since the fitness contribution of the value of the first element ($d_1 = 1$) of this array depends on the values of the second ($d_2 = 0$), fourth ($d_4 = 1$), fifth ($d_5 = 1$), seventh ($d_7 = 0$) and tenth ($d_{10} = 0$) elements, and the business unit now understands the fitness contribution of $d_1$ (given $d_2$, $d_4$, $d_5$, and $d_7$), the only unknown element will be $d_{10}$. As a result, the business unit’s assessment of $c_1$ will be the average of the fitness contributions given the two possibilities for $d_{10}$ (i.e., $d_{10} = \{0, 1\}$). See Table 2 for a numerical example.  

[INSERT TABLE 2 HERE]

**ISD Process as Landscape Search**

Although there are many different approaches to ISD (e.g., sequential waterfall, rapid application development, agile development etc.), in its simplest and most generic form, the ISD process can be characterized as an *incremental* and *iterative* process of requirements determination and systems design (Larman and Basili 2003). During the requirements determination phase, business design choices are made, while technology design choices are made during the systems design phase. These phases are iterated until a satisfactory design configuration is reached. Ultimately, the ISD team seeks to maximize ISD performance. In our model, the organization’s objective is to find the highest value point in the fitness landscape of ISD project configurations. The organization searches for a good configuration of design choices via an incremental and local experiential search process.

---

8 As is illustrated by this numerical example, knowledge overlap affords a more accurate assessment of the fitness contribution of each design choice. Again, we can also infer that the accuracy of each unit’s assessment of fitness contributions will depend not only on the extent of knowledge overlap, but also the relevance of the overlapping knowledge to the decision variable. For example, if there were no knowledge interdependencies across units, then knowledge of the other unit’s design configuration would not matter in the assessment of fitness contributions.
Given the ill-structured nature of ISD (Boland 1978, Curtis et al. 1988, Simon 1973), the ISD search process is assumed to be local (Cyert and March 1963, March and Simon 1958). In other words, organizational units can only effectively consider incremental changes to its current configuration – the units cannot possibly know the performance implications of a random and radically different configuration of design choices (Levinthal 1997). The organizational units are assumed to be able to identify alternative configurations in their immediate neighborhood whose fitness value is superior to their current level of fitness. For example, if the organization’s current ISD project configuration is specified by the array \( d = (1, 0, 0, 1, 1, 1, 0, 1, 0, 0) \), with \( d_{\text{bus}} = (1, 0, 0, 1, 1) \) and \( d_{\text{IS}} = (1, 0, 1, 0, 0) \), then during the requirements determination phase, the business unit may consider 5 (i.e., \( N/2 \)) neighboring locations in the fitness landscape – \( \{(0, 0, 0, 1, 1), (1, 1, 0, 1, 1), (1, 0, 1, 1, 1), (1, 0, 0, 1), (1, 0, 0, 0)\} \) (i.e., permutations of each of the first 5 design choices), while during the systems design phase, the IS unit may consider 5 neighboring locations – \( \{(0, 0, 1, 0, 0), (1, 1, 1, 0, 0), (1, 0, 0, 0, 0), (1, 0, 1, 1, 0), (1, 0, 1, 0, 1)\} \) (i.e., permutations of each of the last 5 design choices). While conducting search, the units are not assumed to evaluate all the neighboring alternatives to find the highest fitness level. Rather, the units will satisfy by adopting an alternative configuration so long as its fitness value is higher than that of its current configuration (March and Simon 1958). Once a unit has moved to a new configuration, we switch the focal search unit. In other words, if the business unit has found a new (higher performing) business configuration during the requirements determination phase, then we move on to the systems design phase where the IS unit can now search for a new (higher performing) technology configuration. If the IS unit finds a new (higher performing) technology configuration during systems design, then the business unit takes over and we repeat the requirements determination phase at the new configuration. This iterative process repeats
until the organization can no longer improve its performance or when a predetermined number of cycles is reached.

**ANALYSIS AND EXPERIMENTAL DESIGN**

To ensure that the results of the simulation reflect the underlying structure of the model rather than a particular realization of a stochastic process, we generate many landscapes with the same underlying pattern of interaction (i.e., a predetermined influence matrix). All results are based on 200 independently generated fitness landscapes for each influence matrix, with 100 ISD projects randomly seeded onto each fitness landscape. The performance of the ISD project is measured as a portion of the highest performance attainable on each landscape.\(^9\)

To generate sufficiently complex fitness landscapes for ISD projects, we set \(N = 16\) so each of business and IS units has, by default, responsibility for 8 configuration choices.\(^10\) Each simulation run is executed for 100 simulated time periods, by which time most ISD projects have reached a stable state (i.e., additional improvements cannot be found).\(^11\) We varied the ISD project complexity parameter \(K\) to represent all levels of overall landscape complexity for \(N = 16\) (i.e., \(K = 0, \ldots, 15\)). For example, when \(K = 0\), each design choice is independent and the

---

\(^9\) In \(NK\) landscapes simulations, the parameter \(K\) determines not only the overall complexity of the fitness landscape (in terms of number of peaks or sticking points) but also the maximum fitness value within the landscape. Smooth landscapes (i.e., low \(K\)) generally have smaller maximum fitness values which are easier to reach compared to rugged landscapes (i.e., high \(K\)), which have larger maximum fitness values, which are albeit more difficult to attain due to the inherent ruggedness of the landscape. As a result, when comparing performance across complexity levels (i.e., across \(K\)), it is appropriate to compare fitness values that are normalized to the maximum attainable performance levels rather than the raw fitness measures.

\(^10\) With \(N = 16\), there are \(2^{16} = 65,536\) possible configurations of ISD projects.

\(^11\) Each period represents one phase (requirements determination or systems design) in the ISD cycle. With an incremental / iterative development process, a single phase may take anywhere from a few days to a few weeks. If we consider, an agile development process where requirements determination and systems design iterations are completed on a weekly cycle, 100 time periods represents 100 weeks, or approximately 2 years.
resulting landscape is smooth. Conversely, when \( K = 15 \), the resulting landscape extremely rugged and search within this landscape is extremely difficult as a single change in one design choice changes the value contributions of all other design choices. When \( K \) is between 0 and 15, the overall complexity of the ISD project landscape is increasing in \( K \). For each level of \( K \) between 0 and 15, we randomly created 50 different influence matrices.\(^{12}\) Figure 4 shows the average number of peaks at different levels of \( K \). As \( K \) increases, the number of peaks in the fitness landscape increases exponentially.\(^{13}\)

[INSERT FIGURE 4 HERE]

For each level of ISD project complexity, we endowed the ISD project team with varying levels of knowledge overlap. We randomly select \( o \) design choice elements from the other unit’s set of design choice elements. Our analyses of the overall impact of knowledge overlap on ISD performance (Section 4.1) are based on 144 experimental configurations when the extent of knowledge overlap is the same across business and IS units – 16 levels of \( K \) (\( K = 0, \ldots, 15 \)) \( \times \) 9 levels of \( o \) (\( o_{bus} = o_{IS} = 0, \ldots, 8 \)) = 144. To investigate the impact of the distribution of knowledge overlap values across business and IS units at all levels of ISD project complexity – 16 levels of \( K \times 9 \) levels of \( o_{bus} \times 9 \) levels of \( o_{IS} = 1296 \).

\(^{12}\) When \( K = 0 \) or \( K = 15 \), there can only be 1 possible influence matrix – i.e., 1’s only on the diagonal when \( K = 0 \), or 1’s filling all cells of the influence matrix when \( K = 15 \). For these cases, 200 randomly generated fitness landscapes were simulated given the respective influence matrices.

\(^{13}\) The overall search complexity of a landscape can be characterized by the number of peaks (i.e., sticking points) within the landscape (Kauffman 1989). When \( K = 15 \), there are on average 3858.4 peaks in the landscape. Given that with \( N = 16 \), there are \( 2^{16} = 65536 \) possible configurations, this level of complexity means that 5.89\% of all possible configurations are sticking points. Put differently, one in every 17 configurations is a local optimum. Since there are 16 local neighbors (neighboring configurations which differ from a local configuration by only value), this level of complexity roughly translates into every move being in jeopardy of reaching a sticking point from which the ISD project cannot escape.
In order to investigate the impact of patterns of interdependencies, we manipulated the number of interdependencies both within and between units while keeping the overall level of ISD complexity fixed at a moderate level \((K = 7)\). These landscapes are denoted \(W7B0\) (i.e., 7 interactions within unit and 0 interactions between units), \(W6B1\), \(W5B2\), \(W4B3\), \(W3B4\), \(W2B5\), \(W1B6\) and \(W0B7\).\(^{14}\) Our analyses of the impact of knowledge overlap for different patterns of knowledge interdependencies (Section 4.3) are based on 72 experiments – 1 level of \(K \times 8\) patterns of within- and between-unit interactions \((W7B0 \sim W0B7) \times 9\) levels of \(o = 72\).

Finally, we also created two additional influence matrices that have a skewed distribution of interdependencies while keeping the average \(K\) for the landscape fixed at 7. One such matrix was unevenly distributed to favor the upper right block (i.e., business design choices depend on IS design choices more than IS design choices depend on business design choices; \(UR\) or Upper Right), and the other to favor the lower left block (i.e., IS design choices depend on business design choices more than business design choices depend on IS design choices; \(LL\) or Lower Left).\(^{15}\) These simulations were run for all combinations of knowledge overlap values across business and IS units and were compared with those where the influence matrix was balanced (i.e., business design choices depend on IS design choices to the same extent as IS design choices depend on business design choices; \(BAL\) or Balanced). Our analyses of the impact of knowledge overlap for different distributions of knowledge interdependencies (Section 4.4) are based on 243 experiments – 1 level of \(K \times 3\) influence matrices for different distribution of

\(^{14}\) For ease of exposition, we named these influence matrices as \(W_{K_{\text{within}}}B_{K_{\text{between}}}\). For instance, \(W7B0\) represents an influence matrix where there are 7 interdependencies within each respective unit and no interdependencies between units, whereas the influence matrix \(W4B3\) would represent one where there are 4 interdependencies within units and 3 interdependencies between units.

\(^{15}\) The influence matrices for the case of skewed distribution of inter-unit knowledge interdependence are shown in Appendix A.
knowledge interdependencies \((UR, LL \text{ and } BAL) \times 9\) levels of \(o_{bus} \times 9\) levels of \(o_{IS} = 243\).

RESULTS

Impact of Knowledge Overlap on ISD Performance

Figure 5 shows how different levels of knowledge overlaps between business and IS units influence the performance of the ISD project over time when the level of ISD project complexity is very low (Figure 5a; \(K = 0\)), moderate (Figure 5b; \(K = 7\)), and very high (Figure 5c; \(K = 15\)).\(^{16}\) As discussed previously, the performance of the ISD project (i.e., fitness \((F)\)) is measured as a portion of the highest performance attainable on each landscape. The results clearly show that ISD performance decreases as the level of ISD project complexity increases (i.e., as \(K\) increase from 0 to 7 to 15), regardless of knowledge overlap levels. When ISD project complexity is low (Figure 5a; \(K = 0\)), all ISD projects, regardless of extent of knowledge overlap, eventually reach global optimum (i.e., fitness \((F) = 1\)) as can be expected given the smooth fitness landscape. However, we note that the rate at which the global optimum is reached is actually faster with lesser levels of knowledge overlap. In fact, the rate at which performance increases initially is greater for lesser levels of knowledge overlap at low levels of complexity (i.e., when \(K < 4\)), even though ultimate performance levels reached at \(t = 100\) is higher for greater levels of knowledge overlap.\(^{17}\) It seems that the overlapping knowledge between units is inducing the ISD projects to explore a greater number of design configurations, which may not be necessary within a smooth fitness landscape. As ISD project complexity increases, even though the global

\(^{16}\) The simulation models were run with all values of \(o_{bus} = o_{IS} = 0, \ldots, 8\) and \(K = 0, \ldots, 15\). Figure 5 only shows the cases when \(o = \{0, 2, 4, 6, 8\}\) and \(K = \{0, 7, 15\}\) for brevity.

\(^{17}\) This result is not shown in Figure 5 but is available in additional exhibit 1 which shows performance over time for different levels of knowledge overlap across all levels of ISD project complexity (i.e., \(K = 0, \ldots, 15\)).
optimum cannot be attained, ISD projects with a greater extent of knowledge overlap are able to reach higher performance levels.

[INSERT FIGURE 5 HERE]

One additional noteworthy observation from the above results is that the differences in final performance at $t = 100$ across different levels of knowledge overlap are found to be larger as ISD project complexity increases. When $K = 0$, there is no difference in eventual performance levels across different levels of knowledge overlap – since all ISD projects, regardless of extent of knowledge overlap, reach the global optimum. However, as ISD project complexity increases, the difference in performance across different levels of knowledge overlap becomes larger. With more complex landscapes, it seems that the additional exploration of alternatives due to the greater extent of overlapping knowledge across units is helping the ISD project team to better navigate the rugged landscape. Figure 6 summarizes the eventual final performance for all levels of knowledge overlap (i.e., $\sigma = 0, \ldots, 8$) across all levels of ISD project complexity (i.e., $K = 0, \ldots, 15$). It clearly shows that, as the ISD project complexity increases, the final performance decreases faster for lesser levels of knowledge overlap than for greater levels of knowledge overlap.

[INSERT FIGURE 6 HERE]

To better understand the substantive impact of knowledge overlap on ISD performance, we developed a measure of search effectiveness ($SearchEffectiveness; SE$), conceptualized as the ratio of eventual performance gains over attainable performance gains: 

---

18 This result is quite evident in additional exhibit 1 which shows performance over time for different levels of knowledge overlap across all levels of ISD project complexity ($K = 0, \ldots, 15$).

19 As mentioned earlier, since the maximum attainable fitness increases with $K$ (Kauffman 1989), normalized initial fitness values (i.e., $F_{\text{init}}$) decrease with $K$. As a result, the amount of attainable performance gains is smaller (larger) when $K$ is small (large). Therefore, when comparing results across
The results in Figure 7, which plots the effectiveness of landscape search for different levels of knowledge overlap across the range of ISD project complexity levels, show that as ISD project complexity increases, overall search effectiveness decreases as it is more difficult to search within increasingly rugged fitness landscapes, but also that the rate at which search effectiveness decreases is less pronounced when the extent of knowledge overlap is great (e.g., $o = 8$ or $6$ compared to $o = 0$ or $2$). For instance, when ISD project complexity is very high (i.e., $K = 15$), ISD projects in which knowledge overlap across units does not exist (i.e., $o = 0$) can only attain 5.2% of potential performance gains, whereas ISD projects with full knowledge overlap (i.e., $o = 8$) can attain 48.2% of potential performance gains. We also note that the difference in search effectiveness amplifies as the extent of knowledge overlap becomes greater (i.e., $SE(K=15, o=2) - SE(K=15, o=0) < SE(K=15, o=4) - SE(K=15, o=2) < \ldots < SE(K=15, o=8) - SE(K=15, o=6)$).

As a point of reference, we may compare the search effectiveness across levels of knowledge overlap to assess the power of knowledge overlap in reducing the effectual complexity of ISD projects. For instance, the search effectiveness within highly complex ISD projects when the level of knowledge overlap is very high (i.e., $SE(K=15, o=8) = 48.2\%$) is actually greater than that when ISD project complexity $K = 10$ and knowledge overlap $o = 6$, or than that when $K = 7$ and $o = 4$, or than that when $K = 6$ and $o = 2$, or than that when $K = 4$ and $o = 0$ (see dotted line in Figure 7). Given that the difficulty of navigating the ISD project landscapes increases exponentially with $K$ (see Figure 4), these results are quite remarkable.

levels of complexity ($K$), a measure of search effectiveness ($SE$) that takes into account the amount of attainable performance gains is appropriate.
These results suggest that an ISD team with a high degree of knowledge overlap (e.g., \( o = 8 \)) may be expected to produce ISD project performance in a highly complex ISD project landscape (i.e., \( K = 15 \)) that is on par with an ISD team engaged in a much simpler project (i.e., \( K = 4 \)) if knowledge overlap was absent.\(^{20}\)

**Impact of Distribution of Knowledge Overlap**

Whilst the previous section focused on cases where the extent knowledge overlap was equal across units (i.e., \( o_{bu} = o_{IS} \)), the analyses in this section relaxes this assumption and also considers cases with unevenly distributed knowledge overlap. In other words, we consider cases, where the business unit may have more technical knowledge \( (o_{bu}) \) than the IS unit has application domain knowledge \( (o_{IS}) \), or vice versa. Figure 8 shows the final performance at \( t = 100 \) at varying levels of ISD project complexity (i.e., \( K = \{1, 7, 15\} \)).\(^{21}\) The results show when ISD project complexity is low, the distribution of knowledge overlap does not significantly influence overall ISD project performance (see flat surface in Figure 8a for \( K = 1 \)). In fact the response surface remains relatively flat, albeit slightly slanted, for \( K < 4 \), indicating that when ISD project complexity is low, different distributions of knowledge overlap does not lead to differences in eventual ISD performance. However, the response surface becomes increasingly convex (see hollow response surface in Figure 8c for \( K = 15 \)), indicating that when ISD project complexity is high, how the overlapping knowledge is distributed also has an impact on ISD performance. Interestingly, we observe that while greater levels of overall knowledge overlap are generally beneficial (see upward sloping planes in Figure 8b and 8c), ISD performance is

\(^{20}\) The average number of peaks (or sticking points) when \( K = 15 \) (for \( N = 16 \)) is 3858.4 whereas that when \( K = 4 \) is only 117.8 (see Figure 4).

\(^{21}\) Although the simulation models were run for all values of ISD project complexity \( (K = 0, \ldots, 15) \), Figure 8 only shows the cases when \( K = \{1, 7, 15\} \) for brevity. The results for all values of \( K = 0, \ldots, 15 \) are available in additional exhibit 2.
actually greater when knowledge overlap is unevenly distributed. For instance, when the total amount of knowledge overlap is held constant (e.g., $o_{IS} + o_{bus} = 8$), we see that ISD performance is greater when the IS unit’s business knowledge is either greater or lesser than the business unit’s IS knowledge (e.g., $o_{IS} = 8$ and $o_{bus} = 0$; $o_{IS} = 7$ and $o_{bus} = 1$; $o_{IS} = 1$ and $o_{bus} = 7$ or $o_{IS} = 0$ and $o_{bus} = 8$), than when they are equal (i.e., $o_{IS} = o_{bus} = 4$). This effect is more pronounced at higher levels of ISD project complexity.

[INSERT FIGURE 8 HERE]

**Patterns of Knowledge Interdependencies and Knowledge Overlap**

While the notion that the *degree* of interdependencies (i.e., the value of parameter $K$ in NK landscapes) is an important determinant of overall landscape search performance has been well established in the NK landscapes literature, recent research has observed that the *patterns* of interdependencies also has a substantial impact on the efficacy of landscape search (Rivkin and Siggelkow 2007). The preceding discussion focused on analyzing the impact of knowledge overlap on ISD project performance given a general view of ISD project complexity by generating numerous influence matrices given a specified value of $K$. In this section, we turn our attention to investigating whether patterns of knowledge interdependencies also influences how the extent of knowledge overlap impacts ISD project performance. To do so, we developed a number of influence matrices where the distribution of within-unit and between-unit interdependencies varies, while holding the total number of interdependencies constant (i.e., $K = K_{within} + K_{between} = 7$).

Figure 9 shows how different degrees of knowledge overlaps ($o = \{0, 2, 4, 6, 8\}$) between business and IS units influence ISD performance at the early stage (i.e., $t = 15$) and end of the ISD project (i.e., $t = 100$) for different patterns of ISD project complexities (i.e., when the
The results show that, regardless of the distribution of interdependencies, knowledge overlap is positively associated with ISD performance, except for the case of W7B0 where all knowledge overlaps lead to the same ultimate performance (at $t = 100$). Furthermore, a greater knowledge overlap results in more consistent performance across different distributions of interdependencies. Interestingly, ISD performance tends to be lower when within-unit interdependencies and between-unit interdependencies are evenly distributed (e.g., W3B4) than when those interdependencies are unevenly distributed (e.g., W6B1, W0B7). The only exception is that when a complete knowledge overlap exists (i.e., $o = 8$), performance does not seem to be impacted by the distributions of decision interdependencies. As a result, the performance gaps among varying levels of knowledge overlap become larger as interdependencies become increasingly evenly distributed. Finally, we also observe that the performance differences among varying levels of knowledge overlap, while they become larger as the ISD project is completed (i.e., at $t = 100$), do seem to materialize early in the course of the ISD project (i.e., at $t = 15$).  

The U-shaped results for the performance as within-unit interdependencies decrease and between-unit interdependencies increase suggest that the specialization-based task partitioning into business and IS domains is less effective when the within- and between-unit interdependencies are evenly distributed (e.g., W4B3 or W3B4). It is quite intuitive to infer why performance would suffer less when there are greater within-unit interdependencies (i.e.,

---

22 The results for all time periods ($t = 0, \ldots, 100$) is available in additional exhibit 3 which shows the performance levels of different levels of knowledge overlap for different distributions of within- and between-unit interdependencies across all time periods.
business (IS) design choices depend more on other business (IS) design choices rather than IS (business) design choices; e.g., W7B0 or W6B1). Since there aren’t many between-unit interdependencies, each unit can focus on optimizing the configuration for its own domain, since design choices of the other unit will not impact the fitness of its design choices. However, that performance suffers less when there are greater between-unit interdependencies (e.g., W0B7 or W1B6) is less intuitive. In such cases, due to the greater level of between-unit interdependencies, each unit is essentially at the mercy of the other unit’s design choices. In other words, rather than taking an active role in searching for the optimal local (i.e., within domain) design configuration, each unit will rely on the other unit’s design choices and make its own design decisions that support the design choices of the other unit. Given that the level of within-unit interdependencies is low, the resulting local landscape will be smoother, and consequently this adaptive decision will be efficient and not require extensive search. However, when within- and between-unit interdependencies are more evenly distributed, search is globally more difficult since each unit must actively search within its rugged local landscape and design choices of the other unit can also drastically alter its own landscape. However, the results suggest that greater knowledge overlap can help to mitigate such difficulties.

**Interaction between Knowledge Overlap and Distribution of Inter-unit Interdependencies**

Figure 10 shows how different levels and distributions of knowledge overlaps influence the final performance of the ISD project when the distribution of between-unit interdependencies varies, holding the average number of interdependencies in the influence matrix constant ($K = 7$). Specifically, Figure 10a and 10b show the results when the between-unit interdependencies are unevenly distributed – business design choices are more dependent on IS design choices (Figure
10a) or IS design choices are more dependent on business design choices (Figure 10b); whereas Figure 10c shows the results when business and IS design choices are evenly interdependent.\textsuperscript{23}

The results show the interactions between knowledge overlap pattern and distribution of between-unit interdependencies and suggest that ISD performance is higher when the knowledge overlap pattern matches the distribution of interdependencies. For example, as shown in Figure 10a, when business design choices are more dependent on IS design choices rather than the other way round, the technical knowledge possessed by the business unit contributes to project performance more than the application domain knowledge of the IS unit. Similarly, we found that when IS design choices are more dependent on business design choices than the other way round, application domain knowledge of the IS unit contributes to performance more than the technology knowledge of the business unit.

**DISCUSSION**

In this paper, we present a novel conceptualization of the ISD process as search within a design space made up of application domain and technical configurations. This conceptualization allows us to explore the dynamic relationship between the structure of ISD projects and the distribution of relevant knowledge across the organization in effecting the efficacy of finding high-quality configurations for information systems. More specifically, our \textit{NK} fitness landscape model of ISD equips us to investigate the effect of knowledge overlap on ISD performance under various conditions (e.g., varying levels of interdependencies, distribution of interdependencies, etc.). This research extends and elaborates on the conventional wisdom

\textsuperscript{23} See Appendix 1 for examples of the influence matrices used for these experiments.
that the knowledge overlaps between business and IS units are generally beneficial for ISD performance. The primary objective of this research was to derive new or nuanced theoretical insights about the effect of knowledge overlap on ISD.

Prior literature has investigated the direct effect of knowledge overlap between the IS unit and the business unit on the performance of the ISD process in terms of coordination and decision control (Espinosa et al. 2007a, Tiwana 2009). Furthermore, one study has investigated the moderating effect of project and process novelty on the relationship between knowledge overlap and outsourcing ISD performance (Tiwana 2004). However, prior research has not yet fully revealed the complex conditions under which knowledge overlap affects ISD performance differently. This research fills the gap. We find that the relative effect of knowledge overlap on ISD performance depends on the overall level of ISD complexity in terms of extent of knowledge interdependencies, patterns of such interdependencies, and also distribution of overlapping knowledge.

**ISD Complexity**

The complexity of an ISD project depends not only on the number of design choice elements but also and more importantly on the extent of interdependencies between design choices (Xia and Lee 2005). Prior research suggests that complex ISD projects are generally associated with lower performance (Banker and Slaughter 2000, Xia and Lee 2004) and that the effect of team knowledge on ISD performance may depend on the level of task complexity (Espinosa et al. 2007b). The results of the simulation are consistent with prior studies but also provide additional insights by suggesting that as the overall level of ISD complexity ($K$) increases, overall ISD performance decreases, regardless of knowledge overlap level (see Figures 5 and 6). We observe however that extent of knowledge overlap across the business and
IS units helps deal with ISD complexity by cushioning the negative impact of ISD complexity on performance, and the moderating impact of knowledge overlap is stronger at higher levels of ISD complexity. The moderating role of knowledge overlap is quite remarkable since eventual performance of highly complex ISD projects (i.e., $K = 15$) when the level of knowledge overlap is high (i.e., $o = 6$ or $8$) is similar to (or even greater than) the performance of moderately complex ISD projects (i.e., $K = 7$–$9$) when knowledge overlap does not exist (i.e., $o = 0$) or when the level of knowledge overlap is low (i.e., $o = 2$). In other words, organizations may use knowledge overlap as a means to reduce the effective level of ISD complexity. These findings are summarized into the following testable propositions:

**Proposition 1:** As ISD complexity increases, ISD performance decreases, regardless of knowledge overlap level.

**Proposition 2:** At all levels of ISD complexity, a higher level of knowledge overlap is associated with greater ISD performance.

**Proposition 3:** As ISD complexity increases, the performance gap between a higher level of knowledge overlap and a lower level of knowledge overlap increases.

How does knowledge overlap help mitigate the difficulties of ISD complexity? Our conceptualization of the ISD process as search provides a plausible explanation. Prior work on organizational search (e.g., Gavetti and Levinthal 2000, Rivkin and Siggelkow 2003) has highlighted the importance of balancing exploration and exploitation (March 1991) especially when the landscape to search is rugged and multi-peaked (i.e., when complexity is high). Exploration is important because it provides the variation necessary to escape local peaks and find other fertile regions within landscape. Exploitation is important for organizations to zero in on identifying the location of highest return once a general region has been identified. Knowledge overlap across units provides an implicit mechanism whereby each unit’s exploratory search efforts are coordinated toward a unified configuration. Although each unit searches its own design space independently, knowledge of the other unit’s design elements
helps to have a more accurate assessment of how its design changes will impact the overall fitness for the ISD project. As a result, when a unit (business or IS) searches its design space, it not only worries about how its changes impact the fitness for its own design choice elements but also considers how these changes impact the fitness with respect to the choices of the other unit. In other words, knowledge overlap impels the independent units to search while being cognizant of the impact that its design choices will have on the other unit’s fitness level. In doing so, the ISD team as a whole can be guided toward better exploring high performance regions within the landscape.

That said, we also observed a negative impact of knowledge overlap in terms of hindering the speed at which the organization converges on its optimal design when complexity is low (see Figure 5a). When there does not exist many interdependencies between design elements (i.e., when complexity is low), each unit’s design choices should not significantly impact the fitness of other unit’s decisions. However, having knowledge of and considering the other unit’s design elements may create unnecessary variation that derails the other units search trajectories. As a result, the overall search process takes longer.

**Patterns of Interdependencies**

While the previous discussion centers on a generalized notion of ISD complexity, our analyses provide additional insights into how the patterns of interdependencies impact the complexity of ISD projects and consequently ISD performance.

Although much of the literature using NK landscapes models have assumed a random distribution of interdependencies to model complexity, prior research has emphasized that it is not the overall extent of interdependencies but the patterns of interdependencies that actually determine the complexity of a system (Rivkin and Siggelkow 2007). Our study extends this line
of work by investigating patterns of interdependencies when there are multiple sub-units with specialized knowledge. Our overall findings can be summarized into the following testable propositions:

**Proposition 4:** When all interdependencies are associated with design choices within the same unit and no interdependencies are associated with design choices between business and IS units, knowledge overlap does not impact ISD performance.

**Proposition 5:** A higher level of knowledge overlap results in more consistent ISD performance across different distributions of within- and between-unit interdependencies.

**Proposition 6:** ISD performance is lower when within-unit interdependencies and between-unit interdependencies are evenly distributed than when those interdependencies are unevenly distributed.

**Proposition 7:** As interdependencies become evenly distributed across within units and between units, the performance gap between a higher level of knowledge overlap and a lower level of knowledge overlap increases.

**Proposition 8:** When interdependencies among design choices exist, performance gap between a higher level of knowledge overlap and a lower level knowledge overlap increases over time.

Similar to Rivkin and Siggelkow (2007), our analyses show that even though the overall extent of interdependencies ($K$) is held constant, the patterns of interdependencies has a significant impact on the overall complexity of the system in terms of how easy (or difficult) it is to search for fertile locations on the landscape. We find that ISD projects are most complex when interdependencies are evenly distributed across within and between knowledge domains (e.g., W4B3 and W3B4 in Figure 9). The flip side is that ISD projects are less complex when interdependencies exist mostly within knowledge domains (e.g., W7B0 and W6B1 in Figure 9).

Our results suggest that when multiple units are independently searching for higher performing design configurations, one unit’s search activities can result in unreliable assessments if the

---

24 Rivkin and Siggelkow (2007) investigated archetypical patterns of interdependencies (e.g., random, centralized, hierarchical, small world, scale free, dependent, etc.) with a single unit that searches within these landscapes. Our study is different in that we have multiple (i.e., two) sub-units of the organization specialized in different domains that collaboratively search the landscape where the patterns of interdependencies relate to how design choices impact the fitness of other design choices within and between specialized domains.
fitness of its design choices heavily depends on the design choices of the other unit. As a result, when interdependencies are evenly distributed both within and between the units’ specialized domains, each unit’s design choices will alter the sub-landscape for the other unit, which makes it difficult to converge upon a high-performing solution. Conversely, when interdependencies are mostly within each unit’s specialized domain (i.e., a modular structure of interdependencies), then while the sub-landscape for each unit may be more complex to search, each unit’s search activities will be impacted less by the design choices of the other unit. In other words, in such cases, each unit can independently search for fertile ground without worrying about the consequences of the other unit’s search activities. Finally, somewhat surprisingly, we find that ISD projects are also less complex when interdependencies exist mostly across knowledge domains (e.g., W0B7, W1B6 in Figure 9). When interdependencies are mostly across the units’ specialized domains, this implies that the local sub-landscapes for each unit are relatively smooth (even though the fitness values of these smooth sub-landscapes may change significantly as a consequence of the design choices of the other units). As a result, even though the other unit’s design choices may alter a focal unit’s sub-landscape, the focal unit’s search activities can efficiently find an optimal solution (albeit local).

Again, knowledge overlaps helps overcome the difficulties of search arising from the structural complexity due to the patterns of interdependencies. Knowledge of the other unit’s design elements allows a more accurate assessment of how one’s own design choices will impact the overall fitness for the ISD project. With a better assessment of how the other unit’s design choices impact one’s own search, each unit need not be blindsided by the search activities of the other units, even when the patterns of interdependencies create opportunities for radical shifts in local sub-landscapes.
Distribution of Knowledge Overlap

One of the more interesting results relate to the distribution of knowledge overlap between units. Contrary to conventional wisdom in the IS literature, the analyses suggest that higher performance can be attained when knowledge overlap is unevenly distributed between the two units. In other words, holding the total amount of knowledge overlap constant, it seems more beneficial to have one unit know a lot about the other unit’s domain rather than have both units equally knowledgeable of each other. This effect was heightened when the patterns of interdependencies mirrored the distribution of knowledge overlaps. This finding is conceptually similar to prior research showing that ISD performance is higher when knowledge overlap patterns match the type of novelty characterizing an ISD project (Tiwana 2004). Our findings are summarized into the following propositions:

Proposition 9: ISD performance is higher when knowledge overlap is unevenly distributed between business and IS units

Proposition 10: As ISD complexity increases, the performance gap between unevenly and evenly distributed knowledge overlap increases.

Proposition 11: ISD performance is higher when the knowledge overlaps pattern matches the distribution of interdependencies; when IS design choices are more dependent on business design choices than the other way round, business knowledge of the IS unit contributes to performance more than IS knowledge of the business unit does, and vice versa.

It seems that organizations are better off if one of the units takes a leadership role in the ISD process. That said, the question of which unit (i.e., business vs. IS) should take the leadership role depends on the patterns of interdependencies for the ISD project. If both units have equal knowledge of each other’s domains, then each unit may presume to know what’s best for the other unit. However, despite each unit’s best efforts to search the landscape that is concurrently favorable to the other unit, such assertive search behaviors may create excessive variation and lead to each unit attempting to go in different directions. Conversely, if one of the
units has relatively more cross-unit knowledge, then this unit can take the role of exploring the overall landscape for high performance regions and the other unit can exploit by seeking the locally optimal position within that region. This interpretation is somewhat consistent with the notion of the structural ambidexterity proposed by Tushman and O’Reilly (1996) in that organizations become more effective when one unit is responsible for exploration and the other unit is responsible for exploitation.

CONCLUSION

This research leverages the power of the NK fitness landscape modeling method to discover new and nuanced insights about knowledge overlap in ISD and identifies a set of propositions that call for future empirical validation to further advance theory and practice. Our study has a number of implications for ISD theory and practice. Our analyses suggest that much of the difficulties in ISD are due to the complexity of the ISD arising from the structure of the interdependencies within and across business and technical domains. ISD organizations would do well to recognize and identify the structure of knowledge interdependencies of the ISD project and match the ISD team with an appropriate level of knowledge overlap so as to ensure adequate performance. However, the literature has yet to explore the notion of ISD complexity from a structural complexity perspective. For instance, the systems development methodologies practiced in the field (and taught in our classrooms) do not incorporate representations, models or tools to allow the systems analyst to adequately model and understand the interdependencies between application domain and technical knowledge. Also, the de facto iterative development process of analysis and design implicitly induces knowledge specialization and task partitioning.
We believe that much work is needed to develop better theories of ISD processes. Our study is a first step in this direction, harnessing the power of the NK landscape modeling method.

This study is not without limitations. Despite the advantages of simulation methods in their ability to incorporate complex dynamics without worrying about analytical tractability and their ability to study constructs of interest that may be difficult if not impossible to manipulate in field studies, simulation models are stylized theoretical models of reality that require rigorous validation through empirical testing. That said, we believe that our NK landscapes model of ISD has allowed us to provide valuable theoretical insights that can guide further theoretical and empirical research.
REFERENCES


FIGURES AND TABLES

**Figure 1. Influence Matrix (N = 10, K = 4)**

\[
\begin{bmatrix}
1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\
\end{bmatrix}
\]

(a) "Smooth" Fitness Landscape

(b) "Rugged" Fitness Landscape

**Figure 2. Fitness Landscapes**

**Figure 3. Examples of Influence Matrices (N = 10)**
Figure 4. Ruggedness of ISD Project Landscapes ($N = 16$)

Figure 5. ISD Performance Over Time at Varying Levels of ISD Project Complexity
Figure 6. ISD Performance at $t=100$ Across Knowledge Overlap and ISD Project Complexity

Figure 7. Impact of Knowledge Overlap on Search Effectiveness
Figure 8. ISD Performance at \( t = 100 \) Across Knowledge Overlap and ISD Project Complexity

(a) \( K = 1 \)

(b) \( K = 7 \)

(c) \( K = 15 \)

Figure 9. ISD Performance for Different Distributions of Within- and Between-Unit Interdependencies

(a) Fitness at time period \( t = 15 \)

(b) Fitness at time period \( t = 100 \)

Figure 9. ISD Performance for Different Distributions of Within- and Between-Unit Interdependencies
(a) When business design choices are more dependent on IS design choices than the other way around (UR)

(b) When IS design choices are more dependent on business design choices than the other way around (LL)

(c) When business and IS design choices are evenly interdependent

Figure 10. ISD Performance with Different Distribution of Knowledge Overlaps and Between-Unit Interdependencies ($t = 100$)

Table 1. Fitness Contribution Given Knowledge Specialization

<table>
<thead>
<tr>
<th>ISD Project Configuration</th>
<th>Fitness Contribution of $d_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 0, 1, 1, 0, 0)</td>
<td>0.894</td>
</tr>
<tr>
<td>(1, 0, 1, 1, 0, 1)</td>
<td>0.954</td>
</tr>
<tr>
<td>(1, 0, 1, 1, 1, 0)</td>
<td>0.127</td>
</tr>
<tr>
<td>(1, 0, 1, 1, 1, 1)</td>
<td>0.723</td>
</tr>
<tr>
<td>(1, 0, 1, 1, ?, ?)</td>
<td>avg = 0.674</td>
</tr>
</tbody>
</table>

Table 2. Fitness Contribution Given Knowledge Overlap

<table>
<thead>
<tr>
<th>ISD Project Configuration</th>
<th>Fitness Contribution of $d_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 0, 1, 1, 0, 0)</td>
<td>0.894</td>
</tr>
<tr>
<td>(1, 0, 1, 1, 0, 1)</td>
<td>0.954</td>
</tr>
<tr>
<td>(1, 0, 1, 1, 0, ?)</td>
<td>avg = 0.924</td>
</tr>
</tbody>
</table>
APPENDIX A - INFLUENCE MATRICES

(a) Influence Matrix UR

<table>
<thead>
<tr>
<th>Business Design Choices</th>
<th>IS Design Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 1 0 0 1 0 1</td>
<td>0 1 1 0 1 1 0 1</td>
</tr>
<tr>
<td>0 1 0 1 1 1 0 1</td>
<td>0 1 0 1 1 1 0 1</td>
</tr>
<tr>
<td>0 1 1 1 0 0 1 0</td>
<td>1 0 1 0 0 1 1 1</td>
</tr>
<tr>
<td>0 0 1 1 1 0 1 1</td>
<td>1 1 1 1 0 1 0 0</td>
</tr>
<tr>
<td>1 0 0 1 1 1 0 1</td>
<td>1 0 0 1 1 1 0 0</td>
</tr>
<tr>
<td>0 1 1 0 1 1 0 1</td>
<td>0 0 1 1 1 0 1 1</td>
</tr>
<tr>
<td>1 1 0 1 1 0 1 0</td>
<td>1 1 1 0 1 1 0 1</td>
</tr>
<tr>
<td>1 0 1 0 0 1 1 1</td>
<td>1 1 1 0 0 1 0 1</td>
</tr>
</tbody>
</table>

(b) Influence Matrix LL

<table>
<thead>
<tr>
<th>Business Design Choices</th>
<th>IS Design Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 1 0 0 1 0 1</td>
<td>0 0 0 1 0 0 0 0</td>
</tr>
<tr>
<td>0 1 0 1 1 1 0 1</td>
<td>1 1 1 0 0 0 1 0</td>
</tr>
<tr>
<td>0 1 1 1 0 0 1 0</td>
<td>1 1 1 1 0 0 1 0</td>
</tr>
<tr>
<td>0 0 1 1 1 0 1 1</td>
<td>1 1 1 0 1 0 1 0</td>
</tr>
<tr>
<td>1 0 0 1 1 1 0 1</td>
<td>1 0 1 1 1 1 0 0</td>
</tr>
<tr>
<td>0 1 1 0 1 1 0 1</td>
<td>0 1 1 0 1 1 0 1</td>
</tr>
<tr>
<td>1 1 0 1 1 0 1 0</td>
<td>1 1 0 1 1 0 0 0</td>
</tr>
<tr>
<td>1 0 1 0 0 0 1 1</td>
<td>1 0 1 0 0 1 1 1</td>
</tr>
</tbody>
</table>

(c) Influence Matrix BAL

<table>
<thead>
<tr>
<th>Business Design Choices</th>
<th>IS Design Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 1 0 0 1 0 1</td>
<td>1 0 0 1 0 0 1 0</td>
</tr>
<tr>
<td>0 1 0 1 1 1 0 1</td>
<td>1 1 0 1 0 0 1 0</td>
</tr>
<tr>
<td>0 1 1 1 0 0 1 0</td>
<td>1 1 1 1 0 0 1 0</td>
</tr>
<tr>
<td>0 0 1 1 1 0 1 1</td>
<td>1 1 1 0 1 0 1 0</td>
</tr>
<tr>
<td>1 0 0 1 1 1 0 1</td>
<td>1 0 1 1 1 1 0 0</td>
</tr>
<tr>
<td>0 1 1 0 1 1 0 1</td>
<td>0 1 1 0 1 1 0 1</td>
</tr>
<tr>
<td>1 1 0 1 1 0 1 0</td>
<td>1 1 0 1 1 0 0 0</td>
</tr>
<tr>
<td>1 0 1 0 0 1 1 1</td>
<td>1 0 0 1 0 1 1 1</td>
</tr>
</tbody>
</table>

Notes: UR: Upper Right; LL: Lower Left; BAL: Balanced.
In the UR matrix, business design choices are more dependent on IS design choices than the other way around (average $K = 7$). In the LL matrix, IS design choices are more dependent on business design choices than the other way around (average $K = 7$). Finally, in the BAL matrix, the business and IS design choices are evenly interdependent ($K = 7$). These matrices (UR, LL and BAL) were used for the experimental results reported in Figures 11a, b and c, respectively.