Links between systems thinking and complex decision making

Kambiz E. Maani* and Vandana Maharaj

Kambiz Maani is an Associate Professor of Management and System Sciences at the University of Auckland and Head of the School of Business and Economics at Tamaki Campus. He has a MS in Engineering and PhD in Operations Research from the University of Illinois at Urbana-Champaign. He has held academic and research positions at a number of institutions including MIT, Boston University, and the Rand Corporation. Dr Maani is an author and speaker on Systems Thinking, Leadership, and Organizational Learning. His research focuses on the integration of management and systems science. His latest book (with Bob Cavana): *Systems Thinking and Modelling—Understanding Change and Complexity* (Prentice Hall, 2000) is used widely at universities and corporations worldwide.

Vandana Maharaj completed her MCom (Hons) in Management Science and Information Systems

Abstract

There is a widely held belief that systems thinking is an answer to the increasing complexity of the world as well as the workplace. Despite strong interest and assertions, however, the relationship between systems thinking and complex decision making has received scant attention in the literature. Using Richmond's (1997) classification scheme as the theoretical base, this paper investigates the link between systems thinking and complex decision making using Verbal Protocol Analysis (VPA) methodology. The findings of the study indicate that while the degree of systems thinking does matter, certain *types* of systemic thinking would be more relevant to performance. Further, evidence shows that the subject's approach to the problem is also a highly pertinent factor in task performance, in that better performers displayed a distinctive pattern of thought that differed from that of the poor performers. Better performing subjects attempted to gain an understanding of the system structure *before* they proceeded to develop strategies and take action. The findings revealed a cyclical thought pattern that was consistently followed by better performing participants. This pattern, termed the CPA cycle, consists of three distinct phases of *conception, planning*, and *action*. This research contributes to the fields of systems thinking and complex decision making by integrating knowledge and methodology from several disciplines including psychology, management and IT. Specific contributions include a novel research methodology and, in particular, operationalization of the systems thinking paradigm, as well as identification of disaggregated factors affecting complex decision making. The managerial and organizational implications of the research are compelling and invite further research in this nascent field.

Syst. Dyn. Rev. **20**, 21–48, (2004)

There is a widely held belief that systems thinking is an answer to the increasing complexity of the environments in which we live and work. However, there is little empirical evidence to support the notion that systems thinking is indeed effectual in dealing with complexity.

Despite some notable research work (see, for example, Booth Sweeney and Sterman 2000; Doyle 1997; Buchner 1995; Pennington *et al.* 1995; Brehmer 1992; Funke 1991; Sterman 1989a, b), there is a curious gap in the literature on the relationship between systems thinking and complex decision making. According to Doyle (1997):

> Many claims have been made concerning the ability of systems thinking interventions to change the nature and quality of thought about complex systems, . . . [yet] important questions about the relationship between systems thinking and basic cognitive processes such as learning, memory, problem solving, decision making, and updating mental models remain unanswered.

* Correspondence to: K. E. Maani, Department of Management Science and Information Systems, School of Business and Economics, The University of Auckland, Private Bag 92019 Auckland, New Zealand. E-mail: k.maani@auckland.ac.nz
This paper is dedicated to the memory of Barry Richmond whose work has inspired this research

System Dynamics Review Vol. 20, No. 1, (Spring 2004): 21–48
Published online in Wiley InterScience
(www.interscience.wiley.com). DOI: 10.1002/sdr.281

Received September 2002
Accepted November 2003

from The University of Auckland in 1999. She also holds a BCom in Marketing and Commercial Law and a DipCom in International Business and Management Science and Information Systems both from the University of Auckland. She is currently responsible for product development at New Zealand's largest electricity and gas retailer.

This paper seeks to investigate the link between systems thinking and complex decision making. Richmond's (1997a, b, c, d) seven-classification scheme is used as the theoretical basis for this research. The Verbal Protocol Analysis (VPA) methodology was used to gather and analyse the empirical data from experiments with a group of subjects. As part of this approach, a coding scheme was developed to operationalize systems thinking. The contrasting performance of superior and poor performing subjects in complex decision tasks was then scrutinized and its theoretical and practical implications studied. Three distinct but related hypotheses, derived from the literature, are addressed in this study:

H_1: *Does (more) systems thinking lead to better performance in complex decision making?*
H_2: *Do certain (disaggregated) dimensions of systems thinking have a greater impact on performance in complex decision making?*
H_3: *Do certain sequences or patterns of systems thinking lead to better performance in complex decision making?*

The first hypothesis addresses the main objective of this study. The second and third hypotheses are related to the theoretical framework, namely, Richmond's Seven Systems Thinking skills.

This research contributes to the fields of systems thinking and complex decision making. Specific contributions include a novel research methodology and, in particular, operationalization of the systems thinking paradigm, as well as identification of disaggregated factors affecting complex decision making.

Systems thinking paradigm

Systemic thinking is rooted in cognitive processes. In this study, we adopt the notion of systems thinking as a *paradigm*. This refers to systems thinking as a "world view"—seeing things holistically and interconnected. However, translating the systems thinking paradigm into "measurable" elements has remained a research challenge. As an attempt to operationalize systems thinking, Richmond (1993) proposed a set of "thinking skills" and later (1997a) added to and further defined these thinking *skills*. To date, these still remain the sole "operational" guide to systems thinking.

Richmond (1997a) suggests that systems thinking requires operating on at least seven thinking tracks simultaneously. His updated seven thinking skills are shown below.

1. Dynamic thinking
2. System-as-cause thinking
3. Forest thinking

4. Operational thinking
5. Closed-loop thinking
6. Quantitative thinking
7. Scientific thinking

He stipulates that the numbering and consequently the sequence of the seven thinking skills is important as this serves as a *process* for using systems thinking, with each thinking skill building on the previous one. As skills 6 and 7 are primarily relevant to system dynamics modeling efforts, in this study we focus on the first five skills.

Dynamic thinking

Dynamic thinking is essentially a mental application of the behavior-over-time graph. It allows a problem or issue to be framed in terms of a pattern of behavior over time. It means that one needs to put a current situation in the context of time scale—"The trajectory should thus have a historical segment, a current state and one or more future paths" (Richmond 1997b, p. 6).

System-as-cause thinking

System-as-cause thinking builds on dynamic thinking. This thinking enables the determination of plausible explanations for the behavior patterns identified with dynamic thinking. System-as-cause thinking "holds that relationships that are not under the control of decision makers within a system should be eliminated from consideration" (Richmond 1997c, p. 6). Essentially, this perspective means viewing a system's behavior as the result of the system and as such under the control of decision makers.

Forest thinking

Forest thinking is seeing the "big picture". "Forest thinking gives us the ability to rise above functional silos and view the system of relationships that link the component parts" (Richmond 1997d, p. 6).

Operational thinking

Operational thinking seeks to identify causality—determining how behavior is generated. Generally people have a tendency to think "correlationally" or to think in terms of "influence". Operational thinking looks at the structure or "physics" of relationships, at *how* one variable affects another; not just that they affect each other. Operational thinking helps to recognize the notion of interdependence; that generally within a system, there is a web of relationships (Richmond 1998).

Closed-loop thinking

Closed-loop thinking helps to identify closed-loop structures. It maintains that causality does not run in just one direction, but rather that an "effect" usually feeds back to change one or more of the "causes", and that the "causes" themselves affect each other. It is important as part of closed-loop thinking not to prioritize "causes" as being most or least important, but rather to understand how dominance amongst them may shift over time (Richmond 1997a).

Systems thinking and complexity

Systems thinking is purported to be highly germane for dealing with complex systems and problems. There is a widely held view that systems thinking is superior to other approaches in dealing with complexity (Richmond 1993). In Checkland's words, it is "the use of a particular set of ideas, systems ideas, in trying to understand the world's complexity" (Checkland 1981, p. 3). It is also argued that today systems thinking is needed more than ever as we are being overwhelmed by complexity (Senge 1990).

As is clear from these comments, systems thinking has increasingly been accepted as a response to complexity because our default understanding in complex situations does not lead to adequate actions (Schaffernicht 1999). However, despite the accepted value of systems thinking for dealing with complex systems, most individuals appear to have a great deal of difficulty thinking systemically. "We've grown up in a reality in which 'local' perspectives enabled us to do just fine, we have developed certain 'habits of thought' which make it difficult to learn in an interdependent reality" (Richmond 1994a, p. 213).

Numerous studies illustrate non-systemic thinking by individuals confronted with complex problems. Explanations of participant behavior "reflect an 'open-loop' conception of the origin of dynamics, as opposed to a mode of explanation in which change is seen as arising from the endogenous interactions of decision makers with their environment" (Sterman 1989b, p. 336). In addition, it has been found that people are insensitive to feedback and underestimate time lags between action and response (Sterman 1989b). This insensitivity to feedback "reflects a failure on the part of the decision maker to assess correctly the nature and significance of the causal structure of the system, particularly the linkage between their decision and the environment" (Sterman 1989a, p. 324).

Call for empirical research

Despite the wide acceptance that systems thinking is highly effective for dealing with complexity, there have been calls from within the field for empirical substantiation of this assertion:

> Many claims have been made concerning the ability of systems thinking interventions to change the nature and quality of thought about complex systems. Yet, despite the increasing number of interventions being conducted in both educational and corporate settings, important questions about the relationship between systems thinking and basic cognitive processes such as problem solving, decision making, . . . remain unanswered. (Doyle 1997, p. 253)

Huz *et al.* (1997, p. 150) have raised similar issues. Further, Cavaleri and Sterman (1997, p. 171) observe that "The relationship between the use of systems thinking and organisational performance remains the province of anecdote rather than rigorous follow up research". More recently, Delauzun and Mollona (1999, p. 364) added their voice: "There has been some concern about the scarcity of reported studies dealing with the actual contribution of system dynamics in enhancing effectiveness or productivity". Clearly, there is an influential body of researchers who recognise that a gap exists with regard to empirical research on the effectiveness of systems thinking.

In recognition of this gap, two articles (Huz *et al.* 1997; Cavaleri and Sterman 1997) addressing the question of the effectiveness of systems thinking interventions were published in 1997 in *System Dynamics Review*. Huz *et al.* (1997) repeated a group model building intervention in four counties in New York State. Four control counties were also selected and all eight counties were observed via pre- and post-intervention measures. The report of their pilot study concludes that the participants' perception of the intervention was that it was productive and worth while and that there were significant shifts in participants' goal structures and change strategies. There was also greater alignment of participant mental models and greater understanding of system structure and behavior.

In another study, Cavaleri and Sterman (1997, p. 171) reported on a "follow-up evaluation of a well known systems thinking intervention designed to improve quality and performance in the claims unit of Hanover Insurance". They found that subjects reported a much greater awareness of their thinking and changes in their behavior, which they attributed to the intervention. Subjects went on to say that their management style became more "systemic" and that this helped in the design of new policies.

To summarise, most of the studies thus far have focused on systems thinking interventions rather than the effectiveness of systems thinking skill (*paradigm*). The study of complex problems has shed little light on the issue of what attributes or skills are best for dealing with such problems. There are few conclusive findings and no established theories. Some consistent characteristics have emerged, however, amongst good and poor performing participants, that show interesting parallels to systems and linear thinking respectively.

In conclusion, there is a curious gap in both systems thinking and complex decision making fields. The gap concerns empirical studies on the value of subscribing to a systems thinking paradigm when faced with complex problems. Thus, despite some rigorous research, the absence of theories on the nature of systems thinking and its causal relationship with complex decision making persists in the literature. This article seeks to address that gap.

Research hypotheses

The aim of the research is to investigate empirically the postulate that systems thinking is effective in dealing with complexity. More broadly, the research questions are related to studies of human cognition (Funke 1991). In this context, systems thinking falls within the area of "cognitive abilities", where task performance is the dependent variable in the complex problem solving/ decision making process. As stated earlier, the research attempts to address three hypotheses:

H_1: *Does (more) systems thinking lead to better performance in complex decision making?*
H_2: *Do certain (disaggregated) dimensions of systems thinking have a greater impact on performance in complex decision making?*
H_3: *Do certain sequences or patterns of systems thinking lead to better performance in complex decision making?*

As stated earlier, no other theoretical guides to assist in the operationalisation of systems thinking were found in the literature. Therefore, in this paper, Richmond's five thinking skills collectively represent subscribing to a systems thinking paradigm.

Thus, H_2 and H_3 are based on disaggregating or "splitting" of the systems thinking paradigm into five constituent skills. H_2 asserts that certain elements of systems thinking are perhaps more relevant and more effectual for performance on complex problems than others. In addition to the disaggregation notion, the impetus for H_3 came from Richmond's (1997a) natural *sequence* to thinking skills, which, he asserts, is essential for complex problem solving. Hence, H_3 postulates that a *sequence* or pattern of use may exist amongst the five skills, which would lead to better task performance.

Research methodology

This research is concerned with exploring individuals' thought processes in complex decision making. The research method adopted in this study is Verbal Protocol Analysis (VPA), a methodology long used in psychology (Ericsson and Simon 1993). A *protocol* is an audio record of the thought processes. Protocols are generated by getting participants to verbalize their thoughts while performing a specific task. This process, known as "think aloud", is one of the most promising techniques in studies of dynamic decision making (Doyle 1997). Video recordings can also be used to garner greater insight through non-verbal indications. Here, VPA is used in connection with a microworld representing a complex problem. While performance data is gathered through the microworld, behavioral and contextual information is provided through observations using VPA.

VPA "has been used extensively as an effective method for in-depth examination of cognitive behaviors" (Schenk *et al.* 1998, p. 32). It captures what occurs between the introduction of a stimulus and the measurement of outcomes, a capability beyond the traditional input–output methodologies. VPA "provides access to what information is examined, the manipulations conducted on the input stimulus and, additionally, what evaluations or assessments are made by the problem solver". As such it "provides the greatest data richness and information value per data point" (Todd and Benbasat 1987, p. 496). In general, verbal reports are a valuable and thoroughly reliable source of information about cognitive processes (Ericsson and Simon 1993).

Participants and sample size

Participants for the study were ten business school graduate students undertaking courses in operations management. The ages of the participants ranged from 21 to 25. Participation in the experiment was voluntary and involved a period of two hours. All the participants had systems thinking training. Of the ten participants, four had completed two university courses in systems thinking and five were in the process of completing the second course. All participants had taken the courses with the same two instructors, ensuring consistency in their systems thinking training. We note that, while training in systems thinking does not necessarily lead to its practice, nevertheless, this selection of participants allowed us to control for systems thinking knowledge.

While considered small in conventional research methodologies, the sample size of ten is within the norms of the VPA. Owing to the high density of data that is found in a single verbalization, VPA samples are typically between two and twenty (Todd and Benbasat 1987). Furthermore, as the VPA process is very labor intensive, the majority of research studies have utilized small sample sizes. Nevertheless, this research takes an exploratory stance towards the research questions and the broader area of investigation.

Microworlds as complex systems

Simulations (microworlds) are perhaps the most widely used tools for representing complex problems (Brehmer 1992; Buchner 1995; Dörner 1980; Funke

1988; Sterman 1989b). Microworlds are *complex* in that they require subjects to consider many different and interacting elements—including possibly conflicting goals. They are *dynamic* as they involve a series of connected decisions, where current decisions are constrained by earlier decisions and the state of the problem changes as a function of the decision maker's actions. Lastly, they are *opaque* in that they do not reveal all their characteristics automatically to the subject, thus requiring him/her to form and test hypotheses about their state and characteristics (Brehmer 1992).

Simulation task

The microworld used in this research represents a fictitious computer technology company called *Computech*; (the *Computech* simulation is part of the *Next Step* CD-ROM, a product of *High Performance Systems, Inc*. 1996). The simulation task requires the participant to act as the CEO for five years made up of 20 quarters. The participant (CEO) can manipulate five levers (decision variables): total sales force headcount, average sales compensation, marketing spending, average price per unit, and capacity order. There was no time limit set on the task but participants typically took around 30–35 minutes to complete the task.

Performance in the simulation was assessed by three objective measures: revenue, profit (as a percentage of revenue), and market share. Participants had to make a decision each quarter using many, few, or none of the five levers at their disposal. All three objectives had specified targets. For example, revenue was to reach $40 million from the starting level of $4 million by the end of the simulation.

Data collection procedure

Data was gathered during the experiments while participants undertook the task. Participants were asked to verbalize their thinking as they worked on the task—"think aloud". The verbalizations were recorded using audiotape. A data collection protocol was created prior to the commencement of data collection. The objective of the protocol was to maintain consistency in data collection. Thus, everything that would be said and done, and the sequence in which it would happen was "scripted" in the data collection protocol.

Before the experiments were started, four practice verbalization exercises, based on Ericsson and Simon's (1993) model, were held to ensure that subjects learn to think aloud as distinguished from "explaining". This is important so that the activity of verbalising does not interfere with the ongoing decision making process.

Coding process

The data collection procedure resulted in a verbal protocol for each participant. Once participants' verbal protocols had been recorded (on audio tape) and transcribed, they were then prepared for coding. For coding purposes the protocols were parsed into quarters. As mentioned earlier, the simulation ran for five years resulting in a total of 20 quarters.

Within each quarter, the transcript was further parsed into "thought" fragments (statements)[1] where each fragment represented a codable unit. Consequently, the quarters had fragments of varying lengths in terms of the number of protocol lines contained in each. On average, statements contained ten words and the protocols averaged 358 statements each. Once the transcripts were parsed into "thought" fragments, they were in a form ready to be coded.[2] The coding of a transcript involved assigning each of the fragments to one of the nine category codes in the coding scheme.[3]

The coding process presented particular challenges, as became apparent from the pilot analysis. The main challenge concerned the overlap amongst the five different systems thinking skills. As mentioned, the purpose of coding was to assign each fragment a single code from the coding scheme. However, because of the definitional overlap within systems categories, while some could be appropriately captured in a single fragment, other categories such as forest thinking could sometimes be evidenced over a series of fragments. This presented a problem as a codable fragment represented the unit of analysis. Thus, if some types of thinking skills were evidenced over a series of fragments, it would be very difficult to create a separable fragment for comparison. This dilemma raised the need for multiple codes for fragments as some fragments could also individually reflect other thinking types, suggesting that they could be assigned more than one category code. This was unacceptable, as it would undermine the uniqueness of fragments for comparative analysis.

To address this problem, a new ranking system was devised for systems thinking categories to eliminate the need for multiple codes. The new ranking system has its theoretical base in Richmond's classification, which denotes interdependence amongst the thinking types. Richmond's (1997a) classification is based on the premise that systems thinking classes unfold in sequence when one approaches a complex problem, implying that their effect is cumulative.[4] With Richmond's *cumulative* notion, the following study ranking system emerged, where 1 denotes the highest (most comprehensive) thinking category and 5 the lowest (most detailed):

1. Forest thinking
2. Closed-loop thinking
3. Operational thinking
4. System-as-cause thinking
5. Dynamic thinking

As discussed earlier, the ranking scheme serves to overcome the problems of multiple codes in that evidence of a higher thinking type implies that lower ranked thinking has also been undertaken. For example, evidence of forest thinking would imply that some or all of the four lower ranked thinking categories have also taken place. Thus, if a fragment exhibits two types of thinking, it is coded with the higher rank of the two thinking categories.

Process and performance measures

Following the coding process, the codes generated from the subjects were converted to percentage frequencies for analysis. This procedure normalises the protocols and allows for comparison among subjects, as not all protocols are of equal length and hence do not contain the same number of fragments. For example, if there were 56 fragments coded with dynamic thinking and a total of 380 fragments in a protocol, then a percentage frequency would be 56/380 = 14.74 per cent. The resulting measure called the *percent frequency measure of occurrence* (Schenk *et al.* 1998) is used as the basis for the quantitative analysis. This measure gives an approximation of the "relative amount of time or energy devoted to an activity" (Pennington *et al.* 1995, p. 180) and is consistent with measures of time and effort used in prior studies (Irwin and Wasalathantry 2000).

As stated earlier, participants were evaluated by their performance on three objective measures: revenue, profit, and market share. Each "raw" performance score was translated into a score that reflected "closeness to goal". For example, if in a quarter a subject achieved market share of 20 per cent this would be divided by the target of 25 per cent, giving 80 per cent. This process was also carried out for the other two performance measures. All three performance measures were considered when determining relative participant performance.

Task structure understanding

In order to determine how well the participants understood the structure of the task and their relationships (system), another measure was developed to capture what the participants thought or *said* and to correlate these to their performance. This is distinct from other measures, which captured what the subjects *did*. This "system understanding measure" was developed based on the relationships within the task system. Table 1 provides a listing of these relationships.

As can be seen in Table 1, the Task Understanding Measure contains four levels. The relationships contained in the system (microworld) were grouped in terms of the nature and complexity of the relationship. This grouping is

Table 1. Definitions of task structure understanding measure

Level 1 (Basic one-to-one relationships—largely intuitive)
1. Price increases, booking rate decreases 2. Price increases, revenue increases 3. Capacity ordering increases, expenses increase 4. Marketing spending increases, booking rate increases 5. Marketing spending increases, expenses increase 6. Sales compensation increases, expenses increase
Maximum score /6 (1 point per relationship)
Level 2 (Complex one-to-one relationships)
1. Sales force increases, sales compensation increases 2. Order booking rate increases, revenue increases 3. Sales force increases, booking rate increases 4. Sales compensation increases, booking rate increases
Maximum score /8 (2 points per relationship)
Level 3 (Three-way relationships)
1. Capacity decreases, booking rate increases, lead time increases 2. Revenue increases, expenses decrease, profit increases 3. Sales force increases, booking rate increases, market share increases 4. Price decreases, booking rate increases, market share increases
Maximum score /12 (3 points per relationship)
Level 4 (Big picture)
1. Understanding that lead-time is the balance between capacity and order booking rate 2. Understanding that price and sales people balance the order booking rate
Maximum score /8 (4 points per relationship)
Total potential score for understanding/34

analogous to systems thinking categories representing a hierarchy and ranking of relationships. Level 1 represents the basic one-to-one relationships in the system. These relationships are considered to be largely intuitive, and it would not require much, if any, time to explore the task of determining them. Next are the Level 2 relationships, which are also one-to-one relationships, but more complex as they are less intuitive and require some understanding of the system. Level 3 relationships are three-way relationships where one factor impacts on another, which in turn affects a third. These relationships are more complex, and require at least some Level 1 and/or Level 2 understanding. Finally, Level 4 represents higher-level relationships as they encompass "big picture" views and multiple variables. Understanding here requires comprehension, at least to some extent, of the relationships in the other three levels.

Again, this implies that in order to understand higher-level relationships, there needs to be some understanding of the lower-level relationships.

The actual measurement of task understanding was done quantitatively. Each relationship was allocated a maximum score. As can be seen in Table 1, the relationships at the different levels have different maximum scores, relative to their level. This scoring assumes a non-linear function in that a participant who captured a "Level 4" relationship had greater understanding than one with understanding of three "Level 1" relationships.

The actual scores were calculated from the protocols by giving the subjects the allotted score according to definitions outlined in Table 1. A total score was then calculated for each participant out of a possible score of 34.

Analysis and results

Overall, the results pertaining to H_1 did not support a *simple* relationship between the level of systems thinking and task performance. The second hypothesis (H_2) postulated that certain types of systems thinking play a more important role than others in affecting performance. The third hypothesis (H_3) tested Richmond's notion that systems thinking categories unfold in sequence. In relation to these hypotheses, a notable trend emerged from the analysis. As a single thinking type, better performers consistently undertook greater forest thinking. What is more interesting, however, is that better performers had utilized more of the three higher ranked thinking types cumulatively. The proportion of each participant's systems thinking that accounted for the top three types, namely, forest, closed-loop, and operational thinking is: 45.33, 33.96, 28.26 and 16.5 per cent, ranked by performance. This suggests that particular *types* of systems thinking could be more related to better task performance.

This supports H_2—the notion that higher ranked thinking types (forest, closed-loop, and operational) would contribute more to the *understanding* of a system and therefore play a greater role in performance. In contrast, lower ranked thinking types (dynamic and system-as-cause) would be expected to be more beneficial in procedural contexts. Thus, they would not be of substantial help in the understanding of the system structure, and consequently not significantly affect performance.

These findings begin to shed light on how one participant, despite having the lowest overall amount of systems thinking (13 per cent), ranked ahead, in task performance, of another participant with a systems thinking score of 30.29 per cent. The former had dedicated a significant amount of his systems thinking to the higher-level types (28.26 per cent) as compared to the sixth ranked participant who had only spent 16.5 per cent of his time on these types. This finding is particularly significant, as it posits that the amount of systems thinking alone does not affect performance, but rather it is the degree of high-level (forest, closed-loop, and operational) systems thinking that counts most.

The utility of systems thinking is further hypothesised to contribute to the understanding of the structure of a complex, dynamic, and opaque system. Thus, high-level systems thinking types are expected to facilitate this to a greater extent than the lower-level types.

However, because of the opaque nature of complex problems, any understanding must be developed over time and thus it denotes a *gradual* and *cumulative* process. This understanding is then utilised to develop strategies that improve performance. This purports that the use of systems thinking would correlate directly with better understanding of system structure, rather than directly with task performance.

In summary, the research offers some evidence to support the hypothesis (H_2) that particular types of systems thinking have greater impact on performance than others. More interestingly, our research findings suggest that systems thinking does not affect performance directly, but it rather affects *understanding*, which can then lead to better task performance. Therefore, systems thinking and task performance are unlikely to correlate *directly* even at a disaggregate level (i.e., individual systems thinking types).

Systems thinking transition patterns

This section addresses H_3, the hypothesis that postulates whether patterns or sequences of systems thinking types have any bearing on performance. To observe any possible recurrent patterns in the type of systems thinking carried out, transition graphs were constructed for each subject's protocol. Transition graphs illustrate shifts amongst different thought processes during a protocol. They show along a time line what type of thought processes the subject was engaged in at various points in time. This allows the researcher to compare visually the protocols of the different participants and identify any consistent patterns that are evident. Transition graphs have been used by other researchers for similar analysis of protocols (see, for example, Irwin and Wasalathantry 2000; Srinivasan and Irwin 1999; Srinivasan and Te'eni 1995).

The premise being examined here is that better performing participants may display a different *pattern* of systems thinking throughout the simulation, or over a series of quarters, than poor performing subjects. This attempts to explain performance not only in terms of quantity and type of systems thinking but also in terms of how thinking patterns are linked together and of the sequence of their occurrence. Each transition graph illustrates *every* statement contained within a subject's protocol. The y-axis shows the systems thinking types (level)[5] while the x-axis contains a scale from 0 to 100 per cent, reflecting the cumulative percentage of statements or fragments in the protocol.

The transition graphs illustrate shifts amongst the five different systems thinking types as well as occurrences of non-systems thinking during each protocol. Gaps in the graphs—when there are no bars—reflect one of the

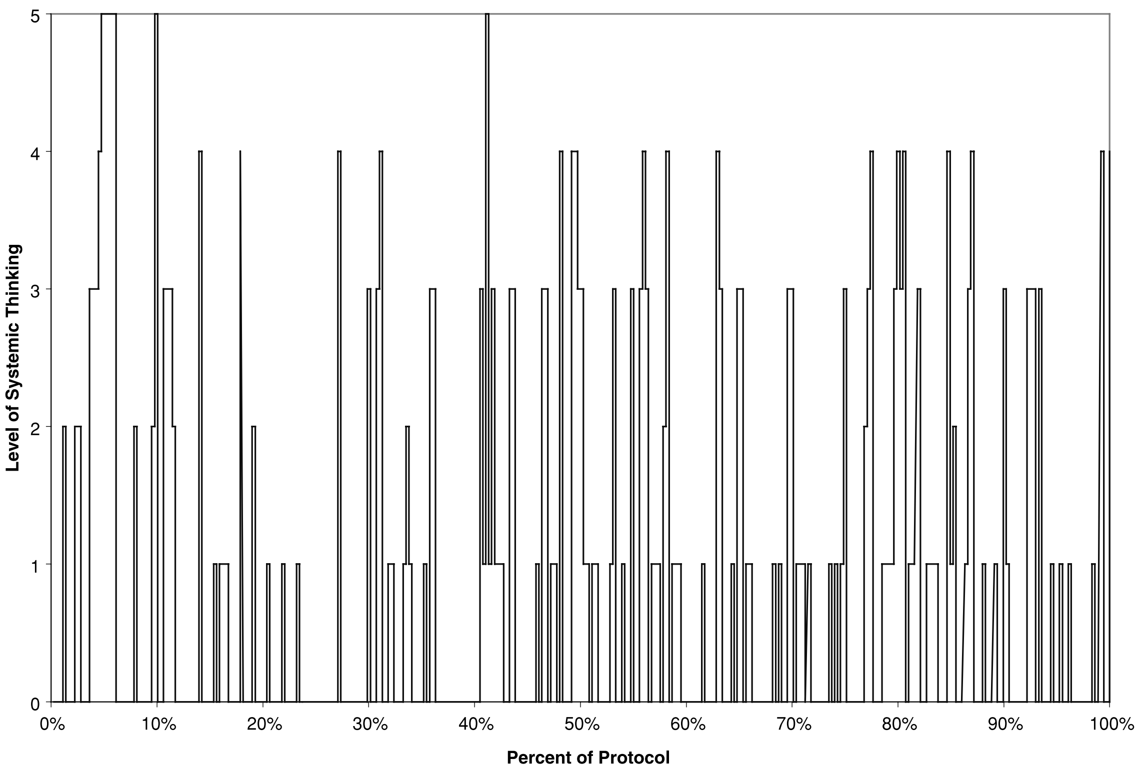

Fig. 1. Systems thinking transition graph for subject IF

non-systems thinking categories, e.g. reflection, motor, etc., as these categories were allocated a ranking of zero.

Again, the findings show consistent patterns overall. Better performing participants repeatedly exhibited transition across multiple levels (especially Levels 3, 4 and 5) throughout the protocol. For example, subject IF's graph (Figure 1) shows transitions throughout the protocol across all five levels. Poor performing participants, on the other hand, display sustained periods at low levels and little or no high-level thinking, unlike their better performing counterparts. In addition, the transition graphs of poor subjects show (Figure 2 for example) many gaps, indicating that no systems thinking took place during these segments of their protocols.

These findings add further support for H_2. The transition graphs correlate well with the findings pertaining to the level and quantity of systems thinking types done by various participants. The results suggest that better performers transited across all five systems thinking levels and did so repeatedly throughout the simulation. Better performers also displayed greater time spent at higher levels of systems thinking on the transition graphs.

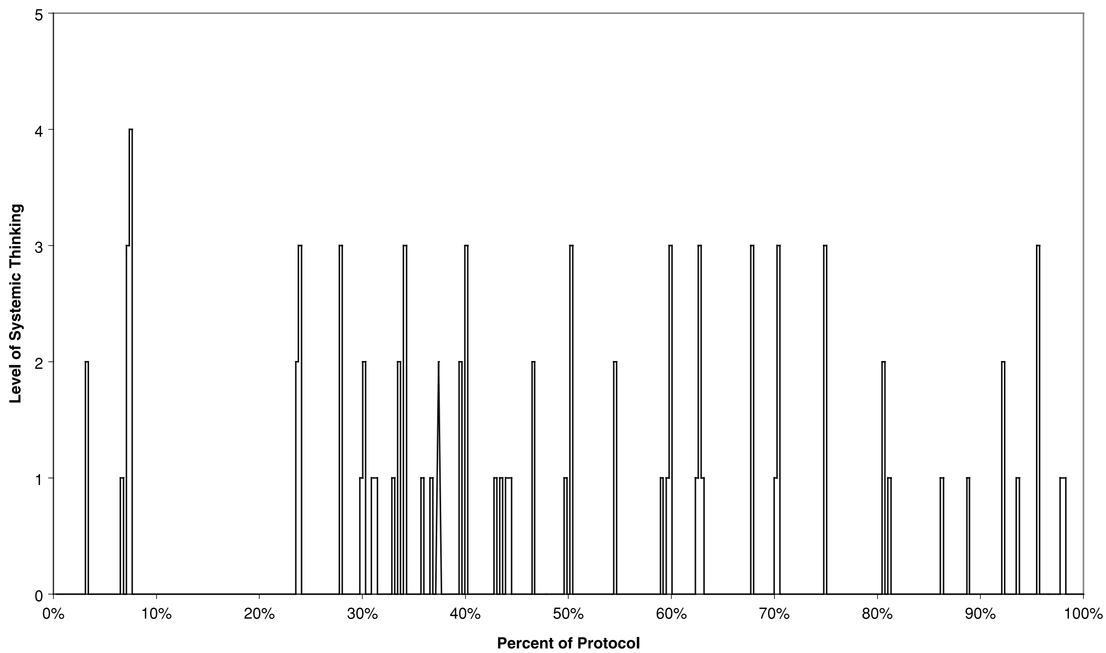

Fig. 2. Systems thinking transition graph for subject DM

The results for another participant (GJ) provide an interesting paradox. His transition graph (Figure 3) shows a sustained period at high levels, but only during the early part of the protocol. This subject, unlike the better performing participants, fails, after about the first 38 per cent of the protocol, to transit across the higher levels of thinking. These findings go some distance towards explaining why this subject did not perform better overall, as would have been expected, given the high-level of his systems thinking. However, while low transitioning may be a contributory factor in this subject's poor performance, it would appear not to be the only factor. The transition graphs for other participants are shown in Figures 4–6.

Decision making pattern

Thus far, the analyses have focused on the coded protocol statements, which revealed the role of systems thinking types and the distinctive patterns of thought processes utilized by various research subjects. Next, our attention was turned to another pattern that emerged when we studied the participants' behavior during the experiments. The focus also shifts here to looking beyond the systems thinking skills towards a more general decision making/problem solving approach.

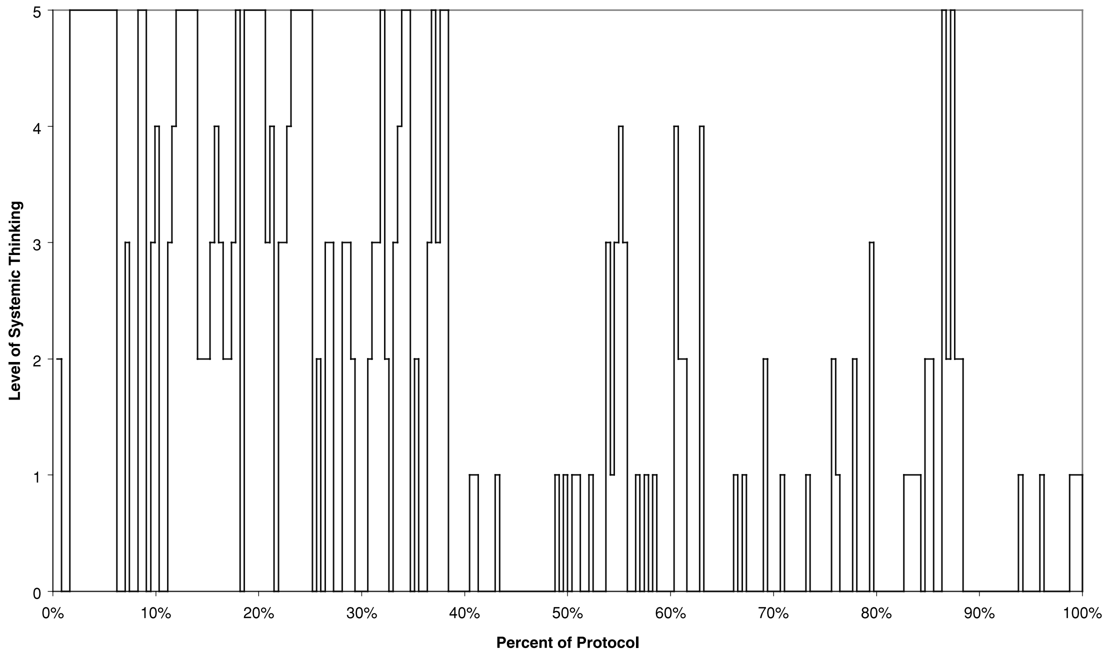

Fig. 3. Systems thinking transition graph for subject GJ

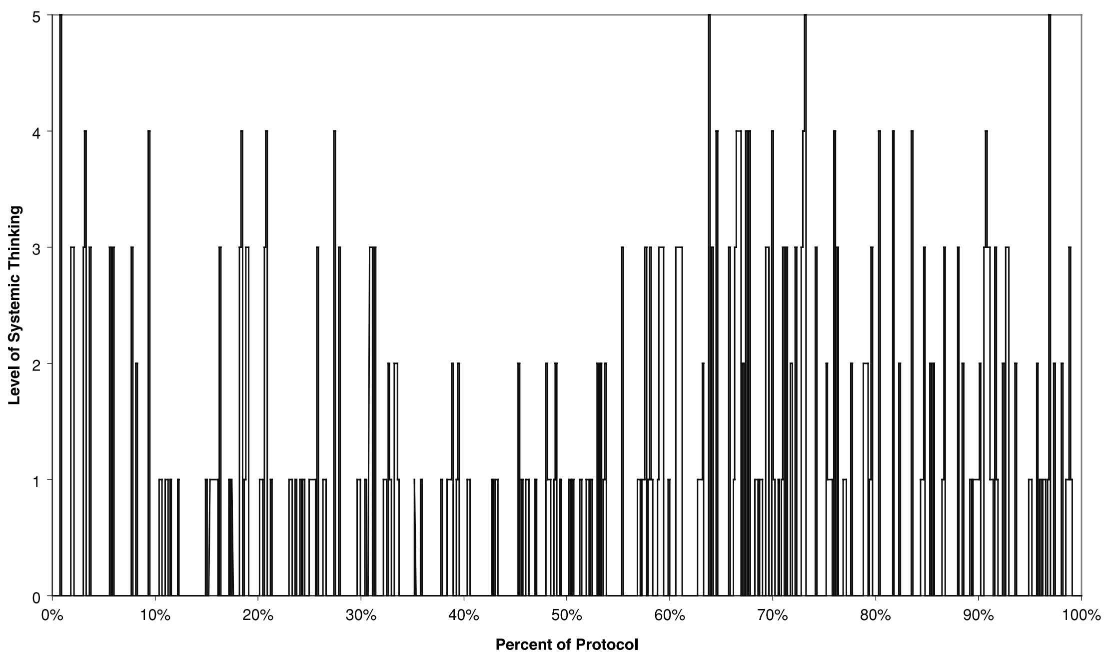

Fig. 4. Systems thinking transition graph for subject PL

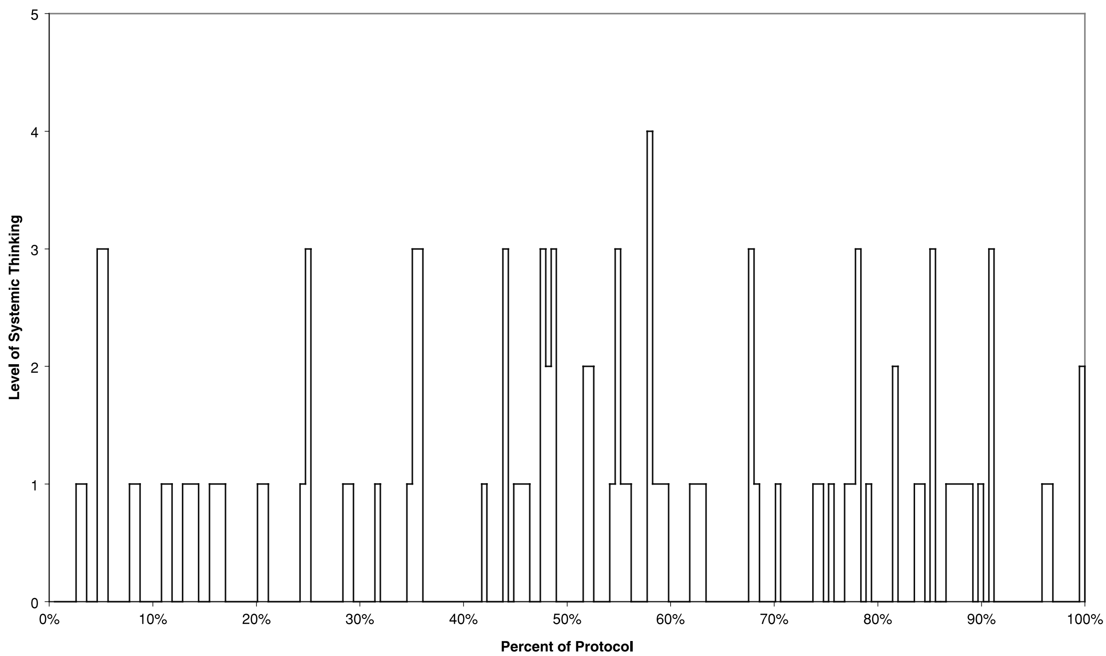

Fig. 5. Systems thinking transition graph for subject DC

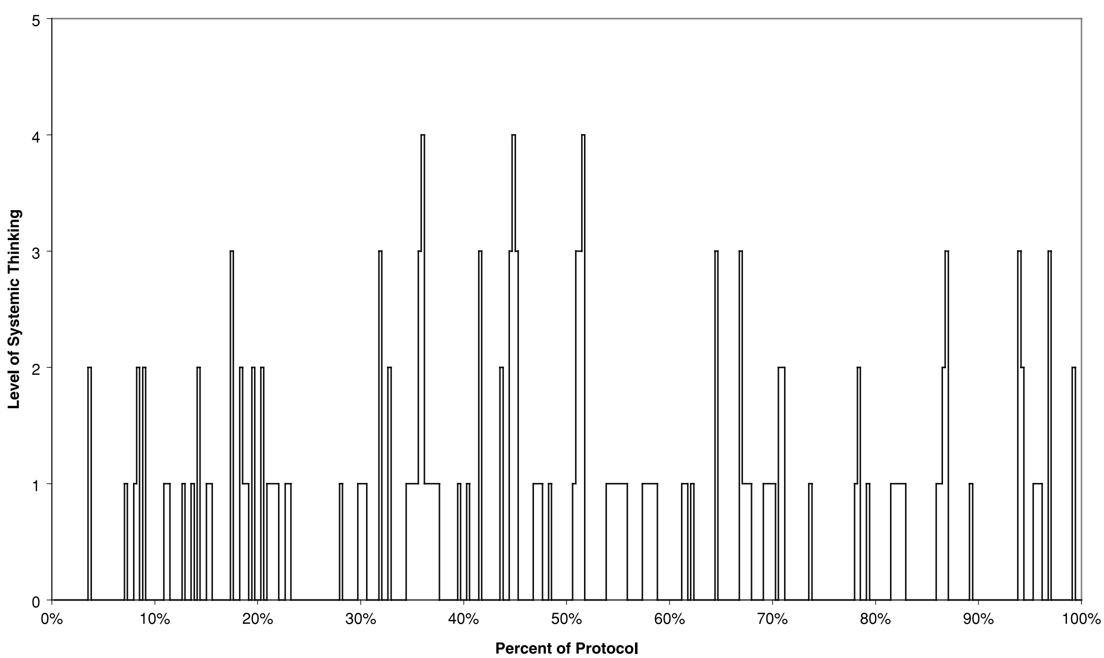

Fig. 6. Systems thinking transition graph for subject AM

To this end, the transcripts were read again and quarter-by-quarter summaries were created. This removed much of the minor detail contained in the transcript, producing in essence an "aggregate" transcript containing the salient points of each protocol. These summary transcripts were then analysed qualitatively for evidence of any patterns or structures. From this analysis it became evident that better performers displayed, in their decision-making approach, a distinct pattern that differed from that of the worse performers. In order to determine the validity of their understanding of the system, superior performers attempted to gain understanding of the system's structure first, develop their strategies next, and finally make decisions, carefully assessing the outcomes of these decisions. We have termed this pattern the CPA cycle.

The CPA cycle, which emerged from the analysis, has three distinct phases—*conception* (C), *planning* (P), and *action* (A). The conception phase of the cycle is where a subject would try to gain understanding of the structure of a problem. This is where systems thinking could be undertaken and would be of value. Examples of conception statements include:

GJ: "okay great, now first of all I'm going to try to see what and how affects each of the factors there"
IF: "first we've got to figure out what's going into demand"
PL: "yeah right the expenses are tracking up because I keep hiring more people"

The next phase is the planning phase. This is where the strategy is developed, ideally based upon the understanding gained in the conception phase. Examples of planning statements include:

PL: "first up sales people: to get more sales we are going to increase sales people"
DC: "we could probably look at reducing the price to get more customers . . ."
IF: "and I actually want to pay these guys a little bit more to give them a bit more encouragement to do something"

The third and final phase is the action phase. Here, the strategy developed is implemented using specific decisions. In the context of the simulation task, this amounts to interventions by adjusting the decision levers and executing the model. Each quarter will therefore end with an action phase. Examples of action statements include:

DC: "I'll hire another person . . ."
DM: "average sales compensation, increase it, by probably 20"
AM: "spend say, spend say $30,000 on marketing"

The action phase should ideally lead to another iteration of the cycle by leading directly onto the conception phase of a new cycle. Alternatively, there would be a "break" between an action phase and a conception phase. A "break" is possible in the cycle at any phase, which means that a new phase does not continue from the prior phase but, instead, the subject performs its function independent of what occurred in the preceding phase.

When an action phase leads directly to a conception phase, a closed loop is formed. The new conception phase will then begin by reviewing the outcomes (i.e., reflection, feedback) of the decisions just made in the previous action phase. The purpose of the review is to evaluate understanding of the system structure that was developed in the last conception phase. The outcome of the conception phase, after each iteration, is a more detailed or clear picture of the structure and the nature of the problem. Examples of reviewing statements as part of the conception phase are:

GJ: "less price should have affected booking rates but hasn't, marketing should have affected it a little and then no"
IF: "I increased sales compensation from 120 to 140 and I increased market spending as well, from zero to 100 and annual revenue per sales person went up from one million to 1.1"
PL: "my recent strategy has sent the revenue into a bigger nose dive. In fact, now I'm losing money and going backwards, . . . the revenues, the problem seems to be the level of expense"

The CPA cycle is particularly pertinent to complex decision making because the structure and behavior of such problems are opaque and intricate. Hence, it is necessary for a person to "build a picture" of the structure of the problem *incrementally*. This can be accomplished through an iterative process of conception, planning, and action. For the cycle to be most effective and value adding, it needs to be iterative. As one cycle is completed, the action phase should lead directly to the conception phase of a new cycle. This is necessary because, as discussed earlier, understanding of complex problems can often only be gained progressively. If the cycles are not iterative, then they can become largely ineffectual in gaining true understanding of problem structure.

To this end, the protocol summaries were analysed in a structured manner, to determine the extent to which participants followed the CPA cycle. The objectives of this analysis were twofold: (1) to see which stages of the cycle were evident in each quarter, and (2) to determine how many complete uninterrupted cycles each participant concluded. The results of this analysis are presented in Figure 7.

As can be seen in Figure 7, better performers, in addition to the amount of high-level systems thinking, were also following the CPA cycle more consistently. One of the key differences between good performers and poor performers was the number of completed cycles. This is illustrated by the number of quarters containing all three phases of the CPA cycle. The other difference was that good performers completed far more continuous iterations of the cycle than did the poor performers. In fact, poor performers often did not follow the

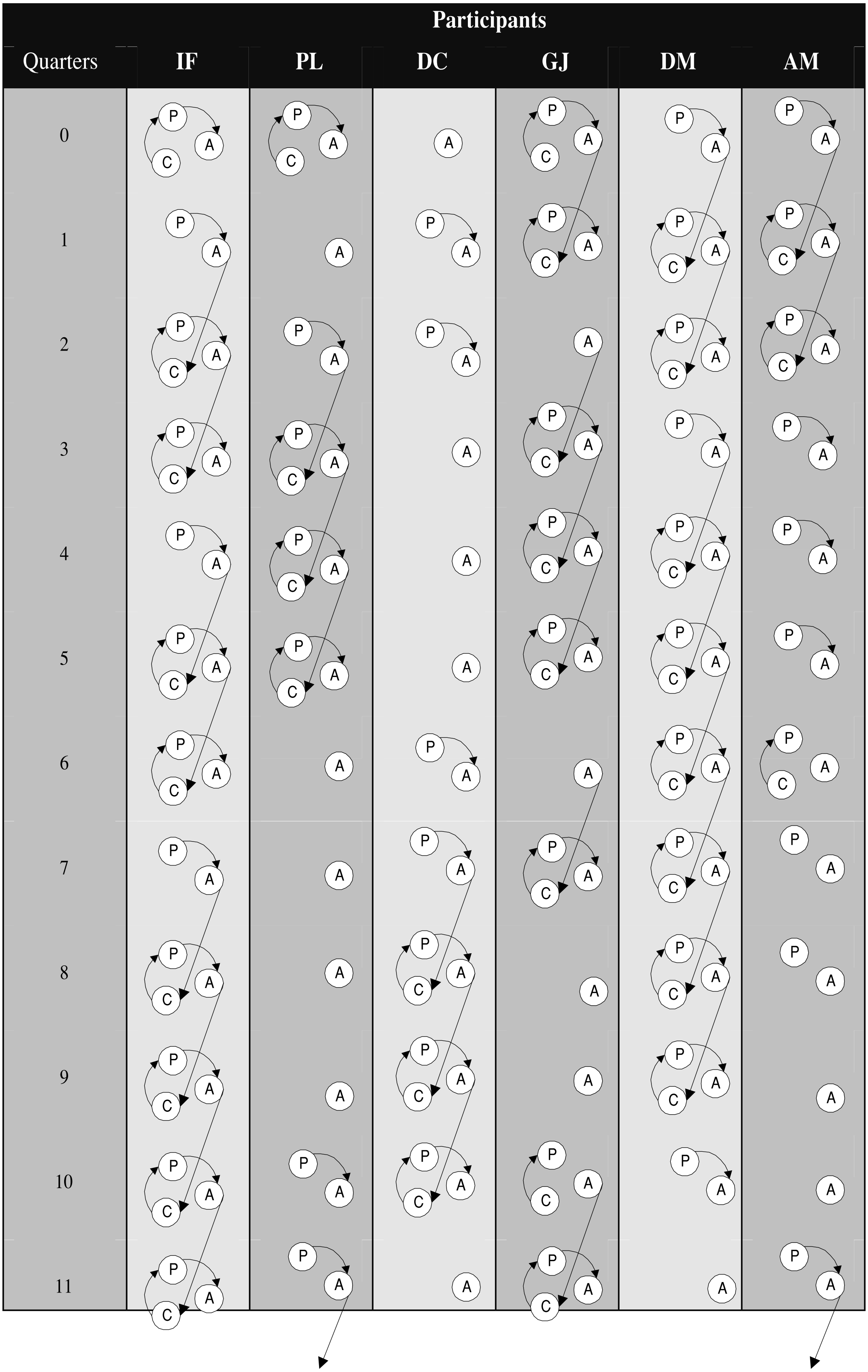

Fig. 7. CPA cycles of subjects' protocols

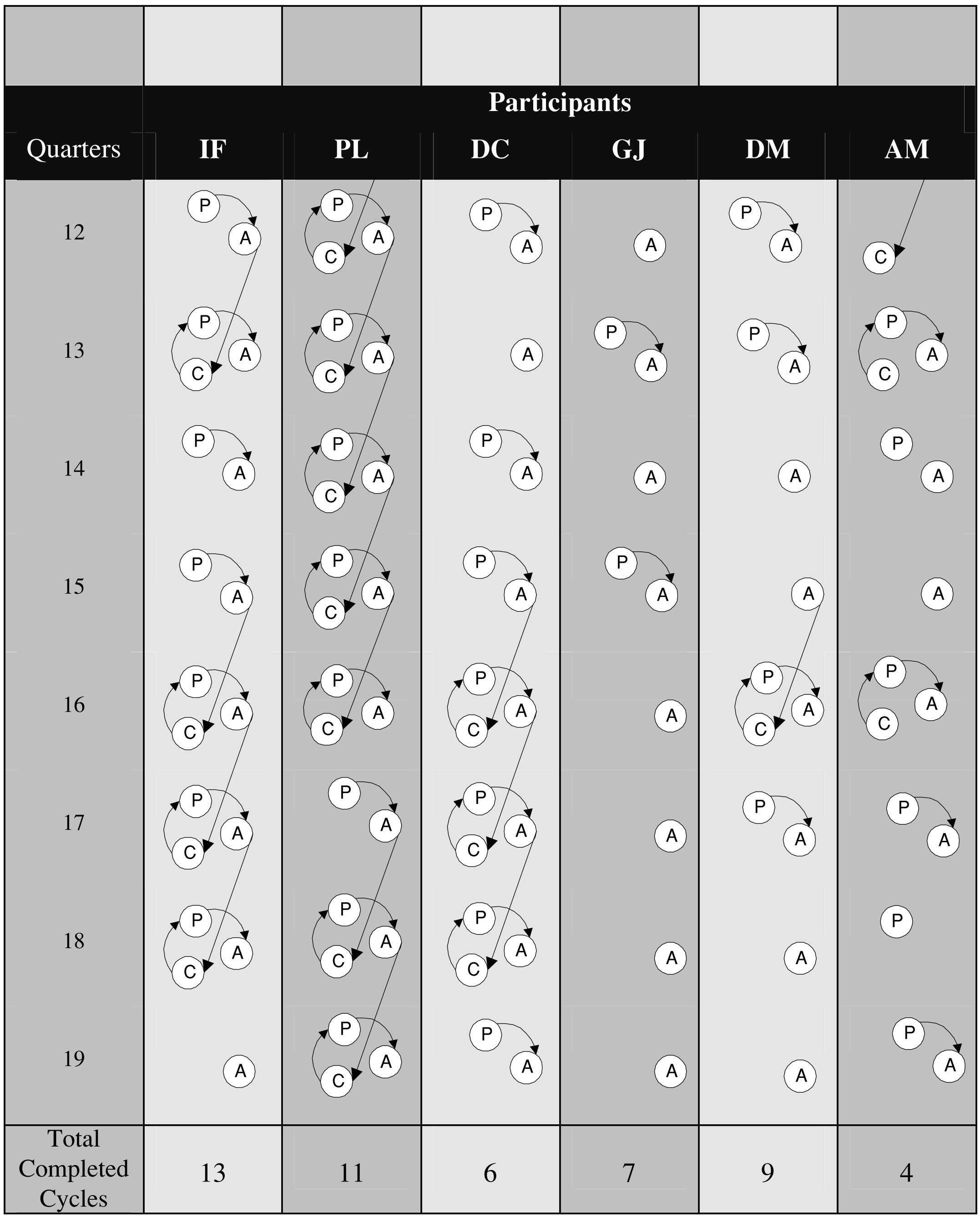

Fig. 7. (*Continued*)

cycle at all or would use it in a rather disjointed fashion. The continuous cycles can be seen by straight line arrows that span across quarters. These show the action phase from one cycle, linking to the conception phase of a new cycle which, as described earlier, would be necessary to gain deeper understanding of system structure.

These results help explain the paradox in participant GJ's performance. Completed and uninterrupted CPA cycles appear to be the deficient factor in his performance, despite the good understanding he demonstrated. From Figure 7, it is clear that until approximately quarter 7, GJ uses a chain of largely complete and uninterrupted CPA cycles, combined with large amounts of high-level systems thinking and sustained transitioning at high levels. However, following this, he fails to complete the cycles as the task progresses—the CPA cycles become principally dominated with just the action phase, which illustrates that decisions are not being planned and understanding is not being developed. The verbalizations shown below, made by GJ from quarter 16 to 19 inclusive, illustrate his failure to use the CPA cycle in his decision making.

Q16: "it drops, why, why people why are you leaving . . . umm okay sales are high, people are high, prices low, market spending I won't do anything now and I will just try it once more"
Q17: "still dropping"
Q18: "still dropping"
Q19: "it's dropping, I'll leave it the way it is because I screwed up badly, yeah"

It is evident that GJ stopped both the conception and planning stages of the cycle as he progressed through the simulation task. Although this finding is not conclusive in explanating why GJ began to utilize only the action phase, it nevertheless provides a significant insight into his overall poor task performance.

Summary

This article began with a simplistic (in retrospect) research question that the more systems thinking a person did, the better their task performance would be. While the overall amount of systems thinking explained the performance of some research subjects (e.g., IF and PL), this premise was not consistent across other participants. Most subjects, in fact, had done similar amounts of systems thinking overall.

Hypothesis two (H_2) shed more light on the issue by showing that high-level systems thinking, and particularly forest thinking, seemed to be highly related to task performance. These findings were consistent across all participants except GJ. GJ had almost the same amount of high-level systems thinking as the best performer IF, with considerably more forest thinking, yet ranked only fourth out of six in task performance. H_2 was further investigated at the disaggregate level where the ensuing findings were found to be inconsistent with those at the aggregate level. Two reasons are suggested for this inconsistency. First, correlating systems thinking to performance in a particular quarter was not feasible. Second, it appeared that systems thinking, and particularly high-level systems thinking, would affect performance *indirectly*, through an intermediary step of task *understanding*. This would indicate that the amount of systems thinking would not necessarily correlate directly with task performance, but rather with *understanding* of task structure.

In order to examine these postulates, the level of each participant's understanding of task structure was measured. As expected, participants who did more high-level systems thinking demonstrated deeper understanding of task structure. Subject GJ, for example, had the greatest comprehension of task structure, which was consistent with expectations. He had approximately the same amount of high-level systems thinking as the best performer IF with considerably more forest thinking—the thinking type that was expected to be most effectual in garnering task structure awareness.

Despite this further insight, a clear explanation had still not emerged to elucidate the findings thus far. What was revealed was that good performing subjects transited across all five systems thinking levels and did this throughout the entire simulation. In contrast, poor performing participants functioned at lower levels of systems thinking and had many gaps in their transitions, indicating no systems thinking.

At this point, research question three (H_3) was revisited with a slightly altered focus. This time the aim was to look for *general* patterns of behavior rather than that of systems thinking per se. What emerged from this analysis was the CPA cycle. The CPA cycle proved to be a reliable differentiator of good and poor performance as better performers completed many more cycles than did the poor performers. Additionally, good performers undertook chains of uninterrupted cycles enabling them to build their understanding of the task structure incrementally and thus perform better on the task. This analysis shed further light on the findings, offering a plausible explanation for contradictory results, such as those of the Subject GJ.

Thought Process Factors

Given the preceding analyses, the performance of individual subjects could be explained in terms of five *Thought Process* factors:

- the amount of overall systems thinking;
- the amount of high-level systems thinking;
- the understanding of task structure;
- the transitions between types of systems thinking;
- the number of complete CPA cycles.

Table 2. Summary of thought process factors affecting performance

Subjects in order of task performance	Amount of overall systems thinking (%)	Systems thinking devoted to high levels (%)	Understanding of task structure (/34)	Transitions between levels of systems thinking	Completed CPA cycles
IF	38.83	45.33	26	Consistent across all levels and throughout protocol	13
PL	31.43	33.96	21	Quite consistent across all levels and throughout protocol except for 1 sustained period at Levels 1–2	11
DC	35.57	20.29	17	Frequent gaps and sustained periods at Level 1 and 3	6
GJ	29.75	44.44	30	Across all five levels for first 40% of protocol, after which largely confined to Level 1	7
DM	13.03	28.26	14	Large gaps with transitions confined to Levels 1–3	9
AM	30.29	16.5	11	Transitions primarily between Levels 1 and 2 and occasionally Level 3, never Level 5	4

Table 2 presents a summary of each participant's five factors, ranked in the order of highest task performance (the dependent variable).

For example, Subject IF topped four of the five process factors and performed best overall. Also, an explanation of the performance of one participant relative to another could be put forward through variation in their thought processes. For example, subject PL performed better than DC on the simulation task despite having less overall systems thinking, because PL did more high-level systems thinking, had better task understanding, better transitioning, and completed more CPA cycles. Likewise, while DM had a considerably smaller amount of overall systems thinking than AM, his task performance was superior to AM. DM's superior scores in other factors could conceivably explain this. This explanation held true for all six participants. That is, their performance relative to each other could be explained by variances in the five factors.

Discussion

This paper investigates the postulate that systems thinking is effective in dealing with complexity. The findings suggest that this may not follow the simple notion that "the more systems thinking, the better the task performance". Our results, however, suggest that, although the overall amount of systems thinking does matter to some extent, the degree of higher-level systems thinking types performed (i.e., forest, closed-loop and operational) matters most. Further, the results show that participants who used all five systems thinking types and did this repeatedly throughout the simulation performed better.

This indicates that performance on a complex problem is an intricate and multi-dimensional process. The degree of high-level systems thinking, and the consistent use of all levels of systems thinking, throughout the decision making exercise appear to have the greatest impact on task performance.

As touched upon previously, in studies of complex decision making/problem solving, characteristics that differentiate good and poor performers display strong parallels to aspects of the systems thinking studied. That is, individuals who display the characteristics of systems thinking, even if they are oblivious to the fact, perform better on complex decision-making tasks.

In earlier studies, the behavior of subjects who performed well reflects the attributes of systems thinking, while the behavior of those who performed poorly often mirrors the direct opposite of systems thinking, or "linear thinking". For instance, Dörner (1980) observes "When solving such complex tasks, most people are not interested in finding out the existent trends and developmental tendencies at first, but are interested instead in the 'status quo' " (p. 91). This is a characteristic of static thinking, the polar opposite of dynamic thinking (Richmond 1997b).

Further, this study adds support to the notion of the "heuristic competence" construct. The construct is described as "a general competence for coping with complex systems" (Brehmer 1992, p. 223). Participants who display "heuristic competence" are described as those "who collect more information, who collect it more systematically, who construct adequate goals, who evaluate the effects of their decisions, and who generally behave in a systematic fashion . . . Subjects who behave in a way that makes it more likely that they will acquire a good model of the task also learn to control the task better" (Dörner *et al.*, quoted in Brehmer, 1992, p. 225). Thus, the heuristic competence behavior would lead to making fewer decisions, collecting more information before making decisions, and checking on results of decisions prior to making new ones.

Heuristic competence is highly analogous with the notions of systems thinking types and the CPA cycle that emerged in this study. The heuristic competence construct, however, is still loosely defined in the literature. The findings of this study propose a more clear description for the construct. Our findings suggest that systems thinking, when used in concert with the CPA cycle, manifests the characteristics of heuristic competence, as it involves understanding of the system structure, developing strategies, making decisions, and carefully assessing the outcomes.

Recent studies suggest that most organizational decision processes do not follow the above cycle, partly because of time pressures as well as established norms and mental models (Moss Kanter 2003). In contrast, in organizations, decisions are often made at the "event" level without adequate attention to patterns of behavior and even less to systemic structures underlying such patterns. This is a critical omission since, as our findings have suggested, understanding of systemic structures underlying organizational dynamics is a crucial prerequisite for the development of robust strategies. This highlights the need for group learning models such as microworlds and learning environments that utilize the elements of the CPA cycle in their conception and pedagogy.

Conclusion

The findings of the study suggest that the notion that systems thinking leads to better decision making and hence to superior performance is rather simplistic. Our experience indicates that the story is much more convoluted. While the degree of systems thinking does matter, the results suggest that certain *types* of systems thinking may be more relevant to superior performance. The "type" of systems thinking, however, is not the only factor responsible for performance in complex problems. In the study reported here, the subject's approach to the problem also appeared to be highly pertinent to task performance, as better performing subjects first attempted to gain understanding of the system structure, then developed and implemented strategies, and carefully assessed the outcomes of their decisions, in order to determine the validity of their understanding of system structure—the CPA cycle. This cyclical pattern is important, because, as a result of their intricate nature, understanding of complex problems appears to develop progressively.

This exploratory research paves the way for further empirical research on systems thinking's nature, role, and effectiveness in complex decision making. A notable contribution of this study to the systems thinking literature is its novel research approach and the methodology developed for operationalizing the systems thinking paradigm.

While this research has shed light on the causal links between systems thinking and complex decision making, a number of fresh questions have come to surface in this study. These include:

- To what extent the systems thinking types individually contribute to task performance.
- What combinations of factors best assist with superior task performance? This is significant as there may be some combinations of factors that would lead to superior task performance.

While the relatively small number of subjects in this study is within the acceptable sample size of the VPA Methodology, the finding of the study should be viewed as suggestive and exploratory. Further research is called for to expand on this critical field of study.

Notes

1. The terms "statement" and "fragment" will be used interchangeably to refer to a codable unit within a subject's verbal protocol.

2. Details of the coding process and results are available from the authors upon request.
3. Only six of the ten verbal protocols generated were coded. One was a pilot subject who was used for refining the procedures and coding scheme. Three other protocols were removed due to poor verbalization and lack of speech clarity for transcription.
4. While this notion may not be viewed as established theory, nevertheless it proved constructive in operationalizing systems thinking because it reflects the relative value that each type is thought to contribute to the decision making process.
5. Here the y-axis represents the *level* (type) of systems thinking, where 1 = dynamic, 2 = system-as-cause, 3 = operational, 4 = closed-loop and 5 = forest thinking.

References

Booth Sweeney L, Sterman JD. 2000. Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review* **16**: 249–286.

Brehmer B. 1992. Dynamic decision making: human control of complex systems. *Acta Psychologica* **81**(3): 211–241.

Buchner A. 1995. Basic topics and approaches to the study of complex decision making. In *Complex Decision Making. The European Perspective*, Frensch PA, Funke J (eds). Lawrence Erlbaum: Hillsdale, NJ, 27–63.

Cavaleri S, Sterman JD. 1997. Towards evaluation of systems thinking interventions: a case study. *System Dynamics Review* **13**(2): 171–186.

Checkland P. 1981. *Systems Thinking, Systems Practice*. Wiley: Chichester.

Delauzun F, Mollona E. 1999. Introducing system dynamics to the BBC World Service: an insider perspective. *Journal of the Operational Research Society* **50**: 364–371.

Dörner D. 1980. On the difficulties people have in dealing with complexity. *Simulation and Games* **11**(1): 87–106.

Dörner D, Kreuzig R, Standel T. 1983. Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität. Bern: Huber.

Doyle JK. 1997. The cognitive psychology of systems thinking. *System Dynamics Review* **13**(3): 253–265.

Ericsson KA, Simon HA. 1993. *Protocol Analysis: Verbal Reports as Data*, revised edn. MIT Press: Cambridge, MA.

Funke J. 1988. Using simulation to study complex decision making: a review of studies in the FRG. *Simulation & Games* **19**(3): 277–303.

——. 1991. Solving complex problems: exploration and control of complex systems. In *Complex Decision Making: Principles and Mechanisms*, Sternberg RJ, Frensch PA (eds). Lawrence Erlbaum: Hillsdale, NJ, 185–222.

Huz S, Anderson DF, Richardson GP, Boothroyd R. 1997. A framework for evaluating systems thinking interventions: an experimental approach to mental health system change. *System Dynamics Review* **13**(2): 149–169.

High Performance Systems Inc. (1996). *Next Step*. CD-Rom High Performance Systems Inc., 46 Centerra Parkway, Suite 200, Lebanon, NH-03766-1487. USA.

Irwin G, Wasalathantry CK. 2000. Reuse in object-oriented modeling: An empirical study of experienced and novice analysts. In *Object Oriented Technology: Challenges and Recommendations* Gibson R (ed). Idea Group Publishing: Hershey, PA, 89–113.
Moss Kanter R. 2003. Leadership and the psychology of turnarounds. *Harvard Business Review* **81**(6): 57–66.
Pennington N, Lee AY, Rehder B. 1995. Cognitive activities and levels of abstraction in procedural and object-oriented design. *Human–Computer Interaction* **10**: 171–226.
Richmond B. 1993. Systems thinking: critical thinking skills for the 1990s and beyond. *System Dynamics Review* **9**(2): 113–133.
——. 1994a. Competing in the 90's: systems thinking, ithink and organisational learning. *Introduction to Systems Thinking and ithink*, High Performance Systems: Lebanon, NH, 1–21.
——. 1997a. The "thinking" in systems thinking: how can we make it easier to master. *The Systems Thinker* **8**(2): 1–5.
——. 1997b. Dynamic thinking: a behavioral context. *The Systems Thinker*, **8**(6): 6–7.
——. 1997c. System-as-cause thinking. *The Systems Thinker* **8**(8): 6–7.
——. 1997d. Forest thinking. *The Systems Thinker* **8**(10): 6–7.
——. 1998. Operational thinking. *The Systems Thinker* **9**(2): 6–7.
Schaffernicht M. 1999. Managing improvement amongst autonomous actors with OMCA: the case of the Chilean educational reform. Paper presented at the International conference of the System Dynamics Society, Wellington, New Zealand.
Schenk KD, Vitalari NP, Davis KS. 1998. Differences between novice and expert systems analysts: what do we know and what do we do? *Journal of Management Information Systems* **15**(1): 9–50.
Senge PM. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday/Currency: New York.
Srinivasan A, Irwin G. 1999. Data abstractions and their use: an experimental study of user productivity. Paper presented at INTERACT, 1999.
Srinivasan A, Te'eni D. 1995. Modeling as constrained decision making: an empirical study of the data modeling process. *Management Science* **41**(3): 419–434.
Sterman JD. 1989a. Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* **43**(3): 301–335.
——. 1989b. Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science* **35**(3): 321–339.
Todd P, Benbasat I. 1987. Process tracing methods in decision support systems research: exploring the black box. *MIS Quarterly* **11**(4): 493–512.

Further Reading

Ford JK, Schmitt N, Schechtman SL, Hults B, Doherty M. 1989. Process tracing methods: contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes* **43**: 75–117.
Frensch PA, Funke J. 1995. Definitions, traditions, and a general framework for understanding complex decision making. In *Complex Decision Making. The European Perspective*, Frensch PA, Funke J (eds). Lawrence Erlbaum: Hillsdale, NJ, 3–25.
Maani K. Cavana R. 2000. Systems Thinking and Modeling—Understanding Change and Complexity. Prentice Hall: Auckland.