USING EYE TRACKING TO EXPOSE COGNITIVE PROCESSES IN UNDERSTANDING CONCEPTUAL MODELS[1]

Palash Bera
Operations and Information Technology Management, Chaifetz School of Business, Saint Louis University,
St. Louis, MO 63103 U.S.A. {palash.bera@slu.edu}

Pnina Soffer
Department of Information Systems, Faculty of Social Sciences, University of Haifa, Carmel Mount,
Haifa, ISRAEL, 31905 {spnina@is.haifa.ac.il}

Jeffrey Parsons
Faculty of Business Administration, Memorial University of Newfoundland,
St. John's, NL A1B 3X5 CANADA {jeffreyp@mun.ca}

Conceptual models are used to communicate information about a domain during the development of information systems. In two experimental studies using business process models, we demonstrate how eye tracking can contribute to understanding the cognitive processes by which readers use conceptual modeling scripts to perform problem solving tasks. In the first study, we compare scripts generated using two process modeling grammars and demonstrate how attention paid to specific parts of scripts generated using grammar variations, and differences in visual association between parts of a diagram, account for task performance. In the second study, we use a combination of eye tracking and verbal protocol analysis to examine how visual association between parts of conceptual modeling scripts can indicate cognitive integration while performing problem solving tasks. The studies show that task performance can be explained with different mental processes, reflected in specific eye tracking behavior, where scripts developed following different rules invoke different cognitive processes. We show that attention can be measured by eye tracking and can explain task performance. In addition, we show that visual association (which is observable) between parts of a modeling script involves cognitive integration (which is not observable). This finding can be used to improve conceptual modeling grammars in several ways, including understanding the effects of alternative visual arrangements of models on how effectively they communicate domain knowledge for particular tasks, and guiding the design of visual modeling notations.

Keywords: Eye tracking, conceptual modeling, process modeling, cognitive processes, attention, visual association, cognitive integration, problem solving

---

Introduction ▰▰▰▰

Conceptual models are abstract representations of some subject matter used to support communication and shared domain understanding among stakeholders, such as information systems (IS) analysts, users, and programmers, thereby enhancing the prospects for successful IS development and use (Wand and Weber 1993). A conceptual model contains a visual arrangement of modeling constructs in the form of graphical symbols and text. Together, the constructs and the rules for arranging them in meaningful ways constitute a *conceptual modeling grammar*, and the diagrams (models) created based on a grammar are called *scripts* (Wand and Weber 2002). Common grammars include the entity–relationship (ER) model and the unified modeling language (UML) for data modeling, and business process model and notation (BPMN) and event-driven process chains (EPC) for process modeling.

A conceptual modeling grammar typically affords flexibility in organizing scripts in various ways, such that these variations might enhance or inhibit domain understanding. Considerable prior research on the effectiveness of grammars has produced insights into the effects of certain grammatical constructs on measures of understanding, including comprehension and problem solving (for a review, see Burton-Jones et al. 2017). However, prior research has paid little attention to the effects of the visual layout of constructs on cognitive processing and understanding. Likewise, extant work has largely treated the process of understanding as a black box: variations in grammars or scripts are generated (usually based on some theory or design principles) and the effects of these variations on understanding are compared. *Therefore, we have limited insight into the cognitive processes involved in reading and understanding scripts generated from conceptual modeling grammars*. As a consequence, we have limited understanding of how to organize the information in scripts created from a particular grammar to promote understanding, and little insight into how the introduction of modeling constructs affects the process by which readers of scripts understand their meaning. Gaining a deeper understanding of these processes can guide efforts to improve the comprehensibility of scripts and provide a foundation for additional rules or guidelines to improve the effectiveness of modeling grammars.

In this paper, we use eye tracking technology to expose the cognitive processes involved in reading and understanding conceptual modeling scripts. Understanding a script entails attending to, assimilating information from, and performing tasks based on the script. While prior research has extensively addressed cognitive aspects of reading conceptual

models (Houy et al. 2014), specific links between users' attention distribution when reading scripts and task performance have not been established. Eye tracking enables us to examine how cognitive processing of information takes place and provides insight into the content of information processing (Rayner 1998). Thus, our research question is

> *How can eye tracking contribute to exposing the cognitive processes of understanding conceptual models?*

In two studies, we demonstrate how eye tracking can help determine the processes by which readers use conceptual modeling scripts to perform problem-solving tasks. In the first study, we compare scripts generated using different process modeling grammars and demonstrate how attention paid to specific parts of a script and visual association of these parts account for task performance. In the second study, we combine eye tracking with verbal protocol analysis to establish a connection between visual association and cognitive information integration, and demonstrate how users integrate information from specific parts of process modeling scripts to perform tasks. Collectively, the studies show that eye tracking technology has a strong potential to shed light on conceptual model understanding, thereby guiding the future development and evaluation of modeling grammars.

Next, we briefly review extant research on evaluating conceptual modeling grammars and discuss limitations of this work in providing insight into the cognitive processes by which grammars enable or impair understanding scripts generated from the grammar. Then we present a theoretical framework for understanding conceptual models with the use of eye tracking technology. Thereafter, we present two eye tracking studies designed based on the framework. We conclude by discussing the implications of our findings and opportunities for using eye tracking in conceptual modeling research.

Evaluating Conceptual Modeling Grammars ▰▰▰▰

Much research on conceptual modeling has focused on identifying desirable properties of modeling grammars. Two independent streams of research are notable. First, work has examined visual properties of scripts, such as layout and use of modeling symbols, as exemplified by work on the "physics of notations" (Moody 2009). Various complexity measures of process modeling scripts have been proposed (e.g., Cardoso et al. 2006), and empirical studies have investigated the

effect of complexity on model understanding (Mendling et al. 2007; Vanderfeesten et al. 2008). In the context of data modeling, Fuller et al. (2010) tested the effect of the visual arrangement of scripts on tasks such as error detection and generation of SQL queries. Findings from such work are useful in providing guidance on how to visually organize models to improve understanding (e.g., Figl et al. 2013; La Rosa et al. 2011; Purchase et al. 2001).

Second, a substantial body of research has proposed and evaluated rules or guidelines that can be added to grammars to produce scripts conforming to certain theoretically derived principles. A prominent line of work in this area is based on the theoretical proposition that good modeling grammars should have constructs that correspond to real-world phenomena: the *theory of ontological expressiveness* (Wand and Weber 1993). Examples include rules for constructing UML scripts that follow ontological criteria (Evermann and Wand 2005) and conceptual modeling rules based on cognitive foundations (Parsons and Wand 1997, 2008). A core idea driving these studies is that scripts generated from grammars based on theoretical prescriptions have clear and understandable semantics, consistent with the real world phenomena that should be represented in the script, and hence are suitable for communicating information.

These theory-based propositions have been tested empirically, based on predictions that a script adhering to guidelines will result in improved task performance compared to one that does not. Example manipulations of the scripts include mandatory versus optional properties of entity types (Bodart et al. 2004; Gemino and Wand 2005), property precedence in UML association classes (Parsons 2011), and representation of external events in EPC (Soffer and Kaner 2011). Task performance in these studies has been measured in various ways, including comprehension, problem solving, error detection, time to complete tasks, and perceived ease of use of scripts.

While such research has shown that variations in representation (whether motivated by notational or theoretical issues) can affect task performance, it does not shed light on the cognitive processes that produce these effects. There is a modest body of literature dealing with cognitive processes in understanding conceptual models, typically using verbal protocols in which participants are asked to think aloud during the course of the task, allowing insight into their thought processes (Newell and Simon 1972). Verbal protocol analysis has been used to analyze the cognitive processes of interpreting scripts and, thus, to understand how different scripts might affect task performance (e.g., Bera et al. 2011; Burton-Jones and Meso 2006; Zugal et al. 2015). It has also been used to investigate the effects of prior knowledge on problem solving performance (Khatri and Vessey 2016).

Notably, two papers provide in-depth discussion and evidence of the cognitive processes of understanding conceptual models based on verbal protocols. Kim et al. (2000) claim that the cognitive process in diagrammatic reasoning consists of perceptual and conceptual processes. They suggest that script readers search and recognize relevant information through perceptual processes, and reason by inferring new information through conceptual processes. In their study, perceptual data was coded by action protocols and transition diagrams, marking the attention given to visual elements. Evidence of conceptual processes was provided by the proportion of returning transitions between diagrams. Such transitions were claimed to indicate integration of information from different elements of a diagram. Shanks et al. (2010) distinguished seven problem-solving phases related to conceptual models. Out of these phases, two were indicated as important: finding the task-relevant information in the diagram and clearly distinguishing the subset of elements on which one should focus for performing the task. Their study analyzed protocols and mapped segments of the text to the suggested problem solving phases.

Similar ideas arise when addressing conceptual modeling scripts as multimedia objects, including visual representations (Moody and van Hillegersberg 2009; Reijers et al. 2011) constructed with pictures and words. Multimedia learning theory (Mayer 1989) suggests three stages of multimedia learning (Mayer 2001). In the first stage, meaningful learning occurs when a learner *selects* relevant information from what is presented. In this process, selected information is added to working memory. In the second stage, the learner *organizes* and *integrates* information in a coherent way. In the third stage, the learner *processes* the selected information using his/her existing knowledge. For problem solving, these three stages have been referred to as two types of mental tasks: *selective* and *integrative* (Grant and Spivey 2003).

While multimedia learning theory, as well as the two studies described above, indicate similar cognitive processes, there is no standard way of identifying and measuring these processes. In this paper, we propose using eye tracking as a quantitative technique for examining the cognitive processes used to understand conceptual models.[2] Specifically, we examine the capability of eye tracking to detect and measure two distinct operations that reflect cognitive processing: (1) focusing on and paying attention to specific regions in a

---

[2]Alternative approaches may be used, including NeuroIS techniques such as fMRI (Riedl et al. 2010). We believe eye tracking is particularly relevant for understanding the cognitive processes of reading diagrammatic representations and focus on this technique. We leave open for future research the use of such techniques to complement the contribution of eye tracking research to this area.

script and (2) integrating information from different regions of a script. We note that the former is well established in the psychology literature, and the latter is further validated in this paper. As opposed to verbal protocols, eye tracking does not capture the *intention* that drives the mental process. Its strength is in objectively, accurately, and unobtrusively measuring the tracks of the cognitive process. We suggest that these tracks can be interpreted in terms of attention to information elements and integration of information elements. We refer to these as *cognitive operations*, to distinguish them from higher-level cognitive processes referred to in the literature.

Research has demonstrated that cognitive processing can be influenced by task type (Glaholt et al. 2010) and visual representation (Desimone and Duncan 1995). In the psychology literature, the visualizations are often pictures, and tasks are generally decision making (e.g., selection of a product). In this circumstance, the focus on cognitive processing of information is generally on *attention*-related processes (Milosavljevic and Cerf 2008; Orquin and Loose 2013).

In contrast, when the visual representations are combinations of words and graphics, users may need to integrate information from multiple parts of the representation to perform a task successfully. Studies show that spatially contiguous representations support integration. The most common task used is learning outcome, such as transfer task (Mayer and Moreno 2003), which is used to measure cognitive processing during learning. An example of such a representation is a diagram of a car's brake pedal system with text descriptions. In this context, an example of a transfer task is, "Suppose you press on the brake pedal in your car but the brakes don't work. What could have gone wrong?" Johnson and Mayer (2012) used eye tracking to show that, when spatially contiguous representations are used with transfer tasks, users engage more in visual association of text and graphics. They assumed this indicates the cognitive operation of integrating information from multiple parts of the representation.

Previous conceptual modeling research (e.g., Gemino and Wand 2005) has treated conceptual modeling scripts as spatially contiguous representations, as they often use words and graphics. Therefore, we expect that script users should engage in the cognitive operation of integrating information from multiple parts of the scripts. To evaluate scripts, open-ended problem solving tasks have been used (e.g., Burton-Jones and Meso 2006). However, depending on the tasks and on the region of the relevant information in the representations, some scripts require information integration and others mainly invoke attention to specific areas. A recent eye tracking paper on process models (Petrusel et al. 2017) identified that viewers focus on relevant parts of the model to

perform comprehension tasks. However, these tasks were based on control flows and not on problem solving. Following an extensive literature review in this area, Zugal (2013) highlighted the importance of cognitive integration in overcoming difficulties associated with fragmentation of models. Cognitive integration capabilities were also associated with high performance of modeling tasks by Martini et al. (2016). Both of these studies emphasized that the need for cognitive integration is characteristic of problem solving tasks.

We refer to attention and integration as two cognitive operations. There is limited knowledge of how these operations interact. Further, we test the effects of these operations on task performance with multiple conceptual modeling scripts. In summary, this paper addresses the following three questions:

(1) *How can we identify and measure the cognitive operations that conceptual modeling script readers use to perform problem solving tasks?*

(2) *How do these cognitive operations affect problem solving task performance?*

(3) *Can we predict an effect of these cognitive operations on problem solving task performance when different scripts are used?*

Using Eye Tracking for Measuring Attention and Association of Information

Several conceptual modeling research studies have used eye tracking technology to study the use of data and process models. In the context of NeuroIS, applying neuroscience theories and tools to IS research, Figl (2017) points out that eye tracking is the only relatively simple neurophysiological tool that has been applied successfully to research on model comprehension. Compared to other NeuroIS tools, such as fMRI (Dimoka et al. 2011), eye tracking is easier to apply and can be considered as psycho-physiological measurement. Porras and Gueheneuc (2010) tested the comprehension performance of a set of UML diagrams. They found that some types of diagrams (such as UML collaboration diagrams) enable users to more efficiently locate specific elements (such as classes) than other types. Yusuf et al. (2007) tested the effect of layout, color, and stereotype usage of UML class diagrams for supporting comprehension tasks. They found that layout with additional semantic information

had a significant effect on how these tasks are performed. Nordbotten and Crosby (1999) tested the effect of graphic styles (use of symbols and annotations) in data model comprehension. They found that data models with high graphic content are difficult to interpret in comparison to those with low graphic content. Petrusel and Mendling (2013) identified factors that influence process model comprehension tasks. They identified specific areas of process models (relevant regions) and observed how viewing and paying attention to these regions affects the task performance. Zimoch et al. (2018) compared novices and experts when reading BPMN and found that experts had fewer fixations and saccades. Chen et al. (2018) used eye tracking to compare fixation times in different ways of integrating the representation of business processes and business rules.

Notably, these studies have used eye tracking primarily to measure specific patterns of behavior in reading diagrams. In contrast, we seek to use eye tracking to understand how attention to specific elements of a diagram and association of information among parts of a diagram can be used to make theory-based predictions of problem solving task performance when different informationally equivalent representations are used.

Measuring Attention

When reading, the eyes make rapid movements to shift attention from one part of the display to another and then remain almost motionless while the brain interprets the material at that location (Rayner 1998). The periods in which the eyes are motionless are called *fixations* (Sharif and Maletic 2010). Fixation information can be used to measure the attention paid to the viewed object (Glaholt and Reingold 2011; Wang et al. 2014). A typical fixation lasts between 200 and 300 milliseconds and is generally understood to indicate where a viewer's attention is directed (Rayner 1998). Fixation duration and fixation count are the most commonly used eye tracking metrics to measure cognitive processing (Just and Carpenter 1976), allowing one to investigate whether individuals mainly scan information or attend to information as they reason or make judgements about it (Glockner and Herbold 2011). Yet, the amount of time spent fixated on certain visual areas needs to be interpreted in relation to performance on a task, as we do in our studies. We return to this issue in the "Discussion and Limitations" section.

In general, two types of attention-related analysis are done with eye movement data: first, on the overall level for viewing the entire area; and second, at a specific level for viewing a subarea. Researchers define "areas of interest" (AOIs) over certain parts of a display and analyze the eye movements that fall within such areas. In this way, the visibility, meaningfulness, and placement of specific elements can be objectively evaluated (Goldberg and Kotval 1998). In the context of this research, elements of conceptual modeling scripts with certain shapes (such as rectangles) are the AOIs we created. As there are many ways of creating these AOIs, the specific AOIs we use depend on the theoretical basis of the conceptual models. For example, organizational units of EPC scripts can be chosen as AOIs, as they depict roles. Table 1 summarizes the definition and interpretation of relevant eye metrics.

Measuring Association[3] of Information

The continuous movement of the eyes between fixations are termed *saccades* (Jacob 1995). Humans use saccades to locate interesting parts of a visual (Yusuf et al. 2007). A number of metrics use saccades for measuring visual association among different areas of the display. These include visit count (Kim et al. 2012), integrative transitions (Johnson and Mayer 2012), corresponding transitions (Johnson and Mayer 2012), and run count (Archibald et al. 2013). The metrics have been studied by various researchers (Archibald et al. 2013; Johnson and Mayer 2012; Kim et al. 2012; Lin and Lin 2014b; Russel et al. 2008), and interpreted as representing integration of information. However, as opposed to a directly observable visual association, this interpretation has not been validated by empirical evidence. Johnson and Mayer (2012) showed that integrative and corresponding transition measures between text and graphics are correlated with high task performance. Ratwani et al. (2008) looked for evidence that saccadic movements reflect cognitive integration through studies that combine eye tracking with verbal protocols. However, they address a limited notion of integration: clustering areas of a graph into aggregated areas. In contrast, we address conceptual modeling scripts, where information from different parts must be integrated by synthesis, comparison, negation, and/or inference to facilitate performance on a task. We aim to establish whether visual association of elements of a diagram, measured by eye tracking metrics, reflects cognitive integration and contributes to task performance.

We use two representative metrics: *run count* and *AOI run count*. Both quantify gaze distribution and reveal the strategy used by participants to solve the task presented to them (Archibald et al. 2013). Run count is defined as the number

---

[3]We use the word association rather than integration because association can be assessed through observation alone using saccadic metrics such as run count. Integration, on the other hand, is a higher-order reasoning and judgment process.

Table 1. Summary of Eye Movement Metrics Related to Attention		
Eye Movement Metric	**Definition (Jacob and Karn 2003)**	**Interpretation**
Fixation count	Total number of fixations on an object of interest	High number of fixations on a particular area indicates it is more important to the viewer than other areas (Poole and Ball 2006)
Fixation duration	Total time of fixations on an object of interest	Longer fixation duration may indicate that the viewed object is engaging or creates interest to the viewer (Cyr and Head 2013; Poole and Ball 2006)

Table 2. Eye Movement Metrics Related to Association of Information	
Metric	**Description**
Run count	The number of times a given area of interest (AOI) is entered and exited. Note that the entry of saccade to the interest area can come from other AOIs or areas that have not been marked as an AOI (e.g., an empty space).
AOI run count	The number of times a given area of interest (AOI) is entered and exited from other specific interest areas (other AOIs).

of times a given area of interest (AOI) is entered and exited (Lin and Lin 2014b; Russel et al. 2008). Researchers view it as related to the "cognitive integration stage," wherein readers visually associate objects (Lin and Lin 2014a).

As per Johnson and Mayer (2012), in addition to run count, we use a derived eye metric, AOI run count, to confirm information association among areas of interest. These metrics are described in Table 2. Appendix A provides details on how run counts are measured.

## Cognitive Theories Explaining Task Performance

Based on prior literature, eye tracking clearly provides the capability to detect and measure the cognitive operation of focusing and paying attention to a specific area. Furthermore, we will show that eye tracking enables detecting the cognitive operation of associating information from various areas of interest. To use these capabilities for understanding the cognitive processes of reading conceptual modeling scripts and predicting task performance, a broader view of these processes is needed. For this, we use two cognitive theories.

The theory of cognitive fit (Vessey and Galletta 1991) can be used to explain the cognitive processes applied by problem solvers in developing and reading conceptual models. To solve a problem, humans create a mental representation of the problem in their working memory, using concepts taken from

both the task representation and the domain (Vessey 1991). According to cognitive fit theory, when the information emphasized in a representation matches the type of information needed for a task, participants are expected to perform the task better than when there is not a match. Cognitive fit theory has been invoked to suggest why scripts developed with some theory-based guidelines lead to better task performance than scripts developed without such guidelines. In particular, the theory holds that different representations of information are suitable for different tasks and different audiences (Moody 2009).

However, fit is difficult to operationalize and can be interpreted in different ways. Hence, different representations can be considered fitting to an equivalent extent. Predictions of task performance for different representations might, therefore, be similar based on cognitive fit considerations. To gain a finer and more concrete explanation of how different representations support task performance, we turn to cognitive load theory (Sweller 1988).

Cognitive load theory deals with the limited capacity of working memory for problem solving. It distinguishes three types of cognitive load: (1) intrinsic load, determined by the complexity of the task; (2) extraneous load, determined by the information representation; and (3) germane load, caused by the need to integrate current information with knowledge from long-term memory to perform a task. Since all of these use the limited capacity of working memory, extraneous load should be minimized to allow more capacity for dealing with intrinsic and germane load.

Figure 1. Use of Eye Tracking in Conceptual Modeling Empirical Studies
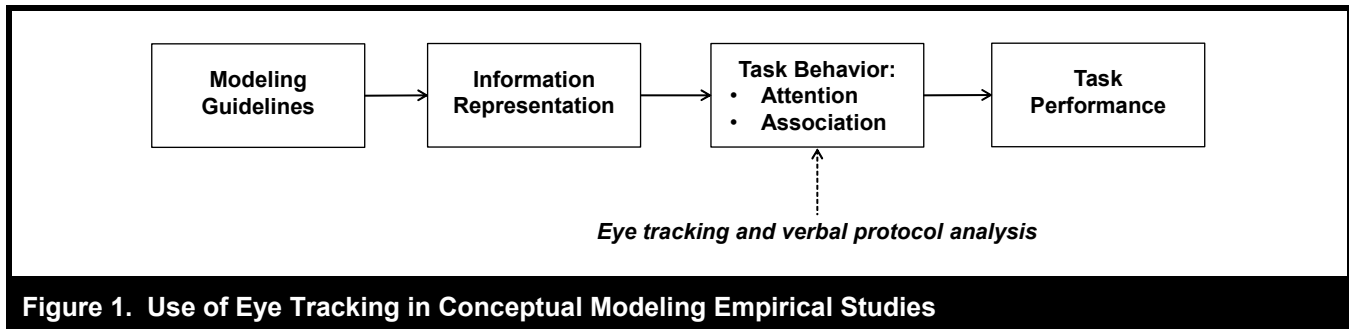
Reducing extraneous load is targeted by multimedia learning, addressing learning tasks that demand more of the learners' cognitive resources (working memory) than they can sustain, a situation known as *cognitive overload* (Mayer 2001). Mayer and Moreno (2003) suggest that several methods can be used to reduce cognitive overload in learning situations. An important one is *signaling*, in which visual cues (such as the organization of concepts) are provided to learners to reduce their cognitive load by helping them to select, organize, and process relevant information.

Using specific theory-based guidelines (versus no guidelines), alternative representations can be created, resulting in different levels of task performance. However, there is a gap in explaining the cognitive mechanisms by which information representation contributes to better task performance. In other words, ***how*** *does the way in which information is represented in conceptual models affect the way the task is performed*? With the capability of eye tracking to detect and measure attention and association of information, insights into the way this effect is achieved can be gained. Milosavljevic et al. (2012) show that visually salient features capture attention more readily than less salient alternatives. Hegarty et al. (2010) note that people fixate more on salient regions of a display; making the relevant area visually salient facilitates the process of filtering task-relevant from task-irrelevant information. Hegarty et al. further state that "visual salience of the task-relevant information affects performance by drawing attention, and the eyes, to the task-relevant locations" (p. 41). Orquin and Loose (2013) note that, in decision tasks, it is expected that decision makers will attend to stimuli with higher task relevance and ignore stimuli with little or no task relevance.

The proposed research model (Figure 1) introduces the task behavior (the cognitive operations involved in the task) as mediating the effect of information representation on task performance. Specifically, the relevant cognitive operations are attention to specific model elements and association of information from different model elements. Eye tracking and

verbal protocol analysis are tools for detecting and understanding these operations.

Based on Figure 1, scripts in which the information representation follows some modeling guidelines are expected to invoke predictable patterns of attention and association of information that differ from the patterns produced from scripts that do not follow these guidelines. These patterns of attention and association will affect task performance.

Using this research model, we can address specific questions concerning different underlying theories. Taking cognitive fit as a theoretical basis, studies that follow Figure 1 can indicate specific cognitive operations that are required by a task and that contribute to the fit between the task and the information representation. Considering cognitive load, eye tracking can be useful in distinguishing between cognitive operations that constitute intrinsic load and those that constitute extraneous load. Furthermore, it might be possible to measure the additional extraneous load resulting from a specific representation as compared to a different one, and gain an understanding of how specific cues contribute to reducing this load.

Note that, when extraneous cognitive load is extremely high, we might not observe the connection of the cognitive operations of attention and association with task performance. This is because, with a very high extraneous load, the remaining cognitive resources might become too low for effective processing of the perceived information (Van Gerven et al. 2002).

To analyze attention and association using eye tracking, specific areas of the scripts have to be considered. To perform tasks related to the scripts, readers need to assimilate and combine information from specific parts of the scripts.

Task performance relates to what the user of the representation must do to isolate and extract the relevant information for the task (Hegarty et al. 2010). To perform a task, a user does not need to focus on the entire model, but only on the

part of the model that relates to the task (Petrusel and Mendling 2013). Thus, with respect to a task, a script can be split into two parts: *task-relevant* and *task-irrelevant*. Such distinction of visuals is common in eye tracking studies (Hegarty et al. 2010). These are demonstrated in the empirical studies in the following section.

Empirical Studies

Prior research in this area has used various theories to predict that a certain form of representation is better than others for specific tasks. We use eye tracking to understand the cognitive operations that explain the impact on performance (see Figure 1). Specifically,

1. Does attention to task-relevant areas explain task performance?

2. Does visual association of information among task-relevant areas explain task performance?

Since, as explained above, there is no direct evidence that visual association of model parts indeed reflects cognitive integration of information, our third question is

3. Does visual association of information reflect cognitive integration of information?

We address these questions by conducting two empirical studies on conceptual modeling. First, we use eye tracking to examine the relationship of attention and visual association with performance, using script variations based on notational considerations. Second, we use eye tracking in conjunction with verbal protocol analysis to determine whether visual association can be used as an indicator of cognitive integration of information within a diagram.

In the first study, we use three versions of business process modeling (BPM) scripts selected based on different visual notation. We compared techniques that differed in the visual syntax used to represent a particular kind of semantic information in business processes: roles. BPM is widely used for organizational purposes, such as process reorganization, activity-based costing, and human resource planning (Becker et al. 2000). From a conceptual modeling perspective, BPM serves to increase domain understanding for the purpose of IS development. For example, Mendling and Recker (2008) state that process models are "required to be intuitive and easily understandable" to be used in initial phases of IS development (p. 1).

In the second study, we selected two of the three versions of BPM scripts used in the first study and used eye tracking in conjunction with verbal protocol analysis to examine whether eye tracking provides evidence of cognitive integration in performing tasks based on diagrams.

In both studies, we tested the effectiveness of alternative scripts for problem solving tasks. In the first study, we use eye tracking techniques to trace attention and visual association during task performance. In the second study, we supplement eye tracking with verbal protocol analysis to trace the relationship between visual association and cognitive information integration in performing tasks. Below, we present the general design, followed by details of each study.

General Study Design, Task, and Participants

Both studies followed a between-subjects design in which each participant received scripts corresponding to the condition to which s/he was assigned. Understanding a domain based on a script is a learning process, the outcomes of which can be measured in terms of retention and transfer (Mayer 2001). To perform a *transfer* task, users need to solve problems that are not directly answerable from a script (Gemino and Wand 2005). Thus, users not only need to comprehend the domain as depicted in the script, but also need to make inferences about related domain concepts not depicted. For this purpose, we used problem solving questions (Burton-Jones and Meso 2006). Participants were asked to answer these questions by referring to a script and making inferences from the script. A high number of correct responses to problem solving questions indicate that information has been integrated into long-term knowledge and a deep level of understanding has occurred (Mayer 2001). A detailed coding manual was developed that contained the correct responses of the tasks (see Appendix B).

In the first study, we recruited 45 students (15 in each of three groups) from a pool of MIS graduate students at a southern U.S. university over a period of two terms. These students were enrolled in Information Systems Analysis and Design courses in which business process modeling was taught. In the second study, 27 business undergraduate students from a midwestern U.S. university were recruited (14 in one group and 13 in the other). The participants had knowledge of business processes through a SAP Enterprise Resource Planning course where business processes were introduced. They were also taught the concepts of business processes for three hours prior to the experiment.

Procedures ■■■■■■■

Participants were randomly assigned to one of the treatment groups. Each participant was informed about the eye tracking equipment and procedures and asked to sign the consent form. Both studies were done in three stages. In the first stage, participants' eyes were calibrated and validated by showing them a series of dots. Then participants answered questions about their domain and modeling familiarity. In the third stage, they were asked to perform two problem solving tasks from each of two scripts, and their eye movements were tracked for this purpose. The second study also captured participants' verbalization while they performed the tasks. The experimental procedure is described in detail in Appendix C.

Study 1: Eye Tracking

We used two grammars (described below), event-driven process chains (EPC) (Scheer 2000) and business process model and notation (BPMN) (White 2004) (see Figure 2). Both are popular in academia and in practice (Recker and Dreiling 2007; Scheer 2000). A BPMN script consists of sets of elements (such as flow objects, activity objects, and swimlanes) that present a graphical model of business process operations (White 2004). EPC uses events and functions as its main constructs, and allows additional information to be represented using constructs such as organizational unit and data element.

The concept of roles of actors (humans or IS) can help in understanding some aspects of business process models. A role consists of actors that have similar functions, responsibilities, and organizational contexts. In BPMN, *pool* and *lane* can be related to roles. A pool represents a participant in a process and a lane is a sub-partition within a pool (White 2004). The construct *organizational unit* (in ovals) in EPC can be considered a role as it indicates a *position*; that is, the role performed by individuals (Davis and Brabander 2007). These constitute alternative ways of visually representing the organizational semantics of role.

BPMN further emphasizes roles in the visual layout of a diagram, as *lanes* and *pools* are used for organizing business activities. Hence, BPMN scripts are appropriate for performing tasks associated with roles. As individuals do not read entire scripts at once, but rather in chunks, they focus on localized areas in the script (Gemino and Wand 2005). This promotes domain understanding if the local information is related to the given task. In contrast, EPC scripts focus on sequences of events and may be more appropriate for sequence-based tasks. According to readability theory

(Kintsch 1979), relevant information that is spread across a script or distributed over different scripts may lead to poor readability. Thus, we expect that a role-related task can be better performed using BPMN than EPC. Consider a task related to an admission department (Task A in Figure 2). In EPC (left panel), the expected cognitive fit with the task is low as information on activities related to admission is scattered in the script (the full script is shown in Appendix D). In BPMN (right panel), the expected cognitive fit with the task is higher as information on activities related to admission is displayed within a single lane.

Although EPC might be more suitable for sequence-related tasks, if an alternative visual representation is used to highlight roles in EPC diagrams, they might still be appropriate for role-related tasks, since highlighting facilitates locating task-relevant areas. Therefore, we designed a variant of EPC in which the organizational units and the associated activities in EPC scripts are highlighted by marking each role with a different color. We term such a script *EPC-highlighted* (EPC-H). Appendix D shows how roles can be highlighted, where all the activities performed by the same role are highlighted in the same color.[4] Extensive research has found that humans cognitively cluster phenomena they perceive to be related (e.g., Bousfield 1953). Focusing on clusters reduces cognitive load and improves humans' abilities to understand the world (Mayer 2010). Moody (2009) claims that humans are highly sensitive to color variations and can detect and remember color differences. Petrusel et al. (2016) found color to be an effective cue in process models. Thus, we expect that highlighting roles in EPC will promote role-related domain understanding. Table 3 summarizes the expected task performance for each BPM grammar.

We used four problem solving tasks to assess domain understanding. To increase generalizability, questions were posed for two domains (with two questions for each domain): *patient treatment* and *grant review*. The complete script of each domain is shown in Appendix D: in EPC-H the AOIs are marked using colored rectangles, while in BPMN the AOI is a swimlane. Each task has a *key role* (Table 4), which is crucial in performing the task. The areas depicted by these roles in the scripts are thus considered to be task-relevant areas. Appendix E provides details on how the task-relevant AOIs were determined.

---

[4]In Appendix D, the rectangles as AOI's were constructed for analysis. These rectangles were not part of the original diagrams shown to participants in the study.
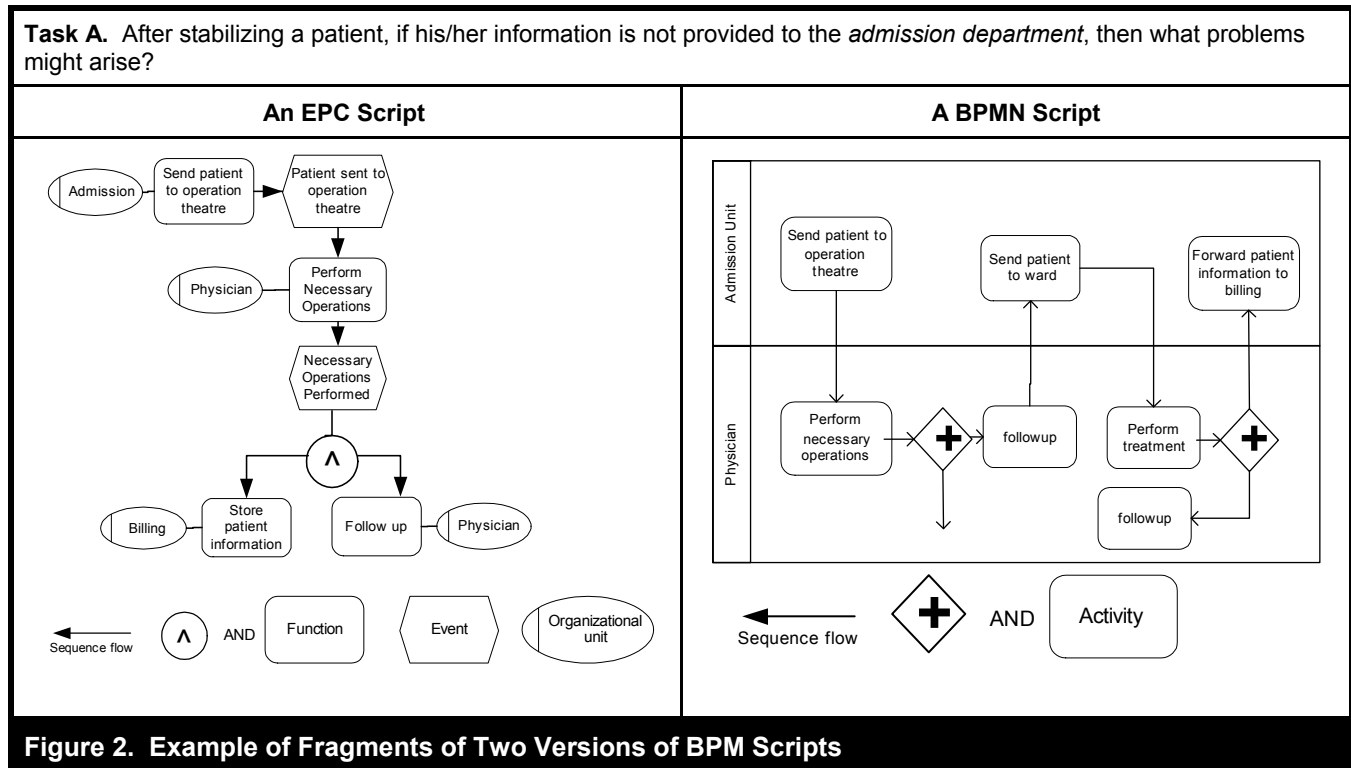
Task A. After stabilizing a patient, if his/her information is not provided to the *admission department*, then what problems might arise?

Figure 2. Example of Fragments of Two Versions of BPM Scripts

Table 3. Expected Role-Related Task Performance for BPM Scripts

Modeling Grammar Used	Expected Cognitive Fit and Cognitive Load for Role-Related Problem Solving Tasks	Expected Performance
EPC	High cognitive load, as the cognitive fit between the problem representation (role-based activities are not organized) and the task (role-based problem solving) is expected to be low	Low
EPC-H (Highlighted)	Moderate cognitive load as the expected cognitive fit between the problem representation (role-based activities highlighted in colors, but physically separated in different parts of the script) and the task (role-based problem solving) is partial (colors provide fit, spatial distribution does not provide fit)	Moderate
BPMN	Low cognitive load as the expected cognitive fit between the problem representation (role-based activities organized in swimlanes) and the task (role-based problem solving) is high	High

Hypotheses

We ask whether, as predicted, participants viewing EPC-H scripts pay more attention to the task-relevant areas than participants viewing EPC scripts, and participants viewing BPMN scripts pay more attention to the task-relevant areas than participants viewing EPC-H scripts. Thus, we propose

> ***Hypothesis 1****: Compared to users of EPC, users of EPC-H scripts will pay more attention to the task-relevant areas to perform a role-related problem solving task.*

> ***Hypothesis 2****: Compared to users of EPC-H, users of BPMN scripts will pay more attention to the task-relevant areas to perform a role-related problem solving task.*

Tying task performance to eye tracking metrics, we now refer to the cognitive operations of attention and visual association.

Table 4. Task-Relevant Areas for Each Problem Solving Task		
No.	**Problem Solving Tasks**	**Task-Relevant Areas (Key Role)**
Patient Treatment Domain		
1	After stabilizing a patient, if his/her information is not provided to the *admission department* then what problems might arise?	• Admission (send patient to operation theater) • Admission (send patient to ward) • Admission (patient information forwarded to billing)
2	What will happen if patients are *diagnosed* immediately after arrival?	• Nursing (check body vitals) • Nursing (stabilize patient) • Nursing (forward to admission) • Nursing (diagnosis) • Nursing (check criticality) • Nursing (discharge)
Grant Review Domain		
3	What will happen if some *review members* are absent in the grant funding meeting?	• Review member/Chair (receive applications) • Review member/Chair (study applications) • Review member (review applications)
4	If the *Review Chair* is not available to a grant funding meeting how will grant applications be evaluated?	• Review member/Chair (receive applications) • Review member/Chair (study applications) • Review member (review applications) • Chair (make decision)

We ask whether, for a role-related problem solving task, attention to and visual association of task-relevant areas explain task performance.

> *Hypothesis 3a: Performance in a role-related problem solving task can be explained by the extent to which attention is paid to task-relevant areas.*

> *Hypothesis 3b: Performance in a role-related problem solving task can be explained by the visual association of information between task-relevant areas.*

Analysis

The correct responses for the problem solving tasks were compiled and compared among the three groups. The maximum attainable score was 3. Compared to the EPC group (mean = 1.50, s.d. = 0.42), task performance was significantly better in the BPMN group (mean = 2.18, s.d. = 0.62, p = 0.001) and marginally better in the EPC-H group (mean = 1.76, s.d. = 0.31, p = 0.08). Task performance was significantly better in the BPMN group than in the EPC-H group. A summary of problem solving task performance is provided in Appendix B. Below we focus on the eye tracking metrics. We first present the overall analysis of the models viewed by the participants. Table 5 presents the fixation analysis of the two domains for the first problem solving task.

From Table 5, we observe that participants in the EPC group spent significantly more time (in terms of fixations and dura-

tion) on the overall models than participants in the EPC-H and BPMN groups (even though they performed poorest of the three groups). Furthermore, participants in the EPC-H group spent significantly more time on the models than participants in the BPMN group.

As a basis for the eye tracking metrics, we defined the task-relevant areas in every script. Considering EPC and EPC-H scripts, we marked task-relevant areas as rectangles around related activities of organizational units viewed by each participant. For example, for the task related to the admission role (Item 1, Table 4), the task-relevant area is the combination of areas shown in activities: *send patient to operation theater, send patient to ward,* and *patient information forwarded to billing*. Appendix D shows the script with the admission-related activities shaded in violet. Total fixation duration and count are calculated by summing the durations and number of fixations of all these areas, including fixations on their bordering lines. For the BPMN script, the entire lane of the admission department is the task-relevant area. To determine the total run count, we calculated the total number of times the relevant areas were entered and exited.[5] In addition, we used the AOI run count as a measure of visual association of AOIs. To calculate AOI run count, we considered the saccadic movements among the relevant areas only.

---

[5]There are two ways of measuring run count for BPMN script. First, the number of times the eyes exited and entered the relevant lanes. Second, the number of times the eyes exited and entered in each activity within a lane. We used the second way as it is more accurate in identifying the eye movement among individual activities.

Table 5. Fixation Analysis of the First Problem Solving Task of the Two Domains					
	EPC M (SD)	EPC-H M (SD)	BPMN M (SD)	t-value EPC-H vs. EPC (p value)	t-value BPMN vs. EPC-H (p value)
Patient Treatment Problem Solving					
Duration (in sec.)	90.5 (21.9)	63.0 (13.8)	36.9 (5.23)	4.1 (< 0.01)	6.85 (< 0.01)
No. of Fixations	266.3 (53.2)	172.3 (36.6)	120.7 (21.5)	5.6 (< 0.01)	4.71 (< 0.01)
Grant Review Problem Solving					
Duration (in sec.)	92.6 (25.6)	63.4 (15.3)	37.8 (10.4)	3.8 (< 0.01)	5.37 (< 0.01)
No. of Fixations	311.8 (81.7)	166.1 (54.4)	106.9 (27.4)	5.8 (< 0.01)	3.76 (< 0.01)

Table 6. Percent Duration, Normalized Duration and Run Counts of Participants Viewing the Task-Relevant Areas							
	EPC	EPC-H	BPMN	t-value EPC-H vs. EPC	p-value EPC-H vs. EPC	t-value BPMN vs. EPC-H	p-value BPMN vs. EPC-H
Area: Admission role	M (SD)	M (SD)	M (SD)				
% duration	9.5 (3.3)	16.3 (5.1)	22.6 (5.2)	3.95	< 0.001	3.38	0.001
Normalized % duration	0.51 (0.19)	0.85 (0.27)	1.18 (0.27)	3.96	< 0.001	3.30	0.001
Run count	11.4 (2.5)	14.9 (2.8)	20.1 (3.4)	2.56	0.02	4.59	< 0.001
AOI run Count	4.3 (1.5)	7.1 (1.4)	9.8 (1.9)	5.37	< 0.001	4.48	< 0.001
Area: Member role							
% duration	22.5 (6.5)	36.4 (6.1)	45.6 (23.3)	5.99	< 0.001	1.94	0.031
Normalized % duration	1.02 (0.29)	1.64 (0.27)	2.15 (1.10)	5.99	< 0.001	1.72	0.048
Run count	19.4 (5.1)	25.6 (6.7)	30.8 (9.3)	2.86	0.004	1.76	0.04
AOI run count	5.8 (1.6)	12.0 (3.1)	14.2 (2.5)	6.96	< 0.001	2.17	0.02

To measure attention, we use the eye metric fixation duration. However, instead of using total fixation duration, we use percentage of fixation duration for the specific areas. This percentage is calculated by dividing the total duration for a specific area by the total duration for the entire script for answering the problem solving question. A high percentage reflects the importance of that area to the viewer (Jacob and Karn 2003). This metric is appropriate for comparison between two groups because individual values for specific areas are standardized.

Table 6 shows that, for the first problem solving task in each domain, the percentages of duration, normalized percentage duration and run counts of the task-relevant areas were consistently higher for EPC-H than for EPC, and higher for BPMN than for EPC-H. For example, for the problem solving task related to admission, the activities performed by

the admission role were viewed 9.5% of the time by the EPC group. The same activities were viewed by the EPC-H group 16.3% of the time and by the BPMN group 22.6% of the time. This trend was also present in the second problem solving task for both domains. We calculated normalized percentage duration as the fraction of time spent viewing an AOI divided by the fraction of the total area of that AOI to the total area of AOIs in the model. To illustrate, consider a viewer who spent 10% of her viewing time on admission related AOI. The total area of AOIs related to admission is 25,577 pixels and the sum of all AOIs is 134,505 pixels. Thus, the fraction of the total area of the admission AOI is 25,577/13,4505 (0.19). Accordingly, normalized percent duration of this viewer for the admission AOI is 0.1 divided by 0.19 which is 0.53.

Table 7 supports Hypotheses 1 and 2. Participants in the EPC-H group paid more attention to the task-relevant areas

Independent Variable	Coefficient B	t Value	P Value	Adj. R Square	Overall P
Table 7. Regression Analysis Using Percentage of Normalized Fixation Time and AOI Run Count as Independent Variables (First Tasks of Both Domains)					
Task 1 (Admission Role – Patient Treatment Domain)					
Group = BPMN					
Constant	-1.17	-1.50	0.16	0.60	< 0.01
Normalized % duration	1.94	2.48	0.03		
Run Count of AOI's denoting central roles	0.10	0.91	0.38		
Group = EPC-H					
Constant	-0.58	-1.28	0.22	0.61	< 0.01
Normalized % duration	0.21	0.56	0.58		
Run Count of AOI's denoting central roles	0.28	3.81	< 0.01		
Group = EPC					
Constant	1.54	2.18	0.05	< 0.01	0.39
Normalized % duration	0.76	0.85	0.40		
Run Count of AOI's denoting central roles	-0.12	-1.10	0.29		
Task 3 (Review Member Role – Grant Review Domain)					
Group = BPMN					
Constant	0.10	0.29	0.77	0.76	< 0.01
Normalized % duration	0.41	2.71	0.02		
Run Count of AOI's denoting central roles	0.07	1.82	0.09		
Group = EPC-H					
Constant	-1.24	-2.75	0.02	0.81	< 0.01
Normalized % duration	0.62	1.84	0.09		
Run Count of AOI's denoting central roles	0.15	4.95	< 0.01		
Group = EPC					
Constant	1.53	1.66	0.12	-0.16	0.96
Normalized % duration	-0.1	-0.14	0.89		
Run Count of AOI's denoting central roles	-0.03	-0.22	0.83		

than those in the EPC group, and participants in the BPMN groups paid more attention to the task-relevant areas than those in the EPC-H group.

To test Hypotheses 3a and 3b (i.e., whether attention to task-relevant areas and visual association of these areas can predict task performance), we conducted regression analysis. For each task, we ran regressions using normalized percentage duration (as a measure of attention) and AOI run count (as a measure of visual association) as independent variables, and number of correct problem solving scores as the dependent variable. Table 7 shows the results for the first problem solving task of both domains. We observed a similar pattern for the other tasks.

Based on these analyses, we found distinct operations in performing tasks using each type of script. For the EPC-H script, visual association of the task-relevant areas (AOI run count) is the significant predictor of task performance. For BPMN, the percentage of normalized time spent on a task-

relevant area is a predictor of task performance. This difference between the predictors for EPC-H and BPMN can be explained by the need to visually associate information scattered over the model in EPC-H, as opposed to being confined in one area as in BPMN. In contrast, in the EPC script neither attention nor visual association predicted task performance.

Study 2: Eye Tracking and Verbal Protocol Analysis

In the EPC-H group, although task performance can be linked with visual association of information (AOI run count), visual association might not automatically result in cognitive integration. Hence, a second study was designed to find evidence of cognitive integration when reading EPC scripts (as visual association was not a predictor of task performance for BPMN scripts, we did not use it in the second study). The focus of this study was to repeat Study 1 (using EPC and EPC-H), testing Hypotheses 1 and 3 from Study 1 using

verbal protocol analysis in combination with eye tracking. During the experimental procedure, participants verbalized their thought processes as they answered the problem solving questions (Khatri and Vessey 2016). The thought processes were analyzed for evidence of cognitive integration.

Results

We first present the eye tracking results, comparable to Study 1. Then, we present the results of the verbal protocol analysis. The analysis for first task of both domains is presented below (the results for the other two tasks are consistent, but not presented to save space). As in the previous study, task performance of the EPC-H group was significantly higher than that of the EPC group. The mean task performance of all the tasks for the EPC-H was 2.50 and for the EPC group was 1.70 (t = 3.14, p < 0.01). The domain and modeling familiarity of the participants was not significantly different between the two groups. Table 8 shows the summary of the fixation count and duration for the first task of each of the two domains.

Table 8 shows that the EPC-H group spent considerably more time on the task than the EPC group. This result runs counter to the previous study, where the EPC group spent more time on the task than the EPC-H group. One possible explanation for this difference is the forced verbalization of the thought process in this study. The salient coloring that formed visual cues in the EPC-H models might have helped the participants elaborate on the answering process (we note that their answers were longer); thus, they spent more time on the task. In the absence of visual cues, EPC users verbalized less and took less time to completing the tasks.

Analysis of the AOIs related to the first task of both domains is presented in Table 9. Note that the AOIs in both the EPC and EPC-H models are the same in size; hence, we did not need to normalize the measures (unlike Study 1). The run count and AOI run count measures for the EPC-H group were significantly higher than the EPC group.

Similar to Study 1 (Table 7), we ran regression models to predict task performance. We found that for EPC-H, task performance depended only on run count (visual association) and not on AOI percentage duration (attention). There were no significant predictors of task performance for the EPC group. This result is similar to that of study 1.

Analysis of Verbal Protocol Data

Cognitive integration is the synthesis of information from two or more sources. This synthesis could be among concepts described in different parts of the model or between concepts in the model and concepts outside the model (prior knowledge). According to Nurgaleeva (2015), cognitive integration includes analysis of conceptual changes, acts of mental selection, and analogies. Ratwani et al. (2008) operationalized cognitive integration as making comparisons between, or forming relationships with, the information extracted. Accordingly, for this study we operationalize cognitive integration as verbal utterances of synthesis of information using logical relations of implication, comparison, contrast, relation, and negation. For example, if a model describes the processes of tasks A and B, then "task A is not the same as task B" or "task A requires more effort than task B" are both verbalizations indicating cognitive integration of concepts from two locations on a diagram (tasks A and B). Table 10 provides examples of the cognitive integration codes. Concepts not mentioned in the diagrams are italicized.

The verbal protocols were transcribed and provided to two coders (who were unaware of the research objectives) for analysis. The coders were also provided with the list of terms from the process models (to relate concepts that are not in the model). A coding document on cognitive integration was prepared and explained to the coders. Both coders independently coded the data. The coders then went through the coded data together and tried to resolve the cognitive integration incidences on which they disagreed. After this process, the coders were able to resolve most of the cognitive integration incidences. The interrater reliability was 84% prior to the reconciliation and 98% after the reconciliation. The number of cognitive integration incidences were summed for each participant and Pearson correlation coefficient was calculated. In view of the high interrater reliability, for the analysis below we used the coding of the first coder.

Table 11 contains the analysis of cognitive integration. The average number of cognitive integration instances was significantly higher in the EPC-H group than in the EPC group. This is consistent with the higher run count and AOI run count measures observed for these tasks.

To connect cognitive integration with visual association, we followed the procedure described by Ratwani et al. (2008). They ran multiple regression models to predict the number of qualitative extractions (as a measure of cognitive integration) from cluster-boundary fixations (similar to our AOI run count measure) and inner fixations (similar to total run counts minus AOI run counts). Table 12 shows that for the EPC-H group, cognitive integration depends only on visual association (i.e., AOI run count). The results are similar for the other two tasks.

Table 8. Fixation Metrics for First Task in Both Domains in Study 2			
	EPC M (SD)	EPC-H M (SD)	EPC-H vs. EPC t-value (p value)
Patient Treatment Domain			
Duration (in sec.)	53.74 (29.80)	103.75 (49.65)	3.02 (0.00)
No. of Fixations	211.57 (89.70)	473.07 (190.06)	4.51 (0.00)
Grant Review Domain			
Duration (in sec.)	44.29 (21.43)	90.61 (33.67)	4.06 (0.00)
No. of Fixations	188.85 (129.60)	401.15 (179.36)	3.50 (0.00)

Table 9. Duration and Run Count Metrics for First Task in Both Domains in Study 2			
	EPC M (SD)	EPC-H M (SD)	EPC-H vs. EPC t-value (p value)
Area: Admission Role			
% of duration	0.21 (0.09)	0.25 (0.10)	0.93 (0.18)
Run counts	36.07 (22.13)	79.07 (40.11)	3.41 (< 0.01)
AOI run count	7.00 (2.93)	19.07 (9.65)	4.32 (< 0.01)
Area: Member Role			
% of duration	0.27 (0.11)	0.32 (0.17)	0.81 (0.21)
Run counts	42 (21.61)	95.76 (44.83)	3.57 (< 0.01)
AOI run count	8.5 (5.48)	22.38 (9.38)	3.14 (< 0.01)

Table 10. Example Coding of Cognitive Integration Cases	
Coding Scheme	**Examples**
Synthesis of concepts	• You may miss a step that is really important • There is simply no way the members can make their own decision • That's why it is so pivotal that the diagnosis step happens in the processes • You could also make an error by not having the right member who could have summarized and entered the application • He [review chair] is the most important person in this process
Comparison of concepts	• So basically the review members make no decision, the review chair does • The review member is performing the same steps as review chair in some processes • Chair actually makes the decision whereas the review members read the application
Integrating with external concepts	• There is no background information about the *history* of the patient • This could be an issue and the hospital may not be *compensated* for the treatment • The patient might be *undercharged or overcharged* • "…and if you have a *quorum* that there must be at least 50% of the review members to be present

Table 11. Cognitive Integration (CI) Analysis Between the Two Groups							
	EPC (Patient Treatment)	EPC-H (Patient Treatment)			EPC (Grant Review)	EPC-H (Grant Review)	
	CI instances (SD)	CI instances (SD)	t-test (p value)		CI instances (SD)	CI instances (SD)	t-test (p value)
Task 1	0.62 (0.76)	1.42 (0.51)	3.08 (< 0.01)	Task 3	1.31 (1.43)	2.50 (1.31)	2.16 (0.02)
Task 2	1.00 (0.91)	1.67 (1.37)	1.42 (0.08)	Task 4	1.31 (1.18)	2.25 (0.86)	2.28 (0.01)

Table 12. Regression Analysis Using Instances of Cognitive Integration as Dependent Variable (Tasks 1 and 3)

Independent Variable	Coefficient B	t value	p value	Adj. R square	Overall p
Task 1 (Admission Role)					
Group = EPC-H					
Constant	0.64	3.39	< 0.01	0.79	< 0.01
Run Count of AOIs denoting central roles	0.04	6.74	< 0.01		
Run Count of AOIs other than central roles	-0.01	-0.85	0.41		
Group = EPC					
Constant	0.30	0.48	0.64	-0.08	0.59
Run Count of AOIs denoting central roles	-0.01	-0.13	0.89		
Run Count of AOIs other than central roles	0.01	0.94	0.37		
Task 3 (Review Member Role)					
Group = EPC-H					
Constant	-0.11	-0.21	0.84	0.71	0.01
Run Count of AOIs denoting central roles	0.11	2.93	0.01		
Run Count of AOIs other than central roles	0.00	0.26	0.79		
Group = EPC					
Constant	1.46	1.72	0.11	-0.17	0.90
Run Count of AOIs denoting central roles	-0.09	-0.45	0.66		
Run Count of AOIs other than central roles	0.02	0.44	0.67		

Discussion and Limitations ▪▪▪▪

Eye movement monitoring is poised to be an increasingly valuable process tracing technique in the next wave of decision making research (Glaholt and Reingold 2011). We utilize eye tracking as a means to expose cognitive processes of users reading conceptual modeling scripts. Two cognitive operations were identified in the literature as potentially detectable by eye tracking. However, to conclusively tie them to eye tracking metrics, two issues had to be resolved.

First, while metrics such as fixation count and duration indicate attention to a certain information element, this can be interpreted in two opposite ways. It can be related to the (successful) processing of an information element (Poole and Ball 2006) or to confusion about it (Goldberg and Kotval 1998). Considerable variation has been shown in fixation time and saccade length depending on the particular search task (Rayner 1998). Some studies have found that information complexity increases fixation durations (Horstmann et al. 2009), whereas others have found that it reduces durations (Chen and Pu 2010; Reutskaja et al. 2011). Fu et al. (2014) note that eye tracking data alone do not tell whether the user is using the information or having trouble processing it. Hence, results should be understood in the context of performance on given tasks (Djamasbi 2014). Our results, linking eye movement analysis with task performance through regression, show that high task performance is explained by

eye metrics related to attention (i.e., high percent normalized duration) or association of information between parts of a diagram (i.e., high AOI run count), rather than confusion caused by viewing parts of the models.

Second, the cognitive operation of integrating information from different model parts has previously been tied to visual association indicated by saccadic movements, but without firm evidence. We provide such evidence by combining eye tracking with verbal protocol analysis. As a result, we can conclude that the cognitive operations of attention to, and association of, information when reading a conceptual model can be detected and measured by eye tracking. Visual information association, in turn, indicates cognitive integration operations, as we have established in the second study. Note that verbal protocol analysis might alter natural eye movement, as users focus on certain areas while verbalizing their thoughts. However, this would bias the results in an opposite direction to our hypothesis, reducing the observed correlation between visual association and cognitive integration, which was nevertheless significant.

Manipulation checks were performed for both studies to ensure that the results are obtained because of significant differences in attention and association concerning task-relevant areas and not due to differences arising from task-irrelevant areas (Appendix F). In addition, it could be argued that participants tend to gaze at the center of the screen and

thus AOI's located in this region might be the reason why users have different fixation durations in the three models. We constructed artificial AOI's using 5% and 10% of the center of the models and compared them. Appendix F provides evidence to reject this alternate explanation.

Our findings also indicate that the measures of attention and association (indicating integration) can be correlated with task performance under certain circumstances. Psychology research has already established that attention to information and information integration are correlated with task performance (e.g., Hegarty et al. 2010; Johnson and Mayer 2012). We contribute to this literature by suggesting specific processes of how problem-solving tasks are performed with conceptual modeling scripts. In particular, Study 1 demonstrates that different cognitive operations are involved in using specific types of scripts for particular tasks. Similarly, different cognitive processes may be invoked for different tasks using the same script. Table 13 shows how task performance can be explained with different cognitive operations invoked by different scripts and tasks. The eye tracking analysis shows clearly when attention to relevant information does or does not affect task performance, highlighting the role of cognitive fit. In Table 13, BPMN and EPC-H fit the role-based problem solving task, as they make the relevant information salient (unlike EPC). Yet, this is not a binary classification, and differences exist in the cognitive fit of BPMN and EPC-H to the given task. These differences can be analyzed based on the differences in the cognitive operations invoked when addressing the task. Overall, the results show that task type and diagram type will determine together what cognitive operations users will adopt. Eye tracking is an effective technique to unravel these operations.

The different performance predictors for EPC-H and BPMN shown in Table 13 are not intuitive. We suggest the following possible explanation. Since EPC-H and BPMN convey the same information, we interpret this finding as indicating that the association operations, which are invoked when reading the EPC-H script, constitute extraneous cognitive load. That is, the task-relevant information, although color-highlighted, is distributed across several parts of the script. This is not intrinsic to the problem (as illustrated by the fact it is not required with a BPMN script), but imposed as a result of an interaction between representation and task. This cognitive load leads to lower task performance in the EPC-H condition than in the BPMN condition.

The regression analysis also indicates that when the information representation in the model does not highlight task-relevant areas, the cognitive processes are unclear. Attention and/or association did not explain task performance in EPC. A possible explanation is cognitive overload on the users

performing the tasks using such scripts. In such cases, individuals might use ineffective strategies to perform the tasks using the scripts. Our findings highlight the need for further research using eye tracking to achieve a better understanding of these strategies and their implications for designing conceptual modeling grammars.

One potential limitation of this study is low external validity. Eye tracking tools are still used primarily in a laboratory that allows a proper set up for eye data collection, so it is unclear whether or how the eye tracking behavior in such studies would generalize to a typical organizational setting in which conceptual modeling scripts are used. However, the focus of this research is on understanding if eye tracking can be used to measure basic cognitive processes associated with using conceptual modeling scripts. With our findings that confirm eye tracking metrics as an indicator of attention and cognitive integration of information, future research can measure integration processes in understanding visual representations.

Conclusions and Research Opportunities

Two important cognitive operations for information processing are attention to specific parts of models and integration of information across several parts of the model relevant to a task (Johnson and Mayer 2012). Our experiments collectively show how these operations can be used to account for performance on tasks that require understanding conceptual modeling scripts.

Prior empirical research has yielded insights into the effect of variations of a conceptual modeling grammar (and resulting scripts) on task performance. However, such research provides limited insight into how cognitive processes can explain task performance. Evidence on the role of attention and information integration is scarce (Johnson and Meyer (2012). Eye tracking promises to be useful in illuminating the cognitive processing differences of modeling grammar variants, differences that can explain task performance differences.

One opportunity arising from this research is to use eye tracking as a tool to test the cognitive implications of different ways of arranging constructs in a modeling script. Work on the physics of notations makes the case that certain choices of notation are preferable to others (Moody 2009), but provides no insight into how such choices affect reader interactions with a script. Similarly, work on the ontological expressiveness of grammars argues that certain choices lead to better representations of an underlying real-world domain, but does

Table 13. Summary of Results on Cognitive Operations and Task Performance		
Process Model Study	**Representation Fits Task (BPMN, EPC-H)**	**Representation Does Not Fit Task (EPC)**
Do users focus on task-relevant areas? (Table 6, Table 9)	Yes	No
Does attention to task-relevant areas explain task performance? (Table 7)	Yes for BPMN No for EPC-H	No
Do users associate information from task-relevant areas? (Table 6, Table 9, Table 11)	Yes	No
Does association of information from task-relevant areas explain task performance? (Table 7)	Yes for EPC-H No for BPMN	No

not provide a basis for understanding effects on model-reading and understanding (Wand and Weber 2002). Using eye tracking techniques, it is possible to identify how users attend to specific parts of the scripts and associate information from multiple sources to perform tasks. In addition, by combining eye tracking data with verbal protocol analysis, we have shown that eye tracking can be used to measure information integration. Thus, future work on evaluating attention and cognitive integration tasks may rely on eye tracking alone to unobtrusively measure integration processes in understanding visual representations. Insights from such work can guide developers and users of modeling grammars in devising rules for constructing grammars (e.g., constraining how constructs are combined), and in arranging constructs in creating scripts. For example, in designing or modifying a conceptual modeling grammar based on theoretical principles, eye tracking can be used to examine the effects of alternative design decisions (e.g., using alternative symbols or suggesting alternative ways of arranging symbols) on how readers attend to and associate information from scripts in performing tasks. Design principles such as those derived from a theoretical foundation may not offer sufficient guidance on how to depict information in a script (Lukyanenko and Parsons 2013). Nevertheless, eye tracking can yield insights on cognitive processing invoked by alternative representations, which can be used to determine best practices for manifesting theoretically derived principles in a grammar or in scripts generated from it.

Using eye tracking effectively will require executing a series of studies examining the effects of alternative grammatical constructs and/or visual arrangements on eye movement. This can be done in combination with measurements using NeuroIS tools such as fMRI (Dimoka et al. 2011) and EEG (Muller-Putz et al. 2015), which can be adapted for use in research to understand conceptual model grammars. Ideally, such research will, in particular, examine the effects of theoretically motivated variations, and of specific choices in notation and layout. In addition, eye tracking can be used to replicate earlier conceptual modeling research to better under-

stand how the results can be explained in terms of cognitive processing revealed via eye movements.

Additional research is also needed to expand the generalizability of the findings reported here. For example, we have conducted studies using small models and limited domains. Future work could systematically vary the size of models, and models from different (familiar and unfamiliar) domains, to determine whether our findings generalize. Likewise, we have examined a limited number of modeling grammars, and our work can be extended to different notations.

Another area for future work is to elucidate how different arrangements of constructs in scripts can be appropriate for different tasks. For example, from EPC and BPMN, we could suggest that highlighting elements scattered across the script is less effective than using a limited area of the script for tasks that involve understanding roles in a process model. Thus, knowledge of the kinds of tasks for which a script will be used can be used to tailor the script layout to the expected task(s).

In summary, eye tracking is a useful tool for conceptual modeling researchers in understanding how information displayed in graphical formats is understood and processed by readers of diagrams. In this paper, we have extended this understanding by validating the connection between metrics of saccadic movements and cognitive integration of information and by demonstrating how cognitive fit between a task and a representation can be analyzed and interpreted. This may, in turn, contribute to the design of modeling grammars and to the construction of scripts in ways that promote understanding. Such work not only can guide research in this area, but also help improve grammars used in practice.

Acknowledgments

References

Archibald, N., Hutton, S., Clarke, M., Mosimann, U., and Burn, D. 2013. "Visual Exploration in Parkinson's Disease and Parkinson's Disease Dementia," *Brain: A Journal of Neurology* (136), pp. 739-750.

Becker, J., Rosemann, M., and Uthmann, C. 2000. *Guidelines of Business Process Modeling*, Berlin: Springer.

Bera, P., Burton-Jones, A., and Wand, Y. 2011. "Guidelines for Designing Visual Ontologies to Support Knowledge Identification," *MIS Quarterly* (35:4), pp. 883-908.

Bodart, F., Patel, A., Sim, A., and Weber, R. 2004. "Should Optional Properties Be Used in Conceptual Modeling? A Theory and Three Empirical Tests," *Information Systems Research* (12:4), pp. 383-405.

Bousfield, W. A. 1953. "The Occurrence of Clustering in the Recall of Randomly Arranged Associates," *Journal of General Psychology* (49), pp. 229-240.

Burton-Jones, A., and Meso, P. 2006. "Conceptualizing Systems for Understanding: An Empirical Test of Decomposition Principles in Object-Oriented Analysis," *Information Systems Research* (17:1), pp. 38-60.

Burton-Jones, A., Recker, J., Indulska, M., Green, P., and Weber, R. 2017. "Assessing Representation Theory with a Framework for Pursuing Success and Failure," *MIS Quarterly* (41:4), pp. 1307-1333.

Cardoso, J., Mendling, J., Neumann, G., and Reijers, H. A. 2006. "A Discourse on Complexity of Process Models," in *International Conference Business Process Management*, J. Eder and S. Dusdar (eds.), Berlin: Springer, pp. 115-126.

Chen, L., and Pu, P. 2010. "Eye Tracking Study of User Behavior in Recommender Interfaces," in *User Modeling, Adaptation, and Personalization*, P. De Bra, A. Kobsa, and D. Cin (eds.), Berlin: Springer, pp. 375-380.

Chen, T., Wang, W., Indulska, M., and Sadiq, S. 2018. "Business Process and Rule Integration Approaches: An Empirical Analysis," in *Business Process Management Forum*, M. Weske, M. Montali, I. Weber, and J. vom Brocke (eds.), Berlin: Springer, pp. 37-52.

Cyr, D., and Head, M. 2013. "The Impact of Task Framing and Viewing Time on User Website Perceptions and Viewing Behavior," *International Journal of Human Computer Studies* (7:12), pp. 1089-1102.

Davis, R., and Brabander, E. 2007. *ARIS Design Platform: Getting Started with BPM*, New York: Springer.

Desimone, R., and Duncan, J. 1995. "Neural Mechanisms of Selective Visual Attention," *Annual Review of Neuroscience* (8:1), pp. 193-222.

Dimoka, A., Pavlou, P., and Davis, F. D. 2011. "NeuroIS: The Potential of Cognitive Neuroscience for Information Systems Research," *Information Systems Research* (22:4), pp. 687-702.

Djamasbi, S. 2014. "Eye Tracking and Web Experience," *AIS Transactions on Human–Computer Interaction* (6:2), pp. 16-31.

Evermann, J., and Wand, Y. 2005. "Ontology Based Object-Oriented Domain Modeling: Fundamental Concepts," *Requirements Engineering Journal* (10:2), pp. 146-160.

Figl, K. 2017. "Comprehension of Procedural Visual Business Process Models," *Business & Information Systems Engineering* (59:1).

Figl, K., Mendling, J., and Strembeck, M. 2013. "The Influence of Notational Deficiencies on Process Model Comprehension," *Journal of the Association for Information Systems* (14:6), pp. 312-338.

Fu, B., Noy, N., and Storey, M. 2014. "Eye Tracking the User Experience: An Evaluation of Ontology Visualization Techniques," *Semantic Web Journal: Interoperability, Usability* (8), pp. 23-41.

Fuller, R., Murthy, U., and Schafer, B. 2010. "The Effects of Data Model Representation Method on Task Performance," *Information and Management* (47), pp. 208-218.

Gemino, A., and Wand, Y. 2005. "Complexity and Clarity in Conceptual Modeling: Comparison of Mandatory and Optional Properties," *Data and Knowledge Engineering* (55), pp. 301-326.

Glaholt, M. G., and Reingold, E. M. 2011. "Eye Movement Monitoring as a Process Tracing Methodology in Decision Making Research "*Journal of Neuroscience, Psychology, and Economics* (4), pp. 125-146.

Glaholt, M. G., Wu, M., and Reingold, E. M. 2010. "Evidence for Top-Down Control of Eye Movements During Visual Decision Making," *Journal of Vision,* (10:5), pp. 1-10.

Glockner, A., and Herbold, A. 2011. "An Eye tracking Study on Information Processing in Risky Decisions: Evidence for Compensatory Strategies Based on Automatic Processes," *Journal of Behavioral Decision Making* (24), pp. 71-98.

Goldberg, J. H., and Kotval, X. P. 1998. "Eye Movement-Based Evaluation of the Computer Interface," in *Advances in Occupational Ergonomics and Safety,* S. K. Kumar (ed.), Amsterdam: ISO Press, pp. 529-532.

Grant, E., and Spivey, M. 2003. "Eye Movements and Problem Solving: Guiding Attention Guides Thought," *Psychological Science* (14), pp. 462-466.

Hegarty, M., Canham, M., and Fabrikant, S. 2010. "Thinking About the Weather: How Display Salience and Knowledge Affect Performance in a Graphic Inference Task," *Journal of Experimental Psychology* (36:1), pp. 37-53.

Horstmann, N., Ahlgrimm, A., and Glockner, A. 2009. "How Distinct Are Intuition and Deliberation? An Eye Tracking Analysis of Instruction-Induced Decision Modes," *Judgment and Decision Making* (4:5), pp. 335-354.

Houy, C., Fettke, P., and Loos, P. 2014. "On the Theoretical Foundations of Research into the Understandabality of Business Process Models," in *Proceedings of the European Conference of Information Systems*, Tel Aviv.

Jacob, R. 1995. *Eye Tracking in Advanced Interface Design, in Virtual Environments and Advanced Interface Design*, New York: Oxford University Press,.

Jacob, R., and Karn, K. 2003. "Eye Tracking in Human–Computer Interaction and Usability Research: Ready to Deliver the Promises," in *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement,* R. Radach, J. Hyona, and H. Deubel (eds.), Oxford, UK: Elsevier Sciences, BV, pp. 573-605.

Johnson, C. I., and Mayer, R. 2012. "An Eye Movement Analysis of the Spatial Contiguity Effect in Multimedia Learning,*" Journal of Experimental Psychology: Applied* (18:2), pp. 178-191.

Just, M. A., and Carpenter, P. A. 1976. "Eye Fixations and Cognitive Processes," *Cognitive Psychology* (8), pp. 441-480.

Khatri, V., and Vessey, I. 2016. "Understanding the Role of IS and Application Domain Knowledge on Conceptual Schema Problem Solving: A Verbal Protocol Study,*" Journal of the Association for Information Systems* (17:12), pp. 759-803.

Kim, J., Hahn, J., and Hahn, H. 2000. "How Do We Understand a System with (So) Many Diagrams? Cognitive Integration Processes in Diagrammatic Reasoning," *Information Systems Research* (11:3), pp. 284-303.

Kim, S., Dong, Z., Xian, H., Upatising, B., and Yi, J. 2012. "Does an Eye Tracker Tell the Truth About Visualizations? Findings While Investigating Visualizations for Decision Making," *IEEE Transactions on Visualization and Computer Graphics* (18:2), pp. 2421-2430.

Kintsch, W. 1979. "On Modeling Comprehension," *Educational Psychologist* (14), pp. 3-14.

La Rosa, M., Ter Hofstede, A. H., Wohed, P., Reijers, H., Mendling, J., and Van der Aalst, W. M. 2011. "Managing Process Model Complexity Via Concrete Syntax Modifications," *IEEE Transactions on Industrial Informatics* (7:2), pp. 255-265.

Lin, J., and Lin, S. 2014a. "Cognitive Load for Configuration Comprehension in Computer-Supported Geometry Problem Solving: An Eye Movement Perspective," *International Journal of Science and Mathematics Education* (12:3), pp. 605-627.

Lin, J., and Lin, S. 2014b. "Tracking Eye Movements When Solving Geometry Problems with Handwriting Devices," *Journal of Eye Movement Research* (7:1), pp. 1-15.

Lukyanenko, R., and Parsons, J. 2013. "Reconciling Theories with Design Choices in Design Science Research. Design Science at the Intersection of Physical and Virtual Design," in *Proceedings of the 8th International Conference on Design Science at the Intersection of Physical and Visual Design*, Helsinki, pp. 165-180.

Martini, M., Pinggera, J., Neurauter, M., Sachse, P., Furtner, M. R., and Weber, B. 2016. "The Impact of Working Memory and the 'Process of Process Modelling' on Model Quality: Investigating Experienced Versus Inexperienced Modellers," *Scientific Reports* (6:25561).

Mayer, R. E. 1989. "Human Nonadversary Problem Solving," in *Human and Machine Problem Solving,* K. J. Gilhooly (ed.), New York: Plenum Press.

Mayer, R. E. 2001. *Multimedia Learning*, New York: Cambridge University Press.

Mayer, R. E. 2010. "Unique Contributions of Eye Tracking Research to the Study of Learning with Graphics," *Learning and Instruction* (20), pp. 167-171.

Mayer, R. E., and Moreno, R. 2003. "Nine Ways to Reduce Cognitive Load in Multimedia Learning," *Educational Psychologist* (38:1), pp. 43-52.

Mendling, J., and Recker, J. 2008. "Towards Systematic Usage of Labels and Icons in Business Process Models," in *Proceedings of the 12th International Workshop on Exploring Modeling Methods in Systems Analysis and Design*, Montpellier, France, pp. 1-13.

Mendling, J., Reijers, H. A., and Cardoso, J. (eds.). 2007. *What Makes Process Models Understandable*?, Berlin: Springer.

Milosavljevic, M., and Cerf, M. 2008. "First Attention Then Intention: Insights from Computational Neuroscience of Vision," *International Journal of Advertising* (27:3), pp. 381-398.

Milosavljevic, M., Navalpakkam, V., Koch, C., and Rangel, A. 2012. "Relative Visual Saliency Differences Induce Sizable Bias in Consumer Choice," *Journal of Consumer Psychology* (22), pp. 67-74.

Moody, D. 2009. "The Physics of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering,*" IEEE Transactions on Software Engineering* (35:6), pp. 756-779.

Moody, D., and van Hillegersberg, J. 2009. "Evaluating the Visual Syntax of UML: An Analysis of the Cognitive Effectiveness of the UML Family of Diagrams," in *Software Language Engineering,* D. Gašević, R. Lämmel, and E. Van Wyk (eds.), Berlin: Springer, pp. 16-34.

Muller-Putz, G., Riedl, R., and Wriessnegger, S. 2015. "Electroencephalography (EEG) as a Research Tool in the Information Systems Discipline: Foundations, Measurement, and Applications," *Communications of the AIS* (37:46).

Newell, A., and Simon, H. A. 1972. *Human Problem Solving*, Englewood Cliffs, NJ: Prentice Hall.

Nordbotten, J. C., and Crosby, M. E. 1999. "The Effect of Graphic Style on Data Model Interpretation," *Information Systems* (9), pp. 139-155.

Nurgaleeva, L. 2015. "Cognitive Integration as the Dynamic Aspect of Modern Educational Practices," *Procedia—Social and Behavioral Sciences* (166), pp. 446-450.

Orquin, J., and Loose, S. 2013. "Attention and Choice: A Review on Eye Movements in Decision Making," *Acta Psychologica* (144), pp. 190-206.

Parsons, J. 2011. "An Experimental Study of the Effects of Representing Property Precedence on the Comprehension of Conceptual Schemas," *Journal of the Association for Information Systems* (12:6).

Parsons, J., and Wand, Y. 1997. "Choosing Classes in Conceptual Modeling," *Communications of the ACM* (40:6), pp. 63-69.

Parsons, J., and Wand, Y. 2008. "Using Cognitive Principles to Guide Classification in Information Systems Modeling," *MIS Quarterly* (32:4), pp. 839-868.

Petrusel, R., and Mendling, J. 2013. "Eye Tracking the Factors of Process Model Comprehension Tasks," in *Advanced Information Systems Engineering*, C. Salinesi, M. Norrie, and O. Pastor (eds.), Berlin: Springer, pp. 224-239.

Petrusel, R., Mendling, J., and Reijers, H. A. 2016. "Task-Specific Visual Cues for Improving Process Model Understanding,*" Information and Software Technology* (79), pp. 63-78.

Petrusel, R., Mendling, J., and Reijers, H. A. 2017. "How Visual Cognition Influences Process Model Comprehension," *Decision Support Systems* (96), pp. 1-16.

Poole, A., and Ball, L. J. 2006. "Eye Tracking in Human–Computer Interaction and Usability Research: Current Status and Future Prospects," in *Encyclopedia of Human Computer Interaction,* C. Ghaoui (ed.), Hershey, PA: Idea Group, pp. 211-219.

Porras, G., and Gueheneuc, Y.-G. 2010. "An Empirical Study on the Efficiency of Different Design Pattern Representations in UML Class Diagrams," *Empirical Software Engineering* (15), pp. 493-522.

Purchase, H. C., McGill, M., Colpoys, L., and Carrington, D. 2001. "Graph Drawing Aesthetics and the Comprehension of UML Class Diagrams: An Empirical Study," in *Proceedings of the*

2001 Asia-Pacific Symposium on Information Visualisation, Australian Computer Society, Inc., pp. 129-137.

Ratwani, R. M., Trafton, J. G., and Boehm-Davis, D. A. 2008. "Thinking Graphically: Connecting Vision and Cognition During Graph Comprehension,*" Journal of Experimental Psychology: Applied* (14), pp. 36-49.

Rayner, K. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological Bulletin* (124:3), pp. 372-422.

Recker, J., and Dreiling, A. 2007. "Does it Matter Which Process Modelling Language We Teach or Use? An Experimental Study on Understanding Process Modelling Languages Without Formal Education," in *Proceedings of the 18th Australasian Conference on Information Systems*, pp. 356-366.

Riedl, R., Hubert, M., and Kenning, P. 2010. "Are There Neural Gender Differences in Online Trust? An fMRI Study on the Perceived Trustworthiness of eBay Offers," *MIS Quarterly* (34:2), pp. 397-428.

Reijers, H. A., Freytag, T., Mendling, J., and Eckleder, A. 2011. "Syntax Highlighting in Business Process Models," *Decision Support Systems* (51:3), pp. 339-349.

Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye tracking Study," *American Economic Review* (101:2), pp. 900-926.

Russel, T., Green, M., Simpson, I., and Coltheart, M. 2008. "Remediation of Facial Emotion Perception in Schizophrenia: Concomitant Changes in Visual Attention," *Schizophrenia Research* (103), pp. 248-256.

Scheer, A. W. 2000. *ARIS: Business Process Modeling*, New York: Springer-Verlag

Shanks, G., Moody, D., Nuredini, J., Tobin, D., and Weber, R. 2010. "Representing Classes of Things and Properties in General in Conceptual Modelling: An Empirical Evaluation," *Journal of Database Management* (21:2), pp. 1-25.

Sharif, B., and Maletic, J. 2010. "An Eye Tracking Study on the Effects of Layout in Understanding the Role of Design Patterns,"in Proceedings of the IEEE International Conference on Software Maintenance, Timisoara, Romania.

Soffer, P., and Kaner, M. 2011. "Complementing Business Process Verification by Validity Analysis: A Theoretical and Empirical Evaluation,*" Journal of Database Management* (22:3), pp. 1-23.

Sweller, J. 1988. "Cognitive Load During Problem Solving: Effects on Learning," *Cognitive Science* (12:2), pp. 257-285.

Van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., and Schmidt, H. G. 2002. "Cognitive Load Theory and Aging: Effects of Worked Examples on Training Efficiency," *Learning and Instruction* (12), pp. 87-105.

Vanderfeesten, I., Reijers, H. A., Mendling, J., Van der Aalst, W. M., and Cardoso, J. 2008. "On a Quest for Good Process Models: The Cross-Connectivity Metric," in *Advanced Information Systems Engineering*, Z. Bellahsène and M. Léonard (eds.), Berlin: Springer, pp. 480-494.

Vessey, I. 1991. "Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature," *Decision Sciences* (22:2), pp. 219-240.

Vessey, I., and Galletta, D. 1991. "Cognitive Fit: An Empirical Study of Information Acquisition," *Information Systems Research* (2:1), pp. 63-84.

Wand, Y., and Weber, R. 1993. "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars," *Journal of Information Systems* (3), pp. 217-237.

Wand, Y., and Weber, R. 2002. "Information Systems and Conceptual Modeling: A Research Agenda," *Information Systems Research* (13:4), pp. 363-376.

Wang, Q., Yang, S., Liu, M., Cao, Z., and Qingguo, M. 2014. "An Eye Tracking Study of Website Complexity from Cognitive Load Perspective," *Decision Support Systems* (62), pp. 1-10.

White, S. 2004. "Business Process Modeling Notation (BPMN)," BPMI.org (https://www.omg.org/bpmn/Documents/BPMN_1-0_(R3).pdf).

Yusuf, S., Kagdi, H., and Maletic, J. 2007. "Assessing the Comprehension of UML Class Diagrams Via Eye Tracking,"in *Proceedings of the 15th IEEE International Conference on Program Comprehension,* Banff, Canada, pp. 113-122.

Zimoch, M., Pryss, R., Layher, G., Neumann, H., Probst, T., Schlee, W., and Reichert, M. 2018. "Utilizing the Capabilities Offered by Eye-Tracking to Foster Novices' Comprehension of Business Process Models," in *Cognitive Computing–ICCC 2018*, J. Xiao, Z-H. Mao, T. Suzumura, and L-J. Zhang (eds.), Berlin: Springer, pp. 155-163.

Zugal, S. 2013. "Applying Cognitive Psychology for Improving the Creation, Understanding and Maintenance of Business Process Models," Faculty of Mathematics, Computer Science and Physics, University of Innsbruck.

Zugal, S., Soffer, P., Haisjackl, C., Pinggera, J., Reichert, M., and Weber, B. 2015 "Investigating Expressiveness and Understandability of Hierarchy in Declarative Business Process Models," *Software & Systems Modeling* (14:3), pp. 1081 -1103.

About the Authors

Palash Bera is an associate professor and endowed Father Davis Professor at the ITM and Operations Department of Chaifetz School of Business, Saint Louis University. Palash's research interests are in conceptual modeling, use of eye tracking methodology, and agile methodology. His current research focuses on the use of conceptual modeling in predictive analytics and conceptual modeling in Artificial Intelligence. He has published in top information systems journals such as *MIS Quarterly, Information Systems Research,* and *Communications of the ACM.*

Pnina Soffer is an associate professor and the former head of the Information Systems Department at the University of Haifa. She received her BSc (1991) and MSc (1993) in Industrial Engineering and Ph.D. in Information Systems Engineering from the Technion–Israel Institute of Technology (2002). Her research deals with business process management, conceptual modeling, and requirements engineering, addressing issues such as goal orientation, flexibility, data and business-focused applications of process mining, as well as cognitive aspects of process modeling. She has published over 120 papers in journals and conference proceedings. She has served as a guest editor of a number of journal special issues and as an editorial board member in various journals. Pnina has held several organizational positions in major conferences, including serving as a program co-chair of BPM and CAiSE.

Jeffrey Parsons is University Research Professor and Professor of Information Systems in the Faculty of Business Administration at Memorial University of Newfoundland. His research interests include conceptual modeling, crowdsourcing, information quality, and recommender systems. His work has appeared in many outlets, including *MIS Quarterly, Management Science, Information Systems Research, ACM Transactions on Database Systems, IEEE Transactions on Knowledge and Data Engineering,* and *Nature*. Jeff is a senior editor for *MIS Quarterly*, a former senior editor for the *Journal of the Association for Information System*, and has served as program co-chair for a number of major information systems conferences.

USING EYE TRACKING TO EXPOSE COGNITIVE PROCESSES IN UNDERSTANDING CONCEPTUAL MODELS

Palash Bera
Operations and Information Technology Management, Chaifetz School of Business, Saint Louis University,
St. Louis, MO 63103 U.S.A. {palash.bera@slu.edu}

Pnina Soffer
Department of Information Systems, Faculty of Social Sciences, University of Haifa, Carmel Mount,
Haifa, ISRAEL, 31905 {spnina@is.haifa.ac.il}

Jeffrey Parsons
Faculty of Business Administration, Memorial University of Newfoundland,
St. John's, NL A1C 5S7 CANADA {jeffreyp@mun.ca}

Appendix A

Measuring Run Counts

To calculate the total run count, the total number of times the task-relevant areas were entered and exited was calculated, whereas for AOI run count, only the saccadic movements among the relevant areas were considered. Consider Figure A1 below. In it, there are nine saccadic movements where blue indicates fixations in nonrelevant areas and red indicates fixations in relevant areas. The red fixations are found in three relevant AOIs. In this figure, the total run count is five—arrows 2, 3, 5, 7, and 9. Arrows 1 and 8 are not considered as they are between fixations in nonrelevant areas. Arrows 4 and 6 are not considered as they are within the same relevant AOIs (as the saccades stay inside the AOI). The AOI run count is a subset of this total run count and in this figure the AOI run count is two (arrows 3 and 5). The arrows 2, 7, and 9 are ignored as they originate or end at nonrelevant AOIs.
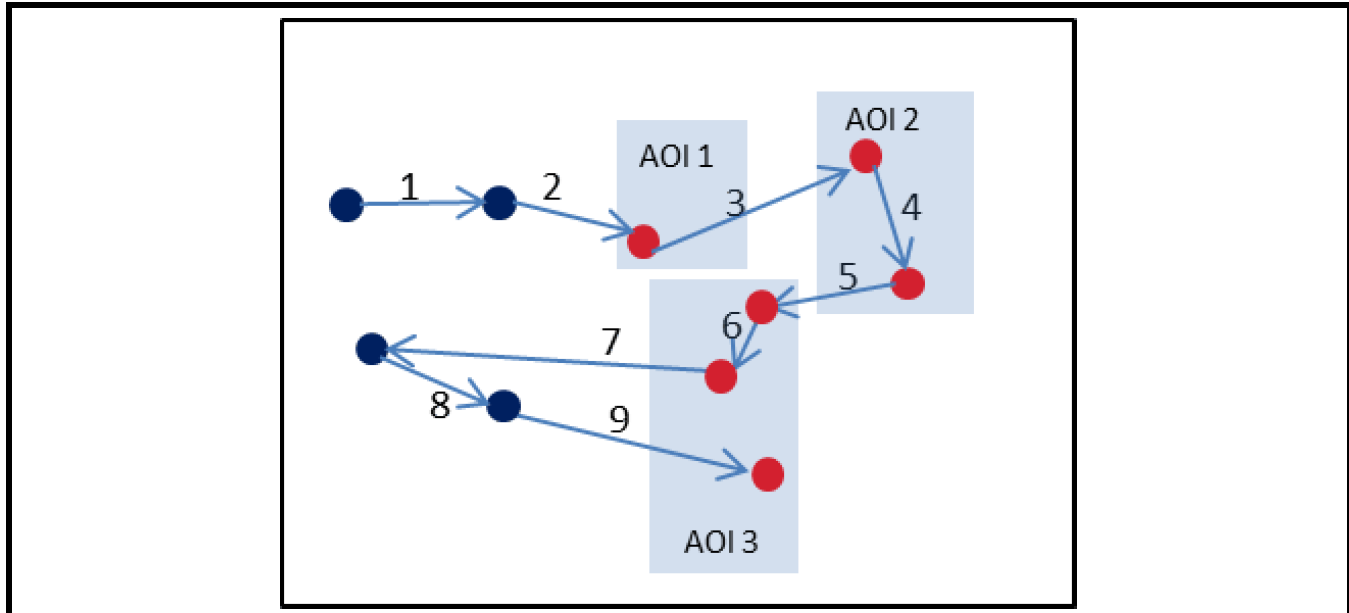
Figure A1. Example Illustrating Run Count Measurement

Appendix B

Coding of Problem Solving Task and Performance Analysis

The dependent variable for measuring domain understanding of the models is operationalized by the number of correct problem solving scores. To measure the number of correct responses, a set of possible correct responses was created by the first researcher. These responses were developed by reading domain descriptions. A university hospital nurse and an ethics manager working in a university were consulted on developing the set of correct responses. Two graduate students unaware of the objective of the study used this set as guidance to mark the participant responses. Table B1 shows the possible set of responses for the admission domain. Note that this set is not exhaustive and the coders used the table as guidance. A participant could provide multiple responses and each response was classified as correct or incorrect by the coder. The total number of correct responses was calculated for each problem solving task. It was possible for a participant to get zero if all the responses were incorrect. If more than one response was correct, then the total number of correct responses was counted. Because the responses were subjective in nature, we used two coders who independently coded the responses. The inter-rater reliability for Study 1 was 89% and 88% for the admission domain and the ethics domain respectively. Given the high inter-rater reliability, the responses coded by Coder 1 were used in the study.

Table B1. Sample Correct Responses of the Problem Solving Tasks for the Admission Domain	
Problem Solving Tasks	**Possible Correct Responses**
After stabilizing a patient, if his/her information is not provided to the *admission department* then what problems might arise?	• The admission department will not know how critical the patient is • The admission department may not know where to send the patient for operation • Difficulty in identifying whether the patient can be discharges • Difficulty in room assignment • No patient history is created • Billing cannot be done • Patient cannot be discharged
What will happen if patients are diagnosed immediately after arrival?	• There may be a misdiagnosis because the patient's body vitals were not checked and the patient was not stabilized • The patient may be misdiagnosed • The condition of the patient may worsen • The patient may die as he/she was not stabilized • No patient record is created

Prior to the analysis of ANCOVA, an analysis was done (Table B2) to determine whether the groups differed in terms of familiarity with the domain and with modeling. No differences were found.

Table B2. Domain and Modeling Familiarity Analysis for BPM Study							
	EPC	**EPC-H**	**BPMN**	**t-value EPC-H vs. EPC**	**p-value EPC-H vs. EPC**	**t-value BPMN vs. EPC**	**p-value BPMN vs. EPC**
	M (SD)	**M (SD)**	**M (SD)**				
Domain Knowledge	3.70 (0.38)	3.82 (0.35)	3.65 (0.37)	-0.87	0.19	0.36	0.36
Modeling Knowledge	4.90 (0.54)	4.80 (0.59)	4.87 (0.52)	0.48	0.32	0.17	0.43

ANCOVA is performed by aggregating the correct scores of the four problem solving questions (two for each domain) (Table 4, body of paper). Performance on the problem solving tasks is presented in Table B3. The results indicate that the effect of BPMN was stronger than the effect of EPC-H models.

Table B3. Analysis of Problem Solving Tasks for BPM Study								
Treatment	**PS mean (EPC)**	**PS SD (EPC)**	**PS mean (EPC-H)**	**PS SD (EPC-H)**	**PS mean (BPMN)**	**PS SD (BPMN)**	**F - value**	**P-value**
EPC vs. EPC-H	1.50	0.42	1.76	0.31			2.17	0.08
EPC vs. BPMN	1.50	0.42			2.18	0.62	12.20	0.001
EPC-H vs. BPMN			1.76	0.31	2.18	0.62	3.13	0.01

PS mean = Average correct number of problem solving tasks; Domain knowledge and modeling knowledge were used as control variables.

Appendix C

Detailed Experimental Procedure ████████████████

Study 1

Eye movements were recorded using *EyeLink 1000* eye tracking software**.** Participants were seated 70 cm from the display monitor (resolution of 1600 × 1200 and refresh rate of 85 Hz). A chin rest was used for head support. The EyeLink 1000 eye tracker records a minimum fixation of 4 milliseconds. The average percentage of rejected observations in the first study was 10.47%. This means that slightly more than 10% of the eye observations were not captured by the eye tracking device. After calibration, gaze-position error was less than 0.5 degree and was sampled at 1000 Hz. Once participants' eyes were calibrated, they were shown the problem solving questions one at a time and asked to read the questions carefully. Following this, they pressed a joystick to see the script (based on the group to which they were assigned) and verbalized the answers. Participants were asked to verbalize rather than write the answers as writing would have taken their eyes off the screen and their eye movements would not be captured properly. This strategy of verbally answering the questions so that the users do not need to type the answer and get distracted was used in Kagdi et al. (2007). If a participant forgot the question, then a research assistant repeated the question. When participants finished answering a question, they pressed the joy stick again to see the next problem solving question. To increase the generalizability, participants answered problem solving questions twice using scripts developed from two domains. The study took approximately 20 minutes to complete.

Study 2

The Tobii Pro X3-120 eye tracker was used in the second study. It has a sampling rate of 120 Hz and provides flexibility for participants to move during the experiment (up to 80 cm). A web-based experiment was setup. Participants answered the problem solving questions in three steps (Table C1). In the first step, participants read the question and clicked on the continue button to view the model. In the second step, while viewing the model, participants verbalized their thought process as they answered the question. The eye tracker recorded the participants' voice. The tasks performed by the participants were recorded by the eye tracker and were available in video. If participants were silent for 10 seconds, a research assistant would prompt the participant to verbalize his/her thought processes. The question and the model were not placed in the same screen to avoid having participants' attention distributed between the diagram and the question. After viewing the model, participants could go back to the question by clicking on the "back to the question" button or continue answering the question by clicking the "continue" button and typing the response at the next screen (step 3).

Fixations and durations were accumulated over visits to the model. This means if a participant visited the model two times by interacting among the three steps (e.g., clicking back to the question from the model and then again visiting the model) the total fixation count is the sum of all the fixations when participants visited the model.[1] The average percentage of rejected observations in this study was 10.35%. Participants averaged 29 minutes to complete the study.

Reference

Kagdi, J., Yusuf, H., and Maletic, J. I. 2007. "On Using Eye Tracking in Empirical Assessment of Software Visualization," in *Proceedings of the 1st ACM International Workshop on Empirical Assessment of Software Engineering Languages and Technologies,* Atlanta, pp. 21-22.

---

[1]In the first study, subjects could not maintain this interactivity, as the eye tracker (Eyelink 1000) did not allow subjects to go back and forth with the model and question (as this would require recalibration). Therefore, in the previous experiment, subjects read the question and then visited the model and then verbalized the answer in sequence. If subjects forgot the question, a research assistant provided a reminder.

Step 1: Participant exposed to the question

After stabilizing a patient, if his/her information is not provided to the admission department then what problems might arise?

Continue

Step 2: Participant exposed to the model
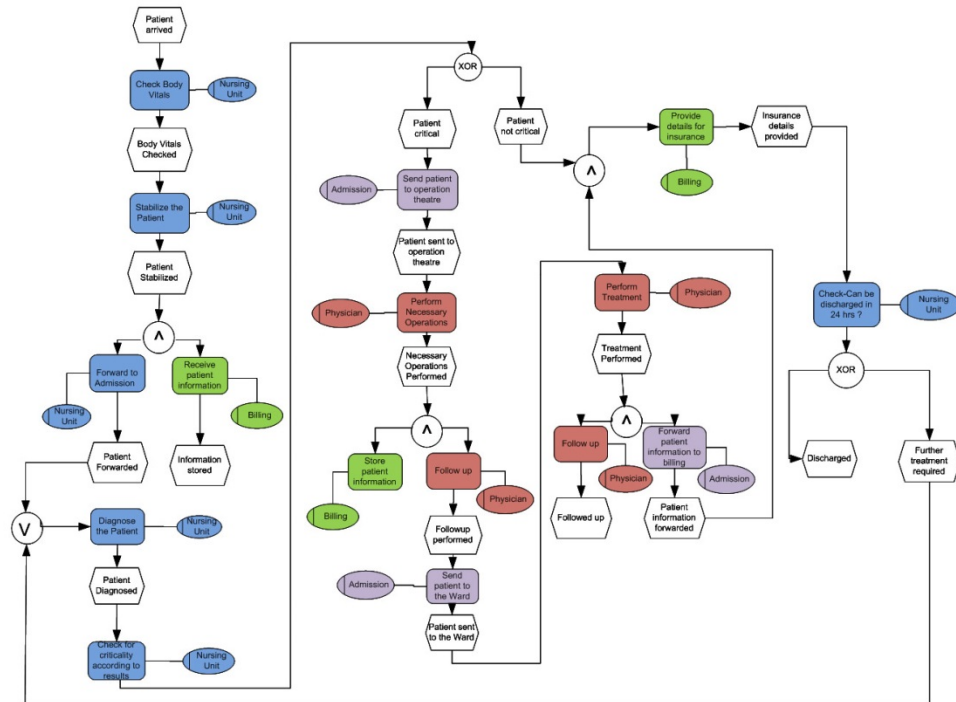
Back to the Question   Continue

Step 3: Participant provides the answer in the following space

Back to the Model   Save   Continue

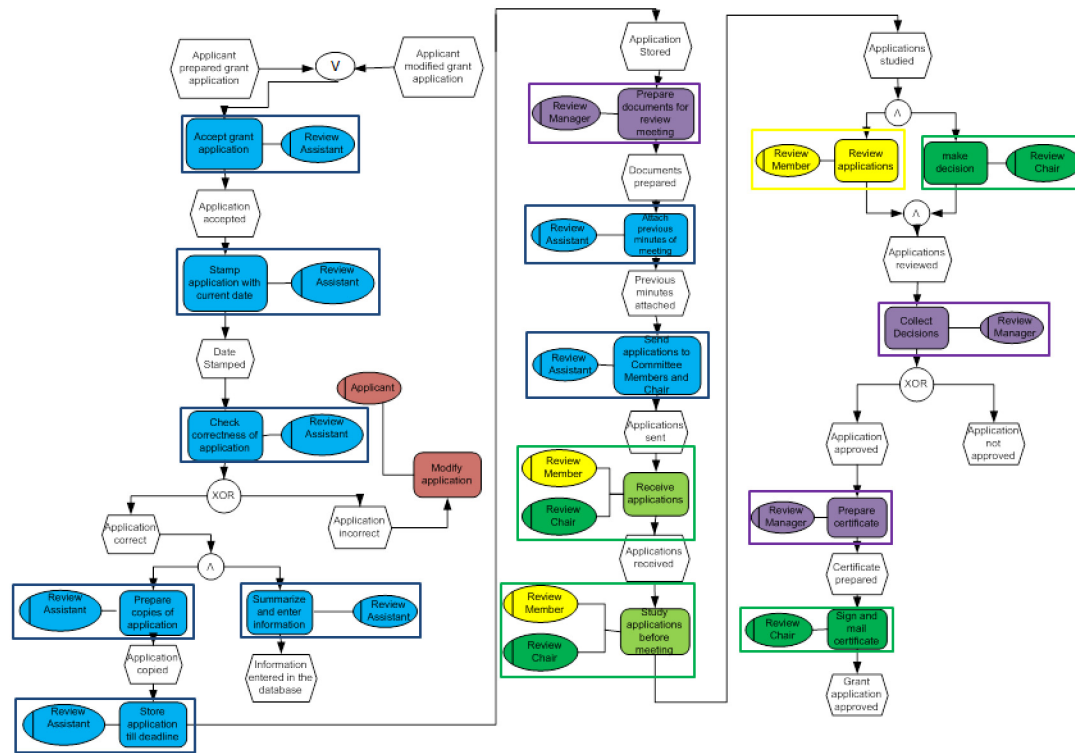**Figure C1. Steps in Answering Problem Solving Questions (Study 2)**

Appendix D

Models Used in the Experiments

Figure D1. EPC-H Model: Hospital Treatment Domain

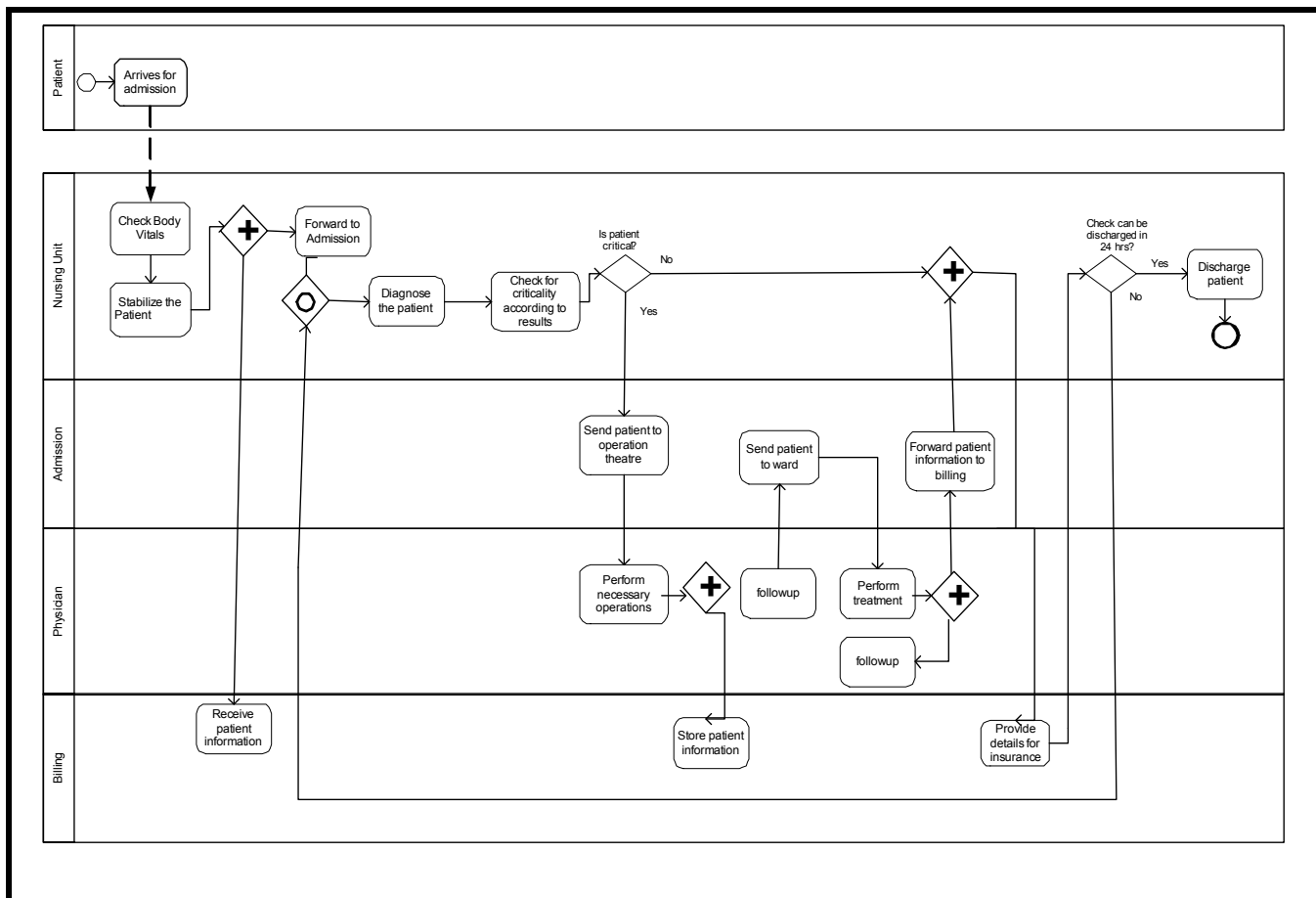**Figure D2. EPC-H Model: Grant Review Domain**

Figure D3. BPMN Model: Hospital Treatment Domain
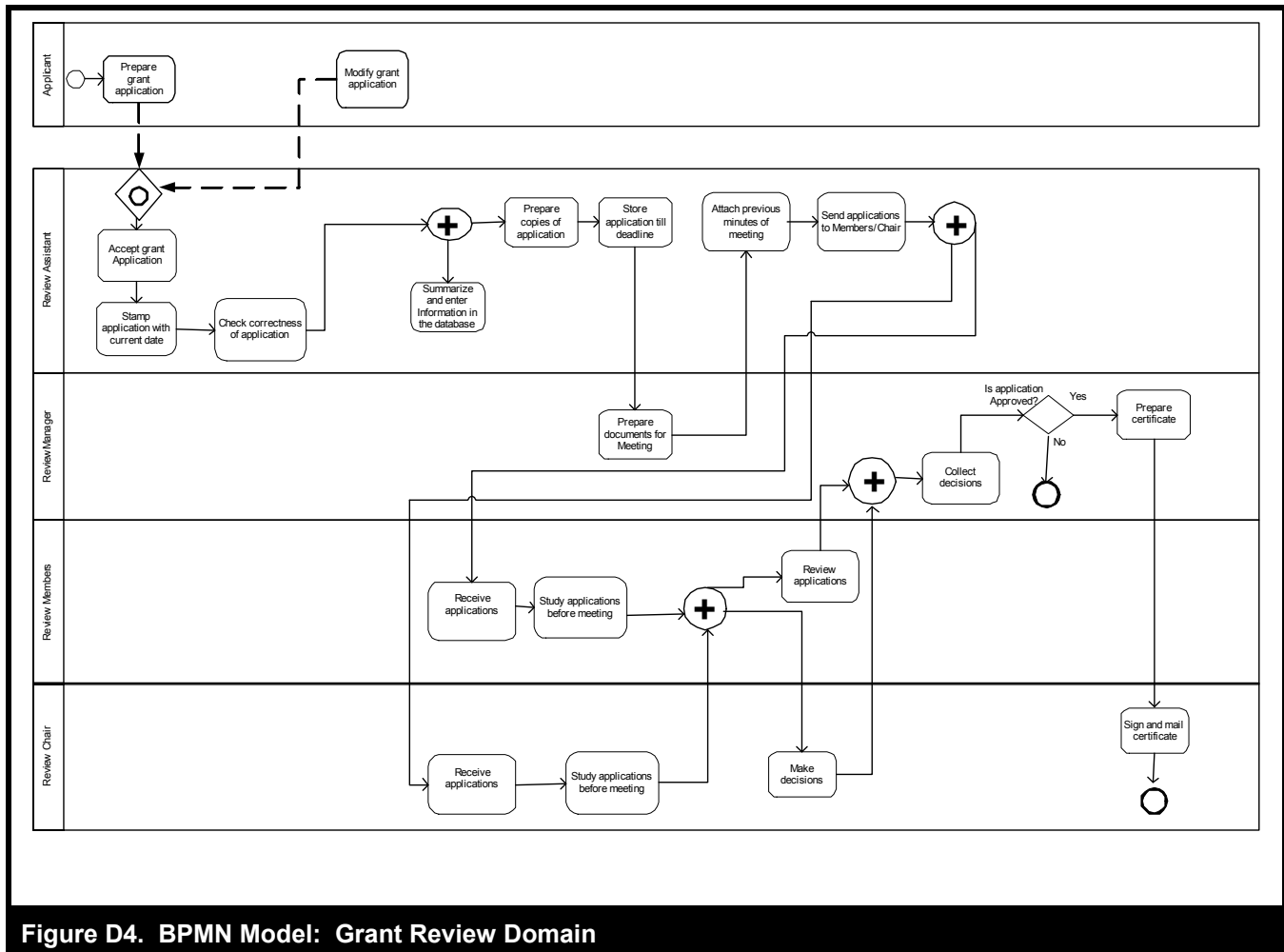
Figure D4. BPMN Model: Grant Review Domain

Appendix E

Procedures for Selecting Task-Relevant AOIs ▰▰▰▰▰▰▰

Task-relevant AOIs depend on the problem solving tasks. A set of steps was followed to select the task-relevant AOIs. These steps and examples of the problem solving tasks are provided below. It is to be noted that the number of possible answers for each task is large and there could be many variations of these answers.

Table E1. Procedure for Selecting Task-Relevant AOIs for the Hospital Treatment Domain		
Step 1	**Step 2**	**Step 3**
Develop possible correct answers for each task	Analyze the answers of the tasks to identify who performs these tasks	Set the role and the corresponding tasks as task-relevant areas
After stabilizing a patient, if his/her information is not provided to the admission department then what problems might arise?		
• The *admission department* will not know how critical the patient is • The *admission department* may not know where to send the patient for operation • *Admission department* cannot forward patient information to Billing • *Admission department* will have difficulty in assigning ward to the patient	Admission department performs these tasks	Admission department and the corresponding tasks performed by this department are the task-relevant areas.
What will happen if patients are diagnosed immediately after arrival?		
• Misdiagnosis by *nurse* because the patient's body vitals were not checked • Misdiagnosis by *nurse* because the patient was not stabilized • Misdiagnosis by *nurse* may lead to worsen patient's condition • *Nurse* does not intimate the admission department	Nursing unit perform these tasks	Nursing unit and the corresponding tasks performed by this department are the task-relevant areas.

Appendix F

Manipulation Checks

Analysis of Task Nonrelevant Areas

In this manipulation check, an analysis of eye metrics on task nonrelevant areas was performed. This was done to ensure that the difference in eye metrics is valid only on task-relevant areas and not on nonrelevant areas. Table F1 supports this for Study 1.

Table F1. Eye Metric for Task Nonrelevant Areas for Grant Review Domain of First Problem Solving Task							
	EPC	**EPC-H**	**BPMN**	**t-value EPC-H vs. EPC**	**p-value EPC-H vs. EPC**	**t-value BPMN vs. EPC**	**p-value BPMN vs. EPC**
Area: Review Manager	**M (SD)**	**M (SD)**	**M (SD)**				
% of duration	9.1 (4.36)	8.89 (3.31)	10.47 (4.48)	0.15	0.44	-0.84	0.21
Run counts	9.40 (1.18)	9.46 (1.30)	10.26 (1.53)	-0.15	0.44	-1.73	0.05
Area: Review Assistant							
% of duration	27.37 (9.38)	28.32 (0.08)	30.81 (5.68)	-0.29	0.38	-1.21	0.12
Run counts	11.80 (1.86)	12.00 (2.45)	12.26 (1.98)	-0.25	0.40	-0.66	0.26

Regression analysis was performed to test whether the percent of time on nonrelevant AOI's contributed significantly to problem solving performance. For this purpose, the percentage time for nonrelevant AOI for the first problem solving on grant review domain (Study 1) was calculated and used as independent variable (Table F2).

Table F2. Regression Analysis Using Percentage of Fixation Time on Nonrelevant Areas as Independent Variable (For the First Problem Solving on Grant Review Domain)				
Independent Variable	**B**	**t**	**P**	**Adj. R square**
Group = BPMN				
Constant	2.86	3.86	0.002	0.05
% of time spent on the AOI's denoting nonrelevant areas	-4.11	-1.30	0.19	
Group = EPC-Highlighted				
Constant	2.51	3.09	0.00	0.02
% of time spent on the AOI's denoting nonrelevant areas	-2.44	-1.14	0.25	

Analysis of Eye Fixations of the Central Regions of Models

It might be possible that users tend to look at the center of the models and the results of the experiment can be explained by users' tendency to focus on the center of the models rather than the task-relevant areas.

To test this proposition, we created two zones (AOIs) at the center of the models. One zone covered 5% of the entire area (pixel size 38,645) and the other covered 10% of the entire area (pixel size 77,292). These areas are shaded in blue in Figure G1. We performed fixation analysis in these zones for all the business process model types (EPC, EPC-H, and BPMN). For this analysis, from the admission domain, we selected the following question: "What will happen if patients are *diagnosed* immediately after arrival?"

Figure F1. Center Regions of the Process Models
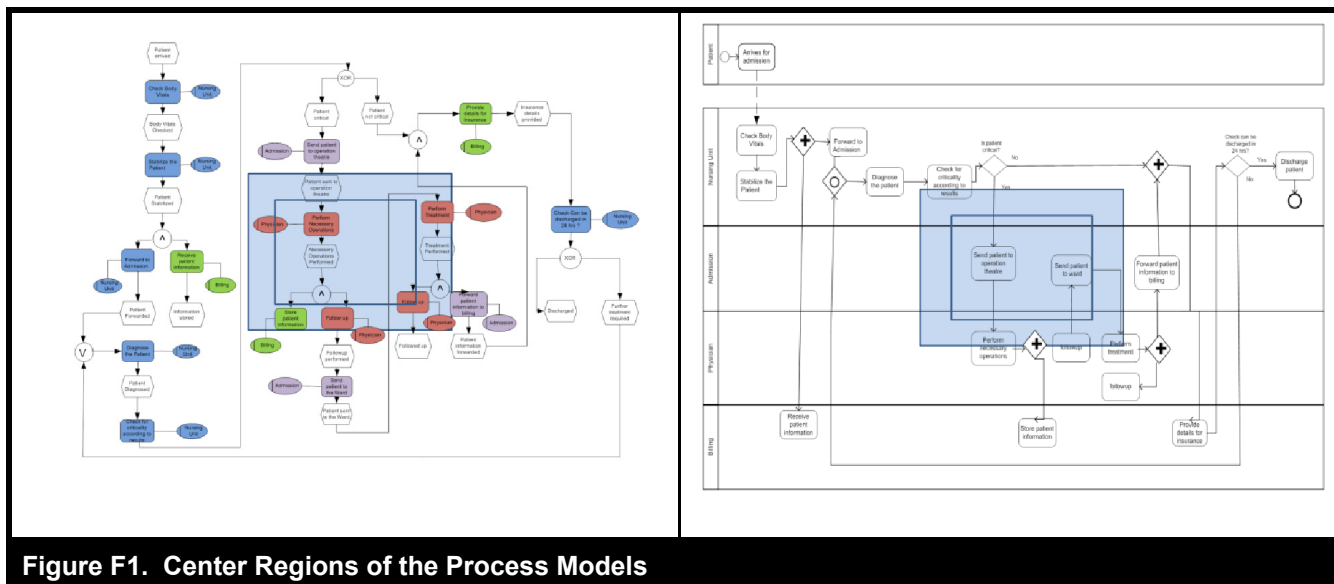
The analysis of the fixation numbers is shown in the table below.

Table F3. Fixation Analysis of the Center Regions of the Process Models					
	EPC M (SD)	EPC-H M (SD)	BPMN M (SD)	EPC-H vs. EPC t score (p value)	BPMN vs. EPC-H t score (p value)
5% AOI	3.2 (1.2)	2.9 (0.8)	2.67 (0.6)	0.71 (0.28)	1.02 (0.16)
10% AOI	10.6 (1.2)	10.1 (0.9)	15.7 (2.3)	1.12 (0.13)	8.93 (0.00)

When the area was selected as 5% of the center of the models, there was no statistical difference in the fixations among the models. However, when the area selected was 10% of the center of the models, there was statistical difference of the fixations between BPMN and EPC-H. This can be explained as part of the task-relevant area (i.e., the nursing swimlane) was present in the BPMN when 10% of the area is considered. These relevant areas (blue activities and roles) are not present in 10% of the center of EPC-H. The above results suggest that the results in the main paper cannot be explained by considering that participants look only at the central area of the model. Rather as the main analysis indicate, the task-relevant areas are responsible for the significant differences in the results.