UNIVERSITEIT GENT

FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

HOVENIERSBERG 24
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

Pattern Recognition of Resource-Event-Agent Conceptual Modelling Structures

Geert Poels [1]

Ann Maes

Frederik Gailly

Roland Paemeleire

February 2006

2006/369

[1] Corresponding author - Geert.Poels@UGent.be
All authors are members of the Management Informatics Research Unit, Department of Management Information, Operations Management and Technology Policy, Faculty of Economics and Business Administration, Ghent University, Hoveniersberg 24, 9000 Gent, Belgium.
E-mail: {Geert.Poels, a.maes, Frederik.Gailly, Roland.Paemeleire}@UGent.be
The authors wish to thank Prof. Maggie Geuens for her useful comments.

D/2006/7012/14

Pattern Recognition of Resource-Event-Agent Conceptual Modelling Structures

Abstract:

Resource-Event-Agent (REA) ontology-based conceptual modelling is a pattern-based approach to structural business process modelling. This paper hypothesizes that diagrammatic Entity-Relationship (ER) representations of REA ontology-based business process models (REA diagrams) are better understood and perceived to be of higher quality by business professionals than informationally equivalent ER diagrams that do not show REA pattern occurrences (Non-REA diagrams). The theoretical background for these hypotheses are (i) cognitive and perceptual psychology theories that explain why pattern recognition with REA diagrams is likely to occur, and (ii) Cognitive Fit theory that predicts performance effects when there is a good match between the mental and conceptual representations of the information required to solve a problem. An experiment with 124 business students showed that REA diagram users were more accurate than Non-REA diagram users when performing model comprehension tasks. Further, participants perceived a REA diagram as easier to interpret than a Non-REA diagram. Given that participants received minimal REA ontology education (though some of them were trained more intensively), the experiment results provide evidence of pattern recognition taking place. The experiment could not show that pattern recognition is stronger or more frequent when users are more familiar with the patterns (because of additional training).

Keywords. conceptual model, business domain ontology, modelling pattern, pattern recognition, task performance, model comprehension, Cognitive Fit theory, Semantic Distance, Human-Computer Interaction research, experiment

Introduction

Enterprise information systems are systems that support organization-wide transaction processing and managerial information provision (Vaassen, 2002). The implementation of the enterprise information system concept, typically by means of standardized software like ERP software packages, has proven to be a difficult undertaking. Mostly, the problems are not technology-related, but people-related. Though ERP systems are designed to operate by codifying a set of business processes, practitioners generally agree that the main barrier to success is exactly the lack of education about the underlying business processes (Dunn et al., 2005a).

The successful implementation of an enterprise information system depends on a thorough understanding of the company's business processes. Acquiring and sharing this understanding requires a tool that supports the description and communication of the policies that govern business processes and the specification of process-related information needs for managerial decision making and management control. Conceptual models may provide the required support as they help analysts understand a domain and sustain communication between users and developers, whilst providing input to the system design process (Wand and Weber, 2002).

To model business process-related policies and information, analysts have relied on common conceptual modelling grammars, methods, and notations such as the Entity-Relationship (ER) model, flowcharting, and the Unified Modelling Language (UML). However, these techniques do not offer specific guidance for modelling enterprise information systems. Although they can be used to describe business reality, they do not tell what objects and properties to include in business process models and how to structure these concepts. A more prescriptive approach is offered in the form of the Resource-Event-Agent (REA) enterprise domain ontology (Geerts and McCarthy, 2002). According to its most widely accepted definition, an ontology is "an explicit specification of a conceptualization: the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them" (Gruber, 1993). In the REA ontology this 'area of interest' is the enterprise. The REA ontology thus provides a normative framework for the conceptual modelling of enterprise information systems. It specifies the objects of interest in the enterprise domain and offers rules to connect these objects into information structures.

The conceptual core of the REA enterprise domain ontology is shown in Figure 1. The structure shown is a reusable pattern of relationships between the three kinds of objects that can be identified in any business process that effectuates an economic exchange or conversion: the events that constitute the process, the resources affected, and the agents involved (for an example of pattern use in business process modelling, see Appendix A).

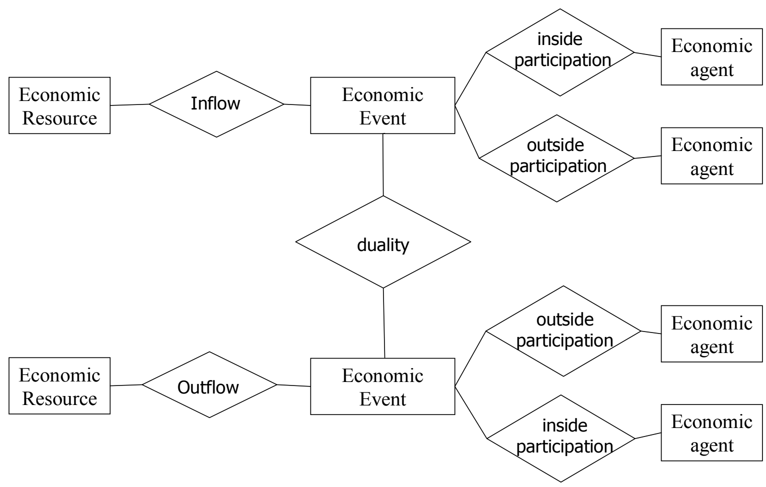


Figure 1. The REA core pattern

The conceptual modelling structure shown in Figure 1 originates in the REA accounting data model (McCarthy, 1982). Over the years, this data model has been extended into a comprehensive modelling framework that was specifically developed for the conceptual design of the enterprise information architecture of a company's accountability and policy infrastructure (Dunn et al., 2005a).

This paper presents an empirical test of the REA ontology. Although this ontology enjoys a remarkable interest of both educators (McCarthy, 1999) and practitioners (McCarthy, 2003), little is known about its usability as a conceptual modelling approach. Almost ten years ago, Dunn and McCarthy (1997) observed that REA-related research has focused on 'design science' (March and Smith, 1995) additions and improvements of the ontology, whereas fundamental questions such as 'is it better than using other approaches (or not using a prescriptive modelling approach at all)?' and 'what underlying mechanisms make the approach better than other approaches?' have not received much research attention. Today, this observation is still valid.

The paper addresses this lack of research by evaluating REA ontology-based conceptual modelling, focusing on the structuring capabilities offered by its patterns. We believe the usability of this ontology for the conceptual modelling of enterprise systems is associated to its pattern-based approach to business process modelling. In particular we postulate that the use of REA ontology patterns in the conceptual model of a business process improves the communication of domain information between analysts and users, as demonstrated by a better user comprehension of the model. The second section of the paper will further define this research question.

To investigate the research question, a research model is proposed in the third section. This model is based on the premise of two mental processes that occur when humans interact with REA ontology-based conceptual models: pattern recognition and learning. Accordingly, hypotheses are developed based on pattern recognition theories from cognitive psychology, and framed with an appropriate information systems theory called Cognitive Fit Theory (Vessey, 1991).

To test the research model, a laboratory experiment with 124 business students was conducted. The main findings are that participants using a conceptual business process model with REA conceptual modelling structures, demonstrated a more accurate understanding of the business processes and policies modelled. Participants also perceived such a model as easier to use when performing model comprehension tasks. The design of the experiment and the analysis of the data collected are presented in the fourth and fifth sections, respectively.

Finally, in the sixth section, the implications of our research findings are discussed, conclusions are drawn, and further research directions are outlined.

Research Question

The use of ontologies can improve communication between people and organizations, can create interoperability between systems, and can improve the reusability and reliability of the systems engineering process (Ushold and Gruninger, 1996). Over the years, a number of ontologies for the business domain have been developed (e.g. TOVE (Fox, 1992), Enterprise Ontology (Ushold et al., 1998), E3-value ontology (Gordijn, 2002), REA ontology (Geerts and McCarthy, 2002), Business Model Ontology (Osterwalder and Pigneur, 2003)). The main difference between these ontologies is the lens through which they look at business reality and that determines their conceptualization of an enterprise (i.e. the enterprise concepts they consider relevant).

What distinguishes the REA ontology from other business domain ontologies is its unique accountability and control perspective (a legacy from its origin as an accounting data model). The view of a business process that is inherent in the core pattern (see Figure 1) is a structural view, i.e. a view emphasizing the elements that constitute a business process and the relationships that provide structure to these elements. Given that the main purpose of a conceptual model is to communicate information about a domain (Gemino and Wand, 2005), a REA ontology-based conceptual model is a structural business process model that conveys information about the business process structure (e.g. what is being done?, what resources are affected?, who is involved?) and the business policies that constrain the process structure (e.g. if we give something up, something must be taken in return).

The accountability and control viewpoint of the REA ontology restricts its domain of application to exchange processes, i.e. processes that involve a 'transfer of ownership' (e.g. sales, purchases, payroll) (taking a predominantly financial accounting perspective with applications like bookkeeping and financial reporting) and conversion processes, i.e. processes that involve a 'transformation of form or substance' (e.g. production) (taking mainly a management accounting perspective with applications such as cost accounting and Activity Based Costing). The specific standpoint chosen also reduces those business processes largely to their transactional, respectively transformational core. Consequently, the main purpose of a REA ontology-based conceptual model of a business process is to convey information about the transactional/transformational core structure of a business process and the policies that govern the transfer of ownership or the transformation of form or substance that occurs within the process.

Business process structure and policies are modelled by analysts and communicated (via the models) to (future) information system users. Users must be able to express their information requirements to analysts and validate the models that are the formal representation of the requirements elicitation. Hence, users must be able to read and interpret conceptual models. According to Antony and Mellarkod (2004) patterns are most useful for non-experts, which they define as functional specialists that contribute towards systems analysis without having much expertise in modelling. This paper therefore evaluates REA-ontology based modelling from the point of view of the business professional in the role of (future) enterprise information system user and investigates whether the REA ontology patterns help the business professional better understand structural business process models.

Evaluating REA Ontology-Based Conceptual Modelling

Conceptual modelling techniques must be evaluated according to the purpose they serve (Parsons and Cole, 2005). Two recent developments are indicative of the REA ontology's importance in the current and future enterprise systems landscape. First, the REA ontology's developers are involved in a number of international standardization efforts for e-collaboration systems (e.g. ISO Open-EDI initiative, UN/CEFACT, ebXML, OAG). This participation has resulted in the adoption of (parts of) the REA ontology as the business process ontology used in the UMM business process and information model construction methodology (UN/CEFACT, 2003) and in the ECIMF system interoperability enabling methodology (ECIMF Project Group, 2003). Currently, the REA ontology is also being considered as the basis for the new ISO 15944-4 standard on the open-EDI business transaction ontology. Second, the REA ontology has been proposed as a theoretical basis for the reference models that underlie ERP systems like those offered by SAP (O'Leary, 2004). At best, these developments can be seen as measures of 'pragmatic success' (Moody and Shanks, 2003). The adoption of REA ontology-based modeling principles in these standards and the possible implication of widespread use, stresses even more the need for a focused evaluation study, clarifying the role of modeling patterns.

Until now, only a few research studies have examined the benefits that the REA ontology offers for the modelling of enterprise information systems. Dunn and Grabski (2000, 2001) showed in two experiments that accounting systems organized according to the REA ontology result in better user performance in information retrieval tasks, as compared to systems based on the traditional Debit-Credit-Account (DCA) accounting model. To answer some of the information retrieval questions in the experiments, the use of source documents (e.g. sales invoices, purchase orders) turned out to be necessary (Summers, 2000), so the evaluation was performed more at the system level than at the model level. A number of other experimental studies by Dunn and colleagues were conducted at the conceptual modelling level, but are intra-grammar (Wand and Weber, 2002) as they investigated alternative modelling formalisms used to represent REA pattern occurrences (ER diagrams versus ER models in textual Backus-Naur form (Dunn and Gerard, 2001), a set of disconnected binary ER diagrams (showing only two entity types and a single relationship type) instead of one comprehensive ER diagram (Dunn et al., 2005b)).

Three empirical studies have evaluated the quality of the REA ontology as a conceptual modelling approach. Poels (2003) presented a controlled experiment that compared the user comprehension of an ER diagram showing a single occurrence of the REA core pattern and an informationally equivalent ER diagram hiding this pattern. The results of this experiment indicate that the correctness of answering comprehension questions about the business process modelled is higher if a diagram that instantiates the REA core pattern is used. Akoka and

Comyn-Wattiau (2004) compared the REA ontology against DREAM, an object-oriented model for developing multidimensional accounting information systems. They showed that the semantic expressiveness of the DREAM model is higher, whereas the REA ontology is less complex. They did, however, not investigate the impact of these differences on the performance of model users or builders. Finally, Gerard (2005) showed that students possessing knowledge structures consistent with the REA ontology's core pattern (as obtained through training) could design more accurate conceptual accounting databases than students with less consistent knowledge structures.

Only the study of Poels (2003) evaluated the usability of the REA ontology from the point of view of the model user (instead of model builder). This study was the starting point for our current study. The observed increased accuracy effect when an ER diagram with a REA core pattern occurrence is used, is again hypothesized and will be assessed. In addition, this new study will also extend the research study presented in Poels (2003) in several ways. The study design and research model are radically altered, as described in the following sections, in order to enhance the validity and the generalizability of the findings. The current research study, amongst others also investigates the influence on perception-based variables like ease of use and user satisfaction. Additionally more significant affecting or confounding variables are included in the new research model and will be empirically tested. Further, although both Poels (2003) and Gerard (2005) demonstrated experimentally that the structuring capabilities of the REA core pattern result in improved performance on tasks that require interacting with the model, neither study attempts to explain the results from a modelling patterns perspective. The current study, on the other hand, provides strong theoretical arguments drawn from appropriate cognitive/perceptual psychology theories (schema theories, production system theories, template matching theories) and IS theory (Cognitive Fit Theory). Based on these theories we postulate that better performance with REA diagrams is a consequence of two mental processes that occur when humans interact with REA ontology-based conceptual models: pattern recognition and learning.

Batra (2005) observes that research in the area of conceptual data modelling patterns has just started. Also, according to Batra and Wang (2004), there have been very few experiments using patterns. Hence, investigating in a controlled manner whether it is the recognition of REA patterns that helps business users understand structural business process models better, may not only provide evidence of the usability of REA-based conceptual modelling; it may also provide information on underlying mechanisms which has implications for modelling patterns research in general.

Research Question Formulation

Guidelines for the evaluation of conceptual modelling techniques state that techniques must be evaluated in relation to other techniques, as there are no absolute evaluation scales (Gemino and Wand, 2003). Here we wish to assess the usability of REA ontology-based conceptual modelling of business processes from the model user point of view. To decide against which to compare this approach, the same solution as in Poels (2003) was chosen, where the prescriptive REA approach (a business domain ontology) was compared against a purely descriptive ER approach (a conceptual meta-model). Although the REA approach is not tied to a particular modelling formalism, the ER model is the preferred and most often used representation format for the REA ontology (Dunn et al., 2005a). Conceptual models of business processes developed using the REA approach are usually presented as ER diagrams (as in Figure 1). But an ER diagram showing business process related information can also be developed without explicit reference to a business domain ontology such as REA. Hence, the research question we formulate is:

> *Is an ER diagram that is used as a conceptual model of a business process better understood if the REA enterprise domain ontology was used to develop the diagram?*

Hypothesis Development

A number of theories from cognitive and perceptual psychology can be used to support the claim that the use of the REA approach is beneficial to the understanding of conceptual models. We frame our theoretical arguments within Cognitive Fit Theory (CFT) (Vessey, 1991), a precursor of the more general Task Technology Fit (TTF) theory (Goodhue and Thompson, 1995). CFT was developed for use in information representation contexts. The theory predicts that performance on information retrieval and interpretation tasks depends on cognitive fit, which is the match between the problem solver's mental representation of the required information needed to perform the task and the conceptual representation of the presented information, as understood from the artefact (e.g. a document or model). If these two representations match, then less cognitive effort will be involved in solving the problem posed by the task, resulting in effective and efficient task performance (Vessey, 1991).

The prediction of CFT was validated in a number of domains, including conceptual modelling (Agarwal et al., 1996; Dunn and Grabski, 2001) and thus provides an appropriate framework for investigating our research question. In the conceptual modelling context, cognitive fit can be related to the notion of semantic distance (Batra, 1993). Semantic distance is the gap between the problem solver's description of the information requirements (i.e. the mental representation) and the equivalent meaning in the modelling language (i.e. the conceptual representation). The smaller the semantic distance, the more cognitive fit there is.

To investigate our research question, the task, model, and fit (or semantic distance) concepts are instantiated as follows:

- Task: to retrieve domain information (and possibly verify against other information), where the domain is a business process;
- Model: a conceptual model of the domain in the form of an ER diagram;
- Fit: the model facilitates the task execution, i.e. the representation of the domain information in the ER diagram matches the mental representation of the information built by the task performer.

Consequently, the research question can be rephrased in terms of CFT as "does an ER diagram with REA pattern occurrences result in a better cognitive fit than an ER diagram without REA pattern occurrences?"

Pattern Recognition Theories

A better cognitive fit means a smaller semantic distance, so the question is whether the mental and conceptual representations match more closely if the conceptual model contains pattern occurrences. We believe the answer can be found in a mental process called pattern recognition. If one or more common patterns of information (i.e. recurring structures of information elements) are recognized in both the task description and the model, then the mental representation that is constructed matches closely the conceptual representation. Hence, semantic distance is small and cognitive fit is high. On the other hand, if no common pattern is recognized, then CFT predicts that more cognitive effort is needed to transform one representation into the other in order to solve the problem posed by the task. More cognitive effort means that executing the task takes longer (i.e. less efficient) and that, in the presence of human cognition limitations, the task outcome is of lower quality (i.e. less effective).

Pattern recognition research in psychology has identified enabling factors and the cognitive and perceptual mechanisms that trigger pattern recognition. Batra (2005) and

Antony and Mellarkod (2004) have suggested a number of psychological theories that might be useful for explaining how conceptual data modelling patterns are recognized. These theories include schema theories or production system theories like the Adaptive Control of Thought (ACT) framework (Anderson, 1996) and template matching theories, which have been used in the field of analogical reasoning and similarity finding. An example of the latter kind of theories is Structure-Mapping Theory (Gentner and Medina, 1998), which proposes three mechanisms by means of which pattern recognition works (Antony and Mellarkod, 2004):

- Literal similarity: finding information elements in the task description that are special cases of information elements in the pattern;
- Abstraction: mapping the task description into a more abstract, deeper structure, that is subsequently compared to the pattern. Compared to literal similarity, abstraction emphasizes more the structure of the information elements than the elements themselves;
- Analogy: finding parallels between the task description and the pattern. Compared to literal similarity and abstraction, analogy emphasizes more the 'overall picture' that is common to the task description and the pattern.

Whereas these notions explain the mechanics of pattern recognition, they assume that the patterns that are used in the comparison are readily accessible. Instead, production system and schema theories focus more on the retrieval of the patterns from long-term memory (LTM) and the matching of retrieved pattern elements with the contents of working memory. These theories infer that, for a pattern to be activated (i.e. make it accessible in working memory), it must first be present in LTM (Batra and Wang, 2004), which brings us to a second mental process involved, i.e. learning. Patterns are stored in LTM by means of a learning process, which can involve instruction, training and experience. Without this learning, pattern recognition cannot occur, so to investigate our research question, we need to consider both processes.

Figure 2 summarizes the previous discussion, showing the interplay between learning and pattern recognition as applied to the REA ontology patterns and from the perspective of CFT. For a good understanding of the figure, a solid arrow leaving the pattern recognition construct should be interpreted as 'a high probability that pattern recognition occurs when creating the representation'. Conversely, a broken arrow should be read as 'a low probability that pattern recognition occurs when the representation is created'. Based on the relative differences between the semantic distances depicted, three cases are distinguished:

(i) If REA patterns are strongly present in LTM and the diagram that is used for the task contains pattern occurrences (North-East corner in the figure), then pattern recognition in both task description and diagram is likely to occur such that a mental representation is created that closely matches the conceptual representation. As a consequence, semantic distance is small, cognitive fit is high and diagram users will perform the task efficiently and effectively.

(ii) If the same diagram with REA pattern occurrences is used, but the presence of REA patterns in LTM is weak (North-West corner), then the probability of users recognizing the patterns in the task description and the diagram is lower, so, on the average, the semantic distance between the mental and conceptual

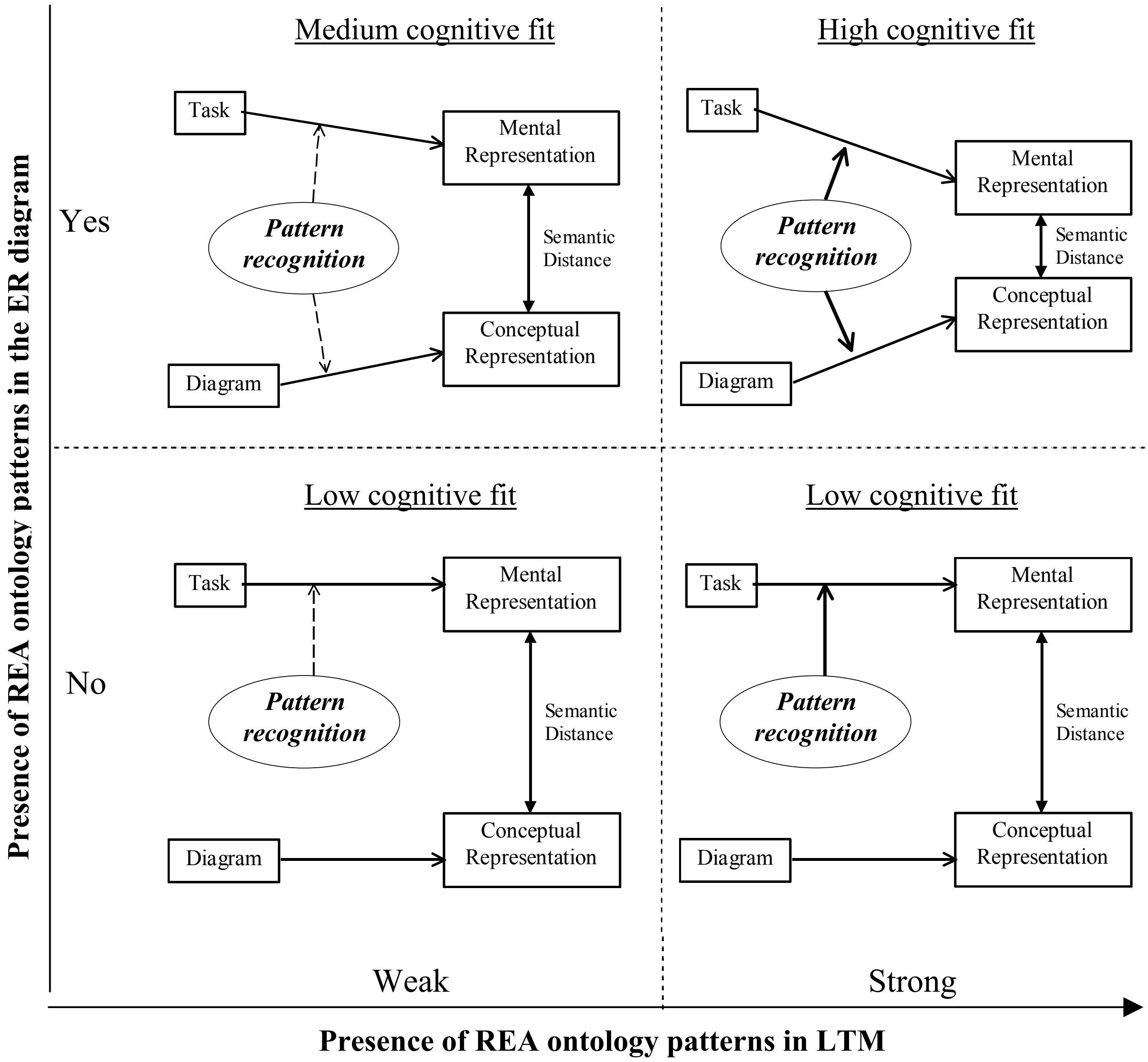


Figure 2. Cognitive fit as a function of the presence of REA ontology patterns in Long-Term Memory (LTM) and Entity-Relationship (ER) diagrams.

representations increases. There might still be some degree of cognitive fit (for those who do recognize (some of the) pattern occurrences), but it is generally lower than with users that can better activate the REA patterns in LTM.

(iii) If the diagram does not show REA pattern occurrences, then the conceptual representation is unlikely to match the mental representation, which means a larger semantic distance and lower cognitive fit. In that case, it does not really matter whether the presence of REA patterns in LTM is weak (South-West corner) or strong (South-East corner) as the patterns can only be recognized in the task description, but not in the diagram. Perhaps semantic distance is largest if the pattern presence in LTM is strong as a mental representation based on pattern recognition will definitely not match the conceptual representation (in case of weak presence the two representations might accidentally be the same), but this line of reasoning is tentative and will not be pursued further.

Note that the condition 'weak presence' of patterns in LTM is used instead of 'no presence'. Weak presence assumes that some learning took place, so there is a chance of pattern recognition. If there was no implicit (e.g. experience) or explicit (e.g. education) learning then patterns cannot be present in LTM and thus pattern recognition cannot occur. In that case the models in the North-West and South-West corners are essentially the same (removing the broken arrows) and no differences in semantic distance or cognitive fit between cases (ii) and (iii) are expected. To investigate our research question in the light of the pattern recognition theories, a setting that employs different intensity levels of learning is more interesting than one that uses a 'no learning' – 'learning' dichotomy.

Localization and Secondary Notation

REA ontology patterns and their representation as a template (i.e. the generic pattern, as in Figure 1) or as an ER diagram fragment (i.e. the pattern occurrence, as in Figure A-1 in Appendix A) exhibit two features that facilitate pattern recognition processes and mechanisms. These features ensure that REA ontology patterns behave like the patterns according to the notion assumed in the pattern recognition theories discussed, i.e. as easily recognizable recurring structures of information elements.

Taking the core pattern (see Figure 1) as an example, a first feature relates to the ontological structuring offered by a modelling pattern. The REA core pattern can be seen as a conceptual topological structure (Dunn and Grabski, 2001). Topology deals with the relative positioning and proximity of the information elements in a structure. In the REA core pattern (as well as in the other REA domain ontology patterns) topology is determined by the semantic relationships between the ontological elements present in the pattern. For instance, the core pattern shows that an entity type representing some type of economic event is related to other entity types representing semantically related concepts such as (types of) dual economic events, resources affected and agents involved (for definitions see Appendix A). If a user familiar with the REA core pattern (i.e. having stored the pattern in LTM) recognizes some diagram element as a type of economic events (e.g. purchases), she can interpret the relationships with that entity type in terms of the REA core pattern semantics (i.e. duality, stock-flow and participation relationships). Next, by following selected relationships, she can locate, in close proximity to the first element, other diagram elements as required by the task and interpret them correctly (e.g. inventory (what is purchased?), purchase agent (who was responsible for the purchase?), vendor (who was purchased from?), cash disbursement (what is done in return for the purchase?); see Figure A-1 in Appendix A).

In terms of the template matching and schema or production system theories discussed before, a matching mechanism like literal similarity (e.g. purchase is a special case of economic event) triggers a pattern recognition process (e.g. the REA core pattern in LTM is activated) that directs the user's attention to the diagram elements that are in close proximity to the first element identified. According to Larkin and Simon (1987), this feature, which they call localization, and the resulting attention direction mechanism lead to better performance.

Knowledge of REA conceptual modelling structures may thus facilitate navigating through the diagram as well as interpreting the presented information. The specific semantics attached to the REA ontology constructs (e.g. the economic duality expressed by a relationship between a pair of 'give-and-take' economic events) might also result in a more accurate understanding (i.e. interpreting the modelled reality as it was intended by the modeller).

The second feature relates to the visual structuring capabilities that are offered by a modelling pattern template. If a fixed format is used for both pattern template and pattern occurrences, then the pattern's topological structure of information elements becomes a spatial topological structure (instead of being a pure conceptual structure). In that case, the pattern not only dictates that related elements can be found in close proximity to each other, but that they are always found, physically, at the same place.

This feature is not shared by all modelling patterns, but is particularly relevant for the REA core pattern. According to Dunn and Gerard (2001), the structuredness of ER diagrams developed using the REA approach is enhanced by the implicit presence of an indexing mechanism in the REA core pattern. This indexing mechanism is formed by a diagram layout where the entity types representing economic resources, economic events, and economic agents are placed in respectively a left, middle, and right column of the diagram. Additionally, the top-down ordering of the entities that represent events may be used to reflect temporal relationships between event occurrences.

This particular placement of entity types on an ER diagram is a diagrammatic convention, not a mandatory rule of REA ontology-based conceptual modelling. Moreover, in integrated diagrams, showing interrelationships between different business processes, this convention is difficult to follow. However, for diagrams showing a single business process, these placement conventions are generally recommended (see e.g. Hollander et al. (2000), Romney and Steinbart (2003), Murthy and Groomer (2005)). They are also adhered to in the REA ontology's main educational textbook (i.e. Dunn et al. (2005a)). Therefore, for single-process models, these readability guidelines are an integral part of the 'good practice' of REA ontology-based modelling. Consequently, regardless of the type of company and business process modelled, users familiar with the REA ontology's core pattern may expect what information to find in which place of the diagram.

In conceptual modelling research, the effect that diagram layout has on comprehension has been postulated (see e.g. Schütte and Rotthowe (1998), Wand and Weber (2002)), but not investigated. In program comprehension research, layout issues (e.g. syntax highlighting, source code spacing and indentation) have received much more attention. Based on research in perceptual psychology, Petre (1995) has argued that layout, or 'secondary notation' as she calls it, enables an analogy mapping process which links perceptual clues to important information. If a reader has learned to look for them, secondary notation clues enable pattern matching (Petre, 1995).

The secondary notation provided by the R-E-A column format obviously helps triggering the analogy template matching mechanism. The 'Gestalt' effect, i.e. the informative impression of the whole that provides insight into the structure (Petre, 1995), is here more important than directing the attention towards individual diagram elements (as with localization and literal similarity). But also the abstraction mechanism might benefice from this layout clue as abstraction depends on analogical processes activated by both surface and structural similarities (Batra, 2005). According to Petre (1995) a person's categorization skills and ability to organize information on the basis of underlying abstractions are reflected in that person's ability to interpret surface features as clues to an underlying structure.

Hypotheses and Research Model

The research question, the CFT framework and the reviewed pattern recognition theories suggest a research model for this study. The research model shown in Figure 3 includes the relevant constructs and relationships that are used as variables and hypotheses in this study. Note that the model is already operationalized for the experiment conducted. However, the discussion of the operational details like treatments and measures is deferred to the Research Method section of this paper.

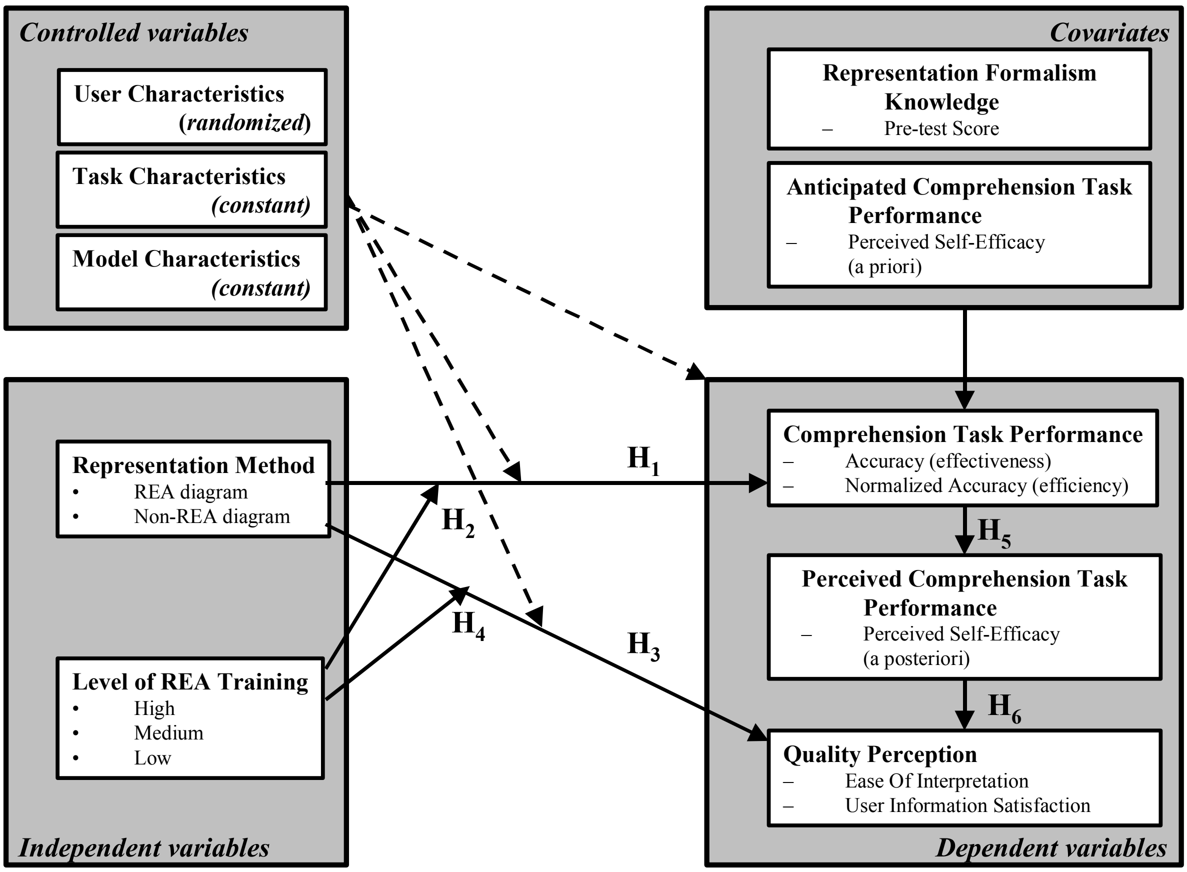


Figure 3. Research model

The main factor under investigation is the *Representation Method* used for conceptual models of business processes, and its effect on the user understanding of the models. Given that for the conceptual modelling of a single business process, the use of the REA approach results in an ER diagram (hereafter called a *REA diagram*) with REA ontology pattern occurrences that can be recognized by users familiar with the REA ontology, we hypothesize that such a diagram is better understood than an ER diagram that was not obtained using the REA approach (hereafter called a *Non-REA diagram*).

The argumentation for this hypothesis is that REA-knowledgeable users confronted with a task requiring information to be retrieved from an ER diagram (i.e. a model comprehension task), will create a mental representation in which the required information is structured according to the learned REA patterns (i.e. pattern recognition in the task). If the ER diagram is a REA diagram, then these users will recognize the pattern occurrences (i.e. pattern recognition in the model) resulting in a good match between the mental and conceptual representations. Following CFT, *Comprehension Task Performance* will be *effective* (i.e. successful retrieval and interpretation of the required information) and *efficient* (i.e. with low cognitive effort involved, thus fast). If on the other hand the ER diagram is a Non-REA diagram, then no REA patterns will be recognized resulting in a larger semantic distance between the mental and conceptual representations, and a less effective and efficient task performance. Hence, we formulate a first hypothesis as follows:

> *H_1: The use of a REA diagram instead of a Non-REA diagram to represent a conceptual model of a business process will have a positive effect on comprehension task performance (in terms of effectiveness and efficiency).*

The pattern recognition theories reviewed stress that patterns must be present in LTM before they can be recognized. The probability that patterns are activated when a user sees a REA diagram increases with the strength of their presence in LTM, which we postulate to be related to the intensity of the learning process. As experience is hard to control, we focus on learning as a result of education. Therefore, the *Level of REA Training* is introduced as a variable in the research model and an interaction effect with *Representation Method* is hypothesized (consistent with Figure 2).

> *H_2: The positive effect that the use of a REA diagram has on comprehension task performance will increase with the level of REA training. The comprehension task performance when using a Non-REA diagram is not affected by the level of REA training.*

It is acknowledged in the research domain that the usability of conceptual modelling techniques should not only be measured in terms of objective performance, but also in terms of users' attitudes towards the techniques, the tasks performed using the techniques, and their own performance (Topi and Ramesh, 2002). Frequently used user attitudes variables in empirical conceptual modelling research are *perception-based variables of model quality* such as perceived ease of use and user satisfaction (Kim and March, 1995; Burton-Jones and Weber, 1999; Dunn and Grabski, 2001; Gemino and Wand, 2005).

CFT predicts that if the conceptual and mental representations do not match, then the mental representation must be reformulated in terms of the conceptual representation, or vice versa. The additional cognitive effort involved will leave the user less satisfied with the diagram used and will create a perception in the mind of the user that the diagram is difficult to use. Using a diagram that has a higher cognitive fit with the comprehension task, users will have a more favourable perception of ease of use and will be more satisfied. Hence, we also hypothesize a direct effect of representation method on quality perception and an interaction effect with the level of REA training. This leads to the formulation of the third and fourth hypothesis as follows:

> *H_3: The use of a REA diagram instead of a Non-REA diagram to represent a conceptual model of a business process will have a positive effect on the user's perception of model quality (in terms of ease of use and satisfaction).*

> *H_4: The positive effect that the use of a REA diagram has on user's quality perception will increase with the level of REA training. Model quality perception when using a Non-REA diagram is not affected by the level of REA training.*

Model quality perceptions depend on cognitive fit and are therefore also contingent on the comprehension task that is performed. Moody (2002) theorized in his Method Evaluation Model that quality perceptions are caused by the actual effectiveness and efficiency of task performance. It is thus plausible that the relationship between *Representation Method* and *Quality Perception* (as well as the interaction effect with training) is indirect, via *Comprehension Task Performance*, implying that if cognitive fit is high and the comprehension task is performed effectively and efficiently, then quality perception will be high.

To verify whether the impact of representation method on quality perception is direct or indirect (or both), a link from comprehension task performance to quality perception is introduced in the research model. If users have not received feedback on their task performance (as in the experiment conducted; confer infra), this link must be implemented using another user attitude, *Perceived Comprehension Task Performance* (Goodhue and Thompson, 1995). Hence, the following relationships are hypothesized:

H_5: The perception of comprehension task performance is related to the actual comprehension task performance.

H_6: The perception of model quality is related to the perception of comprehension task performance.

The research model shows a number of other variables that might impact the dependent variables and thus confound the main effect of representation method as well as the interaction effect between representation method and level of REA training. Most of these variables will be controlled in the experiment to increase the internal validity of the study. They include *User Characteristics* other than level of REA training (e.g. demographic and personality characteristics, domain familiarity), *Task Characteristics* (e.g. task difficulty, nature of the task), and *Model* (or *Representation*) *Characteristics* other than representation method (e.g. size and complexity of the diagram, modelling language and notational system used). In a review of empirical conceptual modelling research, Topi and Ramesh (2002) have identified the user, task and model/representation variables as the three factors impacting user performance and attitudes in conceptual modelling. They further postulated interaction effects between these factors (which we indicated by broken arrows pointing towards our hypotheses). Also in TTF, user, task and technology are the main interacting factors affecting fit (Goodhue and Thompson, 1995).

Two further variables will be controlled in the study by including them as covariates in the data analysis. The first variable is *Representation Formalism Knowledge*, a user characteristic. Knowing how information elements and structures are represented using the constructs of the ER Model is a prerequisite for being able to derive business process structure and policies from ER diagrams. Research has shown that data modelling experience and familiarity with modelling techniques impact model comprehension (Kim and March, 1995; Parsons, 2003). As we cannot preclude that higher levels of REA training are associated with better knowledge of the ER formalism, this variable must be controlled and its effect cancelled out using appropriate data analysis techniques. Using this variable as a covariate also allows assessing its impact on the dependent variables (in particular *Comprehension Task Performance*).

The second variable, also a user characteristic, is the user's belief in her ability to successfully perform the task. This belief has been described using the concept of self-efficacy (Bandura, 1997), which has been shown to be related to actual task performance outcomes (Smith et al., 2003). As users having received more REA training might have higher levels of self-efficacy, the impact of this variable on actual comprehension task performance (and hence its confounding effect) can be controlled by using it as a covariate in the data analysis. To distinguish this variable from the dependent variables *Comprehension Task Performance* (i.e. actual performance) and *Perceived Comprehension Task Performance* (i.e. user perception of the actual performance), the self-efficacy of users *before* the task is referred to as *Anticipated Comprehension Task Performance* in the research model.

Research Method

A laboratory experiment was conducted to test the hypotheses. The strength of a laboratory experiment is that it allows to control confounding variables by creating an artificial research setting, at the expense of generalizability. Topi and Ramesh (2002) postulate in their generic research model for human factors related research in conceptual modelling, complex moderating effects of the user and task variables on the relationship between representation formalism/method and user performance and attitudes. However, according to Parsons and Cole (2005), the relative small amount of theory-based experimental work in the area of conceptual modelling techniques necessitates a focus on simple, theoretically causal relationships involving one or a few independent variables. This observation is especially true in the area of conceptual modelling patterns, where work is just starting (Batra, 2005). We agree with Parsons and Cole (2005) that in the absence of internal validity, external validity is of limited value; hence our choice of controlled experimentation as research method.

The design of the experiment was 2 × 3 between-subjects, with the independent variables as factors. The two levels of the first factor, *Representation Method*, are the experimental treatments (i.e. REA diagram or Non-REA diagram). The second factor, *Level of REA Training*, has three levels (i.e. Low, Medium, High). This experimental design allows assessing the representation method's impact on task performance and user quality perceptions (Hypotheses 1 and 3) as well as the interaction effect of representation method and level of REA training (Hypotheses 2 and 4).

Participants and Allocation to Experimental Groups

The participants were a group of business students enrolled in a junior-level Management Information Systems (MIS) course at an European university. Given the research question and model, the target population of the study are business professionals. As future enterprise information system users, business professionals need to work with conceptual models that are used for communicating with system analysts and developers. The group of business students participating in the study thus only approximates a representative sample of the target population. The experiment participants were generally younger and less experienced than functional domain experts working in business.

The advantage of student participants is that controlling the *User Characteristics* variable becomes easier. In particular, a student's familiarity with the REA ontology patterns is relatively easy to assess, compared to people working in business. Recall from the previous section that before pattern recognition can occur, patterns must be stored in LTM through a process of learning. Learning can be explicit (i.e. education) or implicit (e.g. experience), but implicit learning is much harder to observe and measure. The students participating in the experiment formed a homogeneous group with respect to their educational background and working experience, which would not be the case if business professionals were used. With this students group, the possibility of REA ontology patterns present in LTM because of working experience can practically be ruled out, which facilitates the operationalisation of the *Level of REA Training* independent variable and the control over a possible confound posed by working experience.

Conceptual modelling is a key module of the MIS course from which the study participants were drawn. The primary goal of this module is to give students the knowledge necessary to become future end-users that can effectively interact and collaborate with system analysts during requirements elicitation and validation tasks. According to this educational philosophy, acquiring conceptual modelling skills and in particular being able to understand conceptual models developed by analysts, is essential for business students and the module was conceived with this specific goal in mind.

During the part of the course module that focused on conceptual data modelling, students were first taught the constructs and grammatical rules of the ER model. The notation used for ER diagrams was based on UML as in Connolly and Begg (2002) (see Appendix B for an example). Although UML symbols were used for ER model constructs, the semantics of the UML class diagram constructs that are usually depicted by these symbols (as for instance in object-oriented software modelling) were not studied in the course.

Apart from studying the ER model, students were shown examples of and learned to read ER diagrams of various domains (e.g. university personnel management, hospital operations) with the purpose of understanding the domain information conveyed by the diagrams. The subsequent course module exercises required students to analyze ER diagrams by answering comprehension questions. These questions were chosen to mimic real-life situations where business professionals are confronted with tasks requiring a correct interpretation of the reality shown in the diagrams.

After the course sessions on ER modelling, the students were given a 1-hour lecture on business process reference models in which they were introduced to the main patterns of the REA enterprise domain ontology. Four reference models (sales, acquisition, payroll, and production) were explained with emphasis put on the REA core pattern occurrences contained in the models and, to a somewhat lesser degree, the extension of the business process transactional/transformational core with (optional) commitment events and type images (which are other REA ontology patterns). The lecture also stressed the REA diagram placement conventions and how this diagram layout helps in understanding the reality modelled.

Following the lecture on the REA ontology-based reference models, students could engage in one or two parts of practical course work similar to the previous ER diagram analysis exercises, but now performed on REA diagrams (i.e. conceptual models instantiating the REA ontology-based reference models seen in class). Students had to register for these optional parts of the course, so the identity of the students participating was known to us.

Apart from the lecture on reference models and the subsequent practical course work, no other lecture, practice session or course assignment was devoted to the REA ontology patterns. We thus observe three intensity levels of REA training:

- Low: students not participating in the practical course work on REA modelling;
- Medium: students participating in only the first part of REA practice;
- High: students participating in both the first and second part.

The experiment was conducted after the conceptual modelling module of the course was finished. A total of 124 students from the course participated in the experiment, of which 22 were classified as Low, 69 as Medium, and 33 as High with respect to their level of REA training. This uneven distribution of participants across the three levels of REA training implies that this factor is a measured, rather than a manipulated variable. Both participation in the REA exercises and in the experiment was voluntary making it impossible to ensure that each level of REA training was equally well presented. On the other hand, the allocation of the participants to the REA diagram and Non-REA diagram treatments was random, so representation method is a manipulated variable. The randomization of the *User Characteristics* variable per treatment controls for possible differential influences on the dependent variables. Table 1 summarizes the experimental design and participant allocation.

Table 1. Experimental design and participant allocation

Number of participants in each condition		**Representation Method**		Total
		Non-REA diagram	*REA diagram*	
Level of REA Training	*Low*	11	11	22
	Medium	34	35	69
	High	17	16	33
Total		62	62	124

Instrumentation

The REA and Non-REA diagrams used as experimental objects are included in Appendix C. These diagrams had to be representations of the same business process, allowing us to control the possible confounding effects of domain complexity and familiarity (as demonstrated in Burton-Jones and Weber (1999)).

The business process chosen for the experiment was the hiring of consulting services, which is a process that shares characteristics with both the acquisition and payroll processes that had been studied in the course (via their REA ontology-based reference models). The experiment participants had not looked at this particular business process during the course (neither during the optional course parts with REA exercises).

Another requirement for the experimental objects is that they are informationally equivalent (Siau, 2004). Otherwise, differences in information content may confound attempts to measure the impact of the independent variables on the dependent variables

(Parsons and Cole, 2005). To ensure the information equivalence of the REA and Non-REA diagrams, a same approach as in Poels (2003) was taken, where the Non-REA diagram was derived from the REA diagram by means of two information-preserving transformations:

- Objectification of many-to-many relationship types with attributes: a many-to-many relationship type between entity types A and B is replaced by a new (connecting) entity type C, a one-to-many relationship type between A and C, and a one-to-many relationship type between B and C. The UML classifier for C contains the attributes that were attached (in an association class) to the replaced many-to-many relationship type between A and B. The primary key of C is shown as a composition of the primary keys of A and B, to constrain the number of relationships between the entities of A and B.
- The physical repositioning of the diagram elements.

The first transformation was applied to the duality relationship type (*IsPaymentFor*; see Figure C-1 in Appendix C). The explicit modelling of the duality of 'give' and 'take' events is according to O'Leary (2004) the most distinctive structuring idea of the REA enterprise domain ontology. Also the *Orders* relationship type was objectified. In terms of the REA ontology, the explicit modelling of relationships between commitment events and the economic resources that have been committed to (by a 'reservation' relationship type) is another essential modelling structure, widely used in REA ontology-based reference models (see e.g. Dunn et al. (2005a)), including those shown in the course from which the experimental participants were drawn.

With the second information-preserving transformation we strived for a layout design for the Non-REA diagram that is different from the REA modelling conventions, but without being aesthetically inferior to that of the REA diagram. In the experiment of Poels (2003), the Non-REA diagram 'looked worse' than the REA diagram as the diagram elements were not placed in a logical way, resulting in a highly artificial diagram structure. In the layout for the Non-REA diagram that we created (see Figure C-2 in Appendix C), the entity type representing the outside party (*Consulting Firm*) was selected as the central diagram element. Further, the sequence of event occurrences (*Order Consulting Services – Get Consulting Services – Pay Consulting Services*) was positioned around this central element (in counterclockwise order) such that there is a logical path that can be followed when reading the diagram. Hence, the logic of the Non-REA diagram layout can equally well be justified as that of the REA diagram. The main difference with the REA diagram is that the usual three-column R-E-A arrangement is no longer present and that the sequence of event occurrences is no longer top-down.

The purpose of the two transformations was to create a diagram in which the REA ontology patterns were no longer easily and quickly recognizable to the REA-trained diagram user. Objectifying the duality and reservation relationship types helps in hiding the conceptual topological structures that REA-trained users expect to find when looking at a diagrammatic representation of a business process. The introduction of connecting entity types reduces localization in the Non-REA diagram and thus impedes the matching of diagram fragments to the pattern templates stored in LTM. The physical repositioning of diagram elements removes secondary notation clues from the Non-REA diagram and thus hinders the analogy mapping process. REA-trained users have learned another spatial topological structure than the one shown in the Non-REA diagram, making it unlikely that REA ontology patterns in LTM are activated because of perceptual clues.

The Non-REA diagram continued to show some characteristics of the REA ontology-based reference models. Therefore, as a manipulation check, the two diagrams were pilot tested in studies with graduate business students enrolled in an advanced Accounting Information Systems course and extensively trained in the use of REA ontology patterns (see Poels et al. (2004, 2005)). After answering a list of questions, testing their comprehension of the diagram given, the students in the Non-REA treatment group confirmed in an informal meeting with the experimenters that they were not aware being given a transformed REA diagram, hence demonstrating the effectiveness of the applied transformations.

Experimental Tasks

There were two tasks: a pre-test for measuring knowledge of the ER representation formalism (a covariate in the study) and the experimental task proper, which was a comprehension task.

The pre-test was the same for all participants and comprised 15 questions either literally taken from or derived from a similar test presented in Parsons and Cole (2005). The idea of this test is to assess the user's understanding of the semantics conveyed by ER diagram structural elements. To make this assessment independent of the user's interpretation of the diagram's surface semantics (i.e. the meaning carried over by the labels (names) of the diagram elements (Siau et al. ,1997)), Parsons and Cole (2005) replaced all meaningful labels on the ER diagram fragments used in the test by semantically void Greek letters. This operation also ensures that the possible interplay between a user's knowledge structures (e.g.

stored information patterns) and diagram comprehension is controlled as the questions can only be answered by referring to the diagram structural elements.

Our ER formalism pre-test (included in Appendix D) focused on the participant's understanding of five modelling concepts that are important for the correct representation and interpretation of business policies:

- Questions 2 - 4: structural constraints over a single relationship type;
- Questions 1, 8 - 10: structural constraints over a sequence of relationship types;
- Questions 5 - 7: inferred meaning at the entity/relationship instance level;
- Questions 11 - 13: consistency of relationship types that span a same set of entity types;
- Questions 14, 15: objectification of many-to-many relationship types (which is a relevant concept to test given our operationalization of the Non-REA diagram treatment).

The comprehension task was performed using the diagram given (either the REA or the Non-REA diagram). The task consisted of answering 15 questions about the consulting services process as it was represented in the diagram. As the diagram was the only information source available for answering the questions, participants were 'forced' to make an effort to understand the diagram.

The diagram comprehension questions (also included in Appendix D) required the participants to derive or verify the policies that govern the business process. To answer the questions, participants had to search the diagram for relevant pieces of information (entity types and/or attributes), identify the links between these pieces of information (relationship types), and interpret the constraints that are specified for these links (participation and cardinality constraints) in terms of business policies. Information equivalences ensures that the correct answer to a question can be found in either diagram. However, according to our hypotheses, we expect computational inequivalences (Siau, 2004) to arise between REA diagram users and Non-REA diagram users (that further increase with the level of REA training). In particular, with a REA diagram, finding relevant pieces of information should be easier if the learned spatial topological structure of the REA core pattern is recognized in the diagram. Likewise, the conceptual topological structure of the REA diagram, if recognized, helps selecting relevant relationships that determine the structure of the modelled business process. Finally, the correct interpretation of the constraints that are put on this structure is facilitated if the matching REA ontology pattern elements are activated in LTM.

Our questions were adapted from similar questionnaires for assessing and comparing the user comprehension of conceptual models that are produced via alternative conceptual modelling techniques (in particular we referred to the questionnaires in Bodart et al. (2001), Burton-Jones and Weber (2003), and Gemino and Wand (2005)). Though tailored to the particular case used in the experiment, comprehension questions of the kind we used are the conventional instrument for measuring how well users understand the information that is conveyed by a conceptual model (Gemino and Wand, 2003; Parsons and Cole, 2005).

The list of questions has been compiled such that the entire diagram is covered, though emphasis is put on the business policies governing the transactional core of the process (as well as its extension with commitments). Given that conceptual modelling techniques should be evaluated with respect to the purpose they serve (Parsons and Cole, 2005), our specific choice of questions was naturally guided by the purpose of the REA ontology, which provides modelling patterns for business processes as seen from an accountability and control perspective. As discussed before, the goal of a REA ontology-based conceptual model is to communicate information about the business policies that determine the transactional/transformational core structure of an exchange/conversion process. The relevance of the questions should be evaluated relative to this goal.

Finally, note that the information equivalence of the REA and Non-REA diagrams ensured that exactly the same questions could be used for both treatments. This way we control the *Task Characteristics* variable of the research model.

Measures

The covariate *Representation Formalism Knowledge* was measured by the score on the ER formalism pre-test, calculated as the number of pre-test questions correctly answered.

Anticipated Comprehension Task Performance, the second covariate, was measured via the self-efficacy construct, also called Perceived Self-Efficacy (PSE) (Ajzen, 2002). Since PSE is a task-specific construct, its measurement should also be specific to the particular task under investigation (Bandura, 1997). Although the effect of self-efficacy has been studied in the information systems domain, previous studies were mainly concerned with end-user training and basic computer use rather than information systems development tasks. Only the studies of Smith et al. (2003) and Ryan et al. (2000) are useful for our specific research context. Smith et al. (2003) measured student self-efficacy on declarative and procedural knowledge of Data Flow Diagrams and ER Diagrams and investigated the correlation with actual student performance. Ryan et al. (2000) studied the effect of PSE on data modeling task performance. The PSE measure that we developed is based on the PSE measures used in these two studies. The measure is referred to as the PSE_before measure, to distinguish it from the PSE_after measure used for *Perceived Comprehension Task Performance* (confer infra). The items of the PSE_before instrument are included in Appendix E.

Two dimensions of the *Comprehension Task Performance* dependent variable were measured: effectiveness and efficiency. To measure the effectiveness of a diagram in conveying information to the user, comprehension task accuracy was defined as the number of correctly answered comprehension questions (as in Kim and March (1995), Siau et al. (1997), Bodart et al. (2001), Shoval et al. (2002), Burton-Jones and Weber (2003), Parsons (2003)). To measure the efficiency of a diagram in communicating domain information to the user, the Normalized Accuracy measure of Bodart et al. (2001) was used. This measure relates a participant's comprehension task accuracy and task completion time. It is calculated as the number of comprehension questions correctly answered divided by the time taken to complete the comprehension task. Other research (e.g. Genero et al. (2002), Shoval et al. (2002), and Parsons (2003)) has used task completion time as an alternative, but completion time measures efficiency reliably only if a certain level of accuracy is reached. In practice, a better comprehension may be compromised by a faster comprehension, and vice versa (Bodart et al., 2001). Normalized accuracy should in this context be understood as a productivity measure, i.e. relating an output variable (accuracy of comprehension) to an input variable (comprehension time).

The *Perceived Comprehension Task Performance* was again measured via the PSE items that we defined, but now formulated in the past tense, and referred to as the PSE_after items (see Appendix E). We believed that, when formulated in the past tense, the items would capture the participants' perception of how well they accomplished the task (on condition that the instrument is administered directly after the experimental task).

Finally, the *Quality Perception* construct, and more specifically, the dependent variables Perceived Ease Of Interpretation (PEOI) and User Information Satisfaction (UIS), were assessed using existing measures that have been validated before in empirical studies on conceptual modelling. PEOI was measured using a four-item instrument by Gemino and Wand (2005) that was adapted from Moore and Benbasat (1991) and goes back to the Perceived Ease Of Use (PEOU) measure that Davis (1989) defined for the Technology Acceptance Model (TAM). UIS was measured using a four-item instrument by Dunn and Grabski (2001) that was adapted from Seddon and Yip (1992) and is an overall measure of information satisfaction summarizing more elaborate user satisfaction measures by Ives et al. (1983) and Doll and Torkzadeh (1988). The items of the PEOI and UIS instruments can also be found in Appendix E.

Operational Procedures

The experiment was organized as a class room exercise. The students of the MIS course were informed beforehand that this exercise was also part of a research study and that additional data in the form of questionnaires would be collected. However, no information was given with respect to the research questions that would be tested (to avoid experimenter bias). Participation was strictly voluntary and no course credits could be earned.

Students were motivated to participate in two ways. First, we promised feedback on their performance, suggesting that a similar exercise could be part of the final course exam. Second, four prizes (i-Pod Shuffles and Nanos) were distributed to the best performers. Students were informed that the ranking would be determined based on their scores, and, in case of equal scores, on the time spent. To avoid a ceiling effect, no time limit was set.

When participants entered the class room, they were randomly distributed by a teaching assistant across the two treatment groups and assigned a seat such that neighbours belonged to different treatment groups. The exercise/experiment consisted of four parts, to be executed in the order given (only when a previous part was handed in, the participant received the next part):

- A sheet containing instructions, asking for the participant's name and student number, and containing the pre-experiment questionnaire with the PSE_before items;
- The ER formalism pre-test (15 questions);
- The comprehension task (15 questions) – at this moment the REA or Non-REA diagram was given to the participant and the time was written down by a teaching assistant; when finished, the participant wrote down the time again (projected on a screen in front of the class room) and notified the teaching assistant (who collected the solutions and checked the times);
- The post-experiment questionnaire with the PSE_after items, and the PEOI and UIS items (intermingled on the questionnaire).

Data Analysis and Interpretation

Effect on Comprehension Task Performance (Hypotheses 1-2)

First the hypothesized effect of *Representation Method* on *Comprehension Task Performance* (Hypothesis 1) and the interaction effect with *Level of REA Training*

(Hypothesis 2) are tested. Table 2 presents descriptive statistics for comprehension task Accuracy and Normalized Accuracy.

To test the hypotheses, a MANCOVA was performed with Accuracy and Normalized Accuracy as dependent variables, *Representation Method* and *Level of REA Training* as factors, *Representation Method* × *Level of REA Training* as interaction term, and Pre-test Score (measuring *Representation Formalism Knowledge*) and a priori Perceived Self-Efficacy (PSE_before; measuring *Anticipated Comprehension Task Performance*) as covariates. Results are shown in Table 3.

Table 2. Descriptive statistics of Comprehension Task Performance for each experimental condition

Descriptive Statistics

	Representation Method	Level of REA Training	Mean	Std. Deviation	N
Accuracy	REA	Low	11,0909	2,16585	11
		Medium	11,2857	1,84026	35
		High	11,9375	1,38894	16
		Total	11,4194	1,79752	62
	Non-REA	Low	9,3636	1,96330	11
		Medium	10,5000	1,74512	34
		High	11,1765	1,38000	17
		Total	10,4839	1,77174	62
	Total	Low	10,2273	2,20242	22
		Medium	10,8986	1,82422	69
		High	11,5455	1,41622	33
		Total	10,9516	1,83841	124
Normalized Accuracy	REA	Low	,93645	,383526	11
		Medium	,76267	,285855	35
		High	,79778	,264155	16
		Total	,80256	,301610	62
	Non-REA	Low	,62115	,248058	11
		Medium	,74898	,301293	34
		High	,76573	,238924	17
		Total	,73089	,277240	62
	Total	Low	,77880	,354095	22
		Medium	,75593	,291475	69
		High	,78127	,248023	33
		Total	,76673	,290736	124

The model with Normalized Accuracy as dependent variable is not significant (p = 0.325), so no effect on the efficiency of comprehension task performance can be demonstrated. The model with Accuracy as dependent variable is significant (p = 0.001) with significant effects observed for the factors *Representation Method* (p = 0.003) and *Level of REA Training* (p = 0.019), and the covariate Pre-test Score (p = 0.007). No significant effects are found for the interaction term (p = 0.447) and the covariate PSE_before (p = 0.250).

These results provide partial support for Hypothesis 1: REA diagram users scored significantly higher on the comprehension task than Non-REA diagram users (mean Accuracy score of 11.4 for REA (maximum = 15) versus 10.5 for Non-REA with an observed effect size of 0.39 which represents a medium effect size (Cohen, 1988)). In other words, the use of REA patterns in the ER diagram had a positive effect on the effectiveness dimension of user comprehension, meaning that REA diagram users showed a more accurate understanding of the business process as modelled.

The absence of an interaction effect between *Representation Method* and *Level of REA Training* leads to the rejection of Hypothesis 2. The profile plot in Figure 4 shows that the *Level of REA Training* affects the accuracy of user comprehension, but that this effect is not essentially different for REA and Non-REA diagram users. Post-hoc tests showed a significant difference in Accuracy score between the Low and High training groups (p = 0.005) and a marginally significant difference between the Low and Medium groups (p = 0.058). Hence, regardless of the representation method used for the ER diagram, users with medium to high levels of REA training gave more correct answers to the comprehension questions than users with a low level of REA training.

Table 3. MANCOVA Comprehension Task Performance

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Accuracy	80,020[a]	7	11,431	3,950	,001	,192
	Normalized Accuracy	,686[b]	7	,098	1,170	,325	,066
Intercept	Accuracy	57,791	1	57,791	19,970	,000	,147
	Normalized Accuracy	,333	1	,333	3,983	,048	,033
Pre-test Score	Accuracy	22,144	1	22,144	7,652	,007	,062
	Normalized Accuracy	,004	1	,004	,045	,833	,000
PSE_before	Accuracy	3,876	1	3,876	1,339	,250	,011
	Normalized Accuracy	,108	1	,108	1,287	,259	,011
Representation Method	Accuracy	27,268	1	27,268	9,422	,003	,075
	Normalized Accuracy	,367	1	,367	4,387	,038	,036
Level of REA Training	Accuracy	23,890	2	11,945	4,128	,019	,066
	Normalized Accuracy	,007	2	,004	,044	,957	,001
Representation Method × Level of REA Training	Accuracy	4,692	2	2,346	,811	,447	,014
	Normalized Accuracy	,367	2	,183	2,192	,116	,036
Error	Accuracy	335,690	116	2,894			
	Normalized Accuracy	9,711	116	,084			
Total	Accuracy	15288,000	124				
	Normalized Accuracy	83,293	124				
Corrected Total	Accuracy	415,710	123				
	Normalized Accuracy	10,397	123				

a. R Squared = ,192 (Adjusted R Squared = ,144)

b. R Squared = ,066 (Adjusted R Squared = ,010)

This result is contrary to our expectations, as we postulated that more REA training would only be beneficial for REA diagram users (as Non-REA diagram users would not be able to recognize REA pattern occurrences in the diagram anyway). It is possible that the extra REA training provided to (some of) the students helped them understand the conceptual model better (no matter what representation method used) because of the additional experience gained with solving comprehension tasks of the type required in the experiment. The higher experience with model comprehension tasks might explain the positive effect of REA training on comprehension accuracy, regardless of the representation method used. Also Antony and Mellarkod (2004) proposed that users' ability to find similarities and analogies between the current task and tasks experienced earlier, might provide explanation of their behaviour. This issue should be investigated further in future research.

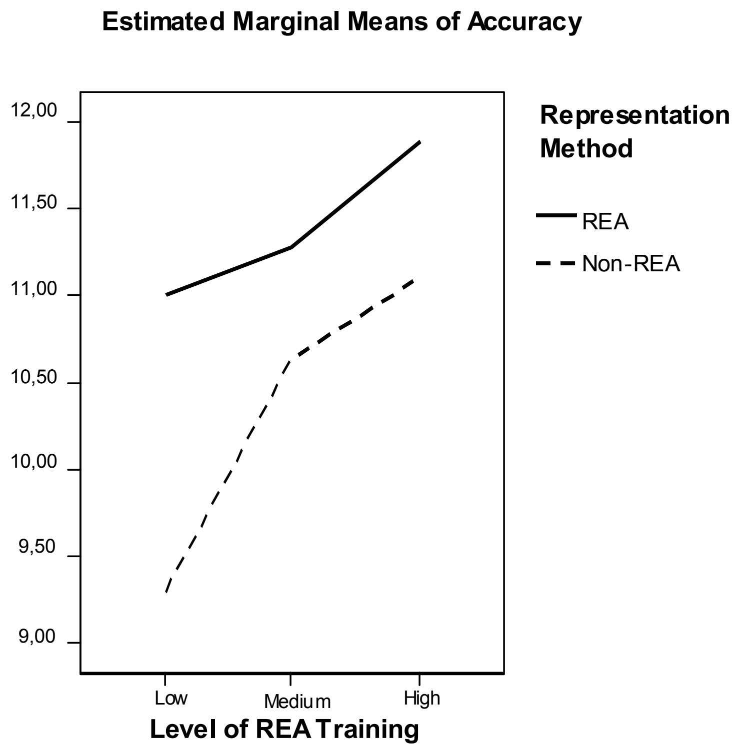


Figure 4. Profile plot Accuracy

As expected, participants that demonstrated a better understanding of ER modelling concepts relevant for modelling business policies (*Representation Formalism Knowledge*, measured in the pre-test), were also more accurate in the subsequent comprehension task. This result is not surprising given that knowledge of the representation formalism is a prerequisite for the correct interpretation of structural business process models.

The inclusion of *Representation Formalism Knowledge* as a covariate in the research model allows controlling this user characteristic and eliminating its effect when testing the main and interaction effects. As a Post-Hoc test we verified that the mean Pre-test Score was not different between the 2 × 3 experimental groups. An ANOVA revealed no significant differences, which is especially relevant for the 'measured' *Level of REA Training* variable. The lack of correlation between the Pre-test Score and the *Level of REA Training* confirms that the additional REA practice helped students acquire business process modelling experience (via the REA-based reference models), but did not deepen their ER formalism knowledge.

Finally, the performance on the comprehension task anticipated by the participants (as measured by PSE_before; Cronbach's alpha 0.848) was not related to their actual performance. In fact, this variable was not significantly correlated to any of the other variables used in the MANCOVA, indicating that the self-efficacy of the participants *prior to the experiment* was effectively controlled in the experiment and that this user characteristic plays no role of interest in the study.

Effect on Quality Perception (Hypotheses 3-6)

First, the hypothesized relationship between actual and perceived task performance is tested (Hypothesis 5). Perceived task performance was measured by the Perceived Self-Efficacy measure (in the 'a posteriori' version, i.e. PSE_after) using an instrument (the post-experiment questionnaire) that included also the items of the Perceived Ease Of Interpretation (PEOI) and User Information Satisfaction (UIS) measures. Whereas PEOI and UIS are existing measures, PSE_after was newly developed for this study. Therefore, and because all three measures capture perceptions about the use of a conceptual model, a reliability and validity analysis was conducted before hypothesis testing took place.

Initial Cronbach alpha's were 0.722 for PEOI, 0.826 for UIS, and 0.837 for PSE_after. A factor analysis revealed a problem of low discriminant validity for PEOI item 2 ("Using the conceptual schema was seldom frustrating"), so it was not further considered in the rest of the analysis (i.e. the average PEOI scores were calculated without item 2 scores). The removal of PEOI item 2 increased the Cronbach alpha value for PEOI to 0.791, well above the usual reliability threshold value of 0.70 (Nunally, 1978). It was further verified that all items of the new PSE_after measure loaded on a single factor, separate from the PEOI and UIS items.

To test Hypothesis 5, two separate regressions were performed (given the significant correlation between the Accuracy and Normalized Accuracy scores). ANOVA results showed significant correlations between Accuracy and PSE_after ($p = 0.015$) and between Normalized Accuracy and PSE_after ($p < 0.001$), leading to the acceptance of Hypothesis 5.

Hence, the hypothesized relationship between actual and perceived comprehension task performance was corroborated.

Next, the hypothesized main effects of *Representation Method* (Hypothesis 3) and *Perceived Comprehension Task Performance* (Hypothesis 6) and the interaction effect *Representation Method* × *Level of REA Training* (Hypothesis 4) on *Quality Perception* were tested. Table 4 presents descriptive statistics for the perception-based variables PEOI and UIS.

Table 4. Descriptive statistics of Quality Perception for each experimental condition

Descriptive Statistics

	Representation Method	Level of REA Training	Mean	Std. Deviation	N
UIS	REA	Low	5,022	,5528	11
		Medium	4,335	1,1261	35
		High	4,843	,9911	16
		Total	4,588	1,0422	62
	Non-REA	Low	4,454	,8277	11
		Medium	4,360	,9578	34
		High	4,485	,9456	17
		Total	4,411	,9200	62
	Total	Low	4,738	,7459	22
		Medium	4,347	1,0389	69
		High	4,659	,9698	33
		Total	4,500	,9830	124
PEOI	REA	Low	4,545	1,2496	11
		Medium	3,619	1,0668	35
		High	4,041	1,2224	16
		Total	3,892	1,1762	62
	Non-REA	Low	3,242	1,0337	11
		Medium	3,607	1,1502	34
		High	3,843	1,2862	17
		Total	3,607	1,1680	62
	Total	Low	3,893	1,3027	22
		Medium	3,613	1,1005	69
		High	3,939	1,2401	33
		Total	3,750	1,1761	124

The hypotheses were tested by means of a MANCOVA with PEOI and UIS as dependent variables, *Representation Method* and *Level of REA Training* as factors, *Representation Method* × *Level of REA Training* as interaction term, and PSE_after as covariate. The inclusion of PSE_after as a covariate in the analysis is legitimate as covariates can be (continuous) independent variables in their own right (Hair et al., 1998) (and given that the other independent variables are categorical). The results of the MANCOVA are shown in Table 5.

Both models are significant (p < 0.001) with a strongly significant effect observed for PSE_after (p < 0.001). In the model with PEOI as dependent variable there is also a marginally significant effect of the *Representation Method* factor (p = 0.083). Further, no other effects of the factors and interaction term on the dependent variables are observed.

Table 5. MANCOVA Quality Perception

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	UIS	34,65 a	6	5,775	8,023	,000	,291
	PEOI	64,14 b	6	10,69	11,80	,000	,377
Intercept	UIS	29,91	1	29,91	41,56	,000	,262
	PEOI	3,479	1	3,479	3,840	,052	,032
PSE_after	UIS	28,12	1	28,12	39,06	,000	,250
	PEOI	51,55	1	51,55	56,90	,000	,327
Representation Method	UIS	,765	1	,765	1,063	,305	,009
	PEOI	2,776	1	2,776	3,064	,083	,026
Level of REA Training	UIS	3,144	2	1,572	2,183	,117	,036
	PEOI	1,971	2	,985	1,088	,340	,018
Represenation Method × Level of REA Training	UIS	,866	2	,433	,601	,550	,010
	PEOI	,450	2	,225	,248	,781	,004
Error	UIS	84,22	117	,720			
	PEOI	105,99	117	,906			
Total	UIS	2629,87	124				
	PEOI	1913,88	124				
Corrected Total	UIS	118,87	123				
	PEOI	170,13	123				

a. R Squared = ,291 (Adjusted R Squared = ,255)

b. R Squared = ,377 (Adjusted R Squared = ,345)

The MANCOVA results allow accepting Hypothesis 6 stating that the perception of model quality is related to the perception of comprehension task performance. Given that also Hypothesis 5 was accepted, diagram users that performed well on the comprehension task (both in terms of effectiveness and efficiency) developed a favourable perception of their task performance (i.e. they realized that they correctly understood the diagram without spending much effort) and accordingly perceived the diagram as easy to interpret, and were satisfied with the information provided for answering the comprehension questions.

Note that by itself, the acceptance of hypotheses 5 and 6 does not imply that REA diagram users perceived the diagram as easier to interpret and were more satisfied with it than Non-REA diagram users. Also Non-REA diagram users with good comprehension task performance (both actual and perceived) are likely to form high-quality perceptions of the diagram. On the other hand, the partial support found for Hypothesis 1 indicates that at least an *indirect* effect of *Representation Method* on the perceived quality of the diagram exists. Hence, following in the research model the relationships established by accepting Hypotheses 1, 5, and 6, we can infer that user comprehension is more accurate with diagrams containing

REA pattern occurrences, that this accuracy is perceived by the users, and that as a consequence, users perceived these diagrams as easier to interpret and were more satisfied with their use.

Empirical evidence of a *direct* effect of Representation Method on the quality perceptions (Hypothesis 3) is found in the marginally significant effect that this factor has on PEOI. The inclusion of PSE_after as a covariate in the analysis adjusts the means on the PEOI dependent variable to what it would be if all participants scored identically on PSE_after. So, independent of the effect that PSE_after has on PEOI, using a REA diagram instead of a non-REA diagram has a (slight) positive effect on the user's perceived ease of interpreting the model.

The direct effect found provides weak support for Hypothesis 3, at least with respect to PEOI. A direct effect of *Representation Method* on UIS was not shown ($p = 0.305$), so the support for Hypothesis 3 is only partial.

Finally, Hypothesis 4 is rejected as no interaction effect is present in the data collected.

Discussion and Conclusions

Summary of Results and Lessons Learned

This paper hypothesized that REA diagrams, i.e. diagrammatic ER representations of REA ontology-based business process models, are better understood by business professionals than informationally equivalent Non-REA diagrams, i.e. ER diagrams showing no REA ontology pattern occurrences. A better understanding is demonstrated by a more effective and efficient performance of tasks that require retrieving and interpreting the information conveyed by the diagrams. It was also hypothesized that the REA diagrams would be perceived as easier to interpret than the Non-REA diagrams and the user information satisfaction with REA diagrams would be higher than with Non-REA diagrams.

The theoretical background for these hypotheses are theories from cognitive and perceptual psychology that can be used to explain the pattern recognition phenomenon. A working assumption of the formulated hypotheses is that diagram users have stored the REA ontology patterns in long-term memory, by means of a learning process. However, the same theories also suggest that there might be different levels of pattern activation. Accordingly, the paper hypothesized interaction effects with the strength of pattern presence in long-term memory (using the intensity of REA ontology training as an operational surrogate).

Specifically, it was hypothesized that REA diagrams would be more beneficial to highly trained users, whereas the level of REA training would have no effect on user comprehension if Non-REA diagrams are used.

A laboratory experiment with 124 business students was conducted to test the hypotheses. For a same comprehension task and informationally equivalent REA and Non-REA diagrams, REA diagram users were more accurate than Non-REA diagram users in terms of the number of comprehension questions correctly answered. When relating the number of correct answers to the time taken to perform the task, no significant differences were found, so the user comprehension effect observed is only an effectiveness effect.

The experiment further showed that the actual comprehension task performance is related to perceptual and satisfaction outcomes (perceived performance, ease of use and user satisfaction). These relationships indicate that participants who performed well on the comprehension task also perceived the diagram as easy to interpret and were satisfied with using the diagram for retrieving the information required by the task. However, apart from these relationships, also a (weak) direct effect of representation method (REA or Non-REA diagram) on perceived ease of interpretation (but not on user information satisfaction) was found. Hence, regardless of the actual and perceived task outcomes (and the effort involved), participants thought that the REA diagram was easier to interpret than the Non-REA diagram.

Given that all the participants received minimal REA ontology education (though some of them were trained more intensively) and were to some extent familiar with the REA ontology patterns' semantics and representation conventions, the experiment results provide evidence of pattern recognition taking place. The less accurate task performance by the Non-REA diagram users and their higher perceived difficulty of interpreting the diagram, are indicative of a larger semantic distance between the mental and conceptual representations (i.e. less cognitive fit). Task performance and ease of interpretation were higher with REA diagram users, suggesting that the reduced semantic distance between mental and conceptual representations (i.e. higher cognitive fit) is caused by the recognition of the REA pattern occurrences in the diagram.

The experiment could not show the hypothesized interaction effect with the level of REA training. Participants who received additional training in REA ontology-based conceptual modelling were significantly more accurate in the comprehension task than participants with minimal REA ontology education. But this effect was observed for both REA and Non-REA diagram users, so there is no direct evidence that pattern recognition is stronger or more frequent when users are (assumingly) more familiar with the patterns. The probable cause of the effect observed is that the additional training helped participants to provide better answers to the comprehensions questions. By the time of the experiment these participants were more experienced in model comprehension tasks than participants with less training.

Our experiment confirms, strengthens and extends the results presented in Poels (2003), the only study found that evaluates REA ontology-based conceptual modelling from the perspective of the model user (instead of model builder). Poels also observed increased accuracy (but no efficiency effect) with respect to comprehension task performance if an ER diagram with a REA core pattern occurrence is used. It is however unclear whether this effect is due to pattern recognition or to the specific diagram layout used for the Non-REA diagram (which was less 'logical' than the REA diagram layout). Furthermore, Poels did not investigate the effect of representation method on perception-based variables like ease of use and user satisfaction. Another difference is that all participants in the experiment of Poels were intensively trained in the use of the REA core pattern (4.5 hours of instruction plus practical course sessions, which is even more than the participants rated as having a 'high' level of REA training in this study). The experiment by Poels was also small-scaled (21 students) compared to our study.

Implications

The results of this study have implications for practitioners, educators and researchers. One of the goals of conceptual models is to improve the communication between analysts and business professionals. It is therefore important that business professionals understand conceptual models (built by the analysts). The correct 'reading' of business process models is not only important to verify whether a required system meets information processing requirements. A thorough understanding of the components that make up a company's integrated value chain is also crucial when a company considers adopting an ERP solution to satisfy its transaction processing and information needs. Our study shows that REA ontology-based conceptual models can make the communication between analysts and business professionals more effective, compared to the use of a purely descriptive approach like ER modelling. Our study also shows that a pattern-based approach grounded in a business domain ontology (like the REA enterprise domain ontology) can be beneficial even with limited levels of instruction or training (though more training helps). These implications for practice go beyond the use of the REA approach proper. They extend to the various methodologies for the modelling of e-collaboration/commerce systems that find their ontological basis in the REA ontology (e.g. UMM, ECIMF).

There are three implications for educators. First, spending a limited amount of time (1 hour of instruction in our case) on the REA ontology, or on REA ontology-based reference models, pays off in terms of students' ability to correctly interpret structural business process models. If REA conceptual modelling structures are learned, students will recognize them in ER diagrams (provided that REA ontology patterns are used to build them) and will score better at model comprehension tasks. Second, it appears that practising with REA diagrams improves students' performance, even with diagrams that do not contain REA pattern occurrences. Third, the study shows that representation formalism knowledge is important for the correct interpretation of modelled business process structure and policies. A course incorporating REA ontology-based conceptual modelling should therefore not neglect representational issues like teaching notational syntax (e.g. UML or ER diagramming) and the semantics of the various modelling constructs used (e.g. the ER Model).

Apart from evaluating REA ontology-based conceptual modelling, our study also contributes to the research on conceptual (data) modelling patterns. Although data modelling patterns have been around for some time, research in this area is only starting (Batra, 2005). The focus of the empirical research to date is on performance effects of pattern use by (novice) designers or analysts and the pattern matching (or retrieval) techniques that are used in this process (see Irwin (2002), Purao et al. (2003), Batra and Wishart (2004)). Our study is original in the sense that modelling patterns are evaluated from the perspective of the model user. Whereas previous research has emphasized pattern recognition in the information requirements put forward by a modelling task, our study also pays attention to pattern recognition in the conceptual model itself. Pattern recognition does not only help building better models (and with less effort); it also helps understanding conceptual models.

One implication for researchers is that attention direction mechanisms and perceptual clues should be considered when designing pattern templates. Especially secondary notation features seem to be useful when pattern recognition is based on matching graphical representations (like diagram fragments and generic pattern templates). Another implication for research is that pattern familiarity (either due to training, experience, or both) is an important research variable. Our study, which is admittedly quite inconclusive with respect to the exact role played by pattern familiarity in pattern recognition processes, indicates that this variable, if not varied (manipulated or measured), should at least be controlled in order to avoid confounding or intervening effects.

Limitations and Future Work

The limitations of our study are due to the choice of research method. It is widely accepted that controlled experimentation sacrifices external validity for internal validity. First of all, task characteristics were held constant by using a single model comprehension task to be performed under all treatments. This task was realistic in the sense that it reflects typical use of structural business process models by business professionals. Moreover, the task was tailored to the purpose of REA ontology-based conceptual models as it involved the interpretation of business process information that is relevant from an accountability and control perspective. However, previous research has shown the impact of task complexity on cognitive fit (Dunn and Grabski, 2001) and pattern-based modelling performance (Batra and Wishart, 2004). Future research may therefore use tasks with different complexity levels.

Apart from the use or no use of REA pattern occurrences, all other model and representational characteristics were held constant. Further, only one case was used in the experiment (i.e. a consulting services business process). These experimental design choices helped controlling variables such as information equivalence, domain familiarity, model size, and notational system used. Future research may wish to replicate this study using other examples of business processes, different model sizes, and other modelling notations and hence increase the generalizability of our results.

A challenging avenue for future research is to conduct a similar study with business professionals (provided that the researcher has access to such a sample). The advantage of students as study participants is that their knowledge of REA ontology patterns is relatively easy to assess. Business professionals with working experience might have implicitly learned some of the REA ontology patterns, just by observing and experiencing business reality. According to McCarthy (2004), the reference models underlying ERP systems are for 60% consistent with the structuring principles of the REA ontology. So implicitly, business professionals might have stored (some of / parts of) REA patterns in long-term memory. It could be interesting to see how well such users understand REA diagrams (compared to Non-REA diagrams). Though potentially interesting (also with respect to generalizability), the big challenge for such a study would be to measure the participants' implicit REA ontology knowledge (or strength of the presence of REA patterns in long-term memory) independently from task performance outcomes. Without such measurement, it would be hard to tell whether possible comprehension performance and quality perception effects are the result of REA ontology pattern recognition.

References

AGARWAL R, SINHA AP and TANNIRU M (1996) The role of prior experience and task characteristics in object-oriented modeling: an empirical study. *International Journal of Human-Computer Studies* 45, 639-667.

AJZEN I (2002) Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior. *Journal of Applied Social Psychology* 32, 1-20.

AKOKA J and COMYN-WATTIAU I (2004) Les systèmes d'information comptables multidimensionnels: Comparaison de deux modèles. In *Proceedings of the 9th Conference of the Association for Information and Management*.

ANDERSON JR (1996) *The architecture of cognition*. Harvard University Press.

ANTONY S and MELLARKOD V (2004) A Methodology for using Data-modeling Patterns. In *Proceedings of the 3rd Symposium on Research in Systems Analysis and Design* (BATRA V, PARSONS J and VENKATARAMAN R, Eds.) St. John's, Canada.

BANDURA A (1997) *Self-Efficacy: The Exercise of Control*. W.H. Freeman.

BATRA D (1993) A Framework for Studying Human Error Behavior in Conceptual Database Modeling. *Information & Management* 25(3), 121-131.

BATRA D (2005) Conceptual Data Modeling Patterns: Representation and Validation. *Journal of Database Management* 16(2), 84-106.

BATRA D and WISHART NA (2004) Comparing a rule-based approach with a pattern-based approach at different levels of complexity of conceptual data modelling tasks. *International Journal of Human-Computer Studies* 61, 397-419.

BATRA D and WANG TW (2004) A Research Agenda for Evaluating and Improving Data Modeling Patterns. In *Proceedings of the 3rd Symposium on Research in Systems Analysis and Design* (BATRA V, PARSONS J and VENKATARAMAN R, Eds.) St. John's, Canada.

BODART F, PATEL A, SIM M and WEBER R (2001) Should Optional Properties Be Used in Conceptual Modelling? A Theory and Three Empirical Tests. *Information Systems Research* 12(4), 384-405.

BURTON-JONES A and WEBER R (1999) Understanding Relationships with Attributes in Entity-Relationship Diagrams. In *Proceedings of the 20th International Conference on Information Systems*, pp 214-228.

BURTON-JONES A and WEBER R (2003). Properties do not have properties: Investigating a questionable conceptual modelling practice. In *Proceedings of the 2nd Annual Symposium on Research in Systems Analysis and Design* (BATRA V, PARSONS J and VENKATARAMAN R, Eds.).

COHEN J (1988). *Statistical power analysis for the behavioral sciences.* Erlbaum, Hillsdale (NJ).

CONNOLLY T and BEGG C (2002) *Database Systems: A Practical Approach to Design, Implementation, and Management, 3rd Ed*. Addison Wesley.

DAVIS F (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13(3), 319-339.

DOLL WJ and TORKZADEH G (1988) The Measurement of End-User Computing Satisfaction. *MIS Quarterly*, 12(2), 259-274.

DUNN CL, CHERRINGTON JO and HOLLANDER AS (2005a) *Enterprise Information Systems: A Pattern-Based Approach, 3rd Ed*. McGraw-Hill.

DUNN CL and GERARD GJ (2001) Auditor efficiency and effectiveness with diagrammatic and linguistic conceptual model representations. *International Journal of Accounting Information Systems* 2(4), 223-248.

DUNN CL, GERARD GJ and GRABSKI SV (2005b) Critical Evaluation of Conceptual Data Models. *International Journal of Accounting Information Systems 6(2), 83-106.*

DUNN CL and GRABSKI SV (2000) Perceived semantic expressiveness of accounting systems and task accuracy effects. *International Journal of Accounting Information Systems* 1(2), 79-87.

DUNN C and GRABSKI S (2001) An Investigation of Localization as an Element of Cognitive Fit in Accounting Model Representations. *Decision Sciences* 32(1), 55-94.

DUNN CL and MCCARTHY WE (1997) The REA Accounting Model: Intellectual Heritage and Prospects for Progress. *Journal of Information Systems* 11(1), 31-51.

ECIMF Project Group (2003) E-Commerce Integration Meta-Framework. Final draft.

FOX MS (1992) The TOVE Project: A Common-sense Model of the Enterprise. In *Industrial Engineering Applications of Artificial Intelligence and Expert Systems* (BELLI F and RADERMACHER F, Eds.) pp 25, Springer-Verlag.

GEERTS GL and MCCARTHY WE (2002) An ontological analysis of the economic primitives of the extended-REA enterprise information architecture. *International Journal of Accounting Information Systems* 3(1), 1-16.

GEMINO A and WAND Y (2003) Foundations for Empirical Comparisons of Conceptual Modelling Techniques. In *Proceedings of the 2nd Annual Symposium on Research in Systems Analysis and Design* (BATRA V, PARSONS J and VENKATARAMAN R, Eds.).

GEMINO A and WAND Y (2005) Complexity and clarity in conceptual modeling: Comparison of mandatory and optional properties. *Data & Knowledge Engineering* 55(3), 301-326.

GENERO M, POELS G and PIATTINI M (2002) Defining and Validating Measures for Conceptual Data Model Quality. *Lecture Notes in Computer Science* 2348, 724-727.

GENTNER D and MEDINA J (1998) Similarity and the development of rules. *Cognition* 65, 263-297.

GERARD GJ (2005) The REA Pattern, Knowledge Structures, and Conceptual Modeling Performance. *Journal of Information Systems* 19(2), 57-77.

GOODHUE DL and THOMPSON RL (1995) Task-technology fit and individual performance. MIS Quarterly 19, 213-236.

GORDIJN J (2002) Value based requirements engineering: Exploring innovative e-commerce ideas. PhD dissertation, Vrije Universiteit Amsterdam.

GRUBER T (1993) A Translation Approach to Portable Ontologies. *Knowledge Acquisition* 5(2), 199-220.

HAIR JF, ANDERSON RE, TATHAM RL, BLACK B (1998) *Multivariate Data Analysis, 5th Ed.*, Prentice-Hall.

HOLLANDER AS, DENNA EL and CHERRINGTON JO (2000) *Accounting, Information Technology, and Business Solutions, 2nd Ed*. McGraw-Hill.

IRWIN G (2002) The Role of Similarity in the Reuse of Object-Oriented Analysis Models. *Journal of Management Information Systems* 19(2), 219-248.

IVES B, OLSON MH and BAROUDI JJ (1983) The Measurement of User Information Satisfaction. *Communications of the ACM* 26(10), 785-793.

KIM YG and MARCH ST (1995) Comparing Data Modelling Formalisms. *Communications of the ACM* 38(6), 103-115.

LARKIN J and SIMON H (1987) Why a diagram is (sometimes) worth ten thousands words. *Cognitive Science* 11(1), 65-99.

MARCH ST and SMITH GF (1995) Design and natural science research on information technology. *Decision Support Systems* 15(4), 251-266.

MCCARTHY WE (1982) The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment. *The Accounting Review* 57(3), 554-578.

MCCARTHY WE (1999) Semantic Modelling in Accounting Education, Practice, and Research: Some Progress and Impediments. In *Conceptual Modelling: Current Issues and Future Directions* (AKOKA J, KANGASSALO H and THALHEIM B, Eds) pp 144-153, Springer Verlag.

MCCARTHY WE (2003) The REA Modelling Approach to Teaching Accounting Information Systems. *Issues in Accounting Education* 18(4), 427-441.

MCCARTHY WE (2004) The Evolution toward REA Accountability Infrastructures for Enterprise Systems. In *Proceedings of the first International Conference on Enterprise Systems and Accounting*, pp 1-2.

MOODY D (2002) Comparative Evaluation of Large Data Representation Methods: The Analyst's Perspective. *Lecture Notes in Computer Science* 2503, 214-231.

MOODY DL and SHANKS G (2003) Improving the quality of data models: empirical validation of a quality management framework. *Information Systems* 28(6), 619-650.

MOORE GC and BENBASAT I (1991). Development of an Instrument to Measure the Perceptions of Adopting and Information Technology Innovation. *Information Systems Research* 2(3), 192-222.

MURTHY US and GROOMER SM (2005) *Accounting Information Systems: A Database Approach*. Cybertext Publishing.

NUNALLY J (1978) *Psychometric Theory*. McGraw-Hill.

O'LEARY DE (2004) On the relationship between REA and SAP. *International Journal of Accounting Information Systems* 5(1), 65-81.

OSTERWALDER A and PIGNEUR Y (2003) Towards Strategy and Information Systems Alignment through a Business Model Ontology. In *Proceedings of the 23rd Annual International Conference of the Strategic Management Society*, Baltimore, USA.

PARSONS J (2003) Effects of Local Versus Global Schema Diagrams on Verification and Communication in Conceptual Data Modelling. *Journal of Management Information Systems* 19(3), 155-183.

PARSONS J and COLE L (2005) What do the pictures mean? Guidelines for the experimental evaluation of representation fidelity in diagrammatical conceptual modeling techniques. *Data & Knowledge Engineering* 55(3), 327-342.

PETRE M (1995) Why Looking Isn't Always Seeing: Readership Skills and Graphical Programming. *Communications of the ACM* 38(6), 33-44.

POELS G (2003) Conceptual Modeling of Accounting Information Systems: A Comparative Study of REA and ER Diagrams. *Lecture Notes in Computer Science* 2814, 152-164.

POELS G, MAES A, GAILLY F and PAEMELEIRE R (2004) User Comprehension of Accounting Information Structures: An Empirical Test of the REA Model. Working Paper 2004/254, Faculty of Economics And Business Administration, Ghent University.

POELS G, MAES A, GAILLY F and PAEMELEIRE R (2005) User Attitudes towards Pattern-Based Enterprise Information Models: A Replicated Experiment with REA Diagrams. Working Paper 2005/315, Faculty of Economics And Business Administration, Ghent University.

PURAO S, STOREY VC and HAN T (2003) Improving Analysis Pattern Reuse in Conceptual Design: Augmenting Automated Processes with Supervised Learning. *Information Systems Research* 14(3), 269-290.

ROMNEY MB and STEINBART PJ (2003) *Accounting Information Systems, 9th Ed*. Prentice Hall.

RYAN SD, BORDOLOI B and HARRISON DA (2000) Acquiring Conceptual Data Modeling Skills: The Effect of Cooperative Learning and Self-Efficacy on Learning Outcomes. *DATABASE for Advances in Information Systems* 31(4), 9-24.

SCHUTTE R and ROTTHOWE T (1998) The Guidelines of Modelling – An Approach to Enhance the Quality in Information Models. *Lecture Notes in Computer Science* 1507, 240-254.

SEDDON P and YIP SK (1992) An Empirical Evaluation of User Information Satisfaction (UIS) Measures for Use with General Ledger Accounting Software. *Journal of Information Systems* 6(1), 75-92.

SHOVAL P, DANOCH R and BALABAN M (2002) Hierarchical ER Diagrams (HERD) – The Method and Experimental Evaluation. *Lecture Notes in Computer Science* 2784, 264-274.

SIAU K (2004) Informational and Computational Equivalence in Comparing Information Modeling Methods. *Journal of Database Management* 15(1), 73-86.

SIAU K, WAND Y and BENBASAT I (1997) The Relative Importance of Structural Constraints and Surface Semantics in Information Modelling. *Information Systems* 22(2/3), 155-170.

SMITH DK, CHANG J and MOORES TT (2003) Comparing Self-Efficacy and Meta-Cognition as Indicators of Learning. In *Proceedings of the 9th Americas Conference on Information Systems*, 3111-3115.

SUMMERS SL (2000) Discussion of perceived semantic expressiveness of accounting systems and task accuracy effects. *International Journal of Accounting Information Systems* 1(2), 88-90.

TOPI H and RAMESH V (2002) Human Factors Research on Data Modeling: A Review of Prior Research, An Extended Framework and Future Research Directions. *Journal of Database Management* 13(2), 3-19.

UN/CEFACT (2003) UN/CEFACT Modeling Methodology (UMM) User Guide. Technical report CEFACT/TMG/N093.

USHOLD M and GRUNINGER M (1996) Ontologies: Principles, Methods and Applications. *The Knowledge Engineering Review* 11(2), 93-136.

USHOLD M, KING M, MORALEE S and ZORGIO Y (1998) The Enterprise Ontology. *The Knowledge Engineering Review* 13(1), 31-89.

VAASSEN EHJ (2002) *Accounting Information Systems: A Managerial Approach*. John Wiley & Sons.

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

WAND Y and WEBER R (2002) Information Systems and Conceptual Modelling: A Research Agenda. *Information Systems Research* 13(4), 363-376.

Appendix A: Core constructs of the REA Enterprise Domain Ontology

The REA core pattern is a conceptual structure of relationships between the resources, events, and agents that constitute the transactional/transformational core of an economic exchange/conversion process. Figure A-1 shows how the REA core pattern is applied for modelling the acquisition process in a retail company. The pattern relates a resource type of the company (e.g. inventory, cash) by a stock-flow relationship type with a type of events occurring in the process that cause resource inflows or outflows (e.g. purchase, cash disbursement). The pattern further shows that each type of events that result in resource inflows (e.g. purchase) is related via a duality relationship type by a type of events that result in resource outflows (e.g. cash disbursement), and vice versa. Further, participation relationship types are used to relate economic event types with agent types, which represent inside parties (e.g. purchase agent, cashier) or outside parties (e.g. vendor) to the economic exchange or conversion process that is modelled.

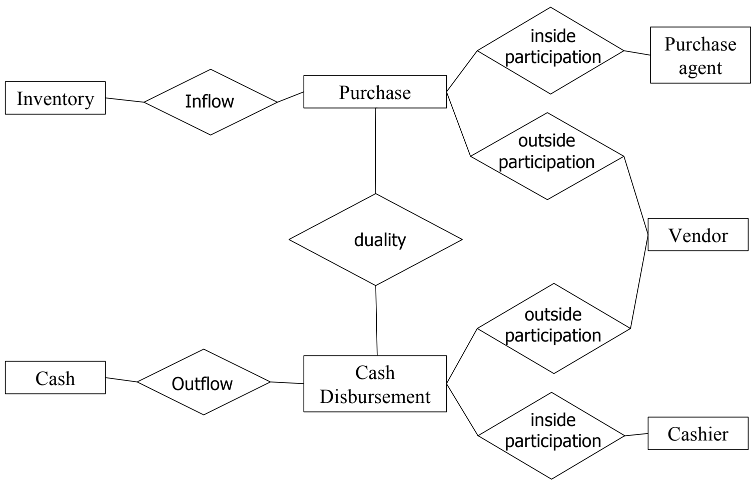


Figure A-1. The acquisition process in a retail company

Appendix B: UML-based notation for ER diagrams

In the experiment and in the course the experiment participants were drawn from, ER diagrams (hence REA and Non-REA diagrams) were notated using UML symbols (following the guidelines of Connolly and Begg (2002)). According to this notation, an entity type is described using an UML classifier with two compartments: a name compartment containing the entity type's name and a list compartment containing the entity type's attributes. Primary key attributes are indicated with a {PK} tag. A relationship type is portrayed as an UML association. The relationship type's name is shown as a name string near the association path along with a reading direction arrow. Relationship type attributes are contained in a list compartment of an UML classifier (called association class) that is attached by a dashed line to the association path. Finally, participation and cardinality constraints are denoted using UML multiplicity specifications, shown as integer intervals of the form *lower-bound* .. *upper-bound*. Figure B-1 presents an example REA diagram in this UML-based notation.

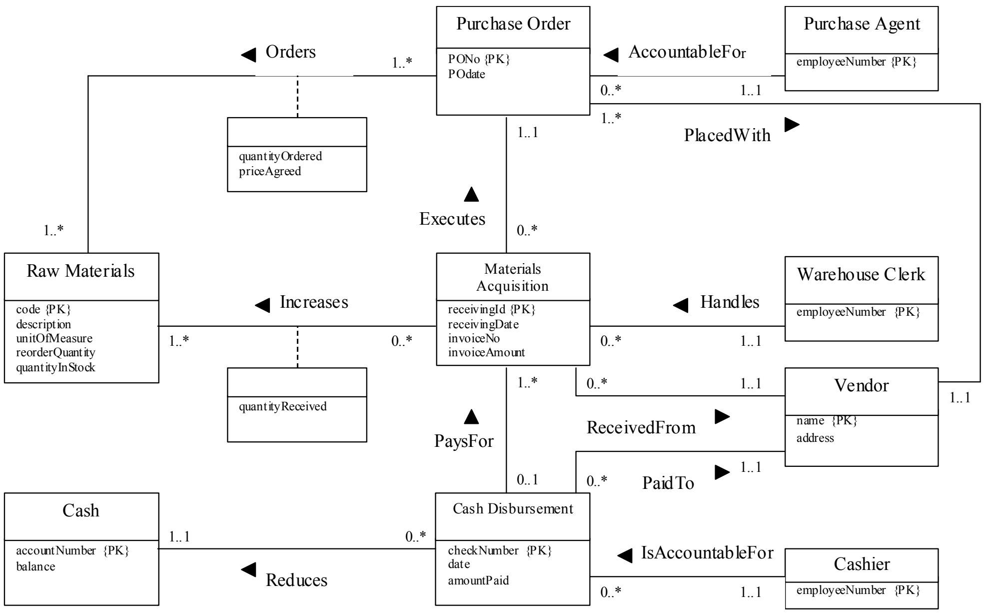


Figure B-1. The raw materials acquisition process in a production company

Appendix C: Experimental objects

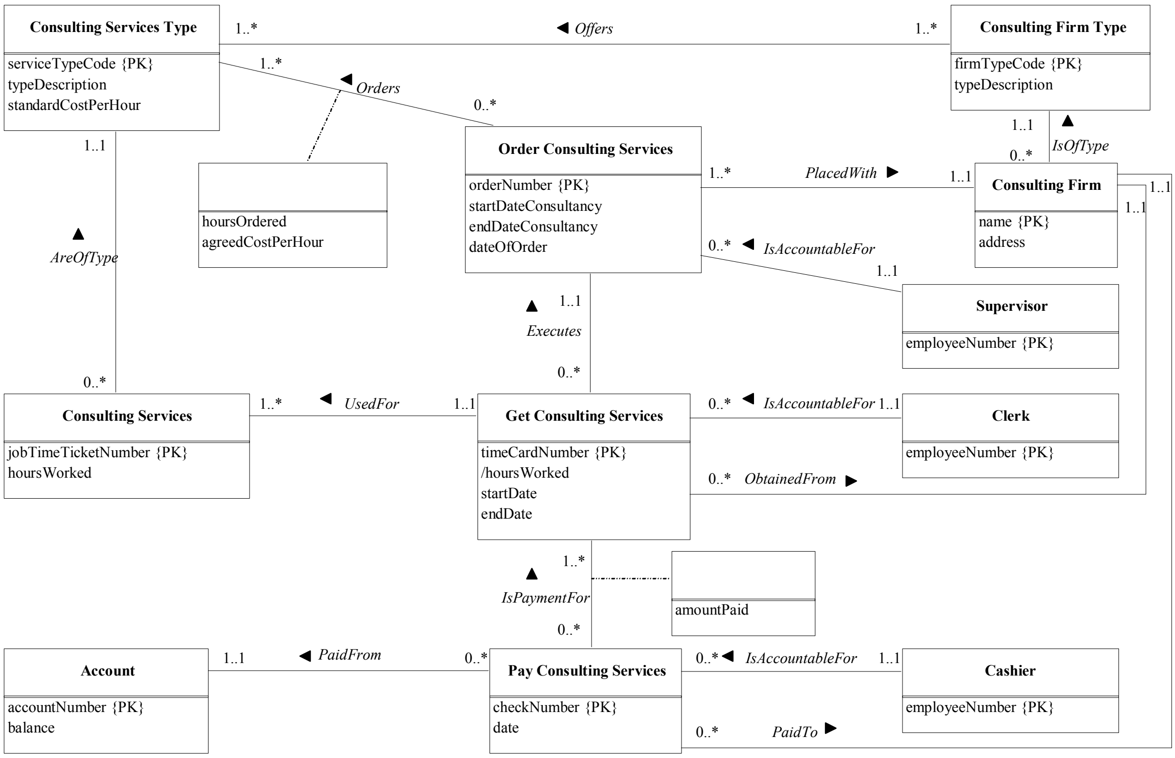


Figure C-1. REA diagram

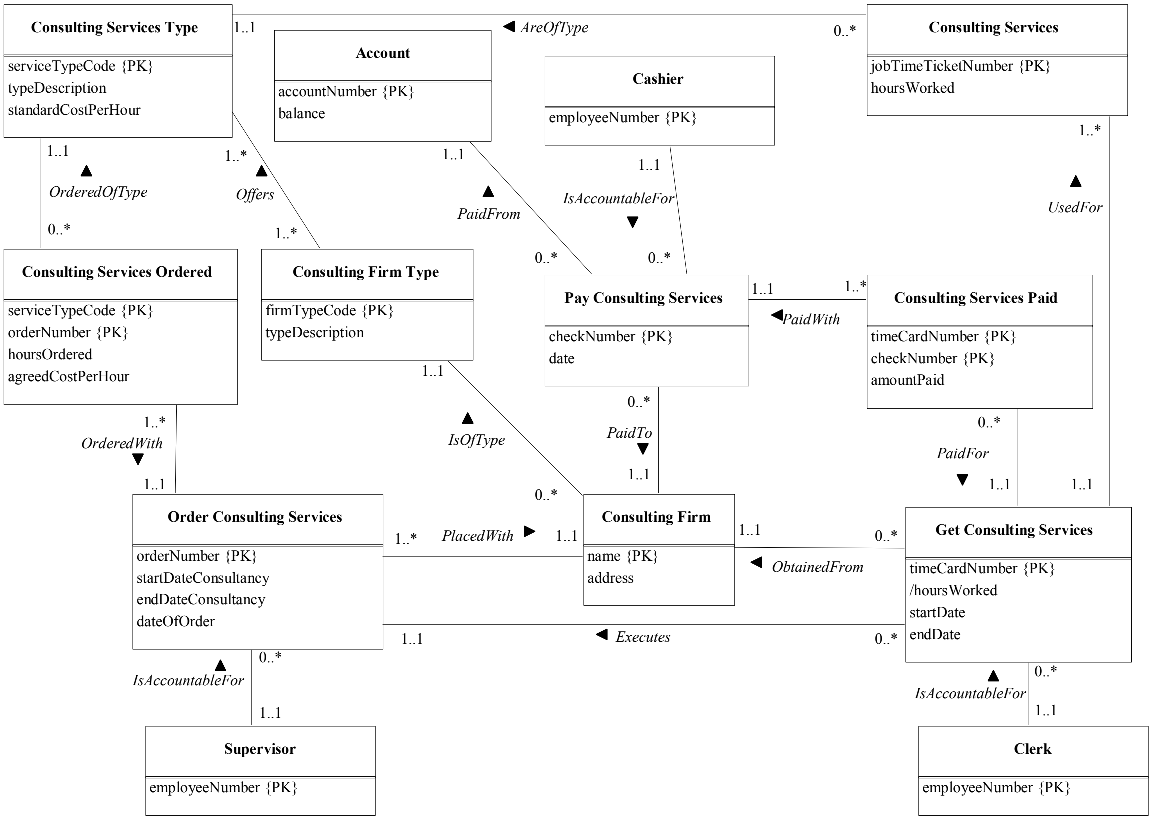
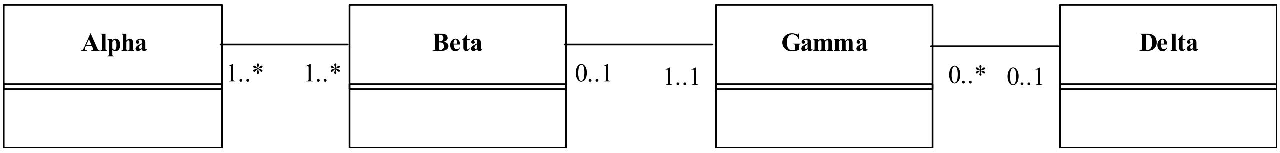


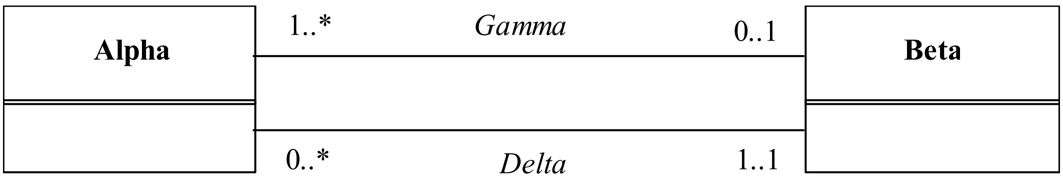
Figure C-2. Non-REA diagram

Appendix D: Experimental Tasks

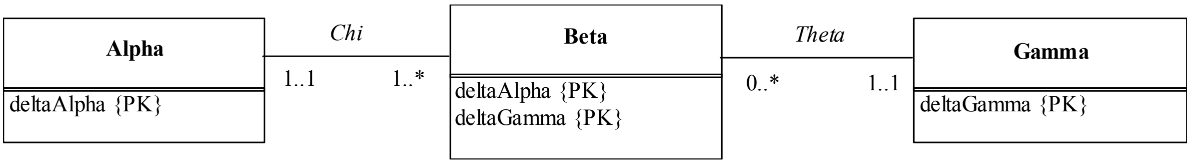
ER Formalism Knowledge Pre-test Questions



1. Can an entity of type Alpha be related to an entity of type Beta that is not related to an entity of type Gamma?
2. Does the diagram specify an upper limit on the number of entities of type Gamma that are related to the same entity of type Delta?
3. Does every entity of type Gamma have to be related to an entity of type Beta?
4. Should an entity of type Gamma be related only once to an entity of type Delta?
5. Should there be a relationship between each entity of type Alpha and each entity of type Beta?
6. Can the entities of type Gamma that are related to an entity of type Beta, be related to different entities of type Beta?
7. Can there be maximum one relationship between some entity of type Gamma and some entity of type Delta?
8. Can an entity of type Gamma be related to more than one entity of type Alpha, via an entity of type Beta?
9. Should an entity of type Beta be related to maximum one entity of type Delta, via an entity of type Gamma?
10. Does every entity of type Alpha have to be related to at least one entity of type Delta, via an entity of type Beta and an entity of type Gamma?



11. Does an entity of type Alpha that is related via a relationship of type Gamma with an entity of type Beta, have to be related via a relationship of type Delta with the same entity of type Beta?
12. Can an entity of type Beta be related to different sets of entities of type Alpha via relationships of types Gamma and Delta?
13. Assume that an entity of type Alpha is related via a relationship of type Delta to some entity of type Beta. Does this entity of type Beta have to be related to the same entity of type Alpha via a relationship of type Gamma?



14. Does there exist a many-to-many entity connection between the types Alpha and Gamma via relationships of the types Chi and Theta with entities of type Beta?
15. Can the entity of type Alpha that has the value “1” for the attribute deltaAlpha1 and the entity of type Gamma that has the value “x” for the attribute deltaGamma be connected to each other more than once via relationships of the types Chi and Theta with entities of type Beta?

Model Comprehension Questions

1. Can consulting services only be obtained from a consulting firm where a supervisor has placed an order?
2. Can a cashier make payments for consulting services which have not been charged to a timecard?
3. Can consulting services with working hours registered on several time cards, be paid for with one single payment transaction?
4. Do all working hours, for jointly obtained and charged consulting services, have to be registered on at least one job-time ticket?
5. Can a supervisor order different types of consulting services with one single order?
6. Can a cashier make a payment to a consulting firm with money drawn from more than one account?
7. Can the consulting working hours that are registered on one and the same job-time ticket, have been obtained on different occasions?
8. Must jointly obtained and charged consulting services belong to a single consulting services type?
9. Must jointly obtained and charged consulting services, be paid with a single payment transaction?
10. Can a cashier make payments for consulting services that have been charged to a timecard, but were not ordered?
11. Must all consulting services obtained from the same consulting firm, be ordered by the same supervisor?
12. Can it be that we do not know which clerk is accountable for a given timecard?
13. Can a type of consulting services be described and be assigned a standard cost per hour without there being any order of this type of services?
14. Must a consulting firm be paid immediately for consulting services charged to a timecard?
15. Can consulting working hours registered on a same job-time ticket be related to more than one order?

Appendix E: Measurement Scales

All items are seven-point Likert scales, anchored at “Strongly disagree” and “Strongly agree”.

Perceived Self-Efficacy

PSE_before: present tense (“I am able to …”)

PSE_after: past tense (“I was able to …”)

PSE_1: I am/was able to correctly interpret the meaning of ER Model constructs.

PSE_2: I am/was able to interpret the meaning of ER Model constructs without much effort.

PSE_3: I am/was able to understand the structure of a business process modelled in an ER diagram (i.e. which activities?, who is involved?, …).

PSE_4: I am/was able to quickly see in an ER diagram the structure of a business process (i.e. which activities?, who is involved?, …).

PSE_5: I am/was able to derive the business policies that govern a business process using an ER diagram.

PSE_6: I am/was able to quickly see in an ER diagram the business policies that govern a business process.

Perceived Ease of Interpretation

$PEOI_1$: It was easy for me to understand what the conceptual schema was trying to model.

$PEOI_2$: Using the conceptual schema was seldom frustrating.

$PEOI_3$: Overall, the conceptual schema was easy to use.

$PEOI_4$: Learning how to read the conceptual schema was easy.

User Information Satisfaction

UIS_1: The conceptual schema adequately met the information needs that I was asked to support.

UIS_2: The conceptual schema was efficient in providing the information I needed.

UIS_3: The conceptual schema was effective in providing the information I needed.

UIS_4: Overall, I am satisfied with the conceptual schema for providing the information I needed.