2003

Representing Things and Properties in Conceptual Modelling: An Empirical Evaluation

Graeme Shanks
University of Melbourne, g.shanks@unimelb.edu

Jasmina Nuredini
University of Melbourne

Daniel Moody
Charles University, Prague, moody@ksint.ms.mff.cuni.cz

Daniel Tobin
University of Melbourne

Ron Weber
University of Queensland, weber@bel.uq.edu.au

Representing Things and Properties in Conceptual Modelling: an Empirical Evaluation

Graeme Shanks
Department of Information Systems
The University of Melbourne
Melbourne, 3010, Australia
Phone: +61 03 8344 1577, Fax +61 03 9349 4596
g.shanks@unimelb.edu.au

Jasmina Nuredini
Department of Information Systems
The University of Melbourne
Melbourne, 3010, Australia
Phone: +61 03 8344 1577, Fax: +61 03 9349 4596

Daniel Tobin
Department of Information Systems
The University of Melbourne
Melbourne, 3010, Australia
Phone: +61 03 8344 1577, Fax: +61 03 9349 4596

Daniel Moody
Department of Software Engineering
Charles University
Malostranske Namesti 25
Prague, Czech Republic 11800
Phone: +420 2 2191 4268, Fax: +420 22191 4323, Mobile: +420 608 378123
moody@ksint.ms.mff.cuni.cz
and
School of Business Systems
Monash University
Melbourne, Australia 3800
dmoody@infotech.monash.edu.au

Ron Weber
Faculty of Commerce
Faculty of Business, Economics and Law
The University of Queensland
Brisbane, Australia, 4072
Phone: +61 03 3365 6585, Fax:+61 03 3365 6788
weber@bel.uq.edu.au

Abstract

The representation of things and properties is a fundamental issue in conceptual modelling. Important theoretical issues surrounding the representation of things and properties remain unresolved. For example, proponents of object-role modelling argue that there should be no distinction between things and properties, while proponents of entity-relationship modelling argue that the distinction is important but provide ambiguous guidelines about how the distinction should be made. In this paper, we use ontological theory to support our arguments about how things and properties should be represented. We describe an experiment that we undertook to test whether an ontologically sound representation of things and properties enabled users to better understand a domain than two other alternative, widely used representations. Our results provide evidence to support the use of ontologically sound representations of things and properties in conceptual modelling.

Introduction

The distinction between things and properties is fundamental to the way humans perceive and understand phenomena. It has been of interest to philosophers concerned with ontology (the nature of the world) (e.g., Bunge 1977) and information systems researchers and practitioners concerned with finding better ways to build conceptual models as a basis for building better information systems. For instance, the representation of things and properties features in early work on conceptual modelling (Chen 1976, Nijssen 1976, Kent 1978) and in recent practitioner books (e.g., Simsion and Witt 2001). More recently representation of things and properties features in object-oriented conceptual modelling approaches, in particular in the Unified Modelling Language approach to object-oriented conceptual modelling (e.g., Rumbaugh et al. 1999).

In the context of conceptual modelling in information systems, the distinction between things and properties and their representation remains problematic for several reasons. Important theoretical issues surrounding the representation of things and properties remain unresolved. For example, in the object-role approach to conceptual modelling, the distinction between things and properties is unimportant (Halpin 1995). Both should be represented using the object symbol in a conceptual schema. In the entity-relationship (ER) model (Chen 1976), things are represented as entity types, and properties are represented as attribute types. However, entity-type symbols are often used to represent both things and properties. For example, an employee's set of skills is often represented as an entity type that is connected to an employee entity type by a relationship. Thus, both a property (a set of skills) and a thing (the employee) have both been represented by the same construct (an entity-type symbol). A deeper understanding of the representation of things and properties is required as a basis for conceptual modelling languages and methods (Wand and Weber 2002).

There is also much confusion between the representation of things and properties in conceptual models and their realization in database designs. For example, Simsion and Witt (2001, p. 104) state: "Attributes in an ER model correspond to columns in a relational model." They further suggest that ER models should be "normalized" and repeating groups of attributes should be removed, forming additional entity types. A conceptual model is used to discover and document user views of an information system and provide a basis for informed discernment, reconciliation, and compromise (Hirschheim et al. 1995). Therefore, the representation of things and properties in conceptual models should be based on a sound underlying theory of representation of phenomena in the world rather than principles of good database design. To the best of our knowledge, however, no rigorous empirical evaluation of alternative representations of things and properties has been undertaken.

Consequently, we undertook to empirically evaluate alternative conceptual-modelling representations of things and properties. Our research was motivated in four ways: improving systems development, testing prior theoretical work, improving user understanding of conceptual models, and improving the practice of conceptual modelling. It is well recognized that the cost of fixing errors grows exponentially the later they are discovered in the system development process (e.g., Boehm 1981). With conceptual modelling work being undertaken early in the system development process, improvements in conceptual modelling practice potentially will lead to high payoffs (Moody and Shanks 1998).

We sought to test previous theoretical work undertaken to predict how well different types of representations facilitate or inhibit human understanding of real-world phenomena. If we can make accurate predictions about what types of conceptual modelling practices are likely to work well, we avoid the high costs associated with learning about the strengths and weaknesses of different practices through experience.

To improve user understanding of conceptual models, it is important to determine which type of representation of real-world phenomena enables humans to understand the phenomena better. When conceptual models are prepared initially (e.g., by systems analysts), users of an information system are asked to evaluate them to determine how accurately and completely the models represent their perceptual worlds. Where users cannot understand the conceptual model clearly, their ability to validate the model is impaired. Moreover, subsequent users may employ conceptual models to understand the functionality provided by an information system. Again, if users cannot understand the conceptual model clearly, their ability to comprehend and use the information system is impaired.

Finally, we sought to contribute to improved conceptual modelling practice. Many different, sometimes ambiguous guidelines for representation of things and properties in the practitioner literature may confuse rather than assist practitioners (Simsion and Witt 2001). If we develop improved conceptual modelling rules for things and properties, we will assist practitioners.

1. Theory and Proposition

The theory used in this study is based on the ontological theory of Bunge (1977). This ontology is particularly suitable for conceptual modelling as it is a realist ontology that is formally defined and has been adapted to information systems modelling (Weber 1997). Weber analyses the representation of things and properties in conceptual modelling and his analysis runs as follows.

1. "The world is made of things that possess properties" (p. 497). Things and properties are the two atomic constructs needed to describe the world.
2. Every thing in the world possesses one or more properties (there are no bare things).
3. Properties themselves cannot have properties. Moreover, properties cannot exist by themselves. They must attach to a thing.
4. Two types of properties that exist in the world are *intrinsic properties*, which depend on one thing only, and *mutual properties*, which depend on two or more things.

5. Two things interact (are coupled) when a history of one thing (manifested as a sequence of the thing's states) would be different if the other thing did not exist.
6. The existence of a mutual property between two things can indicate that they interact with each other. Mutual properties that manifest interactions between two things are called *binding mutual properties*.

In the context of Bunge's (1977) ontological theory, a property can *not* be represented as an entity type. This practice leads to construct (semantic) overload because the same grammatical construct (an entity-type symbol) has been used to represent two ontological constructs (things and their properties).

Figures 1 to 4 show some examples of how entities and properties may be represented in simple ER models. Figures 1 and 2 show things represented as entity types (Employee, Sales Order, Product) and properties represented as entity types (Skill, Sales Order Product). Figures 3 and 4 show alternative representations that we propose where all properties are represented as attributes. Which type of representation is "better"? Does it matter which is used? Weber (1997) contends that Figures 1 and 2 are the poorer representations, while Figures 3 and 4 are better representations.

If the ontological principles are contravened and properties are represented as entities, we argue that the resulting conceptual schema model is limited. Users of the model must use tacit knowledge to determine whether the entity type represents a thing or a property. For example, in Figure 1, Skill could be interpreted as a thing when it is an intrinsic property of Employee. Similarly in Figure 2, Sales Order Product could be interpreted as a thing when it it a mutual property of Sales Order and Product. Research in conceptual modelling indicates that humans distinguish between things and properties as a way of managing complexity in real-world phenomena they are seeking to understand (e.g., Moody 2001, Weber 1996). Including this distinction in conceptual models should therefore help users to better understand the phenomena the models are intended to represent.

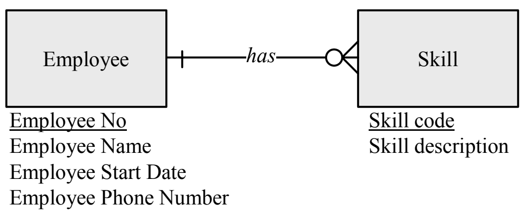


Figure 2: Employee Skill ER Model

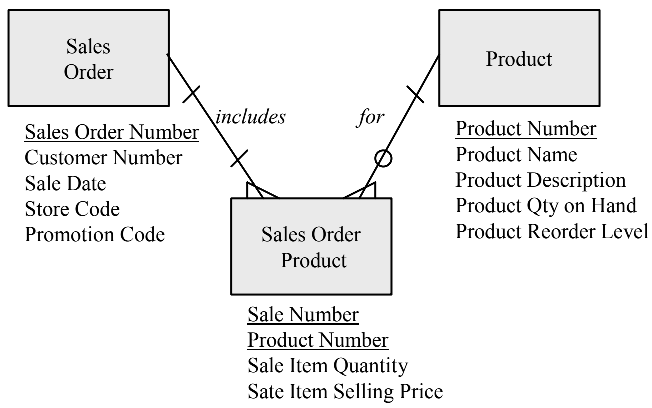


Figure 3: Sales Order Product ER Model

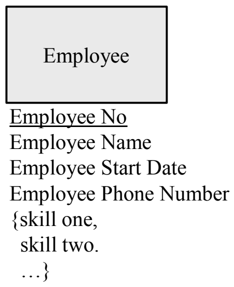


Figure 3: Ontologically Sound Employee Skill ER Model

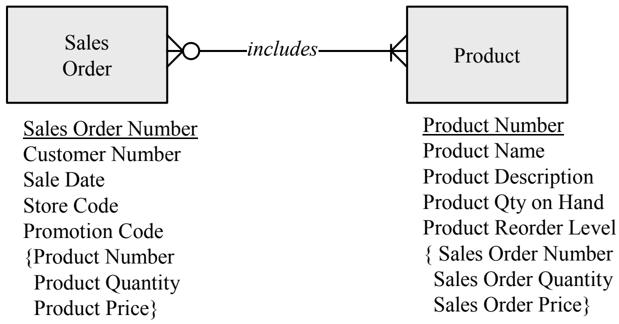


Figure 4: Ontologically Sound Sales Order Product ER Model

We contend that the choice of representation for things and properties is important in terms of users' ability to elicit the meaning of the phenomena described via the representation. Hence, the following proposition motivates the empirical work we undertook:

> *Proposition*: Conceptual models that use an attribute construct to represent properties will enable their users to better understand the semantics associated with the model than conceptual models that use an entity class construct to represent properties.

2. Conceptual Modelling Approaches

To maximize our contribution to conceptual-modelling practice, we decided to base our study on the ER approach to conceptual modelling. The ER model (Chen 1976) is widely used for data modelling in practice, and has been used to design database schemas for over two decades (Thalheim 2000). It distinguishes clearly between the entity-type and attribute-type constructs. In contrast, the object-role modelling approach (Halpin 1995) uses a different notation and does not distinguish between entity types and attribute types. The principles of ontological modelling may be readily applied within the ER modelling approach.

The ER model we used as a control in the study (see Appendix A, the "Practice" ER Model) was the type of model most widely used in practice, where entity types are essentially third normal form relations (Simsion and Witt 2001). In this representation, things are represented as entity types. However, multi-valued attributes (intrinsic properties) are also represented as entity types (known as attributive or characteristic entity types – for example, Customer Contact Person in Appendix A). Similarly, value domains are also represented as entity types (known as classification entity types – for example, Customer Industry Type in Appendix A), and many-to-many relationships (mutual properties) are represented as entity types (known as intersection or associative entity types – for example, Sales Order Item in Appendix A). In ontological terms, many ontological constructs are represented by one modelling construct, entity type, leading to construct overload. In developing the model in Appendix A we first analyzed a typical model from practice to work out the ratios of the different categories of entity types described above. We ensured our model had similar ratios to increase its external validity. We also used a domain, sales order processing, that was widely understood.

When developing the ontologically sound version of this model (see Appendix B, the "Ontologically sound" ER model), we transformed the "Practice" ER Model by first removing the attributive and classification entity types and folding their attributes into the related entity type (e.g., attributes from Customer Contact Person and Customer Industry Type are folded into the Customer entity type). These transformations are consistent with ontological principles for representing intrinsic properties We then removed the associative entity types by folding their attributes into both related entity types (e.g., attributes from Sales Order Item are folded into both the Order entity type and the Product entity type). This transformation is consistent with ontological principles for representing mutual properties. When these transformations are made, minor information losses occur associated with constraints on relationships that were deleted. We were careful to avoid involving these aspects of the models in our comprehension and problem-solving tasks. In practice, these losses would be overcome by the expert data modeller engaging in a dialogue with the user about the application domain and completing the "Practice" version of the ER model.

When developing the version of the model that does not distinguish between things and properties (Appendix C, the "Entity only" ER Model), we transformed the "Practice" ER Model by creating a new entity type for each attribute. This transformation is consistent with the philosophy underlying object-role modelling that no distinction should be made between things and their properties (Halpin 1995). "Facts" that connect things are the key concept. When this transformation was made, a more-complex model resulted. Nonetheless, the constraints on relationships provided clear semantics.

The three categories of model used in this study constitute a continuum (Figure 5) varying from the "Entity only" ER Model, where there is no distinction between things and properties, to the "Practice" ER Model, where some types of property are represented as entity types, through to the "Ontologically sound" ER Model, where a clear distinction is made between things and properties.

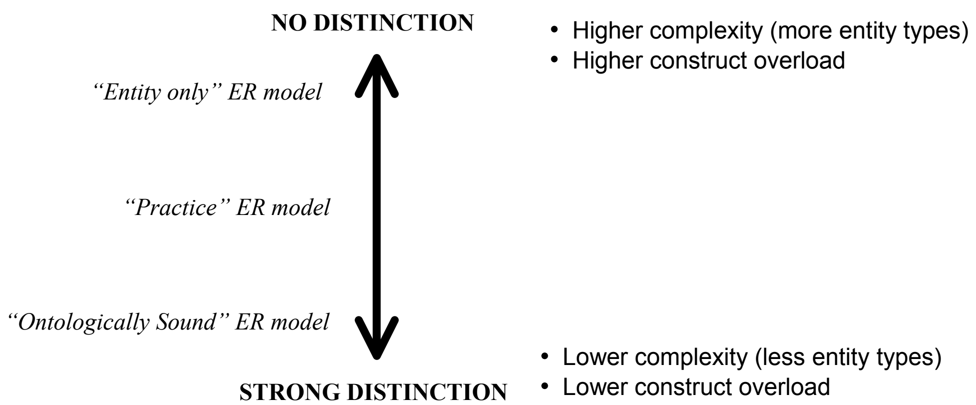


Figure 5: Thing/Property Continuum

Table 1 shows the mapping from ontological concepts to modelling notation constructs for each of the three categories of model.

ONTOLOGICAL CONCEPT	"Entity only" ER	"Practice" ER	"Ontologically sound" ER
THING	Entity	Entity	Entity
INTRINSIC PROPERTY	Entity	Entity or Attribute	Attribute
MUTUAL PROPERTY	Entity	Entity or Relationship	Attribute
VALUE DOMAIN	Entity	Domain	Domain

Table 1 Ontological Mapping

3. Research Method

An experimental setting was chosen for this research to control for extraneous factors that might confound any impacts of alternative representations of things and properties on how well users understand these constructs.

3.1 Design and Measures

A three-group, post-test only experimental design was used with one active between-groups factor. This factor, "type of representation," had three levels. The first level used an "Ontologically sound" ER diagram, the second level used a "Practice" ER diagram, and the third level used an "Entity only" ER model.

The dependent variable, performance, was evaluated using the participants' comprehension and problem-solving performance. Comprehension relates to how well someone understands the "surface-level" features of a domain from a conceptual model. Problem solving provides a better indicator of someone's "deep" understanding of a domain (see, e.g., Mayer 1989). Following Gemino (1999), Bodart et al. (2001) and Shanks et al. (2002), we used comprehension and problem-solving tasks to test how well conceptual models communicate the semantics of a domain to users.

We measured comprehension and problem-solving performance in three ways: (a) accuracy, (b) time taken, and (c) normalized accuracy. Comprehension accuracy was defined as the percentage of comprehension questions correctly answered by each participant. Comprehension time was the time taken by each participant to answer the comprehension questions, expressed in minutes. Normalized accuracy for comprehension was defined as the number of questions answered correctly per hour (calculated by dividing the number of correctly answered questions by the time taken to complete the comprehension task in hours). Problem-solving accuracy was evaluated in terms of whether participants obtained a correct answer to the problem and provided a clear explanation of their rationale. It was expressed as the percentage of problem-solving questions correctly answered by each participant. Problem-solving time was the time taken by each participant to answer the problem-solving questions, expressed in minutes. Normalized accuracy for problem solving was defined as the number of questions answered correctly per hour (calculated by dividing the number of correctly answered questions by the time taken to complete the comprehension task in hours).

3.2 Materials

Five sets of materials were used in the experiment. The first was a summary of the ER symbols used in the diagrams provided to participants in the experiment.

The second set of materials comprised three ER diagrams of a sales order domain: the "Ontologically sound" model (appendix B), the "Practice" model (Appendix A) and the "Entity only" model (Appendix C). The ER diagrams were sufficiently rich to make some problem-solving tasks difficult.

The third set of materials comprised 10 comprehension questions. They were designed to test a user's ability to access and navigate the model for relatively simple tasks. Responses to questions were "yes," "no," or "not sure" (included to minimise guessing). An example is:

> "Can an employee be assigned to manage more than one customer at a time?"

The fourth set of materials comprised 10 problem-solving questions. They were designed to force participants to use the ER diagrams to obtain a correct answer rather than rely on tacit knowledge of the sales order domain. Responses to questions were "possible," "not possible," or "not sure" (included to minimise guessing) with a brief explanation. An example is:

> "An Ontological Plastics supplier wishes to send samples of new and improved hoses to customers who regularly order hoses. Can we determine the number of hoses each customer has had delivered in the previous 3 months and the date of each delivery?"

The fifth set of materials comprised a "personal-profile" questionnaire to obtain information about participants' academic qualifications, industrial experience, and modelling experience.

3.3 Participants

Participants in the experiment were 33 individuals who were either working in industry or were postgraduate students. The former did not play an information technology role in their organizations, nor did they have information systems/technology qualifications. In essence, in the experiment they acted as surrogate end users. Demographic data was collected, but it is omitted here for reasons of brevity. All had at least a Bachelor's degree. Twenty-six had no experience of data models. The remainder had minor experience of one or two modelling techniques like flowcharts or financial models.

3.4 Procedures

Participants were first assigned randomly to one of the three treatments (11 per treatment). They were then run singly or in small groups through the experiment. When they arrived to undertake the experiment, they were asked to complete a consent form and the demographic survey.

Next they were given the document that explained the ER symbols. Participants were permitted to discuss the symbols with the researchers until they indicated they felt confident with the ER symbols. They retained and could refer to the ER summary throughout the experiment.

When participants indicated they were ready to begin, they were then given the "Ontologically sound" ER diagram, the "Practice" ER diagram, or the "Entity only" ER diagram. They retained and could refer to the diagram throughout the experiment. The times they took to answer each comprehension and problem-solving question were recorded. Notes were also made based on participant reactions, queries, and approaches to each question. One researcher conducted the experiment, while another took notes, timed and observed the participant's behaviour during the experiment.

4. Results

Scores for the individual items on the problem-solving dependent measures were calculated. Statistical analyses were performed on the scores for each dependent measure.

4.1 Data Scoring

Scores were awarded as follows:

Comprehension

One mark was given if the answer ("possible" or "not possible") was correct; zero was given if the participant selected "not sure" or their answer was incorrect. Participants were encouraged to answer "not sure" rather than guess an answer.

Problem Solving

Two marks were given if the answer ("possible" or "not possible") was correct; zero was given if the participant selected "not sure" or their answer was incorrect. Explanations were used to amend the score only if the explanation was inconsistent with the answer given. If the answer was correct but the explanation was unclear and did not support the answer, one mark was subtracted from the score. If the answer was incorrect or "not sure" but the explanation indicated the participant was reasoning coherently about the problem, one mark was added to the score. Two researchers independently scored the problem-solving measures on pre-formatted scoring sheets. Few differences arose between the two sets of scores. Where they did occur, they were discussed and reconciled.

These scoring schemes were simple to use and enabled all raw scores to be allocated whole numbers. Final scores were normalised to percentages and are reported in tables below accordingly.

4.2 Data Analysis

Table 2 shows the mean and standard deviation for comprehension scores for each type of model. The accuracy scores are reasonable (approximately 70 percent), and the ontologically sound model scores best on all three measures. In particular, marked differences exist in the time taken and the normalized accuracy scores.

	Accuracy	***Time taken (minutes)***	***Normalized accuracy***
Ontologically sound	73.6 (12.90)	7.26 (2.10)	66.60 (25.2)
Practice	64.5 (18.1)	12.95 (5.12)	34.8 (16.8)
Entity only	63.6 (21.1)	11.89 (5.19)	36.6 (18.0)

Table 2 Comprehension Summary Statistics

Table 3 shows the results of significance testing between the three treatment groups. The "Ontologically sound" group outperformed the "Practice" ER group on both time ($p < 0.003$) and normalized accuracy ($p < 0.002$). It also outperformed the "Entity only" group on both time ($p < 0.012$) and normalized accuracy ($p < 0.005$). There were no significant differences on accuracy. In summary, we obtained strong support for our proposition based on the time and normalized accuracy measures of comprehension performance.

Accuracy

Model	Prac	Sound
Entity Only	t=-0.108 sig=0.915	t=-1.342 sig=0.195
Prac		t=-1.358 sig=0.329

Time

Model	Prac	Sound
Entity Only	t=-0.481 sig=0.636	t=2.747 sig=0.012
Prac		t=3.417 sig=0.003

Normalized accuracy

Model	Prac	Sound
Entity Only	t=0.238 sig=0.814	t=-3.196 sig=0.005
Prac		t=-3.468 sig=0.002

Table 3 Comprehension Differences Between Groups

Table 4 shows the mean and standard deviation for each type of model for problem-solving scores. Overall, the accuracy scores are lower than the comprehension scores and the time taken is considerable longer than the comprehension task, which indicates that this is a more cognitively difficult task. The ontologically sound group scored best on all three measures.

	Accuracy	*Time taken (minutes)*	*Normalized accuracy*
Ontologically sound	63.65 (11.2)	33.32 (12.09)	25.20 (10.2)
Practice	58.20 (16.3)	42.36 (14.27)	19.20 (9.60)
Entity only	55.90 (15.30)	34.55 (12.79)	22.20 (10.2)

Table 4 Problem-Solving Summary Statistics

Table 5 shows the results of significance testing between the three treatment groups. There were no significant differences on accuracy, time, or normalized accuracy. In summary, we obtained no support for our proposition based on measures of problem-solving performance.

Accuracy

Model	Prac	Sound
Entity only	t=0.337 sig=0.740	t=-1.352 sig=0.192
Prac		t=- 0.914 sig=0.372

Time

Model	Prac	Sound
Entity only	t=1.352 sig=0.192	t=0.231 sig=0.819
Prac		t=1.603 sig=0.125

Normalized accuracy

Model	Prac	Sound
Entity only	t=0.777 sig=0.447	t=-0.749 sig=0.463
Prac		t=-1.495 sig=0.151

Table 5 Problem-Solving Differences Between Groups

4.3 Discussion

In this study, we found that use of the ontologically sound representation significantly improved comprehension performance but had no significant effect on problem-solving performance. While this provides partial support for our proposition, it contradicts the findings of Bodart et al. (2001), who found that use of an ontologically sound representation reduced comprehension performance but improved problem-solving performance. Bodart et al.'s study involved the removal of optional properties through sub-typing rather than the thing-property distinction.

A possible explanation for these apparently conflicting findings can be found in theories of human information processing. Psychological studies show that due to limits on short-term memory, humans have a strictly limited capacity for processing information - this is estimated to be "seven, plus or minus two" concepts at a time (Miller 1956). Once the amount of information received exceeds the limits of short-term memory, information overload ensues and comprehension degrades rapidly (Lipowski, 1975). In Bodart et al.'s study, the removal of optional properties increased the complexity of the model, in that optional properties required the addition of subtypes. As a result of the increase in complexity, comprehension performance was reduced.

However in our study, clearly distinguishing between things and properties in the ontologically sound representation reduces complexity compared to the other two representations (7 entities in the ontologically sound representation compared to 25 for the normalised ER model and 71 for the Entity-Only model). Distinguishing between things and properties effectively provides a "chunking" mechanism, which is one of the most-common methods used by the human mind to deal with complexity (e.g., Miller 1956, Cofer 1965). This possibly explains why the ontologically sound model improves comprehension performance compared to the other representations.

Deep-structure understanding (evaluated through the problem-solving task) is less affected by complexity as it involves long-term memory, which is not subject to the same information processing limitations as short-term memory: Effectively, long-term memory capacity is unlimited. Comprehension involves perception of the model and processing in short-term memory rather than reasoning about it in long-term memory as in problem solving, and is most likely significantly affected by complexity.

5. Implications of the Research

For practice, our results support our proposition that things and properties should be modelled explicitly as entity types and attributes. Practitioners should be cautious, therefore, when modelling properties as entity types because they risk undermining users' understanding of the real-world phenomena being represented.

Our results also suggest that practitioners should use different types of model for conceptual modelling and database-design purposes. The "Ontologically sound" model is best for user understanding, while the "Practice" model most likely is more suitable for logical database design. Interestingly, the transformation of the "Practice" model into an "Ontologically sound" model suggests that both types of model can co-exist. The "Ontologically sound" model should be employed with users during the requirements modelling and validation process, while the equivalent "Practice" model should be employed with database designers later in the systems development process.

From a research perspective, our results strengthen a growing body of empirical work that supports the usefulness of ontological theories, particularly Bunge's (1977) ontological theory, as a means of predicting the strengths and weaknesses of conceptual modelling grammars and practices (e.g., Weber 1996, Green and Rosemann 2000, Gemino 1999, Parsons and Wand 2000, Bodart et al. 2001, Shanks et al. 2002). To date, omnibus feature comparisons or case-study comparisons of different grammars and methods (e.g., Olle et al. 1983) have been used to identify problem areas of conceptual modelling grammars. The equivocal results produced using such approaches motivated calls for better theory to guide conceptual modelling research (e.g., Floyd 1986). Ontology provides us with this better theory by allowing us to pinpoint the specific features of conceptual modelling grammars and practices that are likely to be problematical and to then design empirical research to test our predictions.

Furthermore, our research highlights the importance of the ways in which users' understanding of phenomena represented in conceptual models needs to be measured. To test both 'surface-level' and 'deeper-level' understanding, we have used comprehension and problem-solving tasks respectively. The latter have been used because they resemble scenarios or use-cases, which are now widely employed in requirements acquisition.

6. Limitations and Future Research Directions

The major limitations of our research relate to statistical-conclusion validity and external validity. To increase statistical-conclusion validity, we are currently increasing the sample size. Even with our current sample size, however, some of our key tests are statistically significant. Like most experiments the context of the experimental task is limited in scope and somewhat artificial. Nonetheless, our task has enough realism that our results should be robust in other settings involving thing-property representations.

Future research work might examine conceptual-modelling practices, better measures of understanding of conceptual models, and alternative methods of validating conceptual models. Ontological theory also can be used to predict the strengths and weaknesses of other conceptual-modelling practices. For example, approaches to modelling the dynamics of a domain involving alternative ways of representing things and properties may be tested for user understanding.

This research has focussed on the product or output of conceptual modelling. Further research into how user involvement in the conceptual modelling process impacts their ability to understand conceptual models needs to be undertaken.

More work needs to be done to develop valid and reliable measures of user understanding of domain semantics. Our research suggests that comprehension and problem-solving measures have merit. Nonetheless, measures of understanding also need to take into account that (a) users create their worlds (Hirschheim et al. 1995) and (b) shared meaning among a cohort of users may or may not exist.

Finally, alternative methods of having users validate conceptual models as representations of their perceived worlds might be investigated. Our research suggests that methods based on comprehension and problem solving with conceptual models have merit. Nonetheless, a more-systematic articulation of and evaluation of different methods needs to be undertaken.

Acknowledgements

An Australian Research Council Discovery Grant funded this research. We thank Graeme Simsion for helpful comments on our research.

References

Boehm, B (1981), *Software Engineering Economics*, Prentice-Hall, Upper Saddle River, NJ.

Bodart, FM, Sim, Patel, A, Weber, R, (2001), Should optional properties be used in conceptual modelling? A theory and three empirical tests, *Information Systems Research,* vol. 12, no. 4, pp. 384-405.

Bunge, M (1977), *Treatise on Basic Philosophy: Volume 3: Ontology I: The Furniture of the World*, Reidel, Boston.

Chen, PPS (1976), The entity-relationship model: Toward a unified view of data, *ACM Transactions on Database Systems*, vol. 1, no.1, pp. 9-36.

Cofer, CN (1965), On Some Factors in the Organisational Characteristics of Free Recall. *American Psychologist* (April), pp. 261-272.

Floyd, C (1986), A comparative evaluation of system development methods, in TW Olle, HG Sol, AA Verrijn-Stuart, (eds.), *Information System Design Methodologies: Improving the Practice*, North-Holland, Amsterdam, pp. 19-54.

Gemino, A (1999), *Empirical methods for comparing system analysis modelling techniques*, Unpublished PhD thesis, University of British Columbia.

Green, P & Rosemann, M, (2000), Integrated process modelling: An ontological evaluation. *Information Systems*, vol. 25, no. 2, pp. 73-87.

Halpin, TA (1995), C*onceptual Schema and Relational Database Design: A Fact Oriented Approach*, 2nd ed, Prentice-Hall, Sydney.

Hirschheim, R, Klein, H & Lyytinen, K (1995), *Information Systems Development and Data Modelling: Conceptual Foundations and Philosophical Foundations*, Cambridge University Press, Cambridge.

Kent, W (1978), *Data and Reality*, North-Holland, Amsterdam.

Lipowski, ZJ (1975), Sensory and Imaginary Inputs Overload: Behavioural Effects, *Comprehensive Psychiatr*, vol.16, no.3, May/June.

Mayer, RE (1989), Models for understanding, *Review of Educational Research*, vol. 59, pp.43-64.

Miller, GA (1956), The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, *Psychological Review* (March).

Moody, D (2001), *Dealing with complexity: A practical method for representing large entity-relationship models*, Unpublished PhD thesis, University of Melbourne.

Moody, DL & Shanks, G (1998), Improving the quality of entity-relationship models: An action research programme, *The Australian Computer Journal,* vol. 30 pp. 129-138.

Nijssen, GM (1976), *Modelling in Database Management Systems*, North Holland, Amsterdam.

Olle TW, Sol, HL, Tully, CJ (eds.) (1983), *Information System Design Methodologies: A Feature Analysis*, North-Holland, Amsterdam.

Parsons, J & Wand, Y (2000), Emancipating instances from the tyranny of classes in information modelling, *ACM Transactions on Database Systems*, vol. 25, no.2, pp. 228-268.

Rumbaugh, J, Jacobson, I & Booch, G, (1999), *The Unified Modelling Language Reference Manual*, Addison-Wesley, Reading, Mass.

Shanks, G, Tansley, E, Nuredini, J, Tobin, D & Weber, R, (2002), Representing Part-Whole Relationships in Conceptual Modelling: An Empirical Evaluation, in *Proc. International Conference on Information Systems*, Barcelona.

Simsion, G & Witt, G (2001), *Data Modelling Essentials: Analysis, Design and Innovation*, 2nd edition. Coriolis, Arizona.

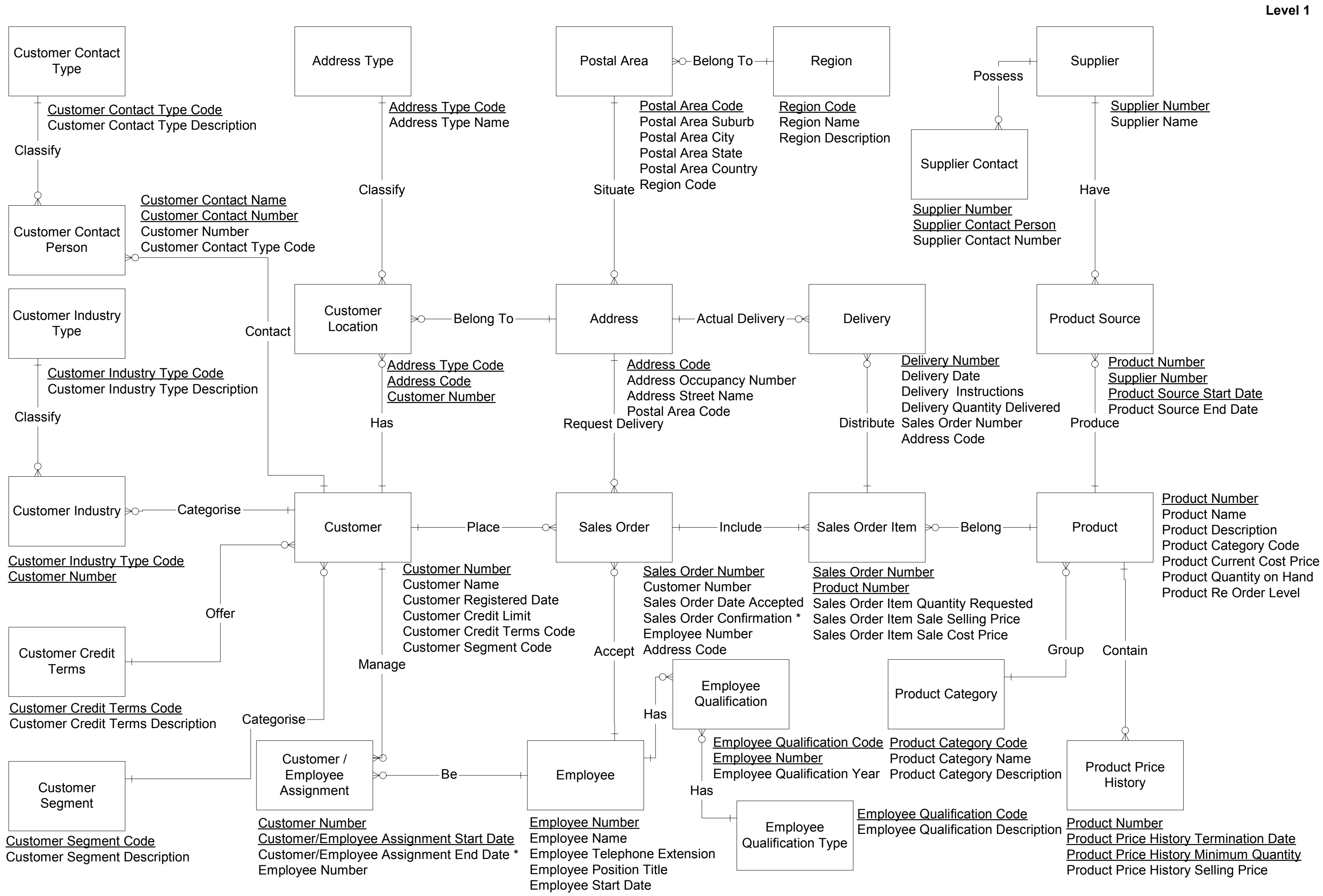
Thalheim, B (2000), *Entity Relationship Modelling: Foundations of Database Technology*, Springer, Berlin.

Wand, Y & Weber, R, (2002), Information systems and conceptual modelling: A research agenda, *Information Systems Research*, vol.13 publication forthcoming.

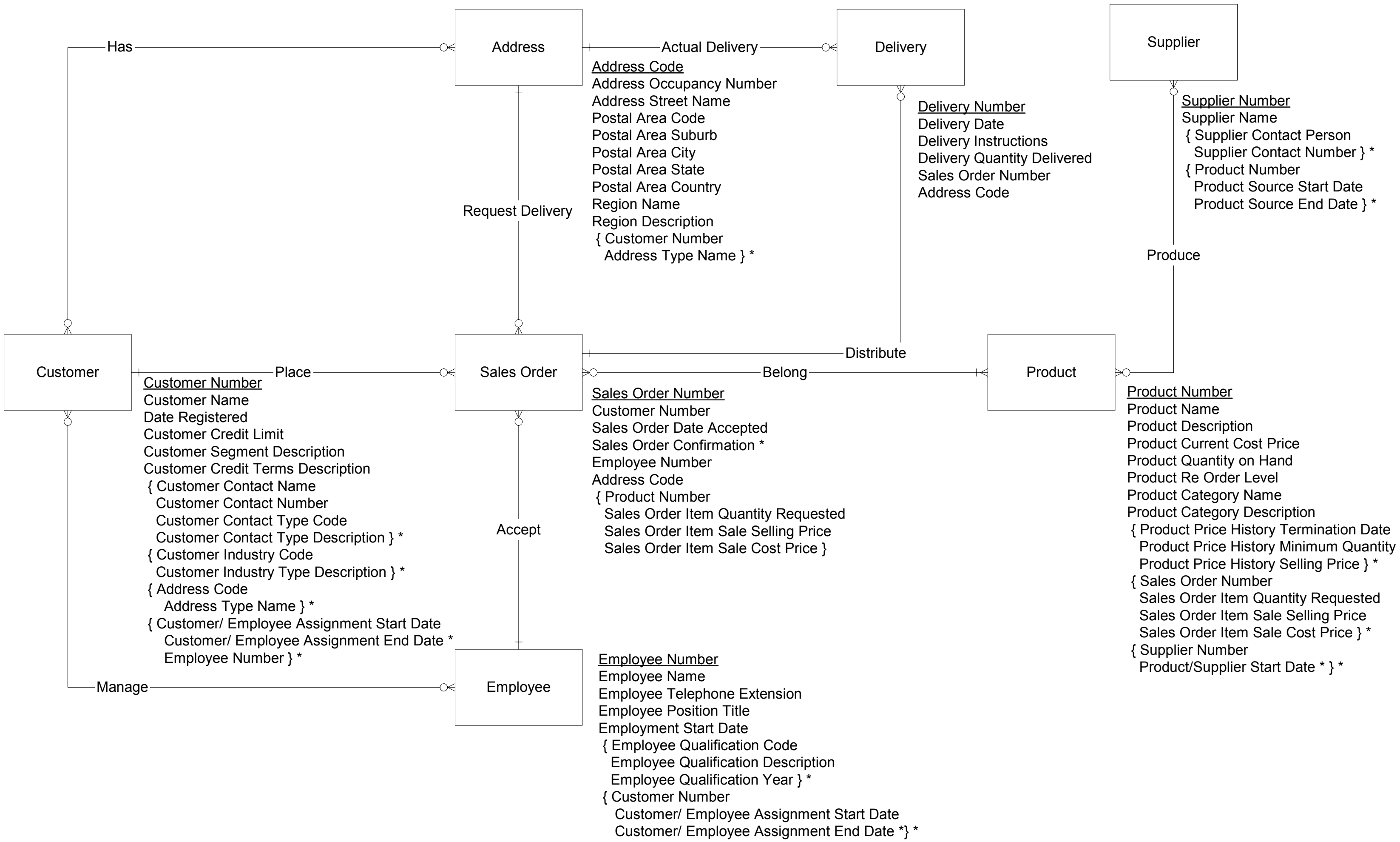
Weber, R (1996), Are attributes entities? A study of database designers' memory structures, *Information Systems Research,* vol. 7. pp. 137-162.

Weber, R (1997), *Ontological Foundations of Information Systems*, Coopers and Lybrand Monograph.

Appendix A "Practice" ER Model



Appendix B “Ontologically Sound” ER Model



Level 3

Appendix C "Entity Only" ER Model