12-31-1999

Understanding Relationships with Attributes in Entity-Relationship Diagrams

Andrew Burton-Jones
University of Queensland

Ron Weber
University of Queensland

UNDERSTANDING RELATIONSHIPS WITH ATTRIBUTES IN ENTITY-RELATIONSHIP DIAGRAMS

Andrew Burton-Jones
PriceWaterhouseCoopers
Australia

Ron Weber
University of Queensland
Australia

Abstract

Conceptual modeling is an important task undertaken during the systems development process to build a representation of those features of an application domain that are important to stakeholders. In spite of its importance, however, substantial evidence exists to show that it is not done well. Designers often provide incomplete, inaccurate, or inconsistent representations of domain features in the conceptual models they prepare. Users often have difficulty understanding the meaning inherent in a conceptual model.

In this paper, we investigate the proposition that part of the difficulties that stakeholders experience with conceptual modeling arises when a conceptual modeling grammar or a representation produced using the grammar lacks ontological clarity. Lack of ontological clarity arises when a one-one mapping does not exist between conceptual modeling constructs and real-world constructs. For example, the grammatical construct of an entity is used to represent both things and events in the real world.

Specifically, we focus on the grammatical construct of a relationship with attributes, which is often used in entity-relationship modeling. We argue that use of this construct produces ontologically unclear representations of a domain. We also report results from an experiment we undertook where we investigated the impact of using relationships with attributes in conceptual modeling representations on the problem-solving performance of users of these representations. Consistent with our predictions, we found that using relationships with attributes undermined problem-solving performance in unfamiliar domains. Contrary to our predictions, however, their use did not undermine problem-solving performance in familiar domains.

Keywords: Conceptual modeling, entity-relationship modeling, ontology, experimental research, problem solving, perceived ease of understanding

1. INTRODUCTION

When designers undertake the task of conceptual modeling during the development of an information system, they seek to articulate a representation that captures those features of an application domain that are important to stakeholders. Their objective is to provide a basis for improved discourse with and among stakeholders, which hopefully leads to a higher-quality system being implemented.

Given the nature of conceptual modeling, we might expect that it occupies an important place within the overall systems development process (e.g., Moody 1998). Indeed, some empirical evidence confirms this view (e.g., Moody and Shanks 1998). Other research shows, however, that conceptual modeling is not always undertaken during systems development work, even when it is deemed to be important (e.g., Hitchman 1995). Moreover, both surveys of practice and laboratory research indicate that conceptual modeling is not always done well (e.g., Batra, Hoffer and Bostrom1990; Batra and Marakas 1995; Goldstein and Storey 1990; Maier 1996; Prietula and March 1991).

The research described in this paper is motivated by the belief that conceptual modeling is critical to high-quality system development work, but that it will remain a problematical task while it lacks sound theoretical foundations. In our view, many extant conceptual modeling methods and tools are rooted primarily in weakly articulated concepts and the intuition, knowledge, and experience of their developers. While often these methods and tools have served designers well, nonetheless ongoing problems associated with their use manifest uncertainties about how to represent the real-world phenomena that must be modeled.

In the research described in this paper, we focus on the conceptual modeling construct of a relationship when it is represented with its own attributes. The reason for our choice is twofold. First, we are seeking to test a theory proposed by Wand and Weber (1993), which predicts that conceptual modeling *grammars* (which provide the foundation for conceptual modeling methods and tools) will be problematical to use when they lack *ontological clarity*. Ontological clarity exists only when the mapping between each modeling construct available in a conceptual modeling grammar and a real-world construct is one-one. We contend that a relationship with attributes violates this condition, thus causing difficulties for its users.

Second, research on the relationship construct is important because (1) it is central to many conceptual modeling approaches and (2) it is used extensively in practice (e.g., Maier 1996). Indeed, we argue that the relationship construct is the linchpin through which humans associate different things in the world, thereby enabling them to conceptualize the subsystems and systems that are fundamental to their being able to reason about the world. When relationships are given attributes, however, our claim is that they will undermine their users' problem-solving abilities.

The remainder of our paper proceeds as follows. First, we present the theoretical background to our proposition that relationships, which have their own attributes, lack ontological clarity and therefore will inhibit their users' problem-solving abilities. Second, we describe an experiment that we undertook to test our proposition. Third, we present the results of our experiment. Finally, we discuss some implications of our work and outline some directions for further research.

2. THEORETICAL BACKGROUND

A growing body of research now indicates that *context* is an important determinant of the ways in which humans assign meaning to symbols or scripts (e.g., Ngwenyama and Lee 1997). This research clearly has implications for our understanding of how the stakeholders in an information system prepare, use, and interpret representations like conceptual models. Users' prior experience with and knowledge of each other or a domain, for example, might overcome interpretation difficulties that they would otherwise encounter because the semantics of a conceptual model are unclear.

Irrespective of the importance of context, however, Wand and Weber (1993) argue that conceptual models generated using grammars that lack *ontological clarity* will be more difficult to understand. Ontological clarity is undermined in three ways:

1. *Construct overload:* A conceptual modeling grammar allows one conceptual modeling construct (e.g., an entity) to be used to represent more than one real-world construct (e.g., a thing and an event).
2. *Construct redundancy:* A conceptual modeling grammar allows one real-world construct (e.g., an event) to be represented by more than one grammatical construct (e.g., an entity and a relationship).
3. *Construct excess:* A conceptual modeling grammar has a construct (e.g., an optional attribute) that has no real-world counterpart (optional attributes are motivated by *implementation* concerns, not conceptual modeling concerns).

In short, ontological clarity exists only if each conceptual modeling construct represents one and only one real-world construct and each real-world construct is represented by one and only one conceptual modeling construct. Hopefully the set of conceptual modeling constructs involved in the mapping is well defined via the rules of the grammar. The set of real-world constructs involved in the mapping, however, will depend on the ontological theory chosen (Weber 1997). The quality of any analysis of ontological clarity depends, therefore, on (1) how well the set of grammatical constructs are specified and (2) the adequacy of the ontological theory chosen in terms of its ability to describe the structure of the world. If both sets are "well-specified," however, when a one-one mapping does *not* exist, Wand and Weber argue that users of a conceptual model will be unsure about what real-world construct is being represented by a conceptual modeling construct. In this light, users' understanding of and ability to reason about the application domain being represented by a conceptual model will be undermined.

3. RELATIONSHIPS WITH ATTRIBUTES AND ONTOLOGICAL CLARITY

The relationship construct is an important component of many conceptual modeling grammars. For example, it is a central construct in the entity-relationship modeling grammar (Chen 1976), which is widely used in practice. It also features in some object-oriented modeling grammars (Embley, Kurtz and Woodfield 1992; Rumbaugh et al. 1991).

Figure 1 shows an example of an entity-relationship (ER) diagrammatic representation of a relationship with attributes. The relationship "married to" has three attributes—namely, "# children," "celebrant," and "date." This type of representation of relationships with attributes is common (e.g., Elmasri and Navathe 1994, p. 67).

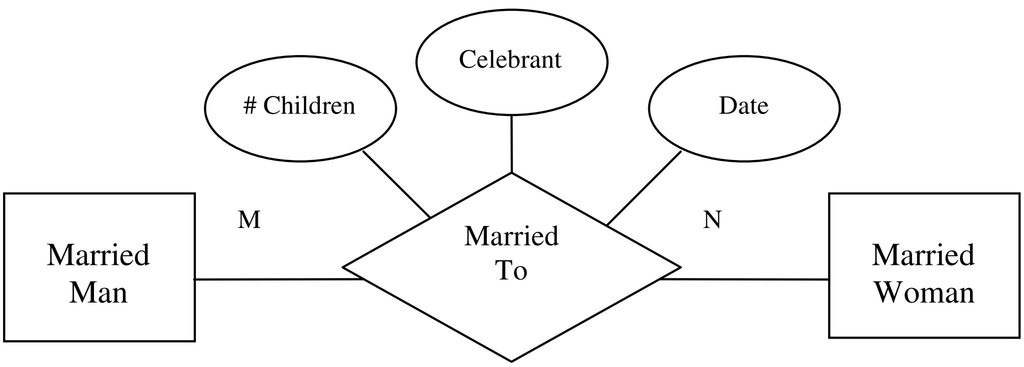

Figure 1. A Relationship with Attributes

We argue that the representation in Figure 1 lacks ontological clarity. The following interpretations of Figure 1 (some of which are only partial) are all plausible:

1. The "married to" relationship represents a *mutual* property (a property that has meaning only in terms of two or more things in the world) that itself possesses three *intrinsic* properties (properties that have meaning in terms of individual things in the world).
2. The relationship and the three attributes are all mutual properties associated with the "married man" and "married woman" entities.
3. The "married to" relationship represents a thing/entity/system in its own right (namely, a marriage) that has three intrinsic properties: "# children," "celebrant," and "date."
4. The "date" attribute represents a *value domain* for the "married to" relationship.
5. The relationship could represent the *event* of being married, where the event has three intrinsic attributes.
6. The "celebrant" attribute could represent a thing/entity.

Each of the above interpretations of Figure 1 is problematical. Under the first interpretation, for a start, like Bunge (1977, pp. 98-99) we reject an ontological position which holds that properties can have properties. Only things have properties. Assigning properties to properties simply means that the first property has not been fully specified. For example, "color" is a property of someone's eyes, but "hue" is not a property of their eye color. It is still a property of their eyes. Moreover, if mutual properties (relationships) can have properties, then presumably they can have mutual as well as intrinsic properties. In other words, designers should be allowed to "attach" both relationships and attributes to relationships—a position we believe most designers would reject. Similarly, intrinsic properties (attributes) should also be able to possess properties, which means that designers should be allowed to attach relationships and attributes to attributes—again, a position we believe most designers would reject. Finally, how many levels of "nesting" of properties do we permit? Can properties build on properties to an infinite extent?

Nonetheless, even if one holds to a position that mutual properties can have intrinsic properties, construct overload exists. We now have two different ontological constructs: an *intrinsic property of a relationship* and an *intrinsic property of a thing*. Yet they are represented by the same grammatical construct (the ER attribute symbol). If ambiguity is to be avoided, users of the diagram must be confident that the person who prepared the diagram did not intend the relationship construct to stand for a thing (see the third interpretation). Moreover, if relationships can have relationships and construct overload is to be avoided, different constructs must be used to represent a *mutual property of a relationship* and a *mutual property of a thing*. Similar arguments apply to the need for different grammatical constructs to represent (1) an *intrinsic property of an attribute*, an *intrinsic property of a relationship*, and an *intrinsic property of a thing* and (2) a *mutual property of an attribute*, a *mutual property of a relationship*, and a *mutual property of a thing*.

Under the second interpretation, construct redundancy and construct overload exist. On the one hand, two different grammatical constructs—a relationship construct and an attribute construct—have been used to represent the single ontological construct of a mutual property (construct redundancy). On the other hand, the attribute construct has also been used to represent two ontological constructs—namely, an intrinsic property and a mutual property (construct overload). As a result, users of the representation must employ extra-model knowledge to resolve any ambiguities that arise. Figure 2 shows the "correct" representation if all the properties are deemed to be mutual properties.

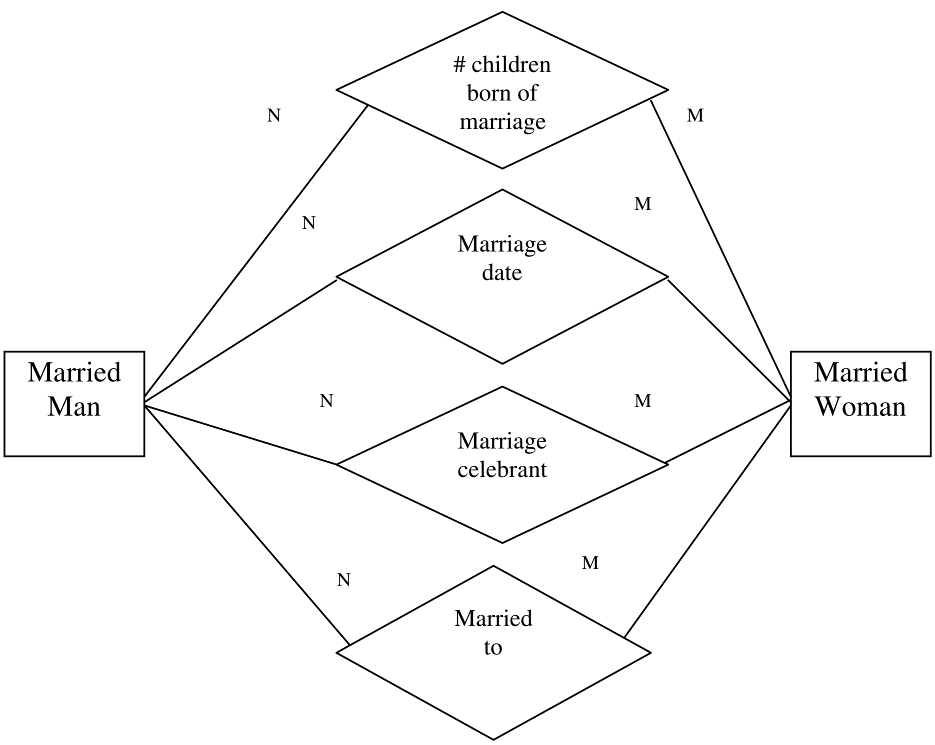

Figure 2. Replacement of Attributes with Mutual Properties

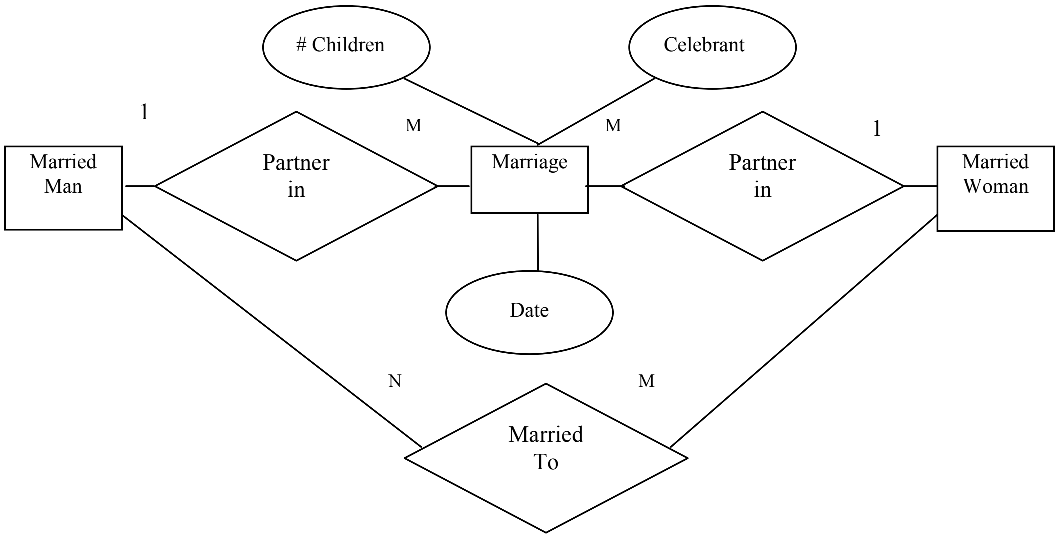

Figure 3. Replacement of a Relationship with an Entity

Under the third interpretation, again construct redundancy and construct overload exist. On the one hand, the entity construct and the relationship construct have been used to represent the ontological construct of a thing (construct redundancy). On the other hand, the relationship construct has also been used to represent the ontological constructs of a thing and a mutual property (construct overload). Once more, users of the representation must employ extra-model knowledge to resolve any ambiguities that arise. Figure 3 shows the "correct" representation if a relationship is deemed to be a thing.

Under the fourth interpretation, construct overload exists. The attribute construct has been used to stand for the ontological constructs of an intrinsic property and the value domain of an intrinsic property. Ambiguity arises because two different ontological constructs have been represented by the same grammatical construct.

Under the fifth interpretation, construct overload again exists. The entity construct has been used to represent the ontological constructs of a mutual property and an event. Moreover, we reject the ontological position that events can have properties. Only things have properties. Events are simply described in terms of the properties of things that change. Nonetheless, if one sustains a view that events can have properties, avoidance of construct overload dictates that different grammatical constructs must be used to represent *properties of events* and *properties of things* (and *properties of properties*) (see the arguments relating to the first interpretation of Figure 1 above).

Under the sixth interpretation, construct overload also exists. The attribute construct has been used to stand for the ontological constructs of an intrinsic property and a thing. Once more, ambiguity arises because two different ontological constructs have been represented by the same grammatical construct.

In the example shown in Figure 1, resolving these ambiguities may be straightforward because many users of the diagram are likely to possess the extra-model knowledge that will allow them to clarify the semantics underlying the representation. In other words, contextual knowledge allows semantic ambiguities to be resolved (Ashcraft 1989). As users' knowledge of an application domain decreases, however, we predict they will be less able to resolve semantic ambiguities that arise. Moreover, we predict that the increased cognitive dissonance that occurs as a result of semantic ambiguities will be manifested as a decline in their ability to solve problems pertaining to the domain being represented. In this regard, we argue that ontological clarity aids problem solving by helping users structure their "problem space" in memory. In Newell and Simon's (1972) problem-solving theory, problem solving difficulty is a function of the size of the problem space. By narrowing the possible interpretations of a model, ontological clarity enables users to construct a smaller search space for solving problems based on the model.

These arguments lead us to propose the following:

> ***Proposition:*** *Relationships with attributes are ontologically unclear. They undermine understanding among the users of a conceptual model and impair their abilities to solve problems about the domain being represented. The effect is greater as users have less contextual knowledge of the domain.*

4. METHOD

To test our proposition, we used a laboratory experiment. Given the exploratory nature of our research, our choice was motivated by the need to exercise careful control over other factors that might impact users' understanding of a conceptual model.

4.1 Design

The experiment used a 2 X 2 mixed design. The between-subjects factor was the ontological clarity associated with a conceptual model. It was manipulated at two levels: ontologically unclear and ontologically clear. The within-subjects factor was the domain knowledge possessed by users of a conceptual model. It was also manipulated at two levels: low knowledge and high knowledge. The order of treatments was varied randomly across participants to control for any order effects. The dependent variables were problem-solving performance and the participant's ease of understanding of the conceptual model shown to them.

4.2 Participants

Seventy-six undergraduate and postgraduate information systems students participated in the experiment. The students were trained in ER modeling in the weeks prior to the experiment. Their undertaking the experiment formed a small part of their subject assessment (4%). Participants were assigned randomly to the two levels of the between-subjects factor.

4.3 Materials and Dependent Measures

ER models were developed for two domains. Participants were expected to be familiar with one domain (a university) and unfamiliar with the other domain (a plant nursery). Two versions of the ER models were created for each domain. One contained ontologically unclear representations by containing six relationships with attributes. The other contained ontologically clear representations by representing the relationships with attributes as either multiple mutual properties (relationships) or things (entities) with their own properties (attributes). Figures 4 and 5 show the ontologically clear and ontologically unclear representations for the nursery domain.

As in Gemino (1998), the problem-solving measure used in our research was a participant's number of acceptable answers to six problem-solving questions asked about the domain. For example, a problem-solving question for the nursery domain was:

> An employee is fired for a serious error during his work that adversely affected many plantlets. From the information provided in the model, write down as many possible errors as you can and explain how each error could have affected the plantlets?

We assessed participants' answers in two ways. The first, based on Gemino, measured participants' total number of acceptable answers. It is denoted "PST" (problem-solving total). Furthermore, because it was also possible to answer problem-solving questions using aspects of the ER models that are the same between groups (e.g., the entities that relate to the relationships with attributes), we also examined participants' answers to problem-solving questions that came from aspects of the models represented differently between the two groups. This latter measure is labeled "PSD" ("problem-solving differential").

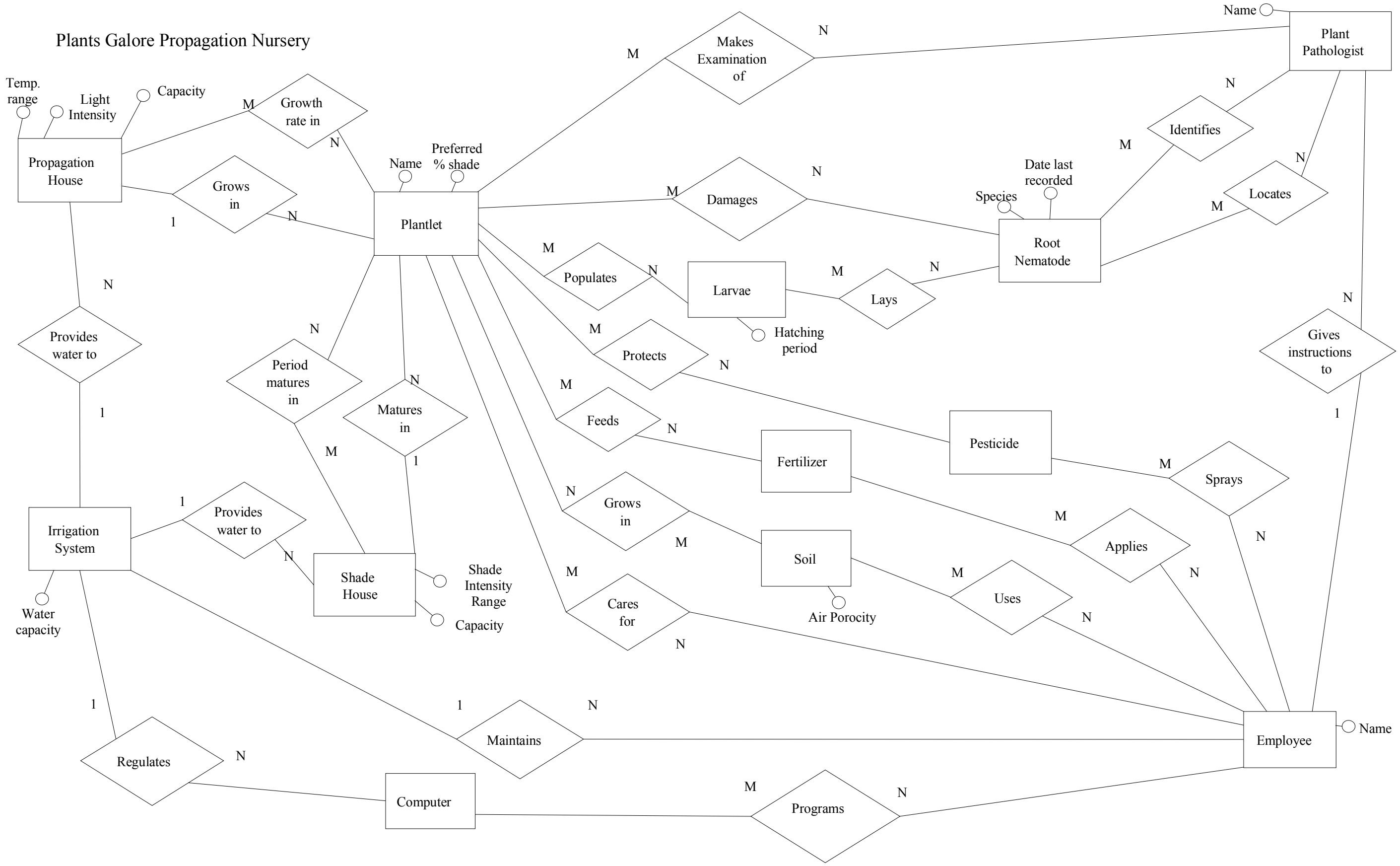

Figure 4. Ontologically Clear ER Model for Nursery

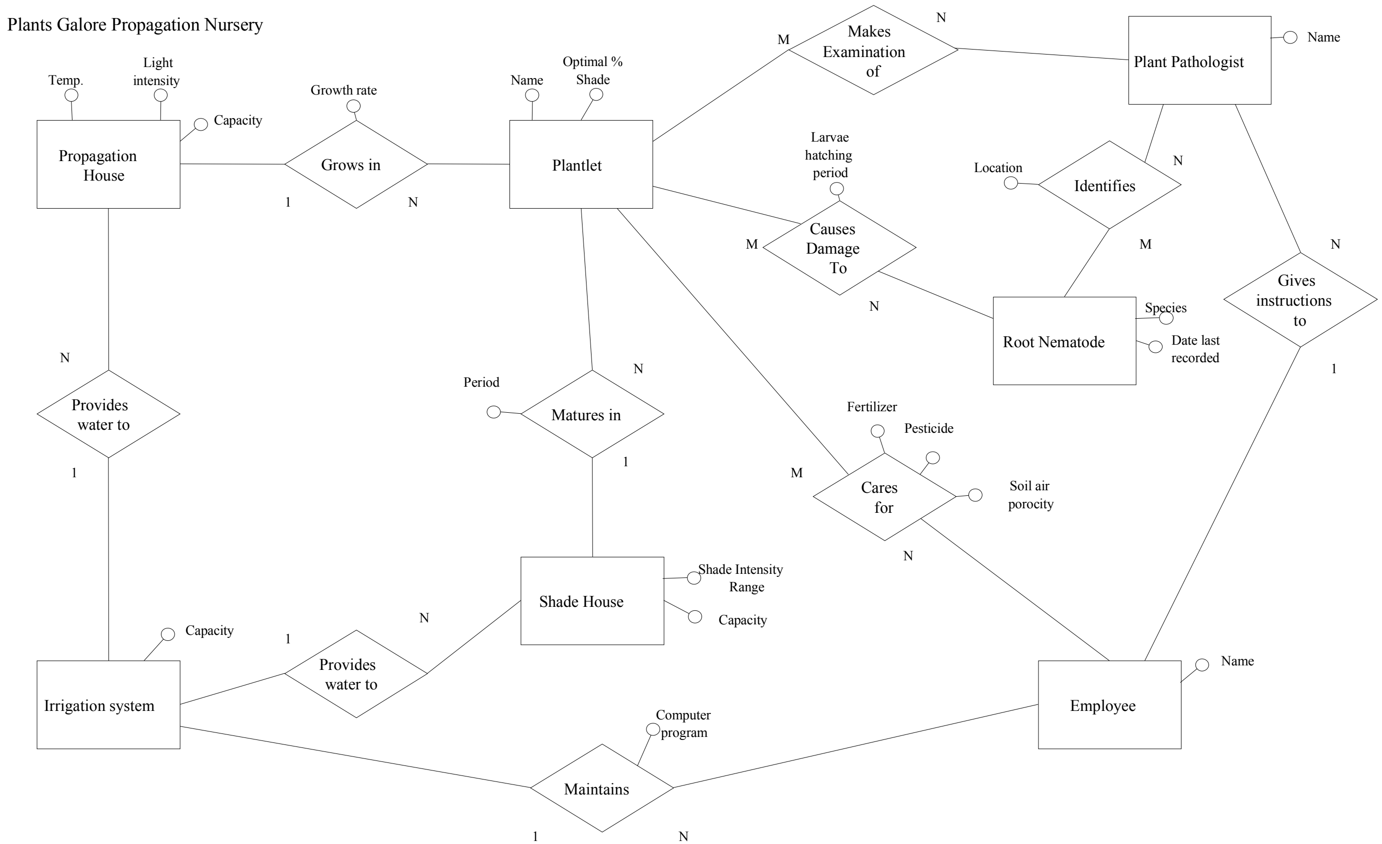

Figure 5. Ontologically Unclear ER Model for Nursery

To measure participants' perceived ease-of-understanding of their ER model ("PEU"), we adapted Davis' (1989) perceived ease-of-use instrument. Six items were chosen from the ease-of-use instruments in Moore and Benbasat (1991) and in Gemino. These were then adapted to ask participants about their model's "ease-of-understanding" rather than "ease-of-use."

4.4 Procedures

Eleven students participated in a pilot test prior to undertaking the experiment proper. Minor changes were made to the experimental materials based on the feedback they provided.

During the experiment proper, participants were first given a one-page document that contained the major constructs used in ER diagrams. They were given one minute to review this document. Next they undertook Case 1. Participants first answered two short questions that asked them about their knowledge of the domain in the case. They rated their level of knowledge on a seven-point scale. Next they reviewed the ER diagram of the domain and answered eight brief comprehension questions requiring "yes/no" responses. These questions were designed to help participants develop a mental model of the domain. When participants signalled they had completed the comprehension questions by raising their hand, their ER diagrams and comprehension questions were collected. They were then provided with the problem-solving questions and the PEU instrument. Participants were given 20 minutes in total to answer both the comprehension questions and the problem-solving questions. After completing the problem-solving questions, participants then completed the PEU instrument.

When Case 1 was completed, participants began Case 2. The procedures used for Case 1 were again used for Case 2.

5. DATA ANALYSIS AND RESULTS

The data analysis proceeded in two steps. First, we undertook some preliminary checks to ascertain whether our treatment manipulations had worked and whether our PEU instruments appeared valid and reliable. Then we undertook tests to determine whether our hypotheses were supported.

5.1 Manipulation of Domain Knowledge

The two short questions asking about domain knowledge were used to conduct a manipulation check to confirm that participants knew more about the university domain than the nursery domain. The results confirmed this prediction (the means for domain knowledge differed at $p < 0.001$).

5.2 PEU Instrument Reliability and Validity

The results of common factor analyses conducted on the PEU instruments for the two domains indicated they had adequate convergent validity. Both eigenvalues were above 3.500. Moreover, in a two-factor solution, all items loaded highest on the first factor and the eigenvalues for the second factor were less than one. The Cronbach alphas for the two instruments were also high (0.890 and 0.900).

5.3 Coding of Problem-Solving Answers

Given the somewhat subjective nature of the problem-solving answers, we had them evaluated by independent coders. We first created a set of acceptable answers for each problem-solving question. Two students and an operator of a propagation nursery were then asked to incorporate additional acceptable answers. The revised set was then given to two independent coders together with participants' responses. The coders initial inter-rater reliability was 73%. After making minor changes to the coding sheets, the coders then worked together with the revised coding sheets and resolved all coding differences.

5.4 Descriptive Statistics

Table 1 provides the descriptive statistics for the data obtained from the experiment.

These statistics indicate that the data are free from significant skewness and kurtosis. Nonetheless, Levene's test showed that one variable (PEU for the nursery domain) violated the homogeneity-of-variance assumption ($p = 0.094$). After removing outliers for this variable, however, it also met the homogeneity-of-variance assumption. An analysis of outliers indicated that they affected the results only for the interaction test on PEU.

Table 1. Descriptive Statistics*

Clarity		Uni: PSD	Uni: PST	Uni: PEU	Nurs: PSD	Nurs: PST	Nurs: PEU
Clear	N	37	37	37	37	37	37
	Mean	6.135	10.541	29.757	5.216	9.622	23.081
	Std. Dev.	2.201	3.610	6.529	2.335	3.968	8.105
	Skewness	.695	.223	-.233	.167	-.473	-.416
	Kurtosis	.170	-.511	-.569	1.247	.689	-.870
Unclear	N	39	39	38	39	39	36
	Mean	5.821	10.385	28.000	3.539	8.103	24.139
	Std. Dev.	2.025	3.551	5.968	2.211	4.217	6.608
	Skewness	.298	.312	-.618	.223	.453	-.735
	Kurtosis	.753	-.377	-.095	-.625	.182	.038
Total	N	76	76	75	76	76	73
	Mean	5.974	10.461	28.867	4.355	8.842	23.603
	Std. Dev.	2.104	3.557	6.272	2.409	4.141	7.373
	Skewness	.521	.263	-.346	.209	.009	-.563
	Kurtosis	.412	-.498	-.311	.194	-.047	-.530

***Clarity**: Unclear (clear) refers to the group receiving ontologically unclear (clear) diagrams.
PSD: Participants' acceptable answers to problem-solving questions that come from aspects of the ER diagrams that differ between groups.
PST: Participants' total number of acceptable answers to problem-solving questions.
PEU: Participants' perceived ease-of-understanding of their ER diagram.
Uni/Nurs: Uni refers to the university case; Nurs refers to the Nursery case.
Cell sizes are lower for PEU because four participants failed to answer the questionnaire.

5.5 Tests of Hypotheses

Table 2 presents the three hypotheses tested in this research.

Because a count of answers for the nursery case does not *necessarily* equate in terms of problem-solving difficulty with a count of answers for the university case, we followed Mayer and Gallini (1990) by *not* examining interaction effects for the problem-solving measures (PSD and PST). Instead, participants' problem-solving performance was tested in each domain separately. The results were then compared across the two domains. As Table 2 shows, we expected ontological clarity to have a positive effect on PSD and PST in *both* domains, but we also anticipated that it would have a greater effect in the nursery domain. Because the PEU instrument asks consistent questions across domains, however, the interaction effect for PEU can be tested statistically.

Table 2. Hypotheses Tested in Research*

Hypotheses for University Domain
H1a: For the university case, participants' scores for PSD will be lower in the ontologically unclear group than the clear group.
H1b: For the university case, participants' scores for PST will be lower in the ontologically unclear group than the clear group.
H1c: For the university case, participants' scores for PEU will be lower in the ontologically unclear group than the clear group.

Hypotheses for Nursery Domain
H2a: For the nursery case, participants' scores for PSD will be lower in the ontologically unclear group than the clear group.
H2b: For the nursery case, participants' scores for PST will be lower in the ontologically unclear group than the clear group.
H2c: For the nursery case, participants' scores for PEU will be lower in the ontologically unclear group than the clear group.

Hypotheses Comparing the University and Nursery Domains
H3a: The positive effect of ontological clarity on participants' scores for PSD will be greater in the nursery case than in the university case.
H3b: The positive effect of ontological clarity on participants' scores for PST will be greater in the nursery case than in the university case.
H3c: The positive effect of ontological clarity on participants' scores for PEU will be greater in the nursery case than in the university case.

*__PSD__: Participants' acceptable answers to problem-solving questions that come from aspects of the ER scripts that *differ* between groups.
PST: Participants' total number of acceptable answers to problem-solving questions.
PEU Participants' perceived ease-of-understanding their ER script.

For each domain, three univariate ANOVAs were estimated with ontological clarity as the independent variable, PSD, PST, and PEU as the dependent variables, and order of tasks as a covariate. The results below also include the test for the interaction between domain knowledge and ontological clarity on PEU. For this latter test, domain knowledge was treated as a within-subjects factor and ontological clarity as a between-subjects factor. As outliers affected the result for this test, the results are presented both with and without outliers.

Tables 3 through 6 summarize the test results. To facilitate comparison between tests, the covariate for order is included, irrespective of whether it is related to the dependent measure. For the test of the interaction effect on PEU, however, the covariate is omitted to facilitate obtaining an understanding of the result. In no test, however, does the omission of the order covariate significantly change the results. Unless indicated, the F-statistics, p-values, and parameter estimates reflect the main effect for ontological clarity in each test.

The results in Tables 3 and 4 show that ontological clarity had a statistically significant effect on the problem-solving differential (PSD) measure for the nursery domain but not the university domain. Table 1 shows that participants who received the ontologically clear representation outperformed those who received the ontologically unclear representation. For both domains, however, ontological clarity had no significant effect on the problem-solving total (PST) measure. As we noted above, however, this measure is less sensitive to the effects of ontological clarity because it includes answers to questions that relate to domain phenomena that are represented the same way in both the ontologically clear and ontologically unclear representations.

Table 3. Univariate ANOVAs for University Domain
Dependent Variable = $\propto_1$ + $\propto_2$Order + $\propto_3$Clarity + ϵ

Dependent Variable	R^2	F-statistic	p-value	$\propto_3$ Estimate
PSD	.046	.435	.512	.315
PST	.011	.036	.851	.155
PEU	.044	1.542	.218	1.783

Clarity: Coded as 1 for the ontologically unclear diagram and 2 for the ontologically clear diagram.
Order: Represents the covariate for order of tasks. It is coded as 1 if participants received the university case first and 2 if they received the nursery case first.

Table 4. Univariate ANOVAs for Nursery Domain
Dependent Variable = $\propto_1$ + $\propto_2$Order + $\propto_3$Clarity + ϵ

Dependent Variable	R^2	F-statistic	p-value	$\propto_3$ Estimate
PSD	.199	11.176	.001	1.667
PST	.159	2.948	.090	1.517
PEU	.034	.354	.554	-1.024

In this test, PEU violates Levene's test for homogeneity-of-variance (p = .094)
After omission of outliers, PST is significant at p = .01.

Table 5. Tests for Interaction Effect on PEU Including Outliers
PEU = $\propto_1$ + $\propto_2$Domain + $\propto_3$Clarity + $\propto_4$Domain*Clarity + ϵ

Effect	Df	Sum of Squares	Mean Square	F-Statistic	P value
Between Subjects					
Intercept	1	99598.113	99598.113	1325.270	.000
Clarity	1	1.724	2.889	.023	.880
Error	70	5260.714	75.153		
Within Subjects					
Domain	1	1003.112	1003.112	48.836	.000
Domain*Clarity	1	70.001	70.001	3.408	.069
Error	70	1437.825	20.540		

In this test, PEU violates Levene's test for homogeneity of variance on the nursery domain (p = .094).
Domain: Coded as 1 for the university domain and 2 for the nursery domain.
Eta squared for Domain*Clarity = .046.

Table 6. Tests for Interaction Effect on PEU Omitting Outliers
$PEU = \propto_1 + \propto_2 Domain + \propto_3 Clarity + \propto_4 Clarity*Domain + \epsilon$

Effect	Df	Sum of Squares	Mean Square	F-statistic	P value
Between Subjects					
Intercept	1	92984.629	92984.629	1451.312	.000
Clarity	1	.191	.191	.003	.957
Error	62	3972.301	64.069		
Within Subjects					
Domain	1	998.550	998.550	51.192	.000
Domain*Clarity	1	45.675	45.675	2.342	.131
Error	62	1209.379	19.506		

Eta squared for Domain*Clarity = .036.

For the PEU measure, the domain factor was statistically significant. Consistent with our expectations, Table 1 shows that participants rated perceived ease of understanding of the representation for the university domain higher than perceived ease of understanding of the representation for the nursery domain. Neither ontological clarity nor the interaction between ontological clarity and domain were significant, however, for the PEU measure, although the interaction term approached significance when outliers were retained ($p = 0.069$).

6. DISCUSSION OF RESULTS

Hypotheses 1a, 1b, 2a, and 2b proposed a positive effect of ontological clarity on participants' problem-solving performance in both domains. Based on the problem-solving differential measure, the results support this proposition for the nursery domain but not for the university domain. The parameter estimates for the university domain are mostly in the expected direction, however, so the results for the university domain may reflect low statistical power.

Hypotheses 3a and 3b proposed that ontological clarity's effect would be greater in the nursery domain than in the university domain. Recall, we cannot test this interaction effect statistically because the problem-solving scores for the two domains are not strictly comparable. Given that ontological clarity had a statistically significant effect on the problem-solving differential measure for the nursery domain but not the university domain, however, we have indirect evidence to support Hypothesis 3a.

Hypotheses 1c, 2c, and 3c were not supported by the results. Ontological clarity had no effect on participants' perceived ease of understanding of the representations they received. The only outcome was that participants perceived the university representation easier to understand than the nursery representation.

We suspect there are two reasons why the PEU results are contrary to our expectations. First, the problem-solving (differential) measure is more sensitive to participants' understanding than the PEU scale. Second, while ontological clarity helps users understand a representation (as shown by their problem-solving differential scores), their initial negative perceptions of the representation remains.

7. IMPLICATIONS

We believe our results have implications for both practice and further research. In terms of practice, we have evidence to support Wand and Weber's (1993) contention that lack of ontological clarity undermines users' ability to understand and work with conceptual models. Clearly, the results of a single study must be treated cautiously. Nonetheless, Kim and March (1995), Weber (1996), and Gemino (1998) have also obtained evidence that shows lack of ontological clarity has detrimental effects on the users of conceptual models. Collectively, therefore, we believe the case against ontologically unclear conceptual modeling grammars and ontologically unclear conceptual modeling representations (diagrams) is building.

We recommend, therefore, that practitioners be circumspect about their choice and use of conceptual modeling grammars. They should avoid grammars that contain an inadequate set of constructs such that designers will be unable to avoid producing ontologically unclear representations of a domain. Furthermore, where the rules of the grammar are sufficiently ill-specified such that designers can produce either an ontologically clear or ontologically unclear representation (for example, as in our experimental representations), they should opt for the ontologically clear representation. Currently we believe that no single factor is the major cause of the problematical use of conceptual modeling in practice that we reported in the introduction to this paper. Nonetheless, we hypothesize that significant improvements will occur if the stakeholders in conceptual modeling recognize that (1) conceptual modeling grammars need to be improved to enable ontologically clear representations of a domain to be generated easily and (2) practices that lead to ontologically unclear representations should be purged from conceptual modeling work.

With respect to the relationship construct, on the basis of our theoretical arguments and our tentative empirical results, we recommend that relationships with attributes *not* be used in conceptual models. Instead, the meaning of a relationship with attributes should be clarified and it should be replaced with constructs that do not undermine ontological clarity.

In terms of research, we see three major avenues for further work. First, more empirical research is needed to determine whether the putative effects of ontological clarity appear to hold generally during analysis and design work. For example, other instances of ontologically unclear conceptual modeling practices could be identified and their effects on users examined via, say, users' ability to query a database. Second, analytical research is needed to determine whether the arguments in support of using ontologically clear conceptual modeling grammars and producing ontologically clear conceptual modeling representations can be sustained on the basis of psychological theories. Such work is needed to enable us to understand whether ontological clarity is likely to be a significant factor impacting users' ability to employ conceptual models successfully. Moreover, it will assist us to understand whether lack of ontological clarity is likely to have an impact on the work that users' undertake with conceptual models. Third, ontological clarity sometimes involves a cost because the ontologically clear diagrams have more elements. At some point, this outcome may lead to information overload, thereby undermining users' understanding of the domain. Future research might address this issue.

8. ACKNOWLEDGMENTS

We are grateful to Paul Bowen and Arvind Patel for their assistance with this research. We also thank the Track Chairs, the Associate Editor, and the Reviewers for their helpful comments on an earlier version of this paper.

9. REFERENCES

Ashcraft, M. H. *Human Memory and Cognition*, Scott, Glenview, IL: Foresman and Company, 1989.

Batra, D.; Hoffer, J. A.; and Bostrom, R. P. "Comparing Representations with Relational and EER Models," *Communications of the ACM* (33:2), February 1990, pp. 126-139.

Batra, D., and Marakas, G. M. "Conceptual Data Modeling in Theory And Practice," *European Journal of Information Systems* (4), 1995, pp. 185-193.

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

Chen, P. P. "The Entity-Relationship Model: Toward a Unified View of Data," *ACM Transactions on Database Systems* (1:1), March 1976, pp. 9-36.

Davis, F. D. "Perceived Usefulness, Perceived Ease-of-Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), September 1989, pp. 319-340.

Elmasri, R., and Navathe, S. B. *Fundamentals Of Database Systems*, 2nd ed., Redwood City, CA: Benjamin/Cummings, 1994.

Embley, D. W.; Kurtz, B. D.; and Woodfield, S. N. *Object-Oriented Systems Analysis—A Model-Driven Approach*, Englewood Cliffs, NJ: Yourdon Press, 1992.

Gemino, A. *To Be or May to Be: An Empirical Comparison of Mandatory and Optional Properties in Conceptual Modeling*, Working Paper, University of British Columbia, 1998.

Goldstein, R. C., and Storey, V. "Some Findings on the Intuitiveness of Entity-Relationship Constructs," in *Entity-Relationship Approach to Database Design*, F. H. Lochovsky (ed.), Amsterdam: Elsevier Science Publishers, B.V., North Holland, 1990, pp. 9-23.

Hitchman, S. "Practitioner Perceptions of the Use of Some Semantic Concepts in the Entity-Relationship Model," *European Journal of Information Systems* (4), 1995, pp. 31-40.

Kim, Y. G., and March, S. T. "Comparing Data Modeling Formalisms," *Communications of the ACM* (38:6), June 1995, pp. 103-115.

Maier, R. "Benefits and Quality of Data Modeling—Results of an Empirical Analysis," in *Proceedings of the Fifteenth International Conference on the Entity-Relationship Approach*, Cottbus, Germany, 7-9 October 1996, pp. 245-260.

Mayer, R. E., and Gallini, J. K. "When is an Illustration Worth a Thousand Words," *Journal of Educational Psychology* (82:4), 1990, pp. 715-726.

Moody, D. L. "Metrics for Evaluating the Quality of Entity Relationship Models," in *Proceedings of the Seventeenth International Conference on Conceptual Modeling*, Singapore, 16-19 November 1998.

Moody, D. L., and Shanks, G. "Improving the Quality of Entity-Relationship Models: An Action Research Programme," *The Australian Computer Journal* (30:4), November 1998, pp. 129-138.

Moore, G., and Benbasat, I. "Development of an Instrument to Measure Perceptions of Adopting an Information Technology Innovation," *Information Systems Research* (2:3), September 1991, pp. 192-222.

Ngwenyama, O., and Lee, A. S. "Communication Richness in Electronic Mail: Critical Social Theory and the Contextuality of Meaning," *MIS Quarterly* (21:2), June 1997, pp. 145-167.

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

Prietula, M. J., and March, S. T. "Form and Substance in Physical Database Design: An Empirical Study," *Information Systems Research* (2:4), December 1991, pp. 287-314.

Rumbaugh, J.; Blaha, M.; Premerlani, W.; Eddy, F.; and Lorensen, W. *Object-Oriented Modeling and Design*, Englewood Cliffs, NJ: Prentice-Hall, 1991.

Wand, Y., and Weber, R. "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars," *Journal of Information Systems* (3:4), October 1993, pp. 217-237.

Weber, R. "Are Attributes Entities? A Study of Database Designers' Memory Structures," *Information Systems Research* (7:2), June 1996, pp. 137-162.

Weber, R. *Ontological Foundations of Information Systems*, Melbourne: Coopers and Lybrand Accounting Research Methodology Monograph No. 4, 1997.