JITTA
JOURNAL OF INFORMATION TECHNOLOGY THEORY AND APPLICATION

COGNITION AND MODELING: FOUNDATIONS FOR RESEARCH AND PRACTICE

CSABA VERES, Norwegian University of Science and Technology
Department of Computer and Information Science, Email: Csaba.Veres@idi.ntnu.no

GITTAN MANSSON, The University of Arizona
Department of Psychology, Email: mansson@email.arizona.edu

ABSTRACT

This paper argues that data modeling for information systems cannot be divorced from human perception, and is therefore marked by the subtle and often unconscious vagaries of cognition. In the absence of a formal semantics for modeling languages, this can result in models that are subjective, ambiguous, and difficult to interpret. Philosophical ontologies that provide a taxonomy of elements in the world have been proposed as a foundation to ground the symbols in various notational systems. Contrary to this view we show that models represent a designer's psychological perception of the world rather than some idealized, philosophical description of that world. A precise ontology of cognitive perceptions is therefore more relevant for the design of diagrammatic notations for use in documenting and unambiguously communicating the analysis of a domain.

INTRODUCTION

This paper reports an experiment which attempts to uncover some deep, universal principles of cognition that can have some impact on the way information systems are perceived and modeled. We focus on analysis models, which represent the current or imagined structure of an information system. These models are supposed to capture a high level, human centered abstraction of the problem domain and are often constructed by teams of analysts consulting with domain experts. As such they must be as transparent as possible to interpretation, not only to unambiguously capture the intended analysis but also to enable interaction and mutual understanding between the various stakeholders. To that end, (Hitchman 2003) argues that an important role for data models is in facilitating interaction between analysts, serving as a structured language for communication about the domain. He points out the importance of considering the linguistic foundations for modeling languages,

Andrew Wenn and Stephen Burgess acted as senior editors for this paper.

Veres, C. and G. Mansson, "Cognition and Modeling: Foundations for Research and Practice", *Journal of Information Technology Theory and Application (JITTA),* 7:1, 2005, 93-104.

in aid of this function. Clearly the modeling process involves a great deal of psychology. It is our contention in this paper that the empirical study of psychological factors is therefore necessary if we are to make generalizations about the way models reflect the "real world" systems that are analyzed and built. We argue that notational systems, or visual modeling languages, should be constructed in a way that complies with our perceptions of reality. This will lead to notations that are more intuitive and less error prone.

It should be noted that our interest in modeling is centered on its use for communication during the analysis phase, encompassing the requirement that the final product is clearly and unambiguously documented for external parties. The model should seem a "natural" representation to analysts and designers alike. We will, however, deny the notion that "the world" imposes rules on the model such that there is a prescribed, correct representation of situations: as most readers will attest from their experience, there are often many different views of any situation, and many equally valid solutions to a problem. Instead we argue that the proper source for modeling constraints is perception. The fundamental structure of all complementary solutions is imposed by underlying psychological representations which give rise to our perceptions and provide the flexibility needed to accommodate multiple world views.

An opposite view is held by the currently dominant approach to the study of the foundations and "proper use" of modeling languages. The received view involves the use of philosophical ontologies (e.g. Wand and Weber, 1995; Shanks, Tansley and Weber, 2003), where ontology is taken to be a branch of philosophy that deals with the description of reality:

> *Ontology, understood as a branch of metaphysics, is the science of being in general, embracing such issues as the nature of existence and the categorical structure of reality. (Honderich 1995, p. 634)*

Note that this differs from the sense of ontology that is typically meant in the research on domain ontologies developed for semantic interoperability in computer science (e.g. Staab and Studer 2004). Such ontologies may be influenced by philosophical analyses, but are typically seen more as an engineering exercise. The premise for using philosophical ontologies as a foundation for analysis is that modeling notations should be designed in a way that establishes some form of isomorphism between the notation and the ontologically analyzed structure of the world. The argument goes something like this: when we do conceptual modeling we do so because we want to capture some significant portions of a domain of interest. But the domain of interest has some fixed, invariant properties by virtue of being part of the "world at large", and we already know a great deal about the "world at large" through the hard work of philosophical ontological analysis. *Ipso facto*,

CONTRIBUTION

This paper contributes to the field of IS in terms of both methodology and practical application. Concerning methodology we show that traditional and rigorous experimental techniques from cognitive psychology can be applied in information systems research. We argue that, since cognition underlies all complex behavior, traditional methods of cognitive research should be employed in developing best practices for IS. We demonstrate that such a rigorous psychological approach can benefit the design of modeling languages.

On the practical side we consider the modeling of ternary relationships, and show psychologically based difficulties in treating all such relationships as formally equivalent. Such equivalence is assumed by current modeling languages as well as prescriptive philosophical models. We propose a new notation for handling at least some scenarios.

Our findings will be interesting to researchers interested in domain modeling and formal foundations for information systems. It will also be interesting to practitioners interested in the way that sub conscious cognitive factors can affect the way situations are perceived and modeled. The methods introduced here open a new area of inquiry into the way that minds, models, and the world interact in information systems analysis.

the ontological analysis should inform the practice of conceptual modeling. Thus, modeling notation will more accurately reflect real world domains, and be more unambiguously interpreted if it is constructed such that it conforms to the structure of the world *as seen through the philosophical ontology*. In addition, it is claimed that an ontological alignment between the world and the model can be used to selected modeling languages that are maximally suitable for a given domain and task, as well as help in validating models with stakeholders (Shanks, Tansley and Weber 2003).

However, (Veres and Hitchman 2002) challenged this general approach. They argued that claims about the intuitiveness of modeling languages constructed by such principles were based on a hidden assumption about psychology. That is, the assumption behind the claim that an “ontologically correct” modeling language will be intuitive is that people's cognition is also “ontologically correct” at some level: if our models represent reality the way it “really is”, then people will find this a natural representation. But this is surely an empirical claim, and should not appear as an assumption. It must be shown that the formal prescriptions that are derived as hypotheses from philosophical analyses of “the world” do in fact help make sense of the fragments of the world involved in modeling scenarios. It is precisely our claim that in general this is not the case and a psychological analysis will do better. Of course, there will be cases in which both analyses agree, and these are uninteresting for purposes of hypothesis testing. For example, it would be odd if some analysis of the world concluded that there were no objects in the world. Our psychological theory therefore serves a secondary role in providing an alternative hypothesis that can be used in deciding the relevant test cases for deciding on the usefulness of the philosophical and psychological models.

Before we can devise a test case we consider a concrete example of the sort of prescriptions the philosophical stance makes, concerning an entity relationship (ER) model (Chen. 1976). The basic ER model uses a relatively simple set of constructs, which enable designers to specify the structure of databases. The primary goal is to capture information about significant elements in a domain, and the relationships between the elements. But despite this relative simplicity, (Wand, Storey, and Weber 1999) stipulate that there is an inherent confusion about the way the relationship construct is used in practice. For instance modelers are often unclear about whether a concept should be represented by an entity or a relationship between entities. For instance if we want to model the fact that students are registered at Universities, we have at least two choices. We could create separate entities for “student” and “university”, and a “registered_at” (or some similar suitable label) relationship between the two. Alternatively, one could draw “registration” as an entity in its own right, such that “university” and “student” both have a relationship with “registration”. Such ambiguities, it is claimed, can result in models that fail to accurately represent the domain and can lead to errors in the interpretation of the model.

Wand, Storey, and Weber (1999) argue that the proper way to resolve these difficulties is by restricting the meaning of notational elements with the use of formal ontologies. Specifically, they argue that an ontology based on the work of Mario Bunge (e.g. Bunge, 1977) reflects the structure of "the world" with sufficient accuracy to be useful in defining the use of modeling constructs. They devise a number of rules (prescriptions and proscriptions) that modelers should abide by, as well as propose a number of new notations that reflect the structure of the world with more accuracy than currently available constructs. They argue that concepts like *registration*, and *assignment* are "really” *mutual properties* that hold between the two primary actors in the relationship (e.g. a University and a Student both have the mutual property *registration* which depends on a relationship between the two). The prescription is that mutual properties have to be modeled as relationships, since entities should only be used to model *things* and not *mutual properties*. This avoids construct overload (where the one modeling symbol represents both things and mutual properties) and therefore confusion.

On the extreme reading the criticism is quite severe: “In short, the theory underlying the nature of and representation of relationships in conceptual modeling is unclear and in our view, problems arise with relationships in conceptual modeling because their nature and underlying meaning are unclear". (Wand, Storey, and Weber 1999, pp. 495-496). In other words, the meaning of the most fundamental constructs in the ER model are ill defined. But (Hitchman 2003) points out that practitioners have little trouble conceiving of many concepts as either entities or relationships during a modeling session. They do not appear confused by the choice, and in fact, use it to arrive at the most parsimonious and cost effective solution. The criticism therefore seems more realistic if construed instead as a problem of *consistency* in which different modelers represent the same concept through a different choice of notation (we are indebted to an anonymous reviewer for this point). A well defined semantics is therefore necessary insofar as it can improve consistency.

Up to this point we have argued against the received view, noting that practice seems to militate against theory on the point of fixing the semantics of entities according to (a specific case of) philosophical ontology. But we also suggested that the human cognitive architecture imposes fundamental constraints on our representations. Surely this ought to result in *some* prescriptions on the representation of entities, based on psychology? In order to answer this question it is necessary to delve a little into the world of psychology and linguistics.

The Role of Psychological Ontology

The present aim is to investigate the way humans perceive abstract entities and their relationships, and to show that these perceptions can guide modeling decisions. On the one hand this goal sounds ridiculously simple minded. Of course we model perceptions. What else would we model? We are grateful to an anonymous reviewer for pointing out that the important role of cognition is widely accepted in the IS community, and does not need to be defended. But, while we are relieved by this observation, we note that wholesale claims about the importance of "cognition" is only half the answer. The real questions about the precise character of that perception are yet to be tackled. No one doubts the importance of cognition, but fewer are in a position to make substantive claims about the precise way in which cognition and modeling interact. (Wand, Storey, and Weber 1999), for example would not deny that we model perceptions of reality, but they would then suggest (without direct psychological evidence we might add) that those perceptions are in some sense underdetermined, vague, and sometimes confused. This is why we need some external guidance, on their account. But instead of relegating perceptions to the status of important but unexplored citizens, in this paper we try to provide a psychologically motivated ontology that describes precisely the way in which perceptions of abstract entities are formed. This will then tell us why ambiguities might arise, what the boundaries of that indeterminacy are, and how to best cope with it. We believe that legislating a solution to uncertainty with recourse to philosophical analysis is not helpful if we want to understand and assist human interpretation and communication. It is not enough to say that our perceptions are confused and fix them with some external frame of reference. We need to know *why* they are confused, and see if we can "fix" the problem without such external assistance.

The linguist Ray Jackendoff (Jackendoff 1983,1990,1997) has taken observations about the correspondence between language, vision and other modalities, and developed a detailed formal description of conceptual structure that can be used to account for a large number of observed facts concerning the syntax to semantics mapping in natural language. This allows us to describe an ontology of conceptual structure that we can subsequently use to make predictions about modeling practice.

Jackendoff (1983) argues that properties of conceptual structures can be conjectured from their interactions with linguistic (syntactic) structures. As a result we can use the large number of existing empirical observations from syntactic theory to generate hypotheses concerning the nature of conceptual structures. In other words, the highly influential and well articulated theoretical constructs from syntactic theory in the Chomskian tradition can be brought to bear on the analysis of conceptual structures. An obvious but important observation is that the conceptual structures themselves are not language dependent, and they are not determined by language use. Quite the contrary; language evolved as a way to externalize the pre existing content of cognition for the purpose of communication (Pinker and Bloom 1990). Thus we gain insight into concepts by observing the correspondences between structural properties of language and conceptual structure. This is not to claim that the two are equivalent; clearly language has properties that would make no sense in the conceptual domain, and vice versa.

Jackendoff stipulates that conceptual structures can be described by a small number of basic ontological categories which form the *major conceptual categories.* These are the semantic "parts of speech" that parallel the status of verbs, nouns, etc. in syntax. We can consequently specify a number of mapping rules between syntactic and conceptual constituents, which clarifies both the relationship between linguistic and conceptual structures, and the nature of those structures themselves.

Jackendoff (1983, 1990) presents linguistic evidence for the identity of the conceptual constituents, and suggests some relatively “standard” categories like [STATE] and [EVENT], but also some novel ones like [PATH] and [AMOUNT]. He further notices that many of the rules and transformations of conceptual structure can be stated independently of the identity of the constituent. Thus, he observes the standard practice of a [PROPERTY] modifying an [OBJECT], (e.g. red dog), but then notes that a [PROPERTY] can modify an [EVENT] in the same way (e.g. long war), and that a [PROPERTY] can even modify a [PATH] (e.g. a long way past my house). The importance of this is that, while no conceptual constituent can be reduced to another, they can be treated as equivalent for certain operations. To see how this helps our understanding of modeling, consider the previous example concerning university registrations. On the “registration” interpretation (where registration is an entity), the model describes a unique [EVENT] or [ACTION] that has taken place, involving a university and a student. This [EVENT] is modified by additional [PROPERTY] concepts like “date”, “office_location” and so on. The end result of the registration is modeled as an entity with relevant attributes. The model which contains the “registered_at” relationship between a “university” entity and a “student” entity describes a [STATE] or perhaps [PROPERTY]. Two observations are relevant. First, the choice of representation in the model can be explained by the presence of the simple ambiguity in the reading of the term “registration”. Both readings are possible because conceptual structures can treat both readings on equal terms. The second observation concerns concept overload in the ER notation. There is a many to one mapping from conceptual structures to ER notation: [OBJECT], [EVENT], etc., are all mapped to an entity. This is the same point made by (Shanks, Tansley and Weber 2003), who point out the problem of *construct overload* where, for instance “... a thing and an event both have to be represented by the same grammatical construct (such as an entity symbol)” p.87. But the psychological explanation shows that some conflations are not necessarily harmful if they are based on the content free rules of conceptual structure. But to the extent that this is a problem, our argument stands, that conflation involving *universally perceived psychological* perceptions of the world is more noteworthy of attention than conflation with respect to a prescriptive analysis of “the world”. A topic for further research is to identify which categories are conflated in which ER expression, and to see if this causes difficulties in given situations?

We will not delve further into the technical details of Jackendoff's theory in this paper. Instead in what follows we use the theory of syntax-semantics correspondences to derive a hypothesis about the representation of certain situations.

Consider the linguistic (syntactic) distinction between *complements* and *adjuncts*. Basically, they both involve the way information might be added to a sentence. Consider the following examples:

1) Sarah robs the 7-11 in New York.

2) Adam robs in Washington in February.

The basic premise in both sentences is that someone commits a robbery. We add information about what, and where these robberies took place. But if we construct a syntactic tree for these sentences we will see that *robs* and *7-11* in sentence 1. are more closely linked than are *robs* and *Washington* in sentence 2. This distinction has certain consequences for the ways in which the sentences can be transformed, as the following examples show (where * denotes an ungrammatical string, and the addition to the original sentence is shown in `courier`).

3) Sarah robs the 7-11 in New York `everyday`.

4) *Sarah robs `every day` the 7-11 in New York.

5) Sarah robs/`*eats`/`*marries`/`*sleeps` the 7-11 in New York.

6) Adam robs in Washington in February `when it is warm`.

7) Adam robs `when it is warm` in Washington in February.

8) Adam robs/`eats`/`marries`/`sleeps` in Washington.

Sentences 3. and 6. simply add an additional phrase to the sentence, with no ensuing difficulties. But look what happens when we try to move the newly introduced phrase in sentences 4. and 7. The move seems fine in 7., but not in 4 which now becomes ungrammatical. The phrasal structure of the sentence won't allow the movement, because the phrase *the 7-11* is too closely bonded to the verb. We say that *7-11* is a *complement* of the verb *robs*, whereas *in Washington* is a less closely linked *adjunct*. Sentences 5. and 8. further show that a verb is quite selective in the complements it will take, but insensitive to its adjuncts.

But how is this distinction reflected in conceptual structure? Following a lexical/conceptual correspondence rule from (Jackendoff 1990) it can be shown that the complement is directly linked to the verb in conceptual structure. On the other hand adjuncts are mapped into some lower position in the recursive definition. The end result is that grammatical subjects and the verbal complement are more closely related in conceptual structure than are the subject and adjuncts, or for that matter the complements and adjuncts. By hypothesis, there are syntactic and conceptual differences between the entities involved in sentences 1. and 2.

So how does this let us make predictions about conceptual modeling? Suppose we were designing a database (for a police application perhaps) that recorded the incidence of robberies, and contained information such as that expressed in sentences 1 and 2. Taken together, the sentences contain information about people, locations, establishments, and months of the year (time periods). For students of introductory database modeling (we will have more to say about choice of subjects later), this would constitute the distinct entities: *Person, Location, Establishment, Time_period*. Each entity would have the capacity to store the provided information (name) for each entity, as well as any additional information. *Person* could have social security number, last known address, etc. *Location* could have city name, country of location, postal/zip code, population, etc. *Establishment* could have name, type, opening hours, etc. *Time_period* could have month, year, day, list of holidays, etc. If we considered the relationships in terms of a simple conceptual model, the situations in 1) and 2) both describe a ternary relationship between three entities: *person-establishment-location* or *person-location-time_period*. And, as (Wand, Storey, and Weber 1999) suggest, all n-ary relationships should be modeled alike. However, we have shown that a conceptual analysis shows that the two scenarios have subtly different underlying representations. Are the two n-ary relationships not alike, after all? The following experiment was designed to test if the previously discussed conceptual difference has any impact on concept modeling, by introducing a modeling symbol that provides an opportunity to model some relationships more faithfully according to the psychological ontology. We investigate if people tend to prefer the conceptually faithful representation in the appropriate situations.

Psychology Experiments and the Real World

In what follows, we will propose a new notational symbol and contrast it with UML to see how it is used to model scenarios involving ternary relationships. We aim to show that the Bunge-Wand-Webber approach (e.g. Wand and Webber, 1989; 1990; 1993; 1995), which has received considerable coverage in the literature, makes the wrong prediction when applying their guidelines. This calls into question their theoretical claims, on empirical grounds.

Our subjects were all undergraduate students studying systems analysis and design. The use of students for modeling experiments can be questioned (e.g. Hitchman, 2003 and references therein) on the grounds that it is unrealistic for practitioner related problems. There is also a related feeling that experiments in this field should employ tasks that mirror as closely as possible those that users are required to perform in practice. On this general view the experiment we will describe in this paper seems somewhat strange and unnatural.

In our defense we remind the reader that our aim is to employ a well controlled experimental task as a way to isolate psychologically salient components of the analysis process. We are interested in probing unconscious cognitive variables which we claim have a strong and general consequence on the mental behavior of students and practitioners alike. Tasks which are performed as part of one's everyday duties may be unsuitable for this purpose because they are likely to be contaminated by conscious strategy and the idiosyncrasies that come from long term individual experience. Practitioners are therefore likely to have learned strategies that obscure the very facts we wish to learn about. Ironically, a “realistic task” might be the worst one to use for the purposes of discovering the underlying facts about modeling. But what sort of task could probe such hidden variables? A great deal of the creative skill in cognitive psychology involves the invention of tasks that manage to tap the hidden processes responsible for behavior. As an example, some of the most illuminating paradigms in cognitive psychology involve tasks that give no subjective clues whatever to questions under investigation. (Swinney, Onifer, Prather, and Hirshkowitz 1979) asked people to simply listen to sentences and to make a lexical decision (word / nonword decision) to a visually presented stimulus (word or pseudoword, e.g. *help* or *selp*) at some random point in the sentence. Unbeknownst to his subjects, they were actually involved in a cross modal priming experiment that used lexical decision latencies to investigate the activation of multiple senses of ambiguous words. The subtle connection between the auditory and the visual stimulus was completely unknown to the subjects. Going one step further, (Forster and Davis 1984) used *masked primes* that were completely unavailable to conscious perception in their experiments on word recognition. They argued that subjects should have no conscious awareness of the prime, and showed the important consequences of ensuring that conscious, strategic factors were eliminated in task performance. Experimental performance was thereby controlled by a stimulus that was completely unavailable for conscious report. In the current experiment we also hope to minimize conscious strategic factors by using a task that is not obviously related to the purpose of the hypothesis under investigation, and subjects who are not affected by years of experience with a particular modeling tradition.

EXPERIMENT DESIGN AND PROCEDURE

The design of the experiment consisted of two sets of sentences: one set included a verbal complement and an adjunct, while the other set contained two adjuncts. By hypothesis the subject and the complement in the first set form a close conceptual relationship, with the remaining adjunct being distantly related to both. The second set contained only the distantly related adjuncts. There were ten verbs tested in each condition. As the example sentences 1 and 2 show, some verbs can take complements optionally; thus many verbs appeared in both conditions. We thought this was a favorable circumstance since it diminished the variability between the two sets in terms of the actual verbs used. Most items were of this sort.

For each of the chosen verbs, six specific examples and one generic description were constructed. For example, the generic sentences for 1. and 2. above were:

9) Person robs establishment in city.

10) Person robs location at time.

The materials were counterbalanced, so each particular example appeared with the old notation for some participants and with the new notation for others. The results for this manipulation were combined in the final analysis. We collected data from 23 participants, who were required to perform a modeling task on the generic sentences. Specific examples of the generic sentence were provided for clarification.

The modeling task itself introduced the new notation, and asked participants to choose between a standard and the new notation for each example. The new notation was designed to reflect the conceptual relationships between entities as explicitly as possible. Figure 1. shows an example of the UML and new notation for a complement - adjunct sentence.

Note that the new notation highlights the close relationship between *Person* and *Establishment*, and shows that these are further modified by the adjunct, *City*.

In each example, participants were shown a generic sentence, six specific example sentences, and were given a choice of the two notations. They were simply asked to choose the model they preferred for the situations described. From a theoretical point of view, the new notation reflects the conceptual relations for sentences with an initial complement, and so it should be preferred in this condition only.

Once the participants had finished their task they were given grammaticality tests on sentences of the type shown in 3 – 8 above. This was to confirm the experimenters' intuitions about the grammatical status of the examples, as well as to ensure that the judgments were shared by the participants themselves. If the participants did not agree with our judgments then it is possible that the conditions were not as well dichotomized as hoped (e.g. some complements may have been adjuncts).

Results

The results were straightforward. Figure 2. shows the total number of preferences for the four conditions. The first two bars show the preferences for sentences in which both modifiers were adjuncts, while the second two show the condition where the first modifier was a complement. The first bar in each pair shows the preference for the new notation while the second bar shows the preference for the standard UML type notation. A chi-square test for independence revealed that the sentence type and notation were not independent in determining preference. There was a disproportionate preference for the new notation for complement-adjunct sentences: Chi-square = 10.23, $p < 0.01$. As expected, people preferred the conceptually faithful notation for sentences where the first modifier was a complement. By hypothesis, this is the situation which is accurately reflected in the new notation.

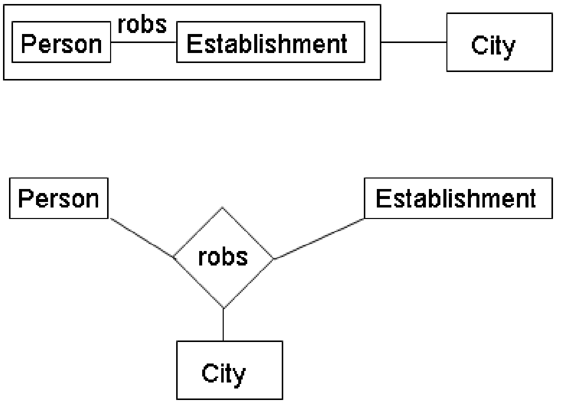

Figure 1. Conceptual notation for ternary relationship using the new notation (on top) and a UML diagram.. Establishment is the complement and city is the adjunct in the example.

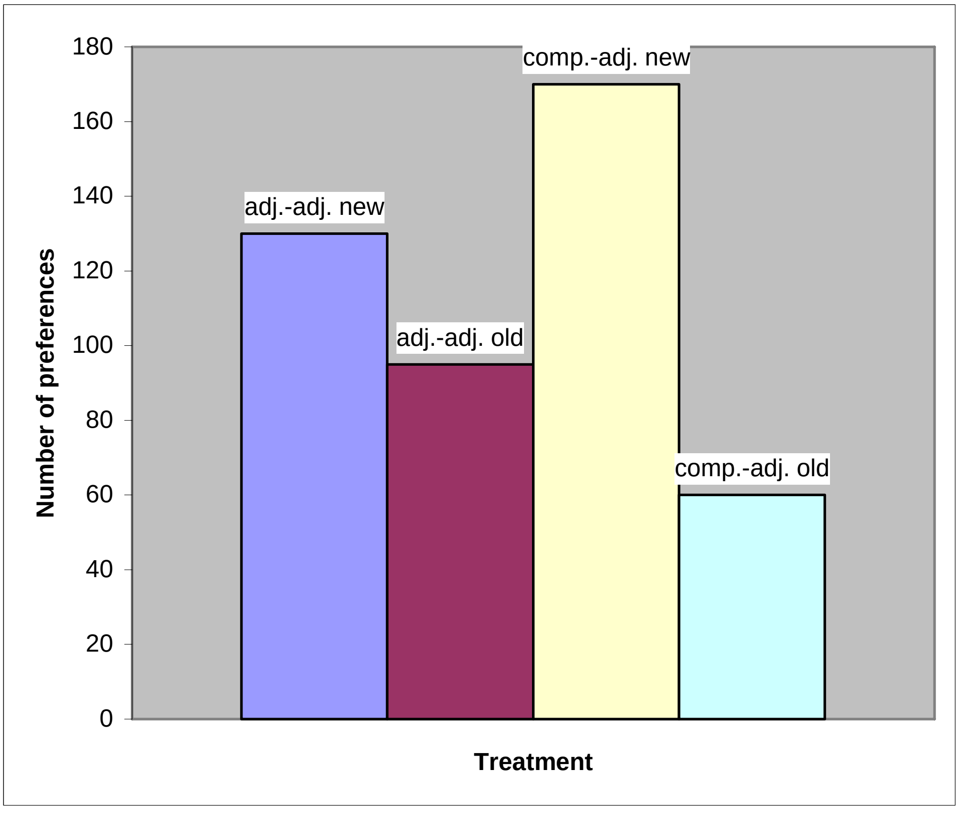

Figure 2. Total number of preferences for each sentence type.

While the results are quite clear, it is possible that the magnitude of the interaction was in fact obscured in this particular study if subjects' intuitions about the grammatical status of the arguments differed from our own. In an additional analysis we correlated the subjects' mean agreement about the grammaticality of the test sentences with the difference in preference, for each sentence, between the two notations. We found a significant correlation for that sub set of sentences that contained two adjuncts, Pearson's R=-0.46, $P < 0.05$. The lower the subjects' agreement (i.e. the less clear the status of the adjuncts), the bigger the difference in preference. In other words, the slight numerical preference for the new notation in the adjunct-adjunct condition may be because some sentences are judged to be complement-adjunct. This difference might disappear altogether if the sentences are selected with maximal subject agreement in mind. This would make the result even more dramatic.

Discussion

The experiment shows that a subtle grammatical difference in sentence structure can influence the preference for the way a scenario is modeled. But our argument is that it is not the language itself that determines modeling preference but the underlying conceptual structures, demonstrating that people do look at ternary relationships in different ways, contrary to the (Wand, Storey and Weber 1999) prescription. This is taken as evidence against the particular philosophical ontology.

This finding is important since it supports our view that cognitive factors, as seen through language structure, determine subtle aspects of modeling. This view complements Hitchman's observations that an important role of conceptual models is in facilitating interaction through discussion about the problem and its solution. The model itself becomes a language that is used in discussing the problem. (Hitchman 2003) argues that the role of a model as an independent conceptual model is less important than its role in interaction: "*Practitioners talk with the notation as well as using the notation to draw a diagram. The entity-relationship model constrains the social interaction because the model provides a way of talking about design. The practitioners use the model to talk about a normalized relational data structure in a way that undermines the idea of the entity-relationship model as an independent conceptual model*." But if this is true, then it is doubly important that the modeling tool works "with" and not "against" the practitioners' conceptual understanding. A notation based on a formal ontology of the world would require the practitioners to constrain their dialogue according to the concepts in that ontology. A notation based on conceptual structures would liberate the dialogue and allow the expression of the practitioners' conceptual understanding. The suggestion we are making is that "conceptual modeling" languages should *really be* conceptual modeling languages in that their design should respect the properties of the cognitive architecture of their users.

In light of the importance of the finding, we address several criticisms that might be leveled at the reported experiment. First, the choice of task might be called into question. We have already mentioned our motivation for using a task that did not obviously reveal the hypothesis under investigation. By asking subjects simply for their preference, we did not draw attention to the fact that we were really interested in the way the models related to the meanings expressed in the sentences. Nevertheless, a further criticism may be that subjects were not given sufficient direction, and therefore the precise research question was not really addressed. We concede that there is no way, in general, to know with certainty how people solve a problem they are given. But that is why we do cognitive science: we find tasks where reliable differences exist between conditions, and hypothesize an explanation for the behavioral difference. The hypothesis is then tested. Here, we have shown such a difference, and provided such an explanation. Any worthwhile criticism of the task owes us an alternative explanation for the significant interaction. As an additional defense, an open ended task such as this stands the further risk of obscuring effects that might otherwise be detected. The fact that we found a significant result is therefore doubly worthy of note.

A further criticism might involve the choice of items. It might, for instance, be proposed that the status of certain entities in the scenarios might be different from others. For instance *time_period* might be better thought of as an attribute than an entity, so sentences involving a *time_period* should not be modeled with three entities at all. But there are two reasons why this criticism of the experiment is not valid. First, none of the subjects complained on this point at post experiment debriefing sessions which indicated that the scenarios posed no special difficulty. Second and more importantly, the experimental result itself mitigates the critique. It must be remembered that every sentence is presented with a choice of "new" or "old" notation, and subjects are asked to indicate their preference. If there was a fundamental problem with any of the scenarios then this should affect both models. Yet again, the significant interaction does emerge. The question then becomes, if neither model makes sense, why do people tend to prefer one model over the other?

It might finally be noted that (Kent 1978) points out that it is possible to conceptualize ternary relationships as two binary relationships. Thus, a relationship between ABC can be seen as A(BC) or (AB)C. This is similar to our finding, suggesting the possibility that the notation simply models this observation. But, he goes on to remark, as a true "formalist" would, that generally there is nothing to decide between the groupings on formal grounds and each is as sensible as the other. But of course the hypotheses presented in this paper give strong reasons to prefer one over the other. Psychology provides principles missing in logic alone. To follow this up, we plan to run the experiment again, but re-arranging the grouping. According to the psychological hypothesis, people should not prefer the new notation if it groups concepts contrary to conceptual structures.

In summary, the work has demonstrated that real world modeling practices can be informed by a deep understanding of cognitive facts. Initially this forces us to consider the way we think about ternary relationships. Future work will involve finding more support for a psychological rather than philosophical basis for modeling languages. Given continued success, we will be able to use the lessons learned in an attempt to find a proper grounding for IS modeling languages. One area of particular interest is in the modeling of workflows in knowledge intensive enterprises. Current approaches are limited in their lack of flexibility and ability to incorporate the various social and cognitive facts that are involved in such tasks (Joergensen, 2004). But once again, (Jackendoff 1992, 1993, 1999) offer detailed analyses of the conceptual structures that underpin behaviors involved with transfer of property, rights and obligations, and other complex social and behavioral activities. We are currently investigating the use of these analyses as a way to inform modeling in the larger socio-cognitive setting. This will bring the relevance of the findings into a significantly larger arena.

Acknowledgments. This work was partly sponsored by the Norwegian Research Council through the WISEMOD project, number 160126V30 in the IKT-2010 program.

REFERENCES

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

Chen, P. “The Entity-Relationship Model Toward a Unified View of Data.” ACM Transactions on Database Systems, 1976, 1:1, 9-36.

Forster, K.I., & Davis, C. “Repetition priming and frequency attenuation in lexical access.” Journal of Experimental Psychology: Learning, Memory, and Cognition, 1984, 10, 680-698.

Hitchman, S. “An interpretive study of how practitioners use entity-relationship modeling in a ternary relationship situation” Communications of the Association for Information Systems , 2003, 11, 451-485

Jackendoff, R. “The Natural Logic of Rights and Obligations”, in R. Jackendoff, P. Bloom, and K. Wynn, (eds.)., Language, Logic, and Concepts: Essays in Memory of John Macnamara, MIT Press/Bradford books, Cambridge, MA, 1999, 67-95.

Jackendoff, R The Architecture of the language faculty, Cambridge, Mass. MIT Press, 1997.

Jackendoff, R. Patterns in the Mind: Language and Human Nature, Harvester-Wheatsheaf, 1993

Jackendoff, R. Languages of the Mind. Cambridge, Mass., MIT Press, 1992.

Jackendoff, R. Semantic Structures, Cambridge, Mass., MIT Press, 1990.

Jackendoff, R. Semantics and Cognition, Cambridge, Mass., MIT Press, 1983.

Joergensen, H. D. Interactive Process Models. Doctoral Dissertation, Norwegian University of Science and Technology (NTNU), Trondheim, 2004

Kent, W.K. Data and Reality. North Holland, 1978

Pinker, S. & Bloom, P. “Natural language and natural selection.” Behavioral and Brain Sciences, 1990, 13:4, 707-784.

Shanks, G., Tansley, E., and Weber, R. “Using ontology to validate conceptual models”. Communications of the ACM, 2003, 46:10, 85-89

Staab, S. & Studer , R. Handbook on Ontologies. Springer, Berlin, Germany. 2004.

Swinney, D.A., Onifer, W., Prather, P., & Hirshkowitz, M. “Semantic facilitation across sensory modalities in the processing of individual words and sentences.” Memory and Cognition, 1879, 7, 159-165.

Veres, C. and Hitchman, S. “Using Psychology to Understand Conceptual Modeling.” In Wrycza, S. (Ed.) Proceedings of the Xth European Conference on Information Systems, Gdansk, Poland, 2002, 473 – 481.

Wand, Y., Storey, V.C., and Weber, R. "An Ontological Analysis of the Relationship Construct in Conceptual Modeling." ACM Transactions on Database Systems, 1999, 24:4, 495-528.

Wand, Y., & Weber, R. "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars." Journal of Information Systems, 1999, 3, 217–237.

Wand, Y. and Weber, R. "On the Deep Structure of Information Systems." *Information Systems Journal,* 1995, 3, 203-223.

Wand, Y., & Weber, R. "An Ontological Model of an Information System." IEEE Transactions on Software Engineering, 1990, 16:11, 1282–1292.

Wand, Y., & Weber, R. "An Ontological Evaluation of Systems Analysis and Design Methods." In E. D. Falkenberg & P. Lindgreen (Eds.), Information Systems Concepts: An In-depth Analysis Elsevier Science Publishers, Amsterdam, 1989, 79–107.

AUTHORS

Csaba Veres works at the Norwegian University of Science and Technology in the department of Computer and Information Science, where he has been for the past two years. Before that he worked at the Department of Information Systems, Melbourne University, Australia, and the Defense Science and Technology Organization in Salisbury, also in Australia. His Ph.D. was awarded at the University of Arizona where he studied in the Department of Psychology. He has published both as a cognitive psychologist and a researcher in the information sciences.

Gittan Mansson is currently studying for a Ph.D. degree in the Department of Psychology at the University of Arizona. Prior to that, she worked as a researcher at the Department of Information Systems, at Melbourne University. She obtained her bachelors degree at California State University, Sacramento. She also received a masters degree at CSUS, specializing in visual cognition.