Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2004 Proceedings

Americas Conference on Information Systems (AMCIS)

December 2004

Understanding Representation Fidelity: Guidelines for Experimental Evaluation of Conceptual Modeling Techniques

B. Jeffrey Parsons
Memorial University of Newfoundland

Linda Cole
Memorial University of Newfoundland

Understanding Representation Fidelity: Guidelines for Experimental Evaluation of Conceptual Modeling Techniques

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

Linda Cole
Faculty of Business Administration
Memorial University of Newfoundland
St. John's, NL A1B 3X5
CANADA
Linda.Cole@xwave.com

ABSTRACT

Recently, there has been a resurgence of interest in experimental research on conceptual modeling in information systems analysis and design. There is a need to explicitly identify the objectives of specific experiments in this area, and the role that assumptions play in experimental design. We provide four guidelines for developing materials for experiments aimed at evaluating conceptual modeling techniques, based on the premise that the primary purpose of conceptual modeling is to facilitate communication between analysts and users in validating domain knowledge relevant to an information system. We offer the guidelines as recommendations to assist the development of experiment materials that support meaningful tests of domain semantics, and present empirical evidence to illustrate the value of two of the guidelines. We also evaluate the degree to which a selection of recent experiments on conceptual modeling adheres to the guidelines, and consider implications of that assessment.

Keywords

Conceptual Modeling, Experimental Design, UML

INTRODUCTION

Hundreds of modeling methods and techniques for systems analysis and design have been proposed (e.g., Oei et al. 1992), although few have seen widespread use. One criticism of modeling techniques is that they lack strong theoretical foundations (Wand and Weber 1993, 2002). As a result, it is difficult to understand if and why methods are useful in supporting specific objectives. This challenge has generated a recent resurgence of interest in research aimed at evaluating techniques, particularly those involving diagrammatic notations that convey semantics of a real world domain.

Wand and Weber (2002) proposed a framework for evaluating aspects of conceptual modeling approaches. Their framework clearly delineates four targets of evaluation: grammars, methods, scripts, and context. The Wand and Weber framework provides valuable guidance on the *general structure* of experimental (and other) studies to evaluate conceptual modeling methods. This paper provides a complementary set of guidelines for developing experimental materials and measures appropriate for research questions that involve the suitability of conceptual modeling techniques for expressing domain semantics.

MOTIVATION

Throughout this discussion, attention is limited to the use of methods for conceptual modeling (capturing semantics of the real world relevant to an information system). We adopt the position of Wand and Weber (1993) that conceptual models represent aspects of the real world, as perceived by humans. Wand and Weber use this premise to argue for the suitability of ontology (the branch of philosophy dealing with what exists), and in particular the ontology of Mario Bunge (1977), as a basis for identifying what needs to be modeled to achieve ontological fidelity in a method. Using that approach, ontological deficits and overloads can be isolated and predictions made about the empirical impacts of these problems. Several experimental studies (e.g., Burton-Jones and Weber 1999, Burton-Jones and Weber 2003, Gemino 1999) have used this foundation to make and test predictions about methods.

The premise adopted by Wand and Weber can be used to explore further the rationale for constructing a model. Various researchers have identified that one of the major purposes of conceptual models is to facilitate communication between clients (or users) and analysts (or developers) (e.g., Kung and Solvberg 1986, Gemino and Wand 2003, Parsons 2003). Analysts construct a (typically graphical) conceptual model to document their understanding of a domain as developed through meetings with clients, examination of documents, and other activities. The model allows users to verify whether or not analyst interpretations reflect reality as perceived by users. From this perspective, research to evaluate modeling methods should focus on their capacity to facilitate this verification. Options for evaluation include: assessing the ability of developers to construct models that capture requirements ('write' tasks), and assessing the ability of readers of models to extract information contained in them ('read' tasks) (Wand and Weber 2002). We focus here on laboratory experiments that involve 'read' tasks.

In addition, research can be classified according to whether comparisons are between different modeling formalisms (intergrammar and multi-grammar) or use different approaches to developing scripts (typically guided by some theoretical basis for alternate forms of representation) within a single formalism (intragrammar) (Wand and Weber 2002). Notably, recent research on ontological analysis of modeling methods has investigated, using an intragrammar approach, whether and how adherence to ontological principles affects the ability of a model to communicate information (e.g., Burton-Jones and Weber 1999, Gemino 1999, Burton-Jones and Meso 2002). In the following, we focus on issues in intragrammar comparisons using alternate scripts (diagrams) motivated by different theoretical considerations.

GUIDELINES

The guidelines consist of four criteria that should be addressed when making choices about key parameters of experimental design – independent variables, dependent measures, participants, and procedures – for experiments comparing the effectiveness of alternate intragrammar scripts in communicating domain semantics. We justify the guidelines, but do not claim that the list is exhaustive.

First, alternative scripts generated to manipulate one or more **independent variables** (e.g., conformance or non-conformance to some ontological guidelines) *should be informationally equivalent* with respect to the dependent variable(s) (Gemino and Wand 2003). That is, it should be possible to answer questions correctly with any of the representational forms used as treatments in an experimental study. Otherwise, internal validity is threatened, since differences in information content may confound attempts to measure differences in the capacity of alternate semantically equivalent representations to support verification. To illustrate, if one form provides enough information to answer selected questions correctly, while a second form does not, it would not be surprising to find that participants receiving the first form outperform those receiving the second form on those questions. Put simply, if one set of materials contains the answers to certain questions while a second set does not, we can only expect the former group to perform better than (or no worse than) the latter on those questions. Information equivalence between diagrams need not be total, but must apply with respect to questions asked related to the dependent variables.

Second, **dependent variables** *should measure performance only with respect to semantics contained in the script(s).* Therefore, questions should be limited to those that test understanding of explicit semantics expressed in a script. Other kinds of questions, such as the problem solving questions used by Gemino (1999) and Bodart et al. (2001), necessitate prior background knowledge to answer meaningfully or correctly, and therefore address issues other than communicating/validating domain semantics. In the next section, we present an example that illustrates this guideline and its implications in more detail.

Third, when selecting **participants**, *subject matter experts (SMEs) should not be used*. This recommendation sacrifices external validity, since a central role of conceptual models to facilitate communication between analysts and users (SMEs). However, given the research objective of assessing the degree to which (and ease with which) information can be extracted from a script, it is critical that participants can answer questions by using only that script, rather than by using background knowledge. Using SMEs to test the capacity of scripts to convey semantics can clearly confound experimental results and thereby threaten internal validity. To illustrate, if we construct a script related to the enrolment of students in courses at a university and conduct an experiment with student subjects (SMEs in that setting), there is no way of distinguishing whether answers are based on information provided in the scripts, or simply on the participants' experience at their own university. This can improperly weaken or strengthen the effects of the treatment, depending on whether or not the participants' knowledge is consistent with the semantics in the script.

Clearly, tests involving SMEs are needed in order to provide the external validity needed to fully understand issues in using conceptual modeling techniques to communicate and verify requirements. However, such tests are valuable only after a reasonably comprehensive and controlled understanding of how different representations facilitate or inhibit the expression of semantics has been developed. We contend that such understanding does not yet exist, and that the design of some prior studies may have inhibited the development of cumulative knowledge in this area.

Fourth, in **experimental procedures** for studies testing the representation capacity of alternative forms of representation, *scripts should be available to participants as they answer questions*. Given that we are interested in testing the ability of the diagrams to convey semantics (rather than, say, ease of memorizing diagram content), there is no reason to remove diagrams in experiments. Moreover, since scripts would be available in a real-world communication/validation situation, there is some external validity to this protocol. Some have advocated removing diagrams on the basis that an objective of studies may be to test 'learning' (Bodart et al 2001, Gemino and Wand 2003). However, that objective is difficult to reconcile with the role of conceptual models in facilitating communication with users/clients who know the domain, and would not learn from reviewing diagrams. Testing learning may lead to important future uses for conceptual models, but we believe our understanding of how methods convey domain semantics remains sketchy to the point that considerable work is needed in understanding more basic issues first.

APPLYING THE GUIDELINES

In this section, we illustrate how failing to adhere to the first three of the four guidelines in Section 2 can interfere with the results obtained in experimental studies.

Figure 1 contains two alternative UML class diagram segments illustrating possible ways of depicting associations or relationships. Figure 1(a) attaches an association class (with attributes) to an association between two entity classes. Figure 1(b) depicts two unrelated associations between the same classes. These representations are not informationally equivalent, according to the first of our four guidelines. Thus, a question such as "Is the chair of a committee a member of that committee?" can be answered correctly strictly from the information provided in Figure 1(a), but not from the information in Figure 1(b). If (a) and (b) were contained in treatments given to two experimental groups, it would not be surprising or interesting to find that the group given 1(a) outperformed the group given 1(b) with respect to this or similar questions.

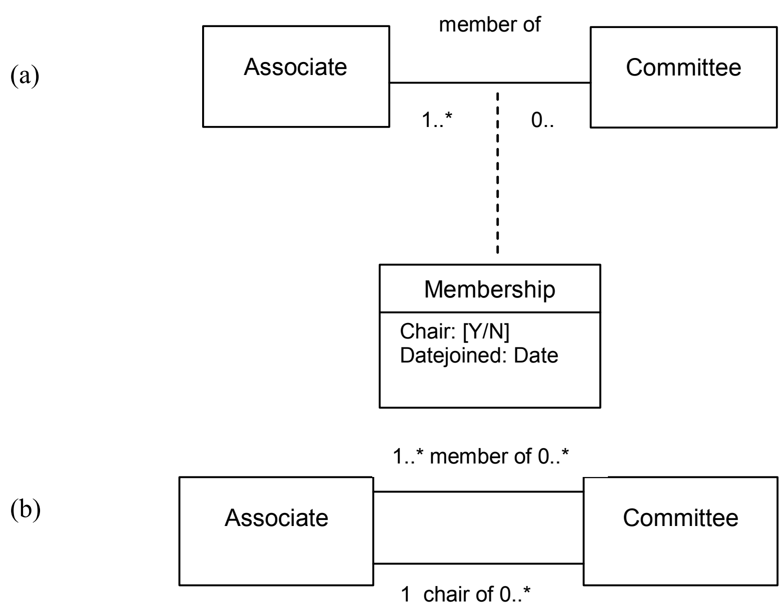

Figure 1: Information inequivalence

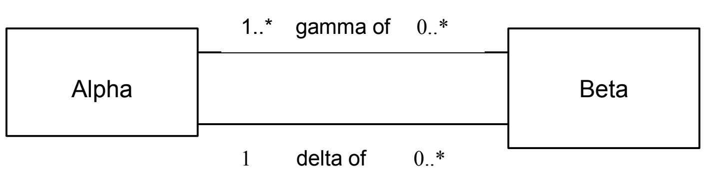

Figure 2: Removing the influence of prior knowledge

The second guideline is illustrated via Figures 1 and 2 jointly: namely, questions should be restricted to those that can be answered from the semantics conveyed by the diagram. Using Figure 1(b) to illustrate, consider a problem solving question such as “What options are available to a committee if the chair resigns?” Answering such a question in a meaningful way depends on interpreting the words in a diagram based on background knowledge. Simply replacing the relevant terms with terms from Figure 2 shows that such questions cannot test the semantics conveyed by a diagram (What options are available to a beta if a delta association between an alpha and a beta is terminated?).

Figures 1 and 2 also illustrate the importance of eschewing the use of subject matter experts in tests of the ability of alternate forms to convey semantics (the third of our guidelines). One could plausibly argue that almost any participant in an experimental study is a sufficient SME to answer a question such as “Is the chair of a committee a member of that committee?” simply by applying a common sense interpretation of the words, regardless of whether this semantics is conveyed by the diagram structure (in Figure 1). This could mitigate any effect of the treatment (i.e., the impact of differences in the way semantics are conveyed by the diagram structure). A simple way to eliminate potential confounds associated with using SMEs is simply to replace semantically laden words with non-words, as in Figure 2. By doing that, prior domain knowledge cannot interfere with the interpretation of the semantics conveyed in the representation embodied by the diagram. For instance, domain knowledge cannot be used to correctly guess an answer to the question “Is the delta of a beta also a gamma of that beta?”

EMPIRICAL EVIDENCE

General Issues

We collected data on the effect of following or violating the second and third guidelines in our framework as part of a larger study examining the effects of alternate ways of representing the concept of ontological precedence on diagram comprehension (additional information about the study design can be found in Parsons and Cole 2004). We report here only the results relevant to assessing these guidelines.

To assess the second guideline, we asked questions dealing with comprehension of diagram semantics, as well as problem solving questions of the type used by Gemino (1999). To assess the third guideline, we used alternate forms of the experimental material. In the first, we used English words from a domain with which participants had some familiarity (subject matter expertise). In the second version, we substitute Greek letters so that participants are able to answer questions only on the semantics conveyed by the diagram structure (no subject matter expertise).

Task Domain and Material

Since the objectives of this study were very focused, we decided not to test entire diagrams. Instead, we used two diagram segments. Figure 3 contains one of the diagram segments used. Figure 3 models logical associations (flight and reservation) between physical objects (passengers and aircraft). The second segment in the study dealt with an application involving reserving tables for customers at a restaurant.

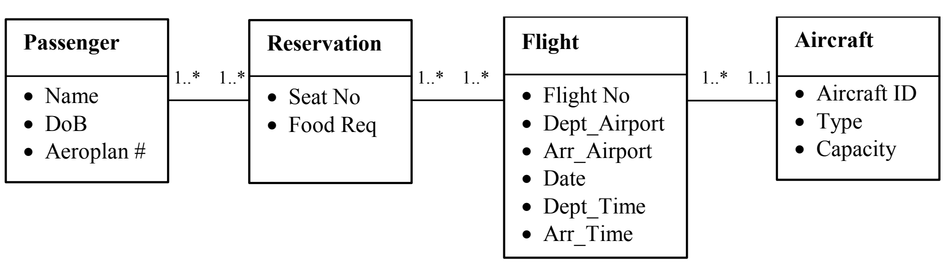

Figure 3: Sample diagram segment used in the study

To test whether background knowledge interfered with the communication of diagram semantics, we constructed alternative representations for which the class and association labels were “semantically void.” Figure 4 contains the segment corresponding to that in Figure 3. In Figure 4, answers to any questions about the diagrams can be based only on the semantics conveyed via the constructs and structure of the diagram, whereas answers to questions about the segment in Figure 3 might be confounded by background knowledge about the general subject domain.

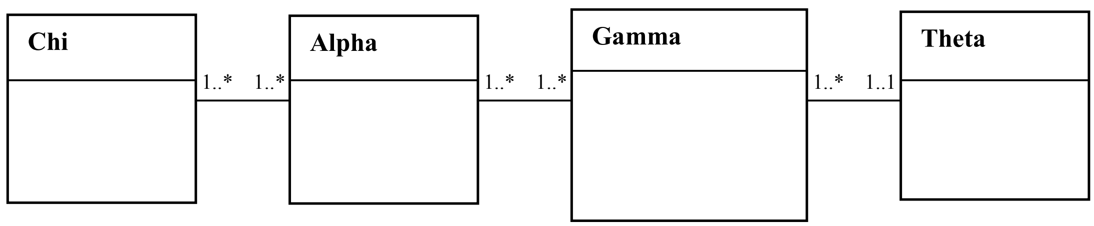

Figure 4: Semantically void equivalent of figure 3

The dependent variables in the study consisted of: (1) the number of correct answers to questions about the semantics conveyed in a diagram segment, and (2) the number of "reasonable" answers to problem solving questions based on the diagram segments. Table 1 contains the comprehension questions asked regarding the diagram segments in Figures 3. Table 2 contains the problem solving questions for the same diagram segments. The "reasonableness" of an answer to a problem solving question was judged with respect to whether a participant provided an answer that made sense given the diagram segment.

Semantic Segment (Figure 3)	Semantically Void Segment (Figures 4)
Can a Passenger have a Reservation with an Aircraft that does not involve a specific Flight?	Can a Chi be associated with an Alpha that does not involve a specific Gamma?
Does the diagram specify an upper limit on the number of Reservations associated with a specific Flight?	Does the diagram specify an upper limit on the number of Alphas associated with a specific Gamma?

Table 1: Sample comprehension questions

Semantic Segment (Figure 3)	Semantically Void Segment (Figure 4)
Under what circumstances would a passenger fail to board the aircraft associated with the reserved flight? List all of the reasons can you think of for this.	Under what circumstances would a Chi fail to board the Theta associated with the alpha-ed Gamma? List all of the reasons can you think of for this.
How would a decision by the airline has to standardize the types of aircraft it flies affect passengers and their reservations? Give as many examples as you can think of.	How would a decision by the Lambda has to standardize the types of Thetas it has affect Chis and their Alphas? Give as many examples as you can think of.

Table 2: Sample problem-solving questions

We propose two research questions based on the guidelines above. First, does an experimental treatment produce differential effects in comprehension versus problem-solving questions, as found in earlier studies (e.g., Bodart et al. 2001)? Second, are responses to questions affected by background knowledge of a domain as indicated by the labels (words) used on a diagram, rather than by the semantics conveyed by the diagram structure?

The experimental materials consisted of four diagram segments: two subject domains (restaurant and airline) and alternate representation of the concept of ontological precedence in which we were interested ('good' and 'bad'), each with a series of questions. In addition, we also varied the use of segments that carried semantics in the words used in class and association labels ('Semantic' segments) with the use of segments that used Greek letters as labels and were therefore semantically void ('Void' segments). Two versions of the material were developed. In version A, the ordering of segments was: Void-Bad; Void-Good; Semantic-Good; Semantic-Bad. In version B, the ordering of segments was: Void-Good; Void-Bad; Semantic-Bad; Semantic Good. In both versions, the Void cases were presented first to eliminate the potential influence of semantics conveyed by the use of words on subsequent interpretation of semantically void segments.

Participants and Procedure

Eighty undergraduate business students with no information systems experience participated in this study. Students were recruited via solicitation in an introductory marketing course. A short training session introduced the basics of class diagram notation, and students participated in a self-test of their understanding of the concepts covered. After the self-test, the administrator reviewed the answers to the questions with the participants.

Next, each participant was randomly assigned to version A or version B (as described above) of the material.

Participants were told that the purpose was to test how people read information from a certain type of diagram used in information systems development. Booklets containing instructions, the diagram segments, and questions for each segment, were distributed. Participants were given unlimited time to complete the exercise and the administrator collected each booklet as it was completed.

RESULTS

To test whether problem solving questions produce different patterns of answers than comprehension questions, we compared dependent measures across treatments using independent samples t-tests. Table 3 shows a general pattern of support for 'good' representations (those following a particular ontological prescription) over 'bad' representations (those violating the prescription) on two types of comprehension questions. Six of eight categories show support in the direction hypothesized in the primary study.

Diagram Segment	Question Type	Semantics	Mean		SD		Sig.
			Good	Bad	Good	Bad	
Restaurant	Existence of precedence [one question]	Semantic	0.93	0.60	0.267	0.496	< .001
	Cardinalities related to preceding mutual properties [three questions]		2.38	2.38	0.838	0.806	n.s.
Airline	Existence of precedence [two questions]		1.28	0.98	0.679	0.800	.04
	Cardinalities related to preceding mutual properties [one question]		0.80	0.55	0.405	0.504	< .01
Restaurant	Existence of precedence [one question]	Void	0.75	0.38	0.44	0.49	< .001
	Cardinalities related to preceding mutual properties [three questions]		2.10	1.93	0.900	0.949	n.s.
Airline	Existence of precedence [two questions]		1.33	0.60	0.701	0.391	< .001
	Cardinalities related to preceding mutual properties [one question]		0.83	0.58	0.385	0.501	< .01

Table 3: Results for diagram comprehension questions

Diagram Segment	Problem Solving Question	Semantics	Mean		SD		Sig.
			Good	Bad	Good	Bad	
Restaurant	Reasons for being late	Semantic	2.50	2.48	1.664	1.853	n.s.
	Ways of increasing capacity		1.53	1.65	1.176	1.122	n.s.
Airline	Reasons for missing association		2.23	2.27	1.833	1.723	n.s.
	Effect of standardization		0.88	1.53	0.723	1.176	.004
Restaurant	Reasons for being late	Void	0.18	0.23	0.501	0.660	n.s.
	Ways of increasing capacity		0.28	0.38	0.506	0.586	n.s.
Airline	Reasons for missing association		0.13	0.18	0.404	0.446	n.s
	Effect of standardization		0.08	0.08	0.267	0.267	n.s.

Table 4: Results for problem solving questions

Table 4 shows parallel results for problem solving questions. The scores show the mean number of "reasonable" suggestions generated under varying conditions. The pattern of significance is quite different from that in Table 3. In particular, the treatments had no substantive effect on the number of correct answers to problem-solving questions. These results confirm that comprehension and problem-solving questions can produce different results, as has been shown in earlier research (Gemino 1999, Bodart et al. 2001). More importantly, they indicate that apparent effects of a treatment on the semantics conveyed by diagram structure (as measured by comprehension questions) can be masked when problem solving questions are asked.

Next, we examined evidence in relation to the third guideline – that subject matter experts should not be used. In our study, participants could bring background (common-sense) knowledge to bear in answering questions about both the restaurant and airline examples when English labels were used on the diagram segments. However, when Greek letters were used, participants had to rely solely on semantics conveyed by the diagram structure and symbols. Thus, by looking at differences in responses between Semantic and Void materials, we can determine the extent to which the use of words may invoke background knowledge and affect responses. Table 5 contains the t-test results of this comparison.

Diagram Segment	Question Type	Representation	Mean		SD		Sig.
			Semantic	Void	Semantic	Void	
Restaurant	Existence of precedence [one question]	Good	0.93	0.75	0.267	0.439	.02
	Cardinalities related to preceding mutual properties [three questions]		2.38	2.10	0.838	0.900	n.s.
Airline	Existence of precedence [two questions]		1.28	1.33	0.679	0.701	n.s.
	Cardinalities related to preceding mutual properties [one question]		0.80	0.83	0.405	0.385	n.s.
Restaurant	Existence of precedence [one question]	Bad	0.60	0.38	0.496	0.490	.02
	Cardinalities related to preceding mutual properties [three questions]		2.38	1.93	0.807	0.944	.01
Airline	Existence of precedence [two questions]		0.98	0.60	0.800	0.591	.01
	Cardinalities related to preceding mutual properties [one question]		0.55	0.58	0.503	0.500	n.s.

Table 5: Comprehension Results for Semantic vs. Void Diagram Segments

Diagram Segment	Problem Solving Question	Representation	Mean		SD		Sig.
			Semantic	Void	Semantic	Void	
Restaurant	Reasons for being late	Good	2.50	0.18	1.664	0.501	< .001
	Ways of increasing capacity		1.53	0.28	1.176	0.506	< .001
Airline	Reasons for missing association		2.23	0.13	1.833	0.404	< .001
	Effect of standardization		0.88	0.08	0.723	0.267	< .001
Restaurant	Reasons for being late	Bad	2.48	0.23	1.853	0.660	< .001
	Ways of increasing capacity		1.65	0.38	1.122	0.586	< .001
Airline	Reasons for missing association		2.17	0.18	1.723	0.446	< .001
	Effect of standardization		1.53	0.08	1.176	0.267	< .001

Table 6: Problem Solving Results for Semantic vs. Void Diagram Segments

The top half of Table 5 shows that, under the ontologically 'good' form of representation, the presence or absence of English words that could affect answers has little impact on the results (the use of Greek letters does not seem to have a substantive negative effect on scores). However, under the 'bad' form of representation, in most cases the use of English words ('semantic' version) leads to significantly higher scores than when Greek letters are used. This indicates that participants receiving semantically void materials are unable to use background knowledge to answer questions correctly. However, this occurs only when the ontological representation is 'bad' (in that case, the words help).

Table 6 contains t-test results analogous to those of Table 5, using problem solving questions as dependent measures. As can be seen, participants scored consistently higher on the 'semantic' segments than on the 'void' segments. However, unlike the results for the comprehension questions in Table 5, there is no difference in the pattern between the version of material that followed the ontological prescription and the version that violated the prescription.

DISCUSSION

Based on a review of prior experimental research on conceptual modeling techniques and the assumption that it is necessary to gain deeper understanding of how humans extract real-world semantics from modeling formalisms, we proposed four criteria that we argue should be followed in such research. The first criterion restates Gemino and Wand's (2003) claim that information equivalence is critical in studies that compare alternate formalisms. The second states that questions should focus on measuring comprehension (semantics conveyed by modeling constructs in a script). We showed empirically how, for a given manipulation, the pattern of results can vary between comprehension questions and problem solving questions that rely on reasoning beyond information contained in a script (by using background knowledge associated with the domain embedded in the words attached to script elements). The third criterion states that subject matter experts should not be used in studies. We showed empirically that domain familiarity can affect results under some treatment conditions. The final criterion states that scripts should be available to participants during a study, since developing a good understanding of how humans extract information from scripts should precede studies of how easy or difficult it is to learn (remember) information from a script that is no longer available to review.

EVALUATION OF SOME PREVIOUS RESEARCH USING THE FRAMEWORK

Table 7 presents a selection of previous research studies in conceptual modeling (word limits preclude an exhaustive review). Each study addresses intragrammar comparisons of scripts based on alternate forms of representation (the independent variable(s)), and assesses the ability of users to use the representation in various ways (the dependent variable).

Reference	Task	Information Equivalence	Answerable from Script	Potential Impact of Prior Knowledge*	Models Available
Bodart et al. (2001)	E1: Reconstruct diagram	n/a**	High	2 domains	N
	E2: Comprehension test	High	High	High & Low	N
	E3: Problem solving test	High	Low		N
Burton-Jones, Meso (2002)	Problem solving test	Moderate	Moderate	Moderate	Y
	Cloze test	Moderate	High		Y
Burton-Jones, Weber (1999)	Comprehension test	Unknown***	Moderate	2 domains	Y
	Problem solving		Moderate	High & Low	N
Burton-Jones, Weber (2003)	Comprehension test	Moderate	High	High	Y
Gemino, Wand (2000)	Comprehension test		High	Low	Y
	Problem solving test	Unknown***	Low		N
	Cloze test		Moderate		N
Parsons (2003)	Comprehension test	High	High	None	Y
Weber (1996)	Reconstruct diagram	n/a**	High	Moderate	N

*Scores on this criterion were subjectively assigned by one of the authors

(e.g., University domain = High, Contracting = Moderate, Plant nursery = Low, Non-words = None).
**Information equivalence not applicable in diagram reconstruction task.
***Unable to determine from the paper (e.g., diagrams and/or questions were not included in the paper).

Table 7: Evaluation of Selected Prior Research Using the Framework

The columns indicate the degree to which the studies comply with our guidelines. The comparison allows us to identify the strengths and weakness of previous research and demonstrates how the framework can be applied to relate a range of independent studies. The first column identifies the study, and the second column briefly summarizes the experimental task. Each of the remaining columns represents a framework criterion as follows:

- Information Equivalence: Did the diagrams used provide the information to answer the questions correctly under any of the representational forms used as treatments in the experimental study?
- Answerable from Script: Could the questions be answered without reasoning beyond the information presented in the script(s)?
- Potential Impact of Prior Knowledge: To what extent could subject matter expertise (formal or common-sense) confound results?
- Script Available: Could participants refer to scripts while answering the questions?

CONCLUSIONS

Researchers involved in experiments to evaluate conceptual modeling techniques need to design materials, develop dependent measures, select subjects, and design procedures that enable the objectives of the experiment to be met. We believe there is a pressing need for basic studies that will improve our understanding of how the modeling constructs used in diagramming techniques convey domain semantics for the purposes of communication and validation.

With respect to the guidelines proposed in this paper, Table 7 highlights a lack of consistency in the selection of materials, questions, subjects, and procedures in previous experimental research. This inhibits the development of a cumulative understanding of how semantics are conveyed in conceptual modeling grammars. The table demonstrates that applying the framework to prior research on conceptual modeling can help in comparing and integrating the findings from previous studies in terms of improving our understanding of how alternate representations facilitate verification of conceptual models.

Moreover, Table 7 suggests there are other research questions for which the guidelines proposed here do not apply. We suggest that researchers need to pay careful attention to articulating their research questions and justifying their experimental design decisions accordingly.

REFERENCES

1. Bodart, F., A. Patel, M. Sim and R. Weber, "Should Optional Properties Be Used In Conceptual Modelling? A Theory And Three Empirical Tests," *Information Systems Research*, 12 (4), 2001, 384-405.
2. Bunge, M.Treatise on Basic Philosphy, *Volume 3: Ontology*, Amsterdam: Reidel, 1977.
3. Burton-Jones, A. and P. Meso, "How Good are these UML Diagrams? An Empirical Test of the Wand and Weber Good Decomposition Model," *Proceedings of the Twenty-Third International Conference on Information Systems*, Barcelona, Spain, 2002, 101-114.
4. Burton-Jones, A. and R. Weber, "Understanding Relationships with Attributes in Entity Relationship Diagrams," In P. De and J. DeGross (Eds.). *Proceedings of the Twentieth International Conference on Information Systems*, Charlotte, NC, 1999, 214-228.
5. Burton-Jones, A. and R. Weber, "Properties do not have properties: Investigating a questionable conceptual modeling practice," in D. Batra, J. Parsons, and V. Ramesh (Eds.), *Proceedings of the Second Annual Symposium on Research in Systems Analysis and Design*, 2003, 14 pp.
6. Gemino, A, Empirical Methods for Comparing Systems Analysis Modeling Techniques. Unpublished Doctoral Dissertation, University of British Columbia, Vancouver, CAN, 1999.
7. Gemino, A. and Y. Wand, "Foundations for Empirical Comparisons of Conceptual Modeling Techniques," in D. Batra, J. Parsons, and V. Ramesh (Eds.), *Proceedings of the Second Annual Symposium on Research in Systems Analysis and Design*, 2003, 29 pp.

8. Gemino, A. and Y. Wand, "Comparing Mandatory and Optional Properties in Conceptual Data Modeling," in P. Bowen and V. Mookerjee (eds.), *Proceedings of the Tenth Annual Workshop on Information Technologies and Systems*, Brisbane, Australia, 2000, 97-102.
9. Kung, C. and A. Solvberg, "Activity Modeling and Behavior Modeling," in T. Olle, H. Sol, A. Verrijn-Stuart (eds.), *Information Systems Design Methodologies: Improving the Practice*, North-Holland, 1985, 145-171.
10. Oei, J.., L. van Hemmen, E. Falkenberg, and S. Brinkkemper , "The MetaModel Hierarchy: A Framework for Information Systems Concepts and Techniques," Technical Report, Department of Informatics, Katholieke Universiteit, Nijmegen, No. 92-17, 1992, 1-30.
11. Parsons, J., "Effects of Local Versus Global Schema Diagrams on Verification and Communication in Conceptual Data Modeling," *Journal of Management Information Systems*. 19(3), 2003, 155-183.
12. Parsons, J. and L. Cole, "Properties of Properties: An Experimental Examination of Property Precedence in Conceptual Modeling," in S. Hartmann and John Roddick (eds.), Conceptual Modelling 2004: *First Asia-Pacific Conference on Conceptual Modelling*, Conferences in Research and Practice in Information Technology, Volume 31, January 2004, 101-110.
13. Parsons, J. and Y. Wand, "Property-based Semantic Reconciliation of Heterogeneous Information Sources," in S. Spaccapietra, S. March and Y. Kambayashi (Eds.), *Lecture Notes in Computer Science 2503 (ER-2002)*, Springer Verlag, 2002, 351-364.
14. Wand, Y. and R. Weber, "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars," *Journal of Information Systems*, 3, 1993, 217-237.
15. Wand, Y. and R. Weber, "Information Systems and Conceptual Modeling: A Research Agenda," *Information Systems Research* , 13(4), 2002, 363-376.
16. Weber, R., "Are Attributes Entities? A Study of Database Designers' Memory Structures," *Information Systems Research*, 7 (2), 1996, 137-162.