Comparing Object-Oriented and Extended-Entity-Relationship Data Models

Bill C. Hardgrave
University of Arkansas

Nikunj P. Dalal
Oklahoma State University

In the past few years, object–oriented (O–O) conceptual data modeling has emerged as an alternative to the traditional technique of entity–relationship modeling. O–O modeling is based on the premise that the resulting models are easier to use and understand. However, most claims of O–O modeling superiority are not empirically verified. Previous studies in this area have focused on a database modeler's ability to create conceptual data models from a written description, but the concept of understanding a completed data model by a database designer has not been investigated. Thus, this study explores a database designer's ability to understand an O–O conceptual data model – the Object Modeling Technique (OMT) – compared to an E–R model – the extended–entity–relationship model (EER). The OMT and EER conceptual data modeling techniques are compared using three modeling performance criteria: (1) model understanding; (2) time to understand; and (3) perceived ease–of–use. Results of this study indicate that the only difference between the two techniques is in the time to understand – OMT is significantly faster for both simple and complex problems.

The field of object–oriented (O–O) technology is growing rapidly. From a $200 million market in 1990, the O–O market is expected to hit $3.5 billion in a few short years (Khoshafian, 1993). Spanning research and practice from programming languages to systems analysis and design to database systems, almost all areas of systems development have been touched by the concepts of O–O. Although research continues to develop new areas within O–O, many recent efforts are aimed at testing and evaluating the fledgling technologies.

One of the newest areas (circa, mid–1980's) within O–O is the concept of O–O databases (Khoshafian, 1993). Most fields of database management – from the development of new forms of databases to new methods of data modeling – are feeling the influence of O–O. In particular, the growing interest in O–O has spawned the growth of several O–O conceptual data models. It is believed that O–O conceptual data models, compared to other conceptual data models (e.g., the entity–relationship model), more closely represent reality and, consequently, provide a higher degree of modeling correctness and understanding (Bock and Ryan, 1993).

However, most claims of O–O modeling superiority are not empirically verified. Additionally, previous studies in this area have focused on a database modeler's ability to create models from a written description (i.e., model correctness), but the concept of *understanding* a completed data model by a database designer has not been investigated. Thus, the purpose of this study is to empirically compare an O–O conceptual modeling technique to a more traditional technique based on the entity–relationship (E–R) model. Specifically, this study investigates an entry–level database designer's ability to understand an O–O conceptual data model, compared to an E–R model.

Background and Related Research

In database design, a primary criterion for evaluating a design is understandability (Blaha, et al., 1988): can end–users, database designers, and original modelers understand

Manuscript originally submitted June 21, 1994; Revised November 30, 1994; Accepted January 21, 1995 for publication.

the structure of the database? The conceptual data model serves as the bridge between users and database professionals; thus, the ability to understand the data model[1], by each party, is particularly important. The ultimate success of the project is dependent on the accuracy of the data model (Jarvenpaa and Machesky, 1989).

It is during the translation of requirements from user to database modeler that many errors occur. Often forgotten, however, are the errors that occur in communication among database professionals via the data model. Potential problems occur because the conceptual data modelers may not be responsible for implementing the design; thus, the designers of the implementable data model must be able to read and understand the conceptual data model. And, because the data model is often used as part of the system documentation, the data model should be clear and understandable to designers that may need the model later (Campbell, 1992).

Related Research

To this point, prior research has primarily investigated a data modeler's or end–user's ability to develop a conceptual or implementable data model from a written problem description. A sample of these articles is discussed next.

Several studies have compared a conceptual data model to an implementable data model. A study by Jih, et al., (1989) compared E–R and relational data models by looking at an end–user's query writing ability as measured by syntax errors, semantic errors, and time needed to write the query. No significant differences were found between the two techniques. Batra, et al., (1990), compared the extended–entity–relationship (EER) and relational models in the areas of modeling correctness and ease–of–use. For the experiment, end–users were asked to create models from a problem description. Results indicated that EER was significantly better on three of the six constructs used to measure modeling correctness. Ease–of–use was not significantly different between the two methods.

Other studies have compared various forms of conceptual models. Using a problem description as the task and end–users for subjects, Sheng and Higa (1990) compared the Structured Object Model (SOM), EER, and the normalization technique. The criteria for evaluation was design accuracy, design speed, and learning speed. Their results indicated that the graph based methods (SOM and EER) provided greater design accuracy and were easier to learn. A recent study by Bock and Ryan (1993) compared modeling correctness between the EER model and Kroenke's O–O model. Entry–level IS professionals created conceptual data models from a written description (the same description used in Batra, et al.'s (1990) study). EER was significantly better on three of the eight constructs used to measure model correctness.

Noticeably absent from this sampling of related research is research which investigates the understanding of the conceptual data model *after* it has been created (i.e., communication between database professionals). Thus, the focus of this study: a comparison of two conceptual models (O–O and E–R) with regard to a database designer's ability to understand the models.

Research Model

The research model used to guide this study is shown in Figure 1. Originally proposed by Jenkins (1982), the model has been modified to accommodate the study of data modeling performance (Batra, et al., 1990; Bock and Ryan, 1993). As seen in Figure 1, the focus of this study is on the data model's impact on performance (i.e., the data model is the independent variable of interest). The human and task variables are controlled.

Data Models

Many different data models exist for both EER and O–O modeling. Thus, specific models from each of these

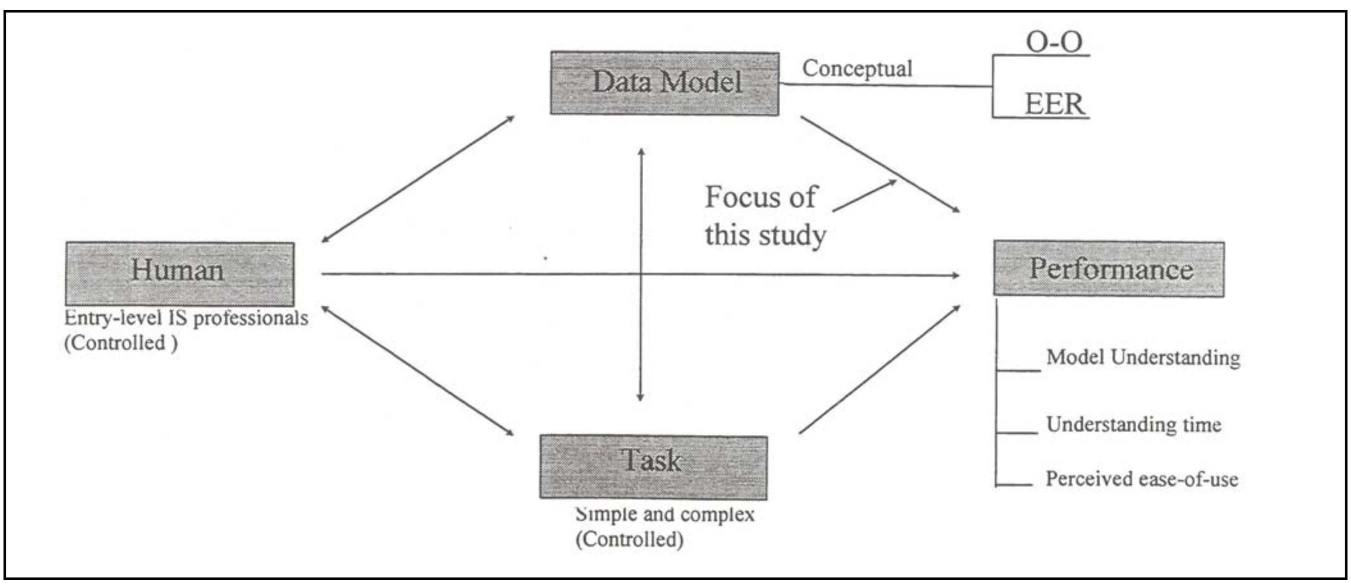

Figure 1: Research Model

domains had to be selected. To represent EER models, the technique as described by McFadden and Hoffer (1991) is used. The Object Modeling Technique (OMT) was chosen to represent the O–O paradigm (Rumbaugh, et al., 1991). Although both EER and OMT are considered semantic data models, they are very different.

The E–R approach, initially proposed by Chen (1976), represents the most common basis for conceptual design (Teorey, et al., 1986). In E–R models, the system is represented by entities, entity attributes, and relationships among the entities. Unfortunately, these simple constructs were not sufficient to model complex real–world situations, specifically generalization/specialization and aggregation. Due to these shortcomings, the original E–R has been modified and extended to produce what is now known as the extended–entity–relationship model. The EER model is now the accepted standard for relational database design (Batra, et al., 1990; McFadden and Hoffer, 1991). The EER technique as described in McFadden and Hoffer's book is similar to most EER techniques in syntax and semantics, and has been used in other studies of this type (e.g., Bock and Ryan, 1993).

The OMT is representative of the emerging O–O techniques of data modeling. OMT is a well–known, comprehensive, methodology that builds upon earlier object–oriented work, and includes modeling, analysis, design, and programming. According to Eckert and Golder (1994), OMT is "an enhanced form of E–R approach adding new concepts and constructs." OMT is said to improve upon the E–R modeling technique in the areas of expressiveness and readability which are key components of understandability (Blaha, et al., 1988). Compared to other leading object–oriented approaches, such as Coad and Yourdon's OOA and OOD methods, OMT provides many more constructs and defines the constructs more precisely (Eckert and Golder, 1994). According to the creators of OMT (Blaha, et al., 1988), OMT offers the advantages of ease of use, ease of understanding, and intuitive richness.

Task

Findings from related research in data modeling have indicated that task complexity is an important variable to consider in evaluating data models (Bock and Ryan, 1993; Brosey and Shneiderman, 1978; Jih, et al., 1989; Sheng and Higa, 1990). Also, proponents of O–O methods contend that O–O is most appropriate for complex problems (Blaha, et al., 1988). Thus, for this study, two tasks will be used: a 'simple' task and a 'complex' task. Task is controlled across data model types for a given task level (i.e., a simple problem situation is modeled using EER and OMT, and a complex problem situation is modeled using EER and OMT).

Modeling Performance

Modeling performance is determined by (1) model understanding, (2) understanding time, and (3) perceived ease–of–use. 'Model understanding' describes the degree of understanding by a subject, and is measured on two levels: facets and overall. Facets, such as entities, relationships, and attributes, are defined as the constructs which comprise a data model (Batra, et al., 1990). The facets used in this study include: (1) binary relationships (i.e., a relationship involving two entities or classes); (2) ternary relationships (i.e., a relationship involving three entities or classes); (3) unary relationships (i.e., a relationship involving only one entity or class); (4) categories (i.e., class/subclass, generalization/specialization); and (5) descriptors (i.e., attributes). Previous studies by Batra, et al., (1990) and Bock and Ryan (1993) have used facets as measures of modeling performance. Facets appear to be a proper construct for measuring understanding. Therefore, Hypothesis 1, stated in the null form, is:

H1:There are no differences in understanding between the EER and the OMT conceptual data models with respect to:

H1a: unary relationships;
H1b: binary relationships;
H1c: ternary relationships;
H1d: categories;
H1e: descriptors;

when testing subjects who have had a very brief exposure to each of the models.

In addition to the individual facets, it is important to measure the overall understanding of the model. Although, Batra, et al., (1990) feel that an overall measure is not necessary or proper, we believe a designer must develop an understanding of the complete model, in addition to the more detailed facets. Overall scores, representing performance for data modeling, have been used in prior studies (e.g., Jih, et al., 1989). Thus, Hypothesis 2, stated in the null form, is:

H2: There is no difference in overall understanding between the EER and the OMT conceptual data models when testing subjects who have had a very brief exposure to each of the models.

Another measure of modeling performance is time (Jarvenpaa and Machesky, 1989; Jih, et al., 1989; Shoval and Even–Chaime, 1987): how long does it take to understand the model? Although time and performance are usually considered a tradeoff, previous studies have shown that this is not necessarily always true (Jarvenpaa and Machesky, 1989). For this study, the time needed to understand the model is a logical measure of modeling performance. Stated in the null form, Hypothesis 3 is:

H3: There is no difference in understanding time between the

EER and the OMT conceptual data models when testing subjects who have had a very brief exposure to each of the models.

The final measure of performance is perceived ease–of–use. Perceived ease–of–use is defined as the "degree to which a person believes that using a particular system would be free of effort" (Davis, 1989, p. 320). In the context of data modeling, ease–of–use indicates the degree of effort needed to use the data model (Batra, et al., 1990). Models that require less mental effort (i.e., easier–to–use) are considered better performers (Batra, et al., 1990). Hypothesis 4, stated in null form, is:

H4: There is no difference in perceived ease–of–use between the EER and the OMT modeling techniques when testing subjects who have had a very brief exposure to each of the models.

Research Methodology

Experimental Design

A lab experiment, using a between–subjects post–test–only design, is used to compare the two data modeling techniques at different levels of task difficulty. The two factors (i.e., the independent variables) are:

1. The type of data model: EER and OMT
2. Task complexity: simple and complex

Task complexity is operationalized as follows. A simple task contains entities, descriptors, and binary relationships. The complex task contains all of the characteristics of a simple task, plus unary relationships, ternary relationships, and categories.

The facets used to determine simple or complex were determined by previous studies. Unary relationships and ternary relationships are included in the complex task because they have been shown to be difficult to model and understand (Batra, et al., 1990; Bock and Ryan, 1993; Sheng and Higa, 1990). Category, a concept added to the original E–R technique, is also included in the complex task. Entities, descriptors, and binary relationships, on the other hand, are very common and appear in almost all real–world conceptual models, thus forming the basis for the simple task.

Subjects

The sample for this study consists of 56 students enrolled in a course in Database Management Systems at a major midwest university. Most of the students were seniors in their last year or semester of coursework, thus serving as surrogates for entry–level IS professionals. Subjects participated as part of the normal coursework for the semester.

Subjects were randomly assigned to one of the four groups (simple OMT, complex OMT, simple EER, complex EER). Because the variance in human characteristics is a threat to valid inference, the human variable is 'controlled' by measuring the potential threat, then using the measures in the data analysis to rule out the threat (Cook and Campbell, 1979). In this particular study, it is desired that subjects have no prior knowledge of the models other than that gained in the database class, and that each of the four groups have equivalent backgrounds. A questionnaire was used to measure the subjects' backgrounds; it was not used to assign the subjects to the groups. Data analysis revealed no significant differences between the four groups due to subject background.

Procedure

The subjects were provided two one–hour lectures on each of the models. The content of the lecture material was designed to be functionally equivalent. To alleviate any novelty effects, subjects were given a homework assignment similar in form to the experimental task, prior to the testing session.

For the testing session, subjects were randomly assigned to one of the four groups and presented with a data model (OMT or EER) for a particular task level (simple or complex). The subjects were asked to examine the model and then answer multiple–choice questions that tested his/her understanding of the data model. The time needed to complete the questions was also noted.

The task variable is controlled by using the same or similar teaching methods to present the models to the subjects, by using equivalent models across modeling techniques for a given level of complexity, and by regulating the amount of time and materials used to teach the subjects.

Dependent Variables

Performance, the dependent variable for this study, is indicated by model understanding, understanding time, and perceived ease–of–use. To measure model understanding, an objective instrument is used. The simple task consists of five multiple–choice questions; the complex task has ten multiple–choice questions. All subjects solving the simple task received the same five questions, but half the subjects used an OMT model, the other half used an EER model. The same method was used for the complex task. Each multiple choice question is designed to measure the understanding of a particular facet. Responses are scored as either correct or incorrect (no partial scores). The facets measured are, for the simple task: binary one–many and binary many–many, and a descriptor (in this case, a link attribute on a binary many–many relationship). For the complex task, the facets are: unary one–many, binary one–many, ternary one–many–many, categories, and descriptors (one descriptor is for an entity/class involved in the ternary relationship, and one descriptor is a link attribute on a ternary relationship). The sum of the correct answers is an indicator of overall understanding.

Understanding time is operationalized as the total time taken to answer the objective questions. Subjects were asked to indicate the starting time when they began the task, and the ending time when they finished the task. Times were verified by the experimenters when the subjects presented the completed tasks.

The instrument used to indicate perceived ease–of–use is adopted from Davis (1989) and Batra, et al., (1990). Each technique, OMT and EER, had an associated ease–of–use questionnaire.

Results and Interpretations

Model Understanding – Facets

A comparison of 'facet understanding' between EER and OMT is shown in Table 1. For the complex problem, there are no significant differences between EER and OMT with respect to the five individual facets. The same is true of the two facets for the simple problem. Although one may infer differences based upon mean scores alone, a conclusion that a difference exists cannot be reached due to a lack of statistical significance. Thus, Hypothesis 1, which states that there are no differences in understanding between the EER and OMT data models with respect to the five different facets, cannot be rejected.

It is interesting to note that the level of understanding of ternary relationships is consistent with the ability to accurately create ternary relationships. Findings in this study indicate a mean percentage correct of 46 percent for both EER and OMT. Subjects in the Batra, et al., (1990) study accurately modeled the ternary one–many–many relationship (using EER) 41 percent of the time. In Bock and Ryan's (1993) study, subjects scored 47 percent for EER and 44 percent for Kroenke's O–O model in modeling a ternary one–many–many relationship.

	Means (% correct)		
Facets	**EER**	**OMT**	**p–value**
Unary (complex)	78.50%	65.35%	0.4030
Binary (complex)	75.00%	65.35%	0.4433
Binary (simple)	83.93%	85.00%	0.9011
Ternary (complex)	46.43%	46.15%	0.9767
Categories (complex)	75.00%	65.35%	0.4826
Descriptors (complex)	25.00%	42.30%	0.1063
Descriptors (simple)	64.29%	80.00%	0.3620

Table 1: Comparison of Model Understanding (Facets)

Another interesting result is the extremely low scores associated with the descriptors for the complex problem. As described earlier, the descriptors for the complex problem consist of an attribute of an entity/class involved in a ternary relationship and a link attribute of a ternary relationship. The low understanding of the descriptors for the complex problem seem to be compounded by the ternary relationship, which has proven to be difficult to understand.

Model Understanding – Overall

Since no significant differences are found at the facet level, it is suspected that the overall score will not show differences. The (partial) ANOVA table (Table 2) is used to indicate significant differences (variation) in the dependent variable (in this case, total score) due to the task, the data model, or the interaction of the task and the data model. The values of interest are highlighted in Table 2. No significant differences are found due to the data model used (TECH; significance level = 0.9300) or the interaction of the data model and task (TECH*TASK; significance level = 0.5014). As expected, the scores between the simple and complex tasks are significantly different (TASK; significance level = 0.0001). Thus, Hypothesis 2, which states that there will be no difference in overall understanding due to the data model, cannot be rejected.

Understanding Time

Hypothesis 3 suggests that there is no difference in understanding time due to the data model. As shown in the (partial) ANOVA table (Table 3), Hypothesis 4 is rejected. Table 3 shows the effects of the data model, the task, and the interaction of the task and the data model on the time it takes to complete the task. A significant difference (significance level=0.0013) exists between the data models (TECH) in the time it takes to understand (i.e., answer the questions). There is no significant difference in understanding time due to the interaction of task and data model (TASK*TECH; significance level=0.9950). As expected, it took significantly longer to perform the complex task than the simple task (TASK; significance level=0.0001).

Dependent Variable: TOTAL SCORE

Source	DF	Type III SS	F–Value	Pr > F
TASK	1	77.37	20.28	**0.0001**
TECH	1	0.03	0.01	**0.9300**
TASK*TECH	1	1.75	0.46	**0.5014**

Table 2: ANOVA Showing the Effects of Task and Technique on TOTAL SCORE

Dependent Variable: TIME				
Source	DF	Type III SS	F–Value	Pr > F
TASK	1	1851652.46	256.81	**0.0001**
TECH	1	82793.71	11.48	**0.0013**
TASK*TECH	1	0.28	0.00	**0.9950**

Table 3: ANOVA Showing the Effects of Task and Technique on TIME

A supplemental analysis of the mean time (in seconds) for each of the four groups is presented in Table 4. For the simple task, EER subjects took an average of 312 seconds, compared to 236 seconds for OMT subjects. The difference is significant (p=0.0183). For the complex task, the average EER time was 677 seconds and OMT was 600 seconds. This difference, also, is statistically significant (p=0.0221). Thus, it appears that for both simple and complex tasks, OMT is faster to use and understand.

Although the time required to understand the model is relatively short, it must be recognized that the time to create the diagram from a problem description would be much more time-consuming. For example, the model produced from the problem description used in the study by Batra, et al., (1990) and Bock and Ryan (1993) took approximately one hour to create(Bock and Ryan, 1993). However, since this study measures understanding time (from a completed model), we would expect the understanding time to be significantly less than creation time (given the same problem situation).

Perceived Ease–of–Use

The null form of Hypothesis 4 suggests that there are no differences in perceived ease–of–use between EER and OMT data models. As presented in Table 5, we fail to reject Hypothesis 4 based upon the available data. Although the scores for EER are slightly better (range=8 to 56, where lower numbers indicate greater ease–of–use), the differences are not significant for either the simple or the complex tasks. Subjects also indicated that the complexity of the task had no significant bearing on the perceived ease–of–use. The instrument used to measure ease–of–use for OMT demonstrated a reliability of 0.88; the reliability for the EER ease–of–use instrument was 0.93. Both scores indicate that the instruments have high reliability.

	EER	OMT	p–value
Simple	312	236	**0.0183**
Complex	677	600	**0.0221**

*cell values are in seconds

Table 4: Understanding Time by Data Model and Task

	EER	OMT	p–value
Simple	25.86	27.33	0.4355
Complex	21.57	24.23	0.6592
p–value	0.2450	0.2996	

Table 5: Comparing Perceived Ease–of–Use

Discussion

Overall, significant differences were not found between EER and OMT with respect to model understanding or perceived ease–of–use. The lack of findings disputes many of the contentions in the literature that O–O models provide an easier to use and easier to understand environment. Similarly, suggestions that O–O would provide better results for complex problems are also not proven in this study.

The one significant finding is that OMT is faster to use and understand than EER. This follows the literature that suggests O–O is more "natural" and thus can be comprehended faster (Booch, 1991). The most plausible explanation for this finding is that, according to Bruegge, et al., (1992), OMT is a "highly expressive tool" (p. 374). OMT uses fewer symbols than EER, but the symbols contain more meaning. Thus, the resulting data models are less cluttered and daunting to the model user and are quicker to comprehend. McGee (1976) calls this attribute (of a data model) "elegance," which means the data model "achieves a given level of modeling capability with the smallest number of structure types, composition rules, and attributes" (p. 374). Although OMT did not prove better for understanding scores, the faster times (without loss ofperformance) can result in increased productivity for database designers.

Needed Research

Unfortunately, we lack a sound theoretical foundation that would allow us to fully understand the factors that influence the understanding and use of conceptual data models (Bock and Ryan, 1993). Thus, research is forced to proceed piece–meal with each experiment incrementally adding to our knowledge of conceptual data modeling. We feel that this study is a positive step in that direction because it uniquely investigates the ability of a database designer to **understand** a data model once it has been created.

However, we have a long way to go. A beginning point is to identify those situations where one model is better than another. The study described herein moves toward that goal by comparing the models for both simple and complex tasks.

Tsichritzis and Lochovsky (1982) suggest that one data model is probably not best universally; instead, we must find those situations where the models are most appropriate. Also, Batra, et al., (1990) recommend that the effect of different data models on modeling performance be considered. As new models are introduced, they must be evaluated. According to Peckham and Maryanski (1988): "Since the complexity of the applications will continue to increase, the designer's requirements of a conceptual model will similarly heighten, and hence new models will continue to emerge" (p. 187).

Conclusion

This study has empirically compared OMT and EER conceptual data modeling techniques along three dimensions: (1) model understanding; (2) time to understand; and (3) perceived ease–of–use. Although the literature would suggest that the O–O technique, OMT, would produce a more understandable and easier–to–use model, the majority of the results of this study do not support these contentions. Instead, results indicate that the only difference between the two techniques is in the time to understand – OMT is significantly faster for both simple and complex problems. However, this is a positive finding. OMT produced a model that is faster to understand without a decrease in the level of understanding, which can result in increases in productivity.

Acknowledgments
The authors are indebted to the associate editor and reviewers for their very valuable comments and suggestions which have greatly improved the quality of the paper.

Endnote
[1] In this article, the generic term "data model" will refer to the conceptual data model. Other forms of data models will be distinguished by using the formal names (e.g., the implementable data model).

References
Batra, D.; J.A. Hoffer; & R.P. Bostrom. (1990). Comparing representations with relational and EER models. *Communications of the ACM* 33(2):126–138.

Blaha, M.R.; W.J. Premerlani; & J.E. Rumbaugh. (1988). Relational database design using an object–oriented methodology. *Communications of the ACM* 31(4):414–427.

Bock, D.B.; & T. Ryan. (1993). Accuracy in modeling with extended entity relationship and object oriented data models. *Journal of Database Management* 4(4):30–39.

Booch, G. (1991). *Object–Oriented Design with Applications.* Redwood City, CA: The Benjamin/Cummings Publishing Company, Inc.

Brosey, M.; & B. Shneiderman. (1978). Two experimental comparisons of relational and hierarchical database models. *International Journal of Man–Machine Studies* 10:625–637.

Bruegge, B.; J. Blythe; J. Jackson; & J. Shufelt. (1992). Object–oriented system modeling with OMT. In *Proceedings of the 1992 OOPSLA Conference*, pp. 359–376.

Campbell, D. (1992). Entity–relationship modeling: one style fits all? *Data Base*, 23(3):12–18.

Chen, P.P. (1976). The entity–relationship model – toward a unified view of data. *ACM Transactions on Database Systems* 1(1):9–36.

Cook, T.D.; & D.T. Campbell. (1979). *Quasi–Experimentation: Design & Analysis Issues for Field Settings.* Boston, MA: Houghton–Mifflin Company.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3):319–340.

Eckert, G.; & P. Golder. (1994). Improving object–oriented analysis. *Information & Software Technology* 36(2):67–86.

Jarvenpaa, S.L.; & J.J. Machesky. (1989). Data analysis and learning: an experimental study of data modeling tools. *International Journal of Man–Machine Studies* 31:367–391.

Jenkins, A.M. (1982). *MIS Decision Variables and Decision Making Performance.* Ann Arbor, MI: UMI Research Press.

Jih, W.J.K.; D.A. Bradbard; C.A. Snyder; & N.G.A. Thompson. (1989). The effects of relational and entity–relationship data models on query performance of end users. *International Journal of Man–Machine Studies* 31:257–267.

Khoshafian, S. (1993). *Object–Oriented Databases.* New York: John Wiley & Sons, Inc.

McFadden, F.R.; & J.A. Hoffer. (1991). *Database Management.* 3rd ed. Redwood City, CA: The Benjamin/Cummings Publishing Company, Inc.

McGee, W.C. (1976). On user criteria for data model evaluation. *ACM Transactions on Database Systems* 1(4):370–387.

Peckham, J.; & F. Maryanski. (1988). Semantic data models. *ACM Computing Surveys* 20(3):153–189.

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

Sheng, O.R.L.; & K. Higa. (1990). An analytical and empirical comparison of end–user logical database design methods. *Journal of Database Administration* 1(1):1–15.

Shoval, P.; & M. Even–Chaime. (1987). Data base schema design: an experimental comparison between normalization and information analysis. *Data Base* 18(3):30–39.

Teorey, T.J.; D. Yang; & J.P. Fry. (1986). A logical design methodology for relational databases using the extended entity–relationship model. *ACM Computing Surveys* 18(2):197–222.

Tsichritzis, D.C.; & F.H. Lochovsky. (1982). *Data Models.* Englewood Cliffs, NJ: Prentice–Hall, Inc.

Bill C. Hardgrave is an assistant professor in the department of Computer Information Systems and Quantitative Analysis at the University of Arkansas. His research interests include object-oriented development, database design, and information systems prototyping.

Nikunj P. Dalal is Assistant Professor of MIS in the Department of Management in the College of Business Administration at Oklahoma State University. His current research interests include: database modeling, intelligent decision support systems, and the national information infrastructure.

Related Content