EFFECT OF DIFFERENT DATABASE STRUCTURE REPRESENTATIONS, QUERY LANGUAGES, AND TASK CHARACTERISTICS ON INFORMATION RETRIEVAL

Pascal A. Bizarro, Bowling Green State University

ABSTRACT

This research paper investigates the impact of different database structure representations, query languages, and task complexity on an information retrieval task. Cognitive fit theory is used to formulate four hypotheses based on different measures of end-user performance. A laboratory experiment is conducted to test the hypotheses. Participants include students majoring in accounting and management information systems (MIS). Analysis of variance and post hoc means comparisons are used to analyze the results. While previous research manipulated only the database structure representation or the query language, this research project extends the current research in accounting and information systems by manipulating both the database structure representation and the query tool. User characteristics, such as professional skills, are explicitly included in the research model, where these previously had been ignored. The findings of the current study cannot be explained by cognitive fit theory. Different combinations of database structure representation and query language are best suited depending on the measure of performance used and on user characteristics. These results have practical implications, because they can help professionals determine the type of database documentation and query tools specific end-users need to access and use to improve their performance in query writing tasks. This study also reveals that user characteristics are important factors to be considered when investigating the end-user's performance. These characteristics should help practitioners and academics designing and implementing customized training.

INTRODUCTION

The accounting and business communities increasingly rely on database applications to convert raw data into useful business information (Hayes and Hunton, 2000). Today, a majority of accounting information systems use relational databases (Hooper and Page, 1996). In the past, information systems (IS) professionals, not end-users, were responsible for query tasks (Borthick, 1992). Recently, this function has shifted to end-users as accounting systems use more easily accessible databases (Hooper and Page, 1996). End-users (e.g., accountants, auditors, and managers) must understand the database structure and be able to use the available query language to transform accounting data into useful business information (Leitheiser and March, 1996).

Accounting end-users must be able to discern whether the potential information is available and then retrieve it (Borthick, 1992). A database structure representation communicates the availability of accounting data (Dunn and Grabski, 2002). This representation details the stored data items and their logical organization. Examples include the entity-relationship (ER) model and the relational model. Accessing data of interest requires knowledge of a database query language. Examples include query-by-example (QBE) and structured query language (SQL).

The IS literature is unclear concerning how database representation type, query tool type, and user characteristic affect end-user performance in query construction tasks. According to Dunn and Grabski (2002, 168), “prior research has not examined these factors or their interactions in a systematic way, as this is a relatively new research field.” They propose that these factors be studied further to learn more about their combined effects on query writing performance.

Past empirical research on information retrieval typically compares one representation model to another, while keeping the language constant (Lochovsky and Tsichritzis, 1977; Jih et al., 1989; Davis, 1990; Chan et al., 1993; Leitheiser and March, 1996), or compares one query language against another without regard to the database representation (Reisner et al., 1975; Chamberlin et al., 1976; Zloof, 1977; Greenblatt and Waxman, 1978; Reisner, 1981; Yen and Scamell, 1993). Only one study manipulated both the database structure representation and the query language (Chan et al., 1993). The current study manipulates both the data model and the query language. Chan et al. (1993), however, did not investigate the interaction between the two factors. The current study solves this problem by analyzing the main effects as well as the interaction of the different factors affecting end-user query performance.

Prior research on query languages, which typically holds the data model constant, suggests that a graphical tool is easier to use, but only for nonprogrammers (Reisner et al., 1975). Prior research on data models indicates overwhelming superiority of the graphical version of the data model for database designing tasks (Batra et al., 1990). Results were inconsistent for those studies that manipulated the data model while investigating user performance in query writing tasks (Jih et al., 1989; Chan et al., 1993). None of these studies examines the use of QBE as a retrieval technique. Therefore, this study includes both SQL and QBE as query languages.

Accountants, auditors, and managers use a variety of query tools, depending on their database management system. In addition, all prior experiments except Amer (1993) used IS students as participants. Therefore, this study uses accounting and MIS students as proxies for potential DBMS end-users. This is the first study that incorporates user characteristics in the research model.

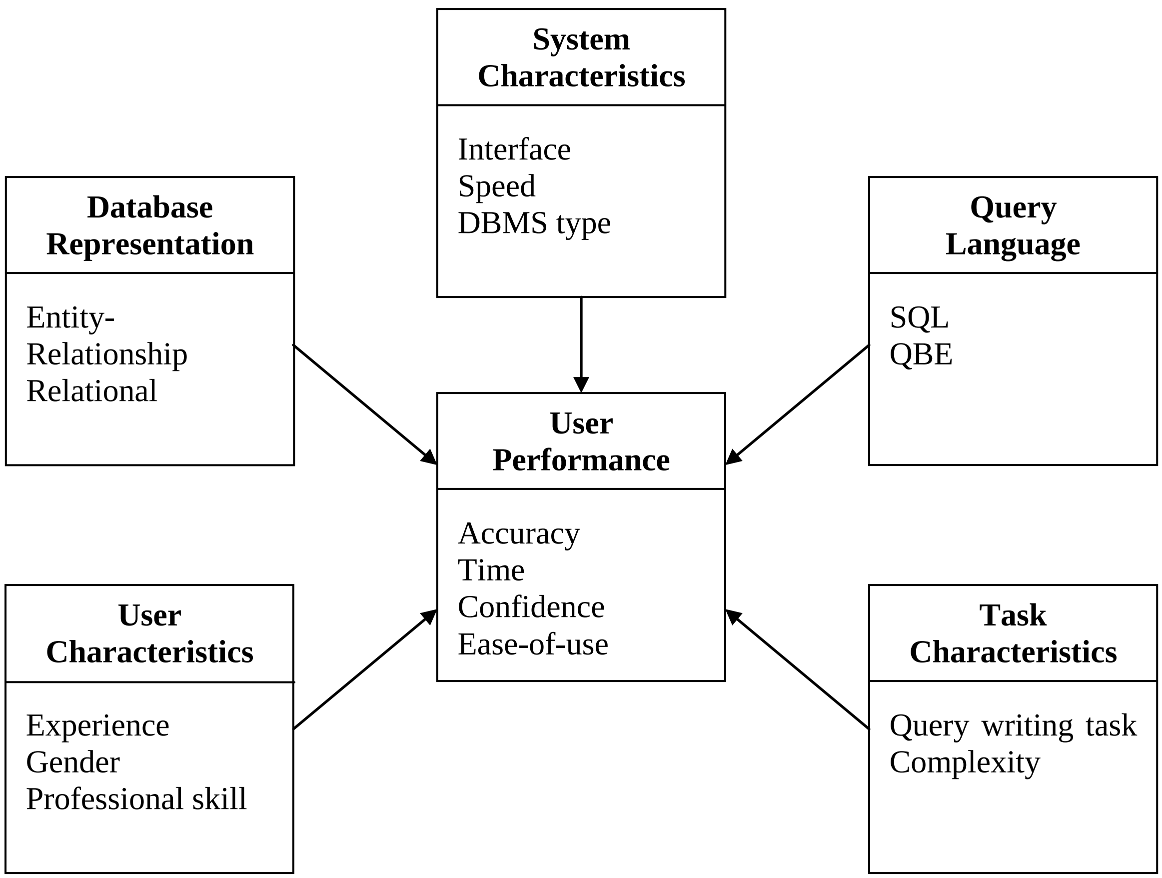
This study investigates the influence of database structure representation, query language, and task characteristics on user performance in the information retrieval process. Organizations often implement formal training programs to teach their new hires and experienced employees the skills they will need, including the ability to efficiently and effectively query a database. Training can be costly to the organization, in both lost production time and costs of the program itself (potential travel costs, instructor’s fees, and materials). By implementing the concepts of this research, organizations can increase the value added from the training and also reduce the cost of the training. By better understanding how the combination of database documentation and query tools can affect the end-users performance in query writing tasks organizations will be able to implement specific trainings that improve the employee’s retention and performance.

The rest of the paper is organized as follows. The next section describes the research model and formulates the hypotheses. The third section explains the research design and methodology. The fourth section reports the experiment results. The last section discusses the study results and conclusions.

HYPOTHESIS DEVELOPMENT

Figure 1 synthesizes the IS data model and query language literature and extends it to a proposed research model for the database user's performance in query writing tasks. The model indicates that the user's performance is influenced by the following five factors: database representation, query language, system characteristics, task characteristics, and user characteristics. A similar version of this research framework is suggested by Dunn and Grabski (2002) as a basis for future research in this field.

Figure 1
Research Model



The majority of the prior research manipulated ony one or two facets of this research model. In the current study, only the system characteristics (such as response time, physical input/output devices, and user interface) are controlled. Data model, query language, user characteristics, and task characteristics are manipulated and relationships between them are investigated.

Cognitive fit literature can be used to explain the relationship between the data model type and the query language type. Cognitive fit theory was originally developed by Vessey (1991) and further improved to explain previously conflicting results regarding relative performance of users presented with information in graphical versus tabular formats (Umanath and Vessey, 1994; Vessey, 1994; Vessey and Galletta, 1991).

According to the theory, when the types of information emphasized in the problem-solving elements (problem representation and task) match, the problem solver can use processes and formulate mental representations that emphasize the same type of information. Consequently, the processes the problem solver uses, both to act on the problem representation and to complete the task, will match. This superior mental representation will result in more effective problem-solving performance (Vessey 1991).

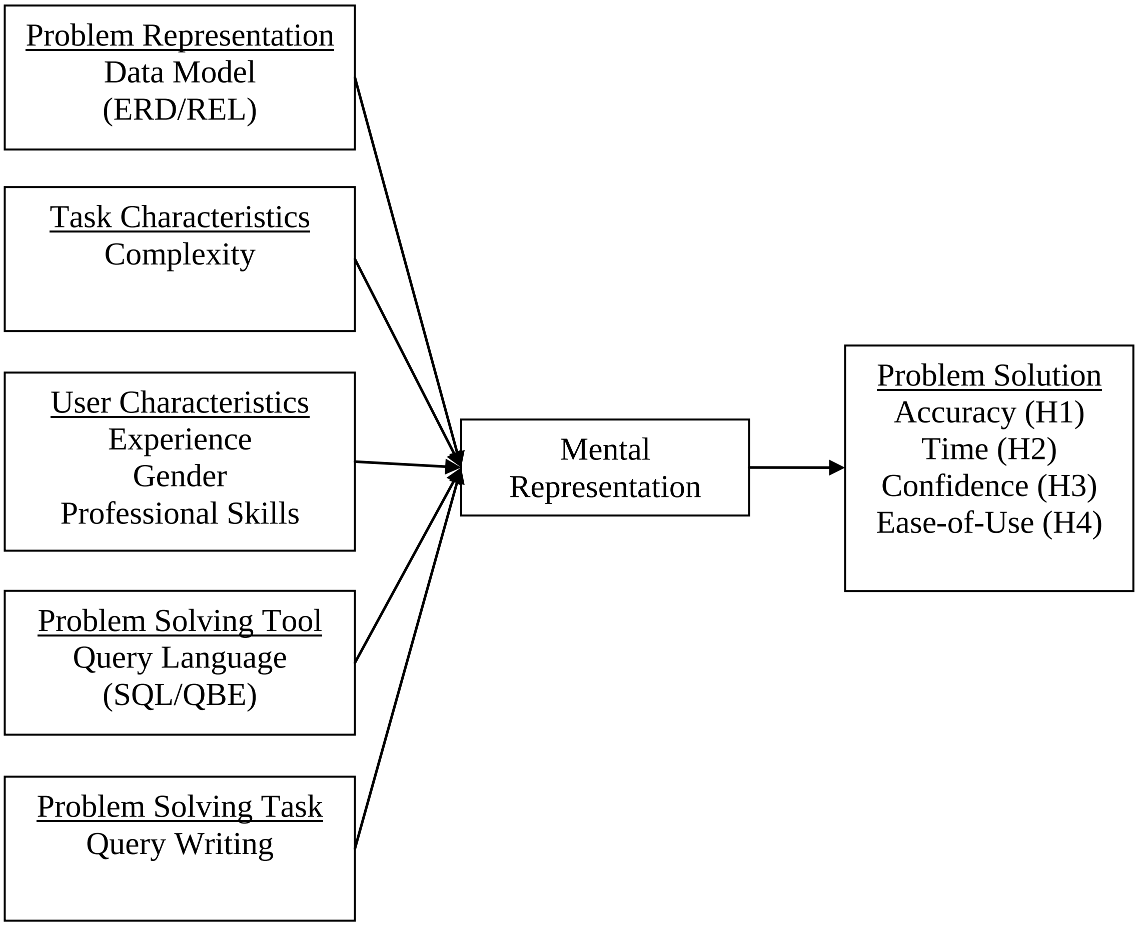
Conversely, when a mismatch occurs between problem representation and task, cognitive fit will not occur, because similar problem-solving processes cannot be used to both act on the problem representation and solve the problem. Because problem solvers form mental representations from materials presented to them, the mental representations likely will be strongly influenced by problem representations. In a mismatch situation, the users must transform their mental representation to derive a solution to the problem. Alternatively, they may formulate mental representations based on the task. In this case, they must transform the data derived from the problem representation into a mental representation suitable for the task solution. In either case, problem-solving performance will deteriorate (Sinha and Vessey, 1992).

Smelcer and Carmel (1997) and Dennis and Carte (1998) applied basic cognitive fit theory to the geographic information systems domain, using it to explain performance differences among users of map and table-based geographic information systems on geographic relationship tasks. Dunn and Grabski (2001) applied cognitive fit to accounting models and integrated cognitive fit theory with the concept of localization to provide additional evidence for how cognitive fit works. Their study compared the traditional debit-credit accounting model with the resources-events-agents (REA) accounting model (McCarthy 1982). The results indicate that task characteristics are important when determining the model representation that best supports the task.

Related research in human problem-solving has examined factors other than task and representation that could affect problem-solving performance. Sinha and Vessey (1992) extended the basic model of cognitive fit to include the problem-solving tool as an additional determinant of problem-solving performance. They argued that matching the type of information provided by the tool, the task, and the problem representation would lead to improved problem-solving performance. Vessey and Weber (1986) conducted an experiment that tested an extended notion of cognitive fit. They found support for the performance effects of matching both the problem-solving tool to the task and the problem representation to the task.

Figure 2 represents the extended cognitive fit theory (Sinha and Vessey, 1992). In this study, the problem representation is characterized by the database structure representation. The ER model has a two-dimensional syntax while the relational model displays the information using a linear syntax. The problem-solving tool in this case is the type of query language. QBE uses a two-dimensional syntax. Conversely, SQL uses a linear syntax. The problem-solving task is query writing based on a particular data model using a specific query language. The authors also add the user characteristics (experience, gender, and professional skills of the query writer) and task characteristics defined by the complexity of the queries as important elements.

Figure 2
Extended Problem Solving Model of Cognitive Fit



Based on the extended cognitive fit theory, the matching of problem representation (data model), the problem-solving tool (query language), and the problem-solving task (query writing) will result in superior user performance in information retrieval (i.e., query accuracy, time to completion, user confidence, and perceived ease-of-use). When the degree of cognitive fit between the database representation and the query language is high, participants will generate more correct queries than when the cognitive fit level is low. The following hypothesis is formulated:

H1 ER/QBE or Relational/SQL participants will generate more correct queries than Relational/QBE or ER/SQL participants.

Cognitive fit theory predicts that users of information that is consistent across problem and task representation will perform more quickly than users of inconsistent information, because less cognitive effort is required to process the information. When the degree of cognitive fit between the database representation and the query language is high, participants will complete the task faster than when the cognitive fit level is low. Thus, the following hypothesis is formulated:

H2 Task completion time will be lower for the ER/QBE or Relational/SQL participants than for the Relational/QBE or ER/SQL participants.

Dunn and Grabski (2001) were the first to investigate the effect of cognitive fit on user confidence. Their results did not support the claim that users' confidence in their responses increases with correctness. Prior research found an inverse relationship between accuracy and confidence (Dickson et al., 1977). Nevertheless, when the degree of cognitive fit between the database representation and the query language is high, participants will be more confident about their answers than when the cognitive fit level is low. The following hypothesis is formulated:

H3 ER/QBE or Relational/SQL participants will be more confident about their answers than Relational/QBE or ER/SQL participants.

If the problem representation, the problem-solving tool, and task do not match, the decision maker must reformulate the mental representation to match with the problem representation, the tool, and the task. Thus, the decision maker will perceive the task to be more difficult. Conversely, when the degree of cognitive fit between the database representation and the query language is high, participants will perceive that the combination is easier to use than when the cognitive fit level is low. The following hypothesis draws on these relationships:

H4 ER/QBE or Relational/SQL participants will perceive the combination easier to use than Relational/QBE or ER/SQL participants.

User performance is evaluated in term of query accuracy, task completion time, user confidence level, and perceived ease-of-use. The matched groups (ER/QBE and Relational/SQL) are expected to outperform the unmatched groups (ER/SQL and Relational/QBE) in each user performance measure based on the cognitive fit theory.

METHOD

A laboratory experiment using a 2x2x2 factorial design was conducted to test the research hypotheses. The general methodological design is straightforward and similar to many previous studies (e.g., Chan et al., 1993; Dunn and Grabski, 2001). Participants are undergraduate accounting and MIS students who participated in the experiment for class credit. These subjects have little, if any, prior exposure to data modeling and query writing.

Participants within both majors were randomly assigned to four groups. The first two groups received training on the ER model. The other two groups received training on the relational model. Each group was trained on the use of a particular query tool: SQL or QBE. Then, they were asked to write a series of queries using the query tool and database representation on which they have been trained. All participants received a description of the same database and one type of database representation according to their prior database structure representation training. Relational and ER documentation were based on the same database domain, and both sets of system documentation contained all information needed to complete the experimental tasks.

Experimental Protocol

The different parts of the experiment were divided into three phases: registering, training, and testing. Both training and testing phases were conducted in a computer laboratory. Participants were required to complete all parts of the experiment to receive full credit.

For the training and testing phases, a web application was designed to administer both parts of the experiment. To develop the training and testing materials, standard database management textbooks were consulted (e.g., McFadden et al., 1999; Pratt, 2001; Pratt and Adamski, 2002). In addition, expert faculty reviewed the experimental materials and their recommendations were reflected in the final draft of the experiment.

Participants completed the training and testing phases during two different sessions separated by a week. During the training phase, participants were given general instructions on how to complete this phase. Then, the students completed a demographic questionnaire. In addition to age, gender, and major, level of experience with databases and query languages was evaluated. A similar demographic questionnaire was used in prior studies (e.g., Leitheiser and March, 1996).

In the first part of the training, participants received instruction on understanding a database structure representation, either ER or relational. The database structure representation training included a description of important concepts in relational databases, such as primary keys and foreign keys. Also, it describes characteristics of one particular database structure representation. At the end of this training, participants were asked a series of multiple-choice questions to measure their understanding of database structure concepts. Participants received explanatory feedback on each of their answers.

After completing the database structure representation training, participants received training on a database query technique (QBE or SQL). To maintain consistency, the same domain (purchasing cycle) was used in the query language training. Topics used in the query language training include: simple retrieval, conditional selection, compound conditions, aggregate functions, sorting, grouping, and joining tables. These topics represent the major parts of select queries. For each of these topics, participants were presented with a sample query used to illustrate the concepts. The QBE grid is adapted from the one used by a popular database software program (Microsoft® Access).

At the end of each query topic, participants were able to practice their procedural knowledge of the query tool and received explanatory feedback in the form of the correct answer and explanations for it. Practice with explanatory feedback has been shown to increase procedural knowledge (Bonner and Walker, 1994). At the end of the query language training, participants received a summary of the database structure representation characteristics and the query language syntax and procedures.

Microsoft® Word was used to calculate readability scores for the content of the training. The Flesch Reading Ease score (Flesch 1948) rates the text on a 100-point scale. A higher score indicates a more understandable document. The Flesch Reading Ease score (Kincaid and McDaniel, 1974) for the training was 51. The Flesch-Kincaid Grade Level score rates the text on a U.S. grade school level. The training received a score of 9.3. This indicates that a ninth grader can understand the document. These scores are appropriate for upper-level undergraduate college students.

In the testing phase of the experiment, the students were asked to construct eight queries based on a new database domain (customers and parts). This test consists of four parts. In the first part, the participants reviewed the training material using a review sheet. The participants were told that they could view this page anytime during the test. In the second part, the participants received a description of a database structure used by a company to store its sales order transactions. Along with the description, the participants had access to the database structure representation they have been trained to understand (ER or relational).

In the third part of the testing phase, the participants were asked to provide their answers to eight different queries using the query language with which they had been trained (QBE or SQL). After each query, the students indicated their confidence level regarding their answer and their opinion about the query complexity. The order of the queries was randomized to avoid any order effect. The last part of the experiment consists of a questionnaire measuring the perceived ease-of-use. A week after the testing phase, all participants received a debriefing on the experiment and the correct answer for each query.

Independent Variables

Query language and *data model* are the between-subjects independent variables manipulated in this study. *Query complexity* is the within-subjects independent variable. Each participant completed eight queries. The queries are classified as simple or complex. *Query complexity* is based on the measurement scheme proposed by Reisner (1981) and the ranking of subject matter experts. A query is classified as simple if it requires the use of simple mapping (one table), simple selection, and/or simple condition. After completing each query question, the participants were asked to indicate their opinion about the level of complexity of the query they just finished to answer. The query complexity level is measured using a 7-point Likert scale question where 1 is very simple and 7 very complex. The complexity data is used as a manipulation check. Four queries were classified as simple because those queries only require one table and simple condition to obtain the information. Also, expert faculty in databases ranked those queries as simple. Every participant ranked those as simple. The other four queries are classified as complex because they require two or more tables, compound criteria, and an expression builder to provide the information. Every participant ranked those as the most complex.

In addition, the means for simple queries and complex queries were compared for both accounting and MIS participants. The difference between simple and complex queries across manipulation groups is significant ($p<0.001$). The accounting and MIS groups agree on the level of complexity for the complex queries ($F = 2.42$, $p = 0.123$). However, their opinions differ about the level of complexity for the simple queries ($F = 25.41$, $p<0.000$). The difference in complexity rating between simple and complex queries for the MIS groups is more defined than for the accounting groups. This finding should not have any impact on the hypotheses testing results because complexity is the within-subject manipulation and both majors displayed a clear cut between the two sets of queries, thus providing evidence for successful manipulation. Age, gender, major, and experience are included as covariates.

Dependent Variables

The dependent variables measured in this study are *query accuracy*, *query task completion time*, *user confidence*, and *perceived ease of use*. To measure *query accuracy*, query solutions were developed, and a specific grading protocol was agreed upon and applied by two independent graders.

The graders scored all participants' solutions for one query and then moved on to the next query. Possible total scores for each query vary because of the different complexity levels, as well as the number of fields, tables, criteria, and links associated with each query. In studies using more than one rater, measuring inter-rater agreement is important. The Kappa statistic was used to evaluate the extent of agreement between raters (Cohen, 1960). Individual query Kappa statistics were above 0.8 and the overall Kappa statistic is 0.91 which reflects almost perfect agreement between the two raters (p=<0.0001). The *query task completion time* is measured as the number of seconds participants spent completing each of the eight queries.

User confidence is measured separately for each query task using an 11-point scale anchored at 0% (extremely unconfident) and 100% (extremely confident). *Perceived ease of use* is measured after the participants have completed all of the query tasks. *Perceived ease of use* is measured using five 7-point Likert scale questions adapted from Davis (1989). The original instrument and adaptations have been used in prior studies with high reported reliability (Cronbach's alpha has ranged from 0.83 in Batra et al. [1990] to 0.93 in Amer [1993]). Cronbach's coefficient alpha (Cronbach, 1951) was computed to assess the level of internal consistency reliability for the perceived ease-of-use construct. The result (0.84) is comparable to the Cronbach's coefficient alpha reported in prior studies.

RESULTS

One hundred sixty-one undergraduate students majoring in accounting (78) and MIS (44) participated in the experiment for class credit. Accounting students were registered in their first introductory AIS course. MIS students were registered in their first introductory database course. Those students had not yet been exposed to any database concepts. After completion of the experiment, data for 123 participants were usable for analysis. The reduction in the number of participants is attributable to technical problems when implementing the experimental materials and to some participants who did not fully complete the experiment nor provide answers to every question.

Demographic Statistics

Demographic data of the accounting participants for the four different groups (ER/SQL, ER/QBE, Relational/SQL, and Relational/QBE) are reported in table 1. The number of participants in the four groups is similar ($\chi^2 = 0.84$, p = 0.358). No differences in gender are found among the four groups ($\chi^2 = 1.28$, p = 0.733). In each of the four groups, female participants form the majority. Concerning age, data for a 47 year-old participant were deleted from the sample because of their large effect. After the elimination, no significant differences in age are found among the treatment groups (F = 3.02, p = 0.087). The means range between 20 and 22 years of age. No significant differences among groups exist based on the number of courses taken prior to the experiment that deal with productivity software, programming languages, databases design, and databases software. The most experience that accounting participants received prior to the experiment is in the number of courses with productivity software as the main topic (one or two courses).

Table 1
Participant Demographic Statistics

Accounting		ER/SQL	ER/QBE	Rel/SQL	Rel/QBE	Test statistic	p-value
Number of participants		22	16	19	21	0.84†	0.358
Gender:	Male [Female]	10 [12]	6 [10]	7 [12]	11 [10]	1.28†	0.733
Age:	Mean (StDev) [Median]	20.5 (1.79) [20]	20.87* (1.13) [21]	21.7 (3.04) [21]	20.5 (1.25) [20]	3.02‡	0.087
Course #1[a]:	Mean (StDev) [Median]	1.4 (0.66) [1]	1.5 (0.97) [1.5]	1.6 (0.83) [2]	1.6 (1.03) [2]	0.24 ‡	0.624
Course #2[b]:	Mean (StDev) [Median]	0.5 (0.86) [0]	0.5 (0.73) [0]	0.2 (0.42) [0]	0.3 (0.56) [0]	0.16 ‡	0.693
Course #3[c]:	Mean (StDev) [Median]	0.3 (0.55) [0]	0.4 (0.72) [0]	0.3 (0.58) [0]	0.4 (0.59) [0]	0.02 ‡	0.894
Course #4[d]:	Mean (StDev) [Median]	0.0 (0.00) [0]	0.2 (0.408) [0]	0.0 (0.00) [0]	0.1 (0.22) [0]	2.05 ‡	0.156
MIS		**ER/SQL**	**ER/QBE**	**Rel/SQL**	**Rel/QBE**	**Test statistic**	**p-value**
Number of participants		12	7	14	11	0.23†	0.632
Gender:	Male [Female]	8 [4]	5 [2]	12 [2]	7 [4]	1.88†	0.598
Age:	Mean (StDev) [Median]	20.3 (1.56) [20]	20.7 (1.50) [20]	22.0 (4.49) [20]	21.2 (2.14) [20]	0.43‡	0.518
Course #1[a]:	Mean (StDev) [Median]	1.1 (0.79) [1]	1.0 (0.82) [1]	1.3 (0.73) [1]	0.9 (1.04) [1]	0.31 ‡	0.582
Course #2[b]:	Mean (StDev) [Median]	2.9 (0.90) [3]	2.3 (0.76) [2]	2.6 (1.22) [3]	3 (0.89) [3]	2.91 ‡	0.096
Course #3[c]:	Mean (StDev) [Median]	0.1 (0.29) [0]	0.3 (0.49) [0]	0.1 (0.36) [0]	0.4 (0.67) [0]	0.00 ‡	0.950
Course #4[d]:	Mean (StDev) [Median]	0.0 (0.00) [0]	0.1 (0.38) [0]	0.1 (0.36) [0]	0.1 (0.30) [0]	1.12 ‡	0.296

[a] Number of courses - main topic: productivity software
[b] Number of courses - main topic: programming languages
[c] Number of courses - main topic: databases design
[d] Number of courses - main topic: databases software
† χ^2-statistic
‡ F-statistic
* One observation with a value of 47 for age was deleted from the sample because of large effect on the sample. Including this data will change the mean (standard deviation) to 22.5 (6.62) and change F (p-value) to 3.96 (0.05)

Table 1 also presents the demographic data for the four MIS groups. Nonparametric tests, to evaluate the equal sample sizes among the four groups, resulted in no significant differences in terms of number of participants ($\chi^2 = 0.23$, $p = 0.632$). No significant differences were found among the four groups based on gender, age, and prior educational experiences.

In contrast to the accounting participants, more MIS participants are males. MIS students have more programming background ($median_{course\#2,\ MIS} = 3$, $median_{course\#2,\ Acc} = 0$). This finding should impact participant performance in completing the query task. MIS groups who used SQL as a query tool may be more comfortable typing the SQL code than using the mouse. These differences in educational experience and gender between the accounting and MIS participants are the reason for separating the two groups when investigating the results.

Results of Hypothesis Testing

The hypotheses and the research question were analyzed using a repeated measures general linear model. The factors data model and query language are crossed factors while complexity is a repeated measures factor. User characteristics defined by age, gender, experience are included in the model as covariates. Experience is the total number of courses with productivity software, programming languages, database design and software as topics. Query accuracy, task completion time, user confidence, and perceived ease-of-use were each analyzed separately.

Statistical analysis was computed first by including both accounting and MIS participants as part of the sample. Major was one of the covariates and was significant for query accuracy, user confidence, and perceived ease-of-use ($p = 0.003$, $p = 0.001$, $p<0.001$, respectively). General linear model was computed to see the existence of a three-way interaction among data model, query language, and major. Only the three-way interaction is significant for perceived ease-of-use ($F = 5.88$, $p = 0.017$). The following subsections present the results for each type of participant treated separately.

Following the ANOVA for each dependent variable, the hypotheses were tested by contrasting combinations of treatments that have similar characteristics. Matched groups' means (ER/QBE and Relational/SQL) were combined and compared with unmatched groups' means combination (ER/SQL and Relational/QBE). The post hoc tests include t-tests on the difference between the two pairs of contrast groups' means. Results of Tukey HSD pairwise contrasts are reported.

H1: Query Accuracy

Hypothesis 1 (H1) states that ER/QBE and Relational/SQL participants will generate more correct queries than Relational/QBE and ER/SQL participants. Table 2, panel A presents ANOVA results with query accuracy as the dependent variable for the accounting participants and the MIS participants.

Table 2
Analysis of Variance

Panel A – Query Accuracy as Dependent Variable

		Task Complexity							
		Simple Queries				*Complex Queries*			
Source	*d.f.*	*F*-Statistic		*p*-value		*F*-Statistic		*p*-value	
Independent Variables:		Acc	MIS	Acc	MIS	Acc	MIS	Acc	MIS
Data Model	1	0.08	1.69	0.78	0.20	1.57	1.54	0.21	0.22
Query Language	1	2.60	3.62	0.11	0.07*	1.10	5.05	0.30	0.03**
Data Model x Query Language	1	0.13	5.63	0.72	0.02**	1.22	3.90	0.27	0.06*
Covariates:									
Age	1	0.69	0.83	0.41	0.37	0.83	0.61	0.37	0.44
Gender	1	2.70	1.38	0.11	0.25	2.90	3.95	0.09*	0.05*
Experience	1	0.42	0.18	0.52	0.68	2.74	1.25	0.10*	0.27

Panel B – Time Completion as Dependent Variable

		Task Complexity							
		Simple Queries				*Complex Queries*			
Source	*d.f*	*F*-Statistic		*p*-value		*F*-Statistic		*p*-value	
Independent Variables:		Acc	MIS	Acc	MIS	Acc	MIS	Acc	MIS
Data Model	1	0.00	0.59	0.97	0.45	0.04	0.56	0.83	0.46
Query Language	1	15.02	27.29	0.00**	0.00**	5.86	11.94	0.02**	0.00**
Data Model x Query Language	1	0.00	8.68	0.97	0.01**	0.97	0.52	0.33	0.47
Covariates:									
Age	1	7.51	0.64	0.01**	0.43	1.62	1.06	0.21	0.31
Gender	1	1.98	0.17	0.16	0.69	0.87	1.09	0.36	0.30
Experience	1	0.41	1.95	0.52	0.17	0.55	0.28	0.83	0.60

Panel C – User Confidence as Dependent Variable

		Task Complexity							
		Simple Queries				*Complex Queries*			
Source	*d.f.*	*F*-Statistic		*p*-value		*F*-Statistic		*p*-value	
Independent Variables:		Acc	MIS	Acc	MIS	Acc	MIS	Acc	MIS
Data Model	1	2.25	0.02	0.14	0.90	1.06	1.29	0.31	0.26
Query Language	1	0.78	0.19	0.38	0.67	3.45	3.05	0.07*	0.09*
Data Model x Query Language	1	0.04	4.45	0.84	0.04**	0.98	1.82	0.33	0.19
Covariates:									
Age	1	0.19	0.37	0.66	0.55	0.62	0.61	0.43	0.44
Gender	1	1.10	0.00	0.30	0.95	0.68	0.45	0.41	0.51
Experience	1	3.00	1.09	0.09*	0.30	0.81	0.55	0.37	0.46

Panel D – Perceived Ease-of-Use as Dependent Variable

Source	*d.f.*	*F*-Statistic		*p*-value	
Independent Variables:		Acc	MIS	Acc	MIS
Data Model	1	0.70	1.08	0.41	0.31
Query Language	1	0.17	0.09	0.69	0.76
Data Model x Query Language	1	2.12	13.03	0.15	0.00**
Covariates:					
Age	1	3.19	0.15	0.09*	0.70
Gender	1	8.26	0.24	0.01**	0.63
Experience	1	4.74	0.77	0.03**	0.39

* Significant at 0.10 level.
** Significant at 0.05 level.

For the accounting participants, no interaction or main effects for data model and query language were found (F = 0.13, p = 0.72). Gender and experience have a marginally significant impact on the complex query accuracy performance (F = 2.90, p = 0.09; F = 2.74, p = 0.10, respectively). For the MIS participants, the data model and query language interaction effect was significant (F = 5.63, p = 0.02 for simple queries; and F = 3.90, p = 0.06 for complex queries).

Only gender was significant at the 0.055 level when MIS participants completed complex queries.

Table 3 reports on contrasts mean comparison of query accuracy scores for the accounting participants. Panel A shows no significant differences for simple queries between the matched groups and the unmatched groups (t = -0.6, p = 0.551). H1 is not supported for the accounting participants. Pairwise mean comparisons show no significant differences between the four groups. Panel B shows similar results for complex queries.

Table 3
Accounting Participants Means Comparison of Query Accuracy Scores[a]

Panel A: Complexity Level – Simple Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_1 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-4.87	8.12	-0.6	74	0.551

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/SQL	Rel/SQL	Rel/QBE	ER/QBE
ER/SQL	86.49 (17.64)	-	0.553	1.476	1.676
Rel/SQL	83.41 (19.00)		-	0.875	1.113
Rel/QBE	78.48 (14.81)			-	0.303
ER/QBE	76.69 (20.02)				-

Panel B: Complexity Level – Complex Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_1 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	6.55	7.04	0.93	74	0.355

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/QBE	ER/SQL	Rel/QBE	Rel/SQL
ER/QBE	57.38 (16.65)	-	1.303	1.463	1.440
ER/SQL	50.78 (17.84)		-	0.188	0.193
Rel/QBE	49.89 (11.58)			-	0.010*
Rel/SQL	49.85 (15.05)				-

[a] Accuracy scores are the average score of the individual queries converted as a percentage.
* The mean difference is significant at the 0.05 level.

Table 4 presents the mean comparison results for the MIS participants. ER/SQL and Relational/QBE groups significantly outperformed the matched group for simple queries and complex queries (t = 2.73, p = 0.01; and t = 2.43, p = 0.02, respectively). ER/SQL group had the highest score for the simple queries and Relational/QBE group for more complex queries. H1 is not supported.

Table 4
MIS Participants Means Comparison of Query Accuracy Scores[a]

Panel A: Complexity Level – Simple Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_1 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-16.34	6.01	2.73	39	0.010*

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/SQL	Rel/QBE	Rel/SQL	ER/QBE
ER/SQL	96.53 (5.21)	-	0.212	0.930	3.034*
Rel/QBE	95.67 (8.57)		-	0.665	2.744*
Rel/SQL	93.05 (7.38)			-	2.327
ER/QBE	82.78 (17.67)				-

Panel B: Complexity Level – Complex Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_1 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-21.40	8.79	2.43	39	0.020*

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: Rel/QBE	ER/SQL	ER/QBE	Rel/SQL
Rel/QBE	75.00 (14.66)	-	2.851*	2.491	3.690*
ER/SQL	57.99 (11.56)		-	0.015	0.780
ER/QBE	57.89 (17.27)			-	0.648
Rel/SQL	53.71 (13.53)				-

[a] Accuracy scores are the average score of the individual queries converted as a percentage.
* The mean difference is significant at the 0.05 level.

H2: Task Completion Time

Hypothesis 2 (H2) states that task completion time will be lower for the ER/QBE and the Relational/SQL groups than for the Relational/QBE and ER/SQL groups. Query task completion time ANOVA is reported in table 3, panel B for accounting participants and MIS participants.

For the accounting participants, no interaction effect was found. A main effect of query language was observed for both levels of query complexity ($F = 15.02$, $p < 0.001$; and $F = 5.86$, $p = 0.02$, respectively). Age was significant for simple queries ($F = 7.51$, $p = 0.01$).

Table 5 reports on task completion time mean comparison for the accounting participants. Both QBE groups took less time to complete the set of queries than the SQL groups. The difference between the Relational/QBE group and the other two SQL groups is significant for simple queries. The difference between Relational/QBE and Relational/SQL groups is significant for complex queries.

Table 5
Accounting Participants Means Comparison of Query Task Completion Time[a]

Panel A: Complexity Level – Simple Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_2 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	0:56	1:18	0.72	74	0.473

Test Statistics and Results of Tukey HSD Contrasts (results sorted in ascending order by mean)

Group	Statistics mean (StDev)	Group: Rel/QBE	ER/QBE	ER/SQL	Rel/SQL
Rel/QBE	05:17 (1:56)	-	-0.345	-2.954*	-3.524*
ER/QBE	05:36 (4:05)		-	-2.395	-2.951*
ER/SQL	07:50 (2:17)			-	-0.685
Rel/SQL	08:27 (3:00)				-

Panel B: Complexity Level – Complex Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_2 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	2:36	2:04	1.26	74	0.212

Test Statistics and Results of Tukey HSD Contrasts (results sorted in ascending order by mean)

Group	Statistics mean (StDev)	Group: Rel/QBE	ER/QBE	ER/SQL	Rel/SQL
Rel/QBE	09:17 (5:01)	-	-0.552	-1.723	-2.896*
ER/QBE	10:07 (4:11)		-	-1.043	-2.162
ER/SQL	11:40 (3:50)			-	-1.249
Rel/SQL	13:26 (4:58)				-

[a] All times are reported in minutes and seconds.
* The mean difference is significant at the 0.05 level.

Table 6 presents similar results for the MIS participants in term of completion time means. Both QBE groups outperformed in term of efficiency the other two SQL groups. The difference between the two QBE groups is not significant for both levels of complexity. H2 is not supported for both majors.

Table 6
MIS Participants Means Comparison of Query Task Completion Time[a]

Panel A: Complexity Level – Simple Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_2 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	3:18	1:08	2.93	39	0.006*

Test Statistics and Results of Tukey HSD Contrasts (results sorted in ascending order by mean)

Group	Statistics mean (StDev)	Group: Rel/QBE	Group: ER/QBE	Group: ER/SQL	Group: Rel/SQL
Rel/QBE	3:52 (0:56)	-	-1.514	-3.235*	-6.012*
ER/QBE	5:12 (2:23)		-	-1.344	-3.766*
ER/SQL	6:20 (1:16)			-	-2.807*
Rel/SQL	8:18 (2:14)				-

Panel B: Complexity Level – Complex Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_2 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	1:13	2:23	0.51	39	0.616

Test Statistics and Results of Tukey HSD Contrasts (results sorted in ascending order by mean)

Group	Statistics mean (StDev)	Group: ER/QBE	Group: Rel/QBE	Group: ER/SQL	Group: Rel/SQL
ER/QBE	09:02 (5:08)	-	-0.396	-1.803	-2.963*
Rel/QBE	09:46 (3:01)		-	-1.546	-2.841*
ER/SQL	12:17 (2:32)			-	-1.308
Rel/SQL	14:14 (4:22)				-

[a] All times are reported in minutes and seconds.
* The mean difference is significant at the 0.05 level.

H3: User Confidence

Hypothesis 3 (H3) states that ER/QBE and Relational/SQL participants will be more confident about their answers than Relational/QBE and ER/SQL participants. Results for the user confidence are similar to those of query accuracy. Those two variables are correlated ($p<0.001$, see table 7 for accounting groups and table 8 for MIS participants). Table 2, panel C reports the ANOVA with user confidence as the dependent variable for both types of participants.

Table 7
Accounting Participants Pairwise Correlation Matrix – Dependent Variables, Independent Variables, and Covariates

	Age	Gender	Experience	Model	Language	Score_S	Score_C	Time_S	Time_C	Conf_S	Conf_C
Gender	0.026										
	0.822										
Experience	0.136	-0.007									
	0.236	*0.952*									
Model	-0.041	-0.029	-0.038								
	0.720	*0.800*	*0.742*								
Language	0.043	-0.045	0.076	0.104							
	0.710	*0.695*	*0.509*	*0.365*							
Score_S	-0.25	0.176	0.022	-0.043	-0.207						
	*0.027**	*0.123*	*0.847*	*0.706*	*0.069*						
Score_C	-0.167	0.194	0.17	-0.12	0.091	0.692					
	0.143	*0.089*	*0.137*	*0.294*	*0.429*	*0.000**					
Time_S	0.124	0.129	0.045	-0.018	-0.437	0.213	0.105				
	0.280	*0.259*	*0.698*	*0.873*	*0.000**	*0.061*	*0.361*				
Time_C	-0.008	0.099	0.055	0.026	-0.303	0.43	0.411	0.732			
	0.945	*0.386*	*0.635*	*0.821*	*0.007**	*0.000**	*0.000**	*0.000**			
Conf_S	-0.072	0.114	0.197	-0.17	0.089	0.493	0.406	0.003	0.197		
	0.529	*0.322*	*0.084*	*0.137*	*0.440*	*0.000**	*0.000**	*0.977*	*0.084*		
Conf_C	0.07	0.067	0.127	-0.103	0.204	0.204	0.291	0.123	0.167	0.712	
	0.543	*0.561*	*0.268*	*0.370*	*0.073*	*0.073*	*0.010**	*0.285*	*0.144*	*0.000**	
Ease	0.068	0.272	0.232	-0.096	-0.073	0.234	0.261	0.254	0.225	0.427	0.529
	0.553	*0.016**	*0.041**	*0.403*	*0.524*	*0.039**	*0.021**	*0.025**	*0.048**	*0.000**	*0.000**

Pearson product-moment correlations are reported with P-value in italic.
Variable definition: Model: data model; Language: query language; Score_S: query accuracy score (simple queries); Score_C: query accuracy score (complex queries); Time_S: query task completion time (simple queries); Time_C: query task completion time (complex queries); Conf_S: user confidence level (simple queries); Conf_C: user confidence level (complex queries); Ease: perceived ease-of-use.
* Significant at 0.05 level.

For the accounting participants, the results do not show any interaction and main effects for simple tasks. Only experience is marginally significant at the 0.1 level. For the MIS participants, the data model and query language interaction effect was found to be significant for simple queries ($F = 4.45$, $p = 0.04$). None of the user characteristics was found to have an effect on the user confidence for writing simple queries. For complex queries, a main effect of query language was observed for both accounting and MIS participants ($F = 3.45$, $p = 0.07$; $F = 3.05$, $p = 0.09$, respectively).

Table 8
MIS Participants Pairwise Correlation Matrix –
Dependent Variables, Independent Variables, and Covariates

	Age	Gender	Experience	Model	Language	Score_S	Score_C	Time_S	Time_C	Conf_S	Conf_C
Gender	-0.071										
	0.650										
Experience	0.591	-0.013									
	*0.000**	*0.933*									
Model	0.156	-0.073	0.097								
	0.319	*0.643*	*0.534*								
Language	0.098	0.103	-0.003	0.091							
	0.532	*0.512*	*0.986*	*0.564*							
Score_S	-0.368	0.188	-0.364	-0.025	-0.265						
	*0.015**	*0.226*	*0.016**	*0.875*	*0.085*						
Score_C	-0.212	0.339	-0.108	0.070	0.259	0.639					
	0.172	*0.026**	*0.491*	*0.655*	*0.093*	*0.000**					
Time_S	-0.116	-0.164	-0.212	0.077	-0.626	0.175	-0.257				
	0.458	*0.294*	*0.172*	*0.624*	*0.000**	*0.261*	*0.097*				
Time_C	0.078	0.062	-0.020	0.085	-0.472	0.402	0.141	0.602			
	0.617	*0.692*	*0.900*	*0.586*	*0.001**	*0.007**	*0.366*	*0.000**			
Conf_S	-0.067	0.038	-0.113	-0.092	-0.019	0.606	0.582	-0.210	0.350		
	0.668	*0.810*	*0.469*	*0.556*	*0.905*	*0.000**	*0.000**	*0.176*	*0.021**		
Conf_C	-0.141	-0.021	-0.168	-0.202	0.255	0.353	0.521	-0.362	-0.153	0.614	
	0.369	*0.892*	*0.281*	*0.194*	*0.099*	*0.020**	*0.000**	*0.017**	*0.327*	*0.000**	
Ease	-0.042	0.160	0.169	-0.232	0.032	0.406	0.494	-0.305	0.113	0.522	0.468
	0.788	*0.307*	*0.278*	*0.134*	*0.840*	*0.007**	*0.001**	*0.046**	*0.469*	*0.000**	*0.002**

Pearson product-moment correlations are reported with P-value in italic.
Variable definition: Model: data model; Language: query language; Score_S: query accuracy score (simple queries); Score_C: query accuracy score (complex queries); Time_S: query task completion time (simple queries); Time_C: query task completion time (complex queries); Conf_S: user confidence level (simple queries); Conf_C: user confidence level (complex queries); Ease: perceived ease-of-use.
* Significant at 0.05 level.

Table 9 compares means for the accounting groups. No significant differences were found regardless of the level of query complexity. Table 10 reports the results for the MIS participants. For simple queries, a significant difference exists between the matched group and the unmatched group but in the opposite hypothesized direction ($t = -2.95$, $p = 0.005$). No significant difference was observed for complex queries. H3 is not supported for both types of participants.

Table 9
Accounting Participants Means Comparison of User Confidence[a]

Panel A: Complexity Level – Simple Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_3 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-2.1	9.722	-0.22	74	0.830

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/QBE	ER/SQL	Rel/QBE	Rel/SQL
ER/QBE	58.59 (18.17)	-	0.497	0.929	1.681
ER/SQL	55.11 (23.70)		-	0.475	1.299
Rel/QBE	52.02 (17.11)			-	0.827
Rel/SQL	46.45 (24.71)				-

Panel B: Complexity Level – Complex Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_3 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-5.99	8.303	-0.72	74	0.473

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/QBE	Rel/QBE	ER/SQL	Rel/SQL
ER/QBE	42.03 (14.67)	-	0.238	0.815	1.991
Rel/QBE	40.60 (14.94)		-	0.620	1.885
ER/SQL	37.16 (22.20)			-	1.302
Rel/SQL	29.74 (18.95)				-

[a] Confidence is measured in a scale anchored at 0% (extremely unconfident) and 100% (extremely confident).
[*] The mean difference is significant at the 0.05 level.

Table 10
MIS Participants Means Comparison of User Confidence[a]

Panel A: Complexity Level – Simple Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_3 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-26.20	8.88	-2.95	39	0.005*

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group			
		Rel/QBE	ER/SQL	Rel/SQL	ER/QBE
Rel/QBE	84.00 (13.03)	-	0.387	2.310	2.172
ER/SQL	81.67 (8.00)		-	2.009	1.902
Rel/SQL	70.54 (16.24)			-	0.247
ER/QBE	68.93 (18.59)				-

Panel B: Complexity Level – Complex Queries - Planned Contrast Test Results

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_3 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-16.04	8.87	-1.81	39	0.078

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group			
		Rel/QBE	ER/QBE	ER/SQL	Rel/SQL
Rel/QBE	58.00 (12.57)	-	0.484	0.913	3.120*
ER/QBE	54.64 (11.68)		-	0.320	2.276
ER/SQL	52.50 (13.73)			-	2.290
Rel/SQL	39.82 (16.18)				-

[a] Confidence is measured in a scale anchored at 0% (extremely unconfident) and 100% (extremely confident).
* The mean difference is significant at the 0.05 level.

H4: Perceived Ease-of-Use

Hypothesis 4 (H4) states that ER/QBE and Relational/SQL participants will perceive the combination as easier to use than Relational/QBE and ER/SQL participants. ANOVA with perceived ease-of-use as the dependent variable is reported in table 3, panel D for accounting and MIS participants.

No interaction effect or any main effect was observed for the perceived ease-of-use for the accounting participants. Gender and experience were found to have a significant effect on the perceived ease-of-use ($F = 8.26$, $p = 0.01$; $F = 4.74$, $p = 0.03$, respectively). Table 2 shows an interaction effect of data model and query language for the MIS participants ($F = 13.03$, $p = 0.001$). None of the user characteristics covariates were found to be significant.

No significant mean difference between the four accounting groups is reported (see table 11). For the MIS participants, ER/SQL and Relational/QBE combinations were found to be easier to use than ER/QBE and Relational/SQL (see table 11). H4 is not supported for accounting and MIS groups.

Table 11
Participants Means Comparison of Perceived Ease-of-Use

Planned Contrast Test Results: Accounting

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_4 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-0.36	0.498	-0.73	74	0.467

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/SQL	Rel/QBE	ER/QBE	Rel/SQL
ER/SQL	3.36 (1.20)	-	0.991	0.922	1.053
Rel/QBE	3.04 (0.95)		-	0.000	0.116
ER/QBE	3.04 (0.98)			-	0.108
Rel/SQL	3.00 (1.19)				-

Planned Contrast Test Results: MIS

Contrast	Means Difference	Standard Error	t	*df*	Prob.
$H_4 : \mu_{ER/QBE} + \mu_{Rel/SQL} > \mu_{ER/SQL} + \mu_{Rel/QBE}$	-2.46	0.59	-4.15	39	0.000*

Test Statistics and Results of Tukey HSD Contrasts (results sorted in descending order by mean)

Group	Statistics mean (StDev)	Group: ER/SQL	Rel/QBE	ER/QBE	Rel/SQL
ER/SQL	4.97 (0.82)	-	0.562	2.796*	3.882*
Rel/QBE	4.74 (1.05)		-	2.210	3.107*
ER/QBE	3.71 (1.14)			-	0.426
Rel/SQL	3.53 (0.86)				-

[a] Perceived Ease-of-Use is measured using the average score from the instrument's 5 questions using 7-point Likert scale where 1 (very difficult) and 7 (very easy)

* The mean difference is significant at the 0.05 level.

DISCUSSION AND CONCLUSION

The results for the hypothesis related to query accuracy are different for the two types of participants. The hypothesis, based on cognitive fit theory, that the combination of database structure representation and query language affect the efficiency of the end-users is not supported for the accounting group. Even though the four different manipulation groups show no significant differences, the two SQL groups outperformed the two QBE groups for simple queries with the group using a graphical representation of the database (ER model) receiving the highest score. When the task is more complex, the graphical combination (ER/QBE) performed the best. Again, the means differences among the four groups are not significant for complex queries. For novice users, the findings suggest that the ER/SQL combination is the most appropriate for a low level of task complexity. But if multiple tables are needed, resulting in higher task complexity, and the users must aggregate information, then the ER/QBE combination will help accountants completing the task. These findings suggest that the accountants should use a graphical representation of the database structure to help them querying the database and use different query languages based on the complexity of the task.

Gender and experience have a marginal impact of the accounting end-user's performance for complex queries. Females outperform males on complex query accuracy (mean [standard deviation]$_{acc,male}$ = 48.3 [15.9], compared with mean [standard deviation]$_{acc,\ female}$ = 54.3 [14.7]). This finding has been observed in prior studies investigating the impact of end-user gender on information system use (Henry and Stone, 1999) and decision analysis (Palvia and Gordon, 1992). Henry and Stone (1999) found that computer self-efficacy and outcome expectancies which have been shown to influence job performance and job satisfaction were higher for females than for males. Palvia and Gordon (1992) concluded that gender, even though this variable does not affect the performance with a particular mode of decision analysis (e.g., tables, trees, or formulas), is significant for the effectiveness in solving the decision analysis problem. This is similar to the current study's result for gender. Finally, the more educational experience the end-user received, the higher his performance is for complex task (with one computer-related course, mean$_{acc,\ complex}$ = 43.1, and with five computer-related courses, mean$_{acc,\ complex}$ = 66.7). This research points to the continuing need to consider user characteristics in query writing tasks.

For MIS end-users, the type of database structure representation combined with a particular query language has an influence on the query accuracy performance. The interaction is significant for simple queries, but marginally significant for complex queries. This can be explained by the fact that query language type is more important for higher level of task complexity. The results for the hypothesis are significant but opposite to the direction predicted. Cognitive fit theory does not explain the findings. The ER/SQL MIS group outperformed the other three groups. For complex queries, the Relational/QBE group provided the most accurate queries. Not surprisingly, the graphical query tool helps in performing more complex task. This finding confirms the conclusion of the query language studies (Greenblatt and Waxman, 1978; Thomas and Gould, 1975; Yen and Scamell, 1993). Interestingly, gender is marginally significant for complex queries. Female MIS users outperformed the male counterparts (mean [standard deviation]$_{MIS,\ male}$ = 55.8 [17.1], compared with mean [standard deviation]$_{MIS,\ female}$ = 68.8 [15.1]). As mentioned for the accounting end-users, knowing the impact of user characteristics such as gender and educational expertise on task performance will help professionals in their job assignment.

The overall results of the current study for query accuracy confirm the results found by Jih et al. (1989), which showed little difference in the performances using the ER representation and the relational model with SQL as query language, only for the accounting group. The MIS participants who used the graphical representation along with the SQL language to write simple queries performed the best. This result was found by Davis (1990), which used the same type of end-users. This combination may be the best for efficiency in writing queries, regardless of the end-user's educational background.

For task completion time, the results were not as predicted using the cognitive fit theory. Both accounting and MIS participants performed similarly, though the accounting groups took more time to complete the queries than the MIS groups. These results are not surprising and confirm what the prior studies on query languages found (e.g., Thomas and Gould, 1975; Greenblatt and Waxman, 1978; Yen and Scamell, 1993). Using a graphical query tool helps complete the task faster than with a non-graphical tool such as SQL.

The only difference between the accounting and MIS participants is that the MIS participants required less time to complete the queries. Running the general linear model, with major as one of the independent variables, produces no interaction between query language and major, and no main effect for major exists. Even though no difference in query accuracy performance exists for the accounting end-users, knowing which tool helps in completing the task faster is very useful.

The results for user confidence level were not as predicted based on the cognitive fit theory. This variable is correlated with query accuracy. The MIS participants consistently across groups reported higher level of confidence compared to the accounting groups. The accounting groups do not differ from one another. The difference between the MIS matched groups and the MIS unmatched groups is significant but in the opposite predicted direction. The unmatched groups submitted higher levels of confidence than the matched groups.

Results for perceived ease-of-use were not as predicted. The accounting groups showed no difference in opinion about how easy to use a particular combination is. This finding is in accordance with those found for the other performance variables of the accounting group. The ER/SQL combination was not as cumbersome and frustrating as the other combinations. Similar results were found for the MIS participants but the difference is significant but in the opposite predicted direction. The ER/SQL combination is the best combination. In terms of users' characteristics, the accounting analytic females with high educational experience find a particular combination easier to use than the other type of users. Knowing the users' characteristics will help professionals better assign user friendly combination of tools to end-users.

While previous studies manipulated only the database structure representation or the query language, this research project extended the current literature in accounting and information systems by manipulating both the database structure representation and the query tool. The previous studies were not theoretically based. The current study uses the cognitive fit theory to elaborate hypotheses. Even though using a theory is a step forward to increase the knowledge about database research, the findings cannot be explained by the chosen theory. For each measure of end-user query writing performance, the results are in the opposite direction or the combination of data model and query language does not affect the end-user performance. Some of the results confirm prior research findings.

Unlike prior research, user characteristics, such as professional field, gender, age, educational experience are explicitly included in the research model. The two types of professional fields (accounting and MIS) performed differently for certain measures of performance indicating that end-users will require access to different tools to complete the query writing task. Females and males performed differently. As prior research has recommended and this study has found, user characteristics are an important factor influencing end-user performance. Future research in this field should include those variables in their model.

The findings have practical implications. While the results from this study do not show that one combination of documentation and query tool is the best of the breed, the methodology of this research and the theories discussed can be utilized to improve both the efficiency and effectiveness of training programs implemented by organizations. Also Professional can increase their performance on an individual level by better understanding what type of documentation and query tool end-users need to improve performance in information retrieval tasks based on the measures of performance (query accuracy, time, confidence, and ease-of-use), and the user characteristics.

REFERENCES

Amer, T. S. (1993). Entity-relationship and relational database modeling representations for the audit review of accounting applications: an experimental examination of effectiveness. *Journal of Information Systems*, 7(Spring), 1-15.

Batra, D., J. A. Hoffer, and R. P. Bostrom (1990). Comparing representations with relational and EER models. *Communications of the ACM,* 33(February), 126-139.

Bonner, S. E. and P. L. Walker (1994). The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review,* 69(January), 157-178.

Borthick, A. F. (1992). Helping users to get the information they want, when they want it, in the form they want it: Integrating the choice and use of information. *Journal of Information Systems*, 6(Fall), v-ix.

Chamberlin, D. D., M. M. Astrahan, K. P. Ezwaran, P. P. Griffiths, R. A. Lorie, J. W. Mehl, P. Reisner, and B. W. Wade (1976). SEQUEL 2: A unified approach to data definition, manipulation, and control. *IBM Journal of Research and Development*, 20(November), 560-575.

Chan, H. C., K. K. Wei, and K. L. Siau (1993). User-database interface: the effect of abstraction levels on query performance. *MIS Quarterly*, 17(December), 441-464.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(February), 37-46.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16(September), 297-334.

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

Davis, J. S. (1990). Experimental investigation of the utility of data structure and E-R diagrams in database query. *International Journal of Man-Machine Studies,* 32(July), 449-459.

Dennis, A. R. and T. A. Carte (1998). Using geographical information systems for decision making: Extending cognitive fit theory to map-based presentations. *Information Systems Research*, 9(June), 194-203.

Dickson, G. W., J. A. Senn, and N. L. Chervany (1977). Research in management information systems: The Minnesota experiments. *Management Science,* 23(May), 913-923.

Dunn, C. L. and S. V. Grabski (2002). Empirical research in semantically modeled accounting systems. In *Researching Accounting as an Information Systems Discipline*, edited by V. Arnold and S. G. Sutton (pp. 157-180). Florida: American Accounting Association.

Dunn, C. L. and S. V. Grabski (2001). An investigation of localization as an element of cognitive fit in accounting model representations. *Decision Sciences,* 32(Winter), 55-94.

Flesch, R. F. (1948). A new readability yardstick. *Journal of Applied Psychology,* 32(June), 221-233.

Greenblatt, D. and J. Waxman (1978). A study of three database query languages. In *Databases: Improving Usability and Responsiveness, Proceedings of the International Conference on Databases: Improving Usability and Responsiveness,* edited by B. Schneiderman. New York: Academic Press.

Hayes, D. C. and J. E. Hunton (2001). When querying databases, you've got to ask the right question. *Journal of Accountancy,* 191(February), 35-45.

Henry, J. W. and R. W. Stone (1999). The impacts of end-user gender, education performance, and system use on computer self-efficacy and outcome expectancy. *Southern Business Review,* 25(Fall), 10-16.

Hooper, P. and J. Page (1996). Relational databases: an accountant's primer. *Management Accounting,* 78(October), 48-53.

Jarvenpaa, S. and J. J. Machesky (1989). Data analysis and learning: An experimental study of data modeling tools. *International Journal of Man-Machine Studies,* 31(September), 367-391.

Jih, W. K., D. A. Bradbard, C. A. Snyder, and 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(September), 257-267.

Kincaid, J. and W. McDaniel (1974). *An inexpensive automated way of calculating Flesch Reading Ease scores*. Patent Disclosure Document No. 031350,215 Patent Office, Washington, DC.

Leitheiser, R. L. and S. T. March (1996). The influence of database structure representation on database system learning and use. *Journal of Management Information Systems,* 12(March), 187-213.

Lochovsky, F. H. and D. C. Tsichritzis (1977). User performance considerations in DBMS selection. In *Proceedings of the 1977 ACM SIG-MOD Conference on Management of Data,* edited by D. C. P. Smith (pp. 128-134). Toronto, Canada.

McCarthy, W. E. (1982). The REA accounting model: A generalized framework for accounting systems in a shared data environment. *The Accounting Review,* 57(July), 554-578.

McFadden, F. R., J. A. Hoffer, and M. B. Prescott (1999). *Modern Database Management*. 5th edition. Menlo Park: Addison-Wesley Longman.

Palvia, S. C. and S. R. Gordon (1992). Tables, trees and formulas in decision analysis. *Communications of the ACM,* 35(October), 104-113.

Pratt, J. (1980). The effects of personality on subject's information processing: A comment. *The Accounting Review,* 55(July), 501-506.

Pratt, P. J. (2001). *A Guide to SQL*. 5th edition. Boston: Course Technology.

Pratt, P. J. and J. J. Adamski (2002). *Concepts of Database Management*. 4th edition. Cambridge: Thomson Learning.

Reisner, P. (1981). Human factors studies of database query languages: A survey and assessment. *ACM Computing Surveys*, 13(March), 13-31.

Reisner, P., R. F. Boyce, and D. D. Chamberlin (1975). Human factors evaluation of two data base query languages - SQUARE and SEQUEL. In *Proceedings of the AFIPS National Computer Conference* (pp. 447-452). Anaheim, California.

Sinha, A. P. and I. Vessey (1992). Cognitive fit: An empirical study of recursion and iteration. *IEEE Transactions on Software Engineering*, 18(May), 368-379.

Smelcer, J. B. and E. Carmel (1997). The effectiveness of different representations for managerial problem solving: Comparing tables and maps. *Decision Sciences,* 28(Spring), 391-420.

Thomas, J. C. and J. D. Gould (1975). A psychological study of query by example. In *National Computer Conference Proceedings* (pp. 439-445). New York: AFIPS Press.

Umanath, N. and I. Vessey (1994). Multiattribute data representation and human judgment: A cognitive fit perspective. *Decision Sciences,* 25(September-December), 795-824.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graph versus tables literature. *Decision Sciences*, 22(Spring), 219-240.

Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information and Management*, 27(August), 103-119.

Vessey, I. and D. Galletta (1991). Cognitive fit: An empirical study of information acquisition. *Information Systems Research,* 2(March), 63-84.

Vessey, I. and R. Weber (1986). Structured tools and conditional logic: An empirical investigation. *Communications of the ACM*, 29(January), 48-57.

Zloof, M. M. (1977). Query-by-Example: A data base language. *IBM Systems Journal*, 16(Winter), 324-343.

Yen, M. Y. and R. W. Scamell (1993). A human factors experimental comparison of SQL and QBE. *IEEE Transactions on Software Engineering*, 19(April), 390-409.