A graphical form of representation for database structure was designed to make explicit aspects of database structure that users must discover in order to compose successful queries. In a laboratory experiment, subjects wrote queries using the graphic style of representation as well as the less explicit textual style of documentation. Graphic and textual database representations were compared in both less complex (i.e., three-file) and more complex (i.e., six-file) databases. Results showed that subjects were faster, more accurate, and generally experience less difficulty in writing queries when using the graphic style of representation than when using the textual style of representation. Style of representation did not interact significantly with the complexity of the database in determining query writing performance. The findings from ratings on the difficulty of using the documentation showed that subjects rated the graphic documentation as more helpful only when it made explicit information for writing queries. Implications of the findings for database administration and directions for future research are discussed.

The database environment in organizations have been continuously evolving over the years. A large number of organizations have successfully established large-scale databases for day-to-day operations in various functional areas such as production, order processing, and inventory. To realize the full potential of information resource management (Guimaraes, 1988), however, it is imperative to leverage the value of these massive “production databases” and extract information from them to support managerial activities. An organization’s database administration function can help accomplish this goal by building separate “information databases” which contain summarized information from, and/or periodic snapshots of the production databases (Goodhue, Quilliard, and Rockart, 1988). End users in middle and upper management can access these information databases to satisfy unanticipated information needs. Indeed, the direct use of database query languages have become increasingly prevalent, as evidenced by the widespread adoption of the SQL standards by many popular mainframe, mini, as well as micro, DBMS vendors.

To use a query language effectively for extracting information from a database, a user must first comprehend the structure of the database. Typically, the only documentation about database structure available on-line in a database management system (DBMS) is a listing of the fields in each relational table. For each field, the documentation usually provides the name, abbreviation, description, data type, and length. The value of the documentation is often limited, however, by the fact that a listing can only be obtained for one table at a time. For complex databases involving more than two or three tables, the available on-line help may not be adequate to help users master the structure of the database.
More research is needed to understand what kind of documentation will make information about database structure most readily comprehensible to users. Is graphical representation better in this regard than the conventional textual documentation? In this study, we will attempt to answer this question by examining the results of a controlled laboratory experiment.

## Previous Studies

Whereas most of the human factors research on database systems has dealt with the query language part of the user interface (see Reisner, 1981 for a review), a second line of research has examined the effect of the data model on the performance of end users. Studies in this tradition (e.g., Lochofsky & Tschritzis, 1977; Brosey & Shneiderman, 1978) have compared hierarchical, network, and relational descriptions of database organization on the ability of users to write and understand queries and to remember the database structure. It can be argued that the recent dominance of SQL and relational databases has rendered the hierarchical vs. network vs. relational comparison moot. However, the issue of what constitutes an appropriate form of documentation about the structure of a relational database for end users who will be writing queries with SQL still remains.

Two approaches to the problem of documenting database structure can be identified. The first approach tries to identify a suitable end-user data model by comparing a variety of existing methods of database documentation. Some of the candidates tested, such as the entity-relationship diagram (Chen, 1976) were originally developed as tools for database design rather than as documentation for users who want to query a database. The motivation of the research in this group is to resolve competing claims among rival data modeling techniques. The study by Jih, Bradbard, Snyder and Thompson (1989) is an example of this approach. Jih et al. (1989) compared the effects on query writing of documentation based on the relational model with that based on the entity-relationship model. The results from the Jih et al. (1989) study showed that users of the relational documentation made fewer syntactic errors in composing queries than did users of the entity-relationship documentation. On the other hand, users of the entity-relationship documentation made fewer semantic errors and were faster in writing queries than were users of the relational documentation. Results from studies that compare specific techniques may help guide selection of a database documentation method, but they will probably not be helpful in understanding why one style of documentation is more effective than another.

A second approach seeks to test the hypothesis that graphical forms of database documentation will be more effective than textual forms of documentation. In this approach, researchers seek to classify existing methods of database representation as either graphic or textual and then to compare representatives of the graphic category with representatives of the textual category. For example, Davis (1990) compared documentation based on data-structure diagrams (Eliason, 1987) and entity-relation diagrams as examples of graphic documentation with ‘conventional on-line documentation’ as an example of textual documentation. Davis (1990) found that scores on a query writing task were higher for subjects using the graphic documentation than for subjects using the textual documentation. Contrasting findings were reported by Leitheiser (1988). In the Leitheiser study, three forms of graphic documentation were used. These were logical data structure diagrams (Carlis, 1985), a tabular display associated with the relational data model, and a second tabular display on which lines were drawn to link the tables with common attributes. A single example of textual documentation was also used. Leitheiser found no difference in query writing success between subjects using the two tabular forms of graphic documentation and the textual documentation. Subjects using the third form of graphic documentation, the logical data structure diagram, performed more poorly on the query writing task than did subjects using the textual and the tabular forms of documentation.

The results on graphic vs. textual forms of database documentation reveal the problems with the simple ‘graphics are better’ hypothesis. They are, in fact, reminiscent of the findings on the merits of graphic vs tabular displays for various decision making tasks. Reviews (e.g., DeSanctis, 1984) of this research have shown that graphic displays are not always better than tabular displays for tasks involving judgment and decision making. Cognitive scientists (e.g., Larkin & Simon, 1987) and researchers in management information systems (e.g., Benbasat, Dexter & Todd, 1986; Vessey, 1991) have concluded that the graphic displays are only better when they make explicit the information needed for a task. For example, when the task involves a judgment about trends, graphic displays are more effective than tabular displays. In contrast, when exact data values must be retrieved, tabular displays are more effective than graphic displays. The implication of the research on table vs. graphs is that graphic forms of database documentation are likely to be more effective than conventional textual forms of documentation when they provide users with the information about the structure of a database that is needed for writing queries.

## Research Objectives

The literature on data models contains suggestions about what features of database documentation will be important to users. Leitheiser (1988) argued that explicitness about ‘storage’ concepts is a crucial feature of database documentation. The following SQL query will retrieve the names of employees in the Accounting Department.

```sql
SELECT EmpName
FROM Employee Department
WHERE DeptNo.Employee = DeptNo.Department
```