An Experimental Investigation of Complexity in Database Query Formulation Tasks

Gretchen Irwin Casterella  
Leo Vijayasarathy  
Computer Information Systems Department  
College of Business  
Colorado State University  
Fort Collins, CO 80525-1277  
Gretchen.Irwin@business.colostate.edu, Leo.Vijayasarathy@business.colostate.edu

ABSTRACT
Information Technology professionals and other knowledge workers rely on their ability to extract data from organizational databases to respond to business questions and support decision making. Structured query language (SQL) is the standard programming language for querying data in relational databases, and SQL skills are in high demand and are taught in most introductory database courses. We examined students’ performance on query formulation tasks, in an experimental setting which varied the complexity of the query and the ambiguity of the information request. Our results confirm the main effects of query complexity and request ambiguity found in prior studies (Borthick et al. 2001). In addition, we found an interaction effect between complexity and ambiguity, namely that low ambiguity is more important as tasks increase in complexity. We also found that students’ confidence with entity-relationship diagrams corresponds to reduced time spent on query formulation, and their ability to evaluate the accuracy of their queries reduces as query complexity increases. We discuss the implications of these findings with some suggestions for future research.

Keywords: Query language, Database management systems (DBMS), Data modeling