



An Empirical Investigation of the Factors Affecting Data Warehousing Success

Author(s): Barbara H. Wixom and Hugh J. Watson

Source: *MIS Quarterly*, Vol. 25, No. 1 (Mar., 2001), pp. 17-41

Published by: Management Information Systems Research Center, University of Minnesota

Stable URL: <https://www.jstor.org/stable/3250957>

Accessed: 03-09-2018 16:41 UTC

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/3250957?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<https://about.jstor.org/terms>



JSTOR

Management Information Systems Research Center, University of Minnesota is collaborating with JSTOR to digitize, preserve and extend access to *MIS Quarterly*

AN EMPIRICAL INVESTIGATION OF THE FACTORS AFFECTING DATA WAREHOUSING SUCCESS¹

By: Barbara H. Wixom
McIntire School of Commerce
University of Virginia
Charlottesville, VA 22903
U.S.A.
bwixom@mindspring.com

Hugh J. Watson
Department of MIS
Terry College of Business
University of Georgia
Athens, GA 30602
U.S.A.
hwatson@terry.uga.edu

Abstract

The IT implementation literature suggests that various implementation factors play critical roles in the success of an information system; however, there is little empirical research about the implementation of data warehousing projects. Data warehousing has unique characteristics that may impact the importance of factors that apply to it. In this study, a cross-sectional survey investigated a model of data warehousing success. Data warehousing managers and data suppliers from 111 organizations completed paired mail questionnaires on implementation factors and the success of the warehouse. The results from a Partial Least

Squares analysis of the data identified significant relationships between the system quality and data quality factors and perceived net benefits. It was found that management support and resources help to address organizational issues that arise during warehouse implementations; resources, user participation, and highly-skilled project team members increase the likelihood that warehousing projects will finish on-time, on-budget, with the right functionality; and diverse, unstandardized source systems and poor development technology will increase the technical issues that project teams must overcome. The implementation's success with organizational and project issues, in turn, influence the system quality of the data warehouse; however, data quality is best explained by factors not included in the research model.

Keywords: Data warehousing, success, IS implementation, Partial Least Squares

ISRL Categories: HA03, FD, AI0610, EL03

Introduction

During the mid- to late 1990s, data warehousing became one of the most important developments in the information systems field. It is estimated that 95% of the *Fortune 1000* companies either have a data warehouse in place or are planning to develop one (META Group 1996). The Palo Alto Management Group predicts that the data ware-

¹Ron Weber was the accepting senior editor for this paper.

housing market will grow to a \$113.5 billion market in 2002, including the sales of systems, software, services, and in-house expenditures (Eckerson 1998). This is not surprising considering that for the past few years, surveys of CIOs have found data warehousing, Year 2000, and electronic commerce to be at the top of their strategic initiatives (Eckerson 1999).

A data warehouse (or smaller-scale data mart) is a specially prepared repository of data created to support decision making. Data are extracted from source systems, cleaned/scrubbed, transformed, and placed in data stores (Gray and Watson 1998). A data warehouse has data suppliers who are responsible for delivering data to the ultimate end users of the warehouse, such as analysts, operational personnel, and managers. The data suppliers make data available to end users either through SQL queries or custom-built decision-support applications (e.g., DSS and EIS).

Data warehousing is a product of business need and technological advances. The business environment has become more global, competitive, complex, and volatile. Customer relationship management and e-commerce initiatives are creating requirements for large, integrated data repositories and advanced analytical capabilities. More data are captured by organizational systems (e.g., barcode scanning, clickstream) or can be purchased from companies like Dun & Bradstreet and Harte Hanks. Through hardware advances such as symmetric multi-processing, massive parallel processing, and parallel database technology, it is now possible to load, maintain, and access databases of terabyte size. All of these changes are affecting how organizations conduct business, especially in sales and marketing, allowing companies to analyze the behavior of individual customers rather than demographic groups or product classes.

Even though there are many success stories (Beitler and Leary 1997; Grim and Thorton 1997), a data warehousing project is an expensive, risky undertaking. The typical project costs over \$1 million in the first year alone (Watson and Haley 1997). While hard figures are not available, it is estimated that one-half to two-thirds of all initial

data warehousing efforts fail (Kelly 1997). The most common reasons for failure include weak sponsorship and management support, insufficient funding, inadequate user involvement, and organizational politics (Watson et al. 1999).

Practitioners and researchers need to better understand data warehousing to ensure the success of these promising, yet risky and costly, IT undertakings. The IT literature contains many studies that investigate the factors that affect the implementation of decision-support applications (e.g., Guimares et al. 1992; Rainer and Watson 1995). While these studies are helpful, a data warehouse is arguably different in that it is an *IT infrastructure* project, which can be defined as a set of shared, tangible IT resources that provide a foundation to enable present and future business applications (Duncan 1995). The capability of such an infrastructure is thought to impact business value by supporting (or failing to support) important business processes (Ross et al. 1996). Few studies have examined the implementation success of infrastructure projects (Duncan 1995; Parr et al. 1999); instead, infrastructure research focuses on the innovation and diffusion of such phenomenon (for several examples, see Chau and Tam 1997; Prescott and Conger 1995).

There is considerable practitioner wisdom on the keys to data warehousing success; however, it is based on anecdotal evidence from a limited number of companies. There has been no academic research that systematically and rigorously investigates the keys to data warehousing success, using data collected from a large cross-section of firms. In this study, we investigate a research model of data warehousing implementation success using data gathered from mail surveys from 111 organizations. The study investigates the implementation of data warehousing in particular, and extends our knowledge of IT implementation in general.

This article first presents a research model for data warehousing implementation success that was developed from a literature review, an exploratory survey, and structured interviews. Next, it describes the cross-sectional survey that was used to collect data and the results from a

Partial Least Squares analysis of the research model. The findings are discussed in the concluding sections.

The Research Model

To develop the research model, the IT implementation, infrastructure, data warehousing, and success literature were reviewed to identify factors that potentially affect data warehousing success. After the literature review, survey data were collected from 126 attendees of a 1996 conference sponsored by The Data Warehousing Institute (TDWI). The survey contained two open-ended questions that asked for a list of critical success factors and obstacles to data warehousing success.² These findings, together with the literature review, were used to create an initial research model and to structure hour-long interviews with 10 data warehousing experts (e.g., book authors, consultants, and seminar speakers). The interviews confirmed that the research model contained appropriate factors and relationships among the model's factors. Minor changes were incorporated into the model based on the interviews.

Figure 1 presents the resulting research model. The rationale for the factors and the relationships among the factors are described in the following sections. Implementation factors, such as management support and user participation, are proposed to influence the success of the data warehouse implementation, which has been broken down into three unique facets. These include success with organizational, project, and technical issues that arise during the lifetime of the warehouse project. Thus, implementation success means that the project team has persuaded the organization to accept data warehousing, completed the warehouse according to plan, and overcome technical obstacles that arose. The success of the implementation in turn

²The actual survey questions were: What are the critical success factors for a data warehousing project? What are the biggest obstacles to a successful data warehousing project?

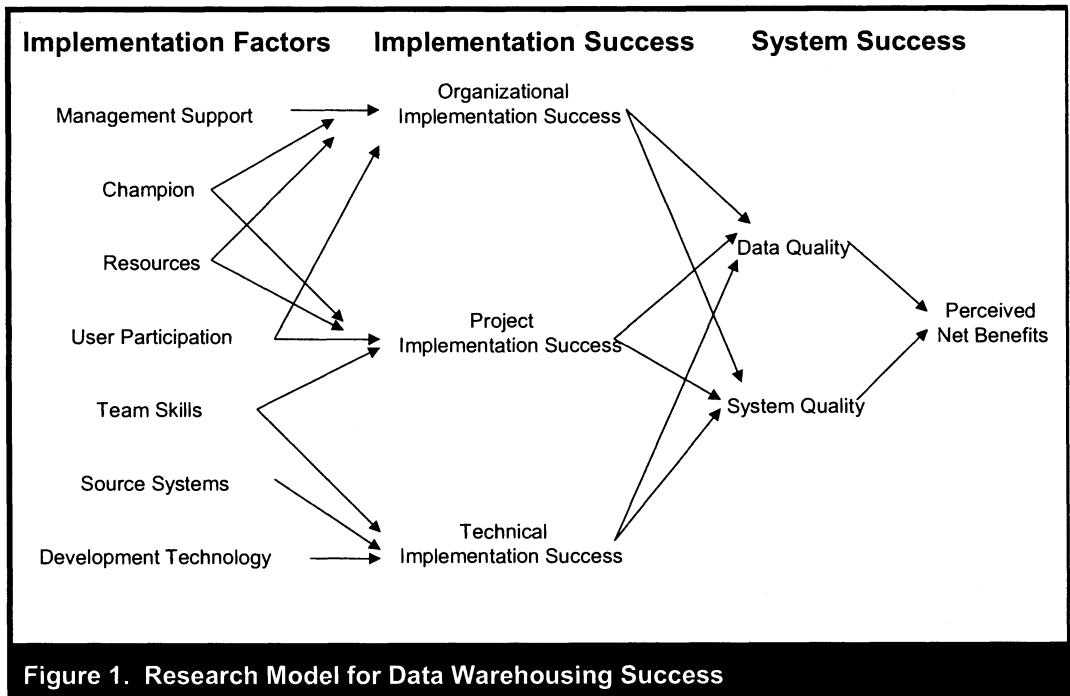
affects the system success, defined as the quality of the data warehouse system and its data. This impacts the perceived net benefits from the use of the warehouse.

Information Systems Success

Researchers have investigated the success of information systems in myriad ways (Garrity and Sanders 1998), such as by measuring the satisfaction of users (Melone 1990), service quality (Pitt et al. 1995), and the perceived usefulness of specific applications (Davis 1989; Moore and Benbasat 1991). Researchers should treat IS success as a multi-faceted construct, choose several appropriate success measures based on the research objectives and the phenomena under investigation, and consider possible relationships among the success dimensions when constructing a research model (DeLone and McLean 1992). Drawing on the work of Seddon (1997), three dimensions of system success were selected as being the most appropriate for this study: data quality, system quality, and perceived net benefits. Empirical studies (e.g., Fraser and Salter 1995; Seddon and Kiew 1994) have found that these three dimensions are related to one another: higher levels of data and system quality are associated with higher levels of net benefits.

Data quality refers to the quality of the data that are available from the data warehouse. This factor has received considerable research attention regarding its definition, component measures, and importance (e.g., Wand and Wang 1996; Wang and Strong 1996). Data quality is frequently discussed in the data warehousing literature as well; providing high-quality data to decision makers is the fundamental reason for a building a warehouse (Watson and Haley 1997). More specifically, data accuracy, completeness, and consistency are critical aspects of data quality in a warehouse (Lyon 1998; Shanks and Darke 1998).

With system quality, the focus is on the system itself. Commonly used performance measures include system flexibility, integration, response time, and reliability (DeLone and McLean 1992). Flexibility and integration are particularly important



for decision-support applications (Vandenbosch and Huff 1997). Systems that integrate data from diverse sources can improve organizational decision making (Wetherbe 1991; Wybo and Goodhue 1995), and flexibility allows decision makers to easily modify applications as their information needs change (Vandenbosch and Huff 1997). System quality (i.e., flexibility and integration) is one of the most important advantages for data warehousing because a warehouse provides the infrastructure that integrates data from multiple sources and flexibly supports current and future users and applications (Gray and Watson 1998; Sakaguchi and Frolick 1997).

A system displaying high data quality and system quality can lead to net benefits for various stakeholders, including individuals, groups of individuals, and organizations (Seddon 1997). It can give users a better understanding of the decision context, increase decision-making productivity, and change how people perform tasks. A data warehouse significantly affects how decision making for end users is supported in the organization because IT professionals no longer have to

extract data and run queries for users as in the past. When supplied with appropriate data access tools and applications, users can perform decision-making tasks faster and more comprehensively (Haley et al. 1999). In general, data warehousing can change the processes for providing end users with access to data and reduce the time and effort required to provide that access (Graham 1996).

Additional success dimensions were not included because they were considered less appropriate for this study than the selected constructs. User satisfaction measures are most often associated with an end-user's perception of a single application, but a data warehouse supports multiple applications rather than being an application itself. Organization-level benefits are difficult or impossible to assess and to isolate from other factors (e.g., actions of competitors) that affect the organization (Lucas 1981; Ragowsky et al. 1996). Extent of implementation has been applied frequently with large-scale systems, such as electronic data interchange (Masseti and Zmud 1996); however, these studies investigate the innovation and diffu-

sion of IT rather than IT implementation success. Also, successful data warehouses may or may not necessarily be implemented widely across an organization. For example, a warehouse may be used by only a few key analysts who are doing critically important work for the company, whereas other companies may find it useful to roll out data warehousing to the entire organization. Use has similar limitations because it is doubtful that frequent or widespread use can accurately identify a successful warehouse.

Data quality, system quality, and perceived net benefits were used in the research model as the three dimensions of data warehousing success. Based on past findings (Fraser and Salter 1995; Seddon and Kiew 1994) and the theoretical foundations developed by DeLone and McLean (1992) and Seddon (1997), we hypothesized:

H1a: A high level of data quality is associated with a high level of perceived net benefits.

H1b: A high level of system quality is associated with a high level of perceived net benefits.

Implementation Success

In the data warehousing literature, from the initial survey, and during interviews, three facets of warehousing implementation success were identified: success with organizational issues, success with project issues, and success with technical issues. These factors were believed to affect the ultimate success of the data warehouse. Of course, there likely are other facets of implementation success; however, to keep the research model to a manageable size, only three implementation success factors that were best supported by the study's model development phase were included. These are described in the following sections.

Organizational Implementation Success

An implementation is not successful unless the system it produces is accepted into the organiza-

tion and integrated into work processes. However, an information system implementation can cause considerable organizational change that people tend to resist (Markus 1983). The likelihood of this resistance increases with the scope and magnitude of the changes that the system creates (Tait and Vessey 1988). Data warehousing, in particular, has profound effects on organizations because it can shift data ownership, use, and access patterns; change how jobs are performed; and modify business processes. It moves data ownership from the functional areas to a centralized group, shifts the responsibilities for data access from information systems personnel to end users, changes how users perform their jobs as a result of having access to warehouse data, and allows businesses to operate differently. These changes potentially lead to resistance from managers, data suppliers, and end users.

Much has been written about how to effectively address issues that result from change (Markus and Robey 1988). For example, Lewin (1951) introduced a popular three-stage model whereby people first are prepared for change (i.e., unfreezing), the change then takes place (i.e., moving), followed by a solidification of the processes and ways of thinking caused by the change (i.e., refreezing). Project teams can encourage the organization to accept data warehousing by arranging for support throughout these three stages. They can put change management programs in place, deal with political resistance effectively when it arises, and encourage people throughout the organization to embrace data warehousing. Without these efforts, data warehousing projects are unlikely to result in high levels of data quality and system quality because key stakeholders are unwilling to support the changes that are required. For example, the consequences can include that subject area database specialists' time is not made available to the project, or changes to operational source systems (to improve data consistency) might be resisted. Thus, we hypothesized:

H2a: A high level of organizational implementation success is associated with a high level of data quality.

H2b: A high level of organizational implementation success is associated with a high level of system quality.

Project Implementation Success

IS projects often include a complex array of tasks and roles that must be managed (Brooks 1975), and data warehouse projects in particular require highly skilled, well-managed teams who can overcome issues that arise during the project (Devlin 1997; Sakaguchi and Frolick 1997). Project teams must be able to focus on critical goals and pertinent issues, and avoid unforeseen circumstances that can put the project at risk. Success with project issues can be measured by how well the team meets its critical time, budgetary, and functional goals (Constantine 1993; Waldrop 1984). In meeting these goals, by definition the team will deliver a data warehouse that provides high-quality data and system features to the client. Thus, we hypothesized:

H3a: A high level of project implementation success is associated with a high level of data quality.

H3b: A high level of project implementation success is associated with a high level of system quality.

Technical Implementation Success

The technical complexity of data warehousing is high because of the large number of diverse and disparate systems that typically need to be understood, reconciled, and coordinated; the large volume of data that must be extracted, transformed, loaded, and maintained; and the complicated analytics that often are applied to the data (e.g., financial profitability models, data mining algorithms). Technical problems may emerge at various points during a data warehousing project, such as when many, heterogeneous data sources must be combined and when new technology for data warehousing must be fit into an existing technical infrastructure. These technical problems may preclude the warehousing team from creating a repository of high-quality data, and the system may not be as flexible or integrated as the

organization requires (Rist 1997). Therefore, we hypothesized:

H4a: A high level of technical implementation success is associated with a high level of data quality.

H4b: A high level of technical implementation success is associated with a high level of system quality.

Implementation Factors

There is no generic model for IT implementation success and, on the whole, the implementation literature is filled with conflicting results (Markus and Robey 1988). One reason for equivocal results is that different IT implementations possess unique qualities that alter the importance or effect of implementation factors. Vatanasombut and Gray (1999) surveyed the data warehousing literature and found nine success factors that are unique to data warehousing, such as cleanse the data to meet the data warehouse quality standard and choose loading intervals that keep data timely. Bischoff and Alexander (1997) indicate that the amount of complexity involved is what makes a data warehousing project different from traditional software engineering or systems development initiatives. As was mentioned earlier, data warehousing is not an application, which has been the research focus of many implementation studies, but is rather an enabler of many different current and future applications. It shares similar characteristics with other infrastructure projects like enterprise networking and enterprise resource planning. Few studies have addressed the implementation success of these kinds of projects.

On the other hand, there are aspects of a data warehousing project that are similar to application-level IT implementations that have been studied thoroughly. For example, project teams must learn new technologies, work with users to gather requirements, select and use appropriate development methodologies, and anticipate and respond to political problems. Therefore, it is reasonable to expect that implementation factors that have consistently been found to affect IT implementation success are relevant to warehousing as

well. Seven implementation factors were included in the research model because of their potential importance to data warehousing success: management support, champion, resources, user participation, team skills, source systems, and development technology. Each factor is theorized to affect one or more of the implementation success factors.³

Management Support

Management support is widespread sponsorship for a project across the management team and consistently is identified as one of the most important factors for data warehousing success. It motivates people in the organization to support the data warehousing initiative and the organizational changes that inevitably accompany it (Curtis and Joshi 1998; Watson et al. 1998). Management support can overcome political resistance and encourage participation throughout the organization (Markus 1983), and it has been found to be important to the success of many kinds of IT implementations, such as decision support systems (Guimares et al. 1992; Igbaria et al. 1997). Users tend to conform to the expectations of management, and they are more likely to accept a system that they perceive to be backed by the management of their organization (Karahanna et al. 1999). Therefore, we hypothesized:

H5: A high level of management support is associated with a high level of organizational implementation success.

Champion

A champion actively supports and promotes the project and provides information, material resources, and political support. Champions are important to data warehousing (Barquin and

Edelstein 1997; Watson et al. 1998), as well as to other IT projects (Beath 1991; Reich and Benbasat 1990). Champions exhibit transformational leadership behavior when they strongly support a project, and they possess the skills and clout needed to overcome resistance that may arise within the organization (Howell and Higgins 1990). Like management support, champions can help data warehousing projects with organizational issues; however, a champion is likely to have even closer ties to the daily actions and goals of the project team. It can be expected that champions not only help data warehousing projects achieve success at an organizational level, but also that they help teams meet their project-level goals. We hypothesized:

H6a: A strong champion presence is associated with a high level of organizational implementation success.

H6b: A strong champion presence is associated with a high level of project implementation success.

Resources

Resources include the money, people, and time that are required to successfully complete the project (Ein-Dor and Segev 1978). Studies have found that resource problems have a negative effect on successful system design and implementation (Tait and Vessey 1988). Resources are likely to be important to data warehousing projects because data warehouses are expensive, time-consuming, resource-intensive initiatives. The presence of resources can lead to a better chance of overcoming organizational obstacles and communicating high levels of organizational commitment (Beath 1991; Tait and Vessey 1988). Resources also can help project teams meet their project milestones. Once tasks are identified, the project timeline is influenced by the amount of time and the people assigned to do the work, so better resources should affect the accomplishment of milestones during implementation (McConnell 1996). Thus, we hypothesized:

H7a: A high level of resources is associated with a high level of organizational implementation success.

³The IT implementation literature shows that the implementation factors have impacts other than the ones described in this study. For example, user participation can help manage user expectations and improve user acceptance of IT. However, for the purposes of this study, we have measured the impacts that were best supported by the study's model development phase.

H7b: A high level of resources is associated with a high level of project implementation success.

User Participation

User participation occurs when users are assigned project roles and tasks, which leads to a better communication of their needs and helps ensure that the system is implemented successfully (Hartwick and Barki 1994). It is particularly important when the requirements for a system are initially unclear, as is the case with many of the decision-support applications that a data warehouse is designed to support. The data warehousing literature indicates that user participation increases the likelihood of managing users' expectations and satisfying user requirements (Barquin and Edelstein 1997; Watson and Haley 1997). When users participate on warehousing projects, they have a better understanding of what the warehouse will provide, which makes them more likely to accept the warehouse when it is delivered. Users also can help the project team stay focused on the requirements and needs of the user community if they participate on the project team throughout the implementation. Thus, we hypothesized:

H8a: A high level of user participation is associated with organizational implementation success.

H8a: A high level of user participation is associated with project implementation success.

Team Skills

People are important when implementing a system and can directly affect its success or failure (Brooks 1975). In particular, the skills of the data warehousing development team have a major influence on the outcomes of the warehouse project (Barquin and Edelstein 1997). Team skills include both technical and interpersonal abilities, and a team with strong technical and interpersonal skills is able to perform tasks and interact with users well (Constantine 1993; Finlay and Mitchell 1994). The skills of development teams have been

traced to IT implementation success (Ancona and Caldwell 1992); only a high quality, competent team can identify the requirements of complex projects (Maish 1979). This mix of skills should help warehouse projects more successfully meet their objectives at a project level, and it should be of great value when technical obstacles need to be overcome. A highly skilled project team should be much better equipped to manage and solve technical problems. We hypothesized:

H9a: A high level of team skills is associated with project implementation success.

H9b: A high level of team skills is associated with technical implementation success.

Source Systems

Past studies have found that the quality of an organization's existing data can have a profound effect on systems initiatives and that companies that improve data management realize significant benefits (Goodhue et al. 1992; Kraemer et al. 1993). A primary purpose of data warehousing is to integrate data throughout the organization; however, data often resides in diverse, heterogeneous sources. Each unique source requires specialized expertise and coordination to access the data. Further, the data that exist often are defined differently across sources, making it challenging for the project team to reconcile and load the data into the warehouse properly. Goodhue et al. (1988) found that the lack of data standards was a "major underlying problem with data, often making it difficult or impossible to share or interpret data across application systems boundaries" (p. 389). Standardized data can result in easier data manipulation, fewer problems, and, ultimately, a more successful system (Bergeron and Raymond 1997). Thus, the quality of data sources depends on the standardization of their technology and data, and we hypothesized:

H10: High-quality source systems are associated with technical implementation success.

Development Technology

Development technology is the hardware, software, methods, and programs used in completing a project. The development tools that a project team uses can influence the effectiveness of the development effort as much as other factors, such as people. The tools can impact the efficiency and effectiveness of the development team, especially if they are not well understood or easy to use (Banker and Kauffman 1991). The development tools needed to build a data warehouse are different from those used with operational systems because warehousing requires sophisticated extraction, transformation, and loading software; data cleansing programs; data base performance tuning methods; and multidimensional modeling and analysis tools. If the development technology does not meet the needs of the project team or work well with the legacy systems, the data warehouse implementation will suffer (Rist 1997; Watson et al. 1998). Therefore, we hypothesized:

H11: Better development technology is associated with technical implementation success.

Research Method

Data Collection

Initial versions of two survey instruments were developed based on the data warehousing, implementation, and success literature. The first instrument was created to measure the implementation factors and the second to measure data warehousing success. Whenever possible, previously tested questions were used, and generally accepted instrument construction guidelines were followed (Converse and Presser 1986; Dillman 1978; Fox et al. 1988). Both surveys were reviewed by the University of Georgia Center for Survey Research; by academics with specific expertise in data warehousing, database, data integration, and survey construction; and by data warehousing experts, such as the head of Arthur Andersen & Co.'s data warehousing practice and the president of The Data Warehousing Institute. The multiple phases of instrument development

resulted in some restructuring and refinement of the survey and established its face and content validity (Nunnally 1978). The resulting surveys were then pilot-tested by 10 organizations to identify problems with the instruments' wording, content, format, and procedures. Pilot participants returned written comments about the survey instruments, and each was telephoned for a more detailed discussion.

Data were collected from two types of respondents at each participating organization to measure perceptions of implementation factors and success factors separately. This approach ensured that the appropriate person provided perceptions for the study (Hufnagel and Conca 1994); otherwise, "halo effects" or other biases could result from one person providing information for both the independent and dependent constructs. A total of 225 survey packets were mailed to the data warehousing managers of operational data warehouses⁴ listed in the researchers' data warehousing database.⁵ The survey that included implementation factor questions was completed by the data warehousing manager or the person most familiar with the data warehousing implementation. This contact was instructed to distribute the success factor survey to one or two data suppliers (two people were encouraged to further reduce single-source response bias), who were clearly defined as the managers of end-user computing or people responsible for an application that uses data from the warehouse (e.g., the executive information system manager). It was felt that data suppliers would be best qualified to assess the success of the data warehouse, as opposed to end users who only see the warehouse through the lens of the data access tool (e.g., managed query environment) or application (e.g., DSS) that they are given.

⁴An operational data warehouse is the result of a data warehouse implementation. It is a data warehouse that has been rolled out to the organization and put into operation.

⁵This database contains more than 350 warehousing companies, consultants, and vendors that have been compiled from The Data Warehousing Institute's conferences, past data warehousing studies, vendor contacts, Web interest, and personal contacts. Of these organizations, 225 have operational data warehouses.

Several rounds of follow-up phone calls and e-mails were used to remind the participants to return the surveys, and 111 companies responded with usable pairs of surveys (an implementation survey and at least one success survey) for an overall response rate of 49%. A total of 55 organizations returned two success surveys, and we examined the level of participant agreement on the success items using a one-way ANOVA with team variation as the independent variable (Amason 1996). In each case, the between-team variation was significantly larger than the within-team variation, suggesting that the scores for each organization could be combined into a single organizational response. Thus, the average of the individual responses was used as the success measures for each organization.

The participating organizations represent the different regions of the United States: 24 from the Northeast, 29 from the South, 34 from the Midwest, and 12 from the West. Also, 12 organizations located in South Africa, Canada, or Austria participated in the study. These organizations ranged in size, with mean gross revenues of \$5.8 billion (minimum = \$150,000; maximum = \$40 billion) and a mean number of employees of 23,571 (minimum = 35; maximum = 300,000). Table 1 shows the industries that are represented. All of the companies had operational data warehouses when answering the surveys, and nearly all of them considered their initiative successful⁶ (26% = "a runaway success"; 72% = "an up and coming system"; 2% = "potentially in trouble").

Most respondents to the first questionnaire were data warehousing managers (65%). Others were people who had significant knowledge of the data warehousing implementation, such as data warehousing staff members (11%) or employees holding some other position in the organization (e.g., IS manager, CIO (24%)). Of the respondents, 91% were actively involved in the project. The respondents to the second survey included functional area managers and professionals

⁶The data were analyzed both with and without the observations that assessed the warehouse as "potentially in trouble." There were no significant differences in the results; therefore, all 111 observations were included in the final dataset.

(45%), IS managers (25%), IS staff members (24%), and other members of the organization (6%). All of these people were responsible for providing warehouse data to end users.

Operationalization of Constructs

All items were developed based on items from existing instruments, the data warehousing literature, and input from data warehousing experts. Existing items were not used unless the measures were well supported by the latter two sources. Items were measured based on a seven-point Likert scale ranging from (1) "strongly disagree" to (7) "strongly agree." Table 2 defines the constructs used in the study and lists their respective survey items. Four items were reverse scaled, and they are noted accordingly.

Success factors. Data quality was operationalized as the accuracy, comprehensiveness, consistency, and completeness of the data provided by the warehouse. These dimensions are common measures of data quality for information systems in general (DeLone and McLean 1992), and data warehousing in particular (Lyon 1998; Shanks and Darke 1998). Flexibility and integration have been shown to be important dimensions of system quality; therefore, system quality was measured by four items that asked about the level of flexibility and integration of the data warehouse. Perceived net benefits was operationalized using three items that measured the change in the jobs of data suppliers and the reduction of time and effort required to support decision making in the end-user community (Graham 1996; Seddon 1997).

Organizational implementation success. This construct was measured using three questions that captured the extent that political resistance in the organization was dealt with effectively, change was managed effectively, and support existed from people throughout the organization (Markus 1983). Management support, champion, resources, and user participation are believed to help project teams overcome organizational issues (Beath 1991; Reich and Benbasat 1990; Steinbart and Nath 1992; Tait and Vessey 1988).

Table 1. Respondents by Industry

Industry	Number of Respondents	Percent of Respondents
Manufacturing	16	14
Healthcare	15	13
Retail/ Wholesale	13	12
Telecommunications	13	12
Financial Services/ Banking	11	10
Insurance	9	8
Government	8	7
Utilities	6	5
Education/ Publishing	3	3
Petrochemical	2	2
Other ^a	15	14

^aOther industries included Transportation, Market Research, Reseller, Travel, Defense, Distribution, and Consumer Products

Project implementation success. This construct included questions that asked how well the project was completed on time, on budget, while delivering the right requirements. A champion, resources, user participation, and team skills have been associated with such outcomes (Finlay and Mitchell 1994; Lawrence and Low 1993; Reich and Benbasat 1990; Yoon et al. 1995).

Technical implementation success. This construct was measured by asking about the technical problems that arose and technical constraints that occurred during the implementation of the warehouse. Poor team skills, source systems, and inadequate development tools have been found to affect the complexity of using technology, resulting in greater technical problems (Finlay and Mitchell 1994; Tait and Vessey 1988). Technical implementation success was defined as the ability to overcome these problems, and its questions were worded with help from data warehousing experts.

Implementation factors. Management support was operationalized as the overall support management showed for data warehousing and their interest in user satisfaction (Yoon et al. 1995). Two items for assessing the project champion were developed to measure whether a

champion existed from a functional area and from the IS area. User participation was measured using three items that assessed the IS-user relationship, the users' responsibilities on the project, and hands-on activities performed by the users (Barki and Hartwick 1994). Based on the work of Waldrop (1984), two items measured the data warehousing team's interpersonal and technical skills. The quality of the source systems was measured based on Wybo and Goodhue (1995) and suggestions from data warehousing experts. The items asked about the diversity of the data source platforms and the data standards that they supported. Development technology items were created to reflect the compatibility of the data warehousing tools with existing technology (Leonard-Barton and Sinha 1993) and the team's experience with the new tools (McFarlan 1981).

Data Analysis

The research model was tested using Partial Least Squares (PLS), a structural modeling technique that is well suited for highly complex predictive models (Wold and Joreskog 1982). PLS has several strengths that made it appropriate for this study, including its ability to handle formative constructs and its small sample size

requirements.⁷ The technique concurrently tests the psychometric properties of the scales used to measure the variables in the model (i.e., the measurement model) and analyzes the strengths and directions of the relationships among the variables (i.e., the structural model) (Lohmoller 1989). (For overviews of PLS, see Barclay et al. [1995] or Chin [1998]).

The test of the measurement model includes the estimation of internal consistency and the convergent and discriminant validity of the instrument items; however, reflective and formative measures should be treated differently. Reflective items represent the effects of the construct under study (Bollen 1984) and, therefore, "reflect" the construct of interest; eight constructs in this study are reflective. Table 2 lists the reflective measures and their internal consistency reliabilities, as defined by Fornell and Larcker (1981). All reliability measures were well above the recommended level of .70, thus indicating adequate internal consistency (Nunnally 1978). These items also demonstrated satisfactory convergent and discriminant validity. Convergent validity is adequate when constructs have an Average Variance Extracted (AVE) of at least .5 (Fornell and Larcker 1981). For satisfactory discriminant validity, the AVE from the construct should be greater than the variance shared between the construct and other constructs in the model (Chin 1998). Table 3 lists the correlation matrix, with correlations among constructs and the square root of AVE on the diagonal. Convergent validity also is demonstrated when items load highly (loading > .50) on their associated factors. Table 2 shows that all of the reflective measures have significant loadings that load much higher than the suggested threshold.

Formative measures are items that cause the construct under study (Bollen 1984). Thus, different dimensions are not expected to correlate or demonstrate internal consistency (Chin 1998). For example, the presence of a champion is caused by having a high-level supporter from the IS area

and/or having a high-level supporter from a functional area. The fact that an IS champion exists does not necessarily ensure that a functional area champion exists, and vice versa. Although internal consistency reliability is inappropriate for formative measures, the item weights can be examined to identify the relevance of the items to the research model (see Table 2). The formative constructs also were carefully reviewed to make sure that they performed as expected in the research model and that they were well supported by past studies and data warehousing resources.

Because this was a cross-sectional study that included data warehousing projects that had been operational for different periods of time, t-tests were conducted to test for the potential influence of time on success. Means were compared for the perceived net benefits items for data warehouses that had been operational for a year or less (N = 44) versus data warehouses that had been operational for more than two years (N = 57). This was done to confirm that data warehouses that were in place longer were not experiencing different benefits from newly implemented ones. None of the null hypotheses (t-tests) could be rejected at the .05 level, suggesting that time did not significantly influence the findings.

The test of the structural model includes estimating the path coefficients, which indicate the strengths of the relationships between the dependent and independent variables, and the R² value, which represents the amount of variance explained by the independent variables. Together, the R² and the path coefficients (loadings and significance) indicate how well the model is performing. R² indicates the predictive power of the model, and the values should be interpreted in the same manner as R² in a regression analysis. The path coefficients should be significant and directionally consistent with expectations.

PLS Graph version 2.91 (Chin and Frye 1996) was used for the analysis, and the bootstrap resampling method (100 resamples) determined the significance of the paths within the structural model. The sample size of 111 exceeded the recommended minimum of 40, which was adequate for model testing. The results are presented in Figure 2.

⁷PLS requires a minimum sample size that equals 10 times the greater of (1) the number of items comprising the most formative construct or (2) the number of independent constructs influencing a single dependent construct.

Table 2. Survey Items[†]

Management Support: widespread sponsorship for a project across the management team. REFLECTIVE				
Fornell = .76	Mean	Std. Dev.	Loading ^{††}	t-Stat
Overall, management has encouraged the use of DW.	5.36	1.33	.91	21.2***
User satisfaction has been a major concern of management.	5.09	1.47	.59	4.00***
Champion: a person within the organization who actively supports and promotes the project. FORMATIVE				
Fornell = .47	Mean	Std. Dev.	Weight	t-Stat
A high-level champion(s) for DW came from IS.	4.31	2.18	.94	5.35***
A high-level champion(s) for DW came from a functional area(s).	5.01	1.87	.87	5.37***
Resources: the money, time, and people required to successfully implement a data warehouse. FORMATIVE				
Fornell = .87	Mean	Std. Dev.	Weight	t-Stat
The DW project was adequately funded.	5.05	1.63	.14	0.50
The DW project had enough team members to get the work done.	4.54	1.80	.38	1.82*
The DW project was given enough time for completion.	4.45	1.65	.60	3.77***
User Participation: when users are assigned project roles and tasks during implementation of the data warehouse. FORMATIVE				
Fornell = .80	Mean	Std. Dev.	Weight	t-Stat
IS and users worked together as a team on the DW project.	5.66	1.60	.82	4.16***
Users were assigned full-time to parts of the DW project.	4.35	2.20	.36	1.30
Users performed hands-on activities (e.g., data modeling) during the DW project.	4.34	2.00	.06	0.20
Team Skills: the technical and interpersonal abilities of members of the data warehousing team. FORMATIVE				
Fornell = .90	Mean	Std. Dev.	Weight	t-Stat
Members of the DW team (including consultants) had the right technical skills for DW.	4.84	1.56	.62	3.86***
Members of the DW team had good interpersonal skills.	5.19	1.39	.46	2.46***

Table 2. Continued				
Source Systems: the quality (e.g., standardization, readiness, disparity) of the source systems that provide data to the warehouse. FORMATIVE				
Fornell = .60	Mean	Std. Dev.	Weight	t-Stat
Common definitions for key data items were implemented across the source systems	4.52	1.83	.63	2.82***
The data sources used for DW were diverse and disparate applications/systems. ^R	2.38	1.71	.05	.21
A significant number of source systems had to be modified to provide data for DW. ^R	4.56	1.93	.65	2.82***
Development Technology: effective hardware, software, methods, and programs to build the data warehouse. REFLECTIVE				
Fornell = .83	Mean	Std. Dev.	Loading	t-Stat
The DW technology that the project team used worked well with technology already in place in the organization.	4.71	1.58	.79	7.04***
Appropriate technology was available to implement DW.	5.34	1.36	.89	17.8***
Organizational Implementation Success: implementation-level success in addressing organizational issues, such as change management, widespread support, and political resistance. REFLECTIVE				
Fornell = .91	Mean	Std. Dev.	Loading	t-Stat
Any political resistance to DW in the organization was dealt with effectively.	4.61	1.44	.90	27.2***
Change in the organization created by DW was managed effectively.	4.20	1.5	.89	33.1***
The DW had support from people throughout the organization.	4.41	1.59	.86	27.4***
Project Implementation Success: implementation-level success in completing the project on time, on budget, with the proper functionality. REFLECTIVE				
Fornell = .84	Mean	Std. Dev.	Loading	t-Stat
The DW project met its critical project deadlines (eg., rollout deadline, initial development deadline).	4.60	1.85	.78	15.3***
The cost of the DW did not exceed its budgeted amount.	4.59	1.79	.79	17.7***
The DW project provided all of the DW functionality that it was supposed to provide.	4.83	1.51	.83	27.6***

Table 2. Continued

Technical Implementation Success: implementation-level success in overcoming technical problems. REFLECTIVE				
Fornell = .91	Mean	Std. Dev.	Loading	t-Stat
Many technical problems arose during the DW implementation. ^R	4.51	1.71	.89	23.1***
Numerous technical constraints were imposed on the DW implementation. ^R	3.86	1.76	.94	53.1***
Data Quality: The quality of data that are provided by the data warehouse. REFLECTIVE				
Fornell = .84	Mean	Std. Dev.	Loading	t-Stat
Users (or applications) have more accurate data now from DW than they had from source systems (e.g., transaction systems).	4.96	1.43	.80	5.50***
DW provides more comprehensive data to users (or applications) than source systems provided.	5.66	1.19	.67	4.84***
DW provides more correct data to users (or applications) in respect to source systems.	4.62	1.40	.70	4.23***
DW has improved the consistency of data to users (or applications) over that of source systems.	5.47	1.22	.82	8.80***
System Quality: the flexibility and integration of the data warehouse. REFLECTIVE				
Fornell = .86	Mean	Std. Dev.	Loading	t-Stat
DW can flexibly adjust to new demands or conditions.	4.86	1.14	.77	19.1***
DW effectively integrates data from systems servicing different functional areas.	5.40	1.18	.73	12.0***
DW is versatile in addressing data needs as they arise.	4.97	1.07	.85	24.0***
DW effectively integrates data from a variety of data sources within the organization.	5.47	1.08	.76	17.9***
Perceived Net Benefits: the benefits of the data warehouse as perceived by a data supplier. REFLECTIVE				
Fornell = .88	Mean	Std. Dev.	Loading	t-Stat
DW has changed my job significantly.	5.25	1.46	.75	11.2***
DW has reduced the time it takes to support decision making to the end-user community.	5.68	1.06	.91	60.1***
DW has reduced the effort it takes to support decision making to the end-user community.	5.44	1.15	.86	29.0***

[†]The variables were measured using seven-point Likert-type scales ranging from strongly disagree to strongly agree.

^{††}Loadings have been provided for reflective measures. They represent the extent to which the variables are related to the underlying construct. Weights have been provided for formative measures. They represent the extent to which the variables are related to the underlying construct.

^RThis item was reverse coded.

- * Indicates that the item is significant at the $p < .05$ level.
- ** Indicates that the item is significant at the $p < .01$ level.
- *** Indicates that the item is significant at the $p < .001$ level.

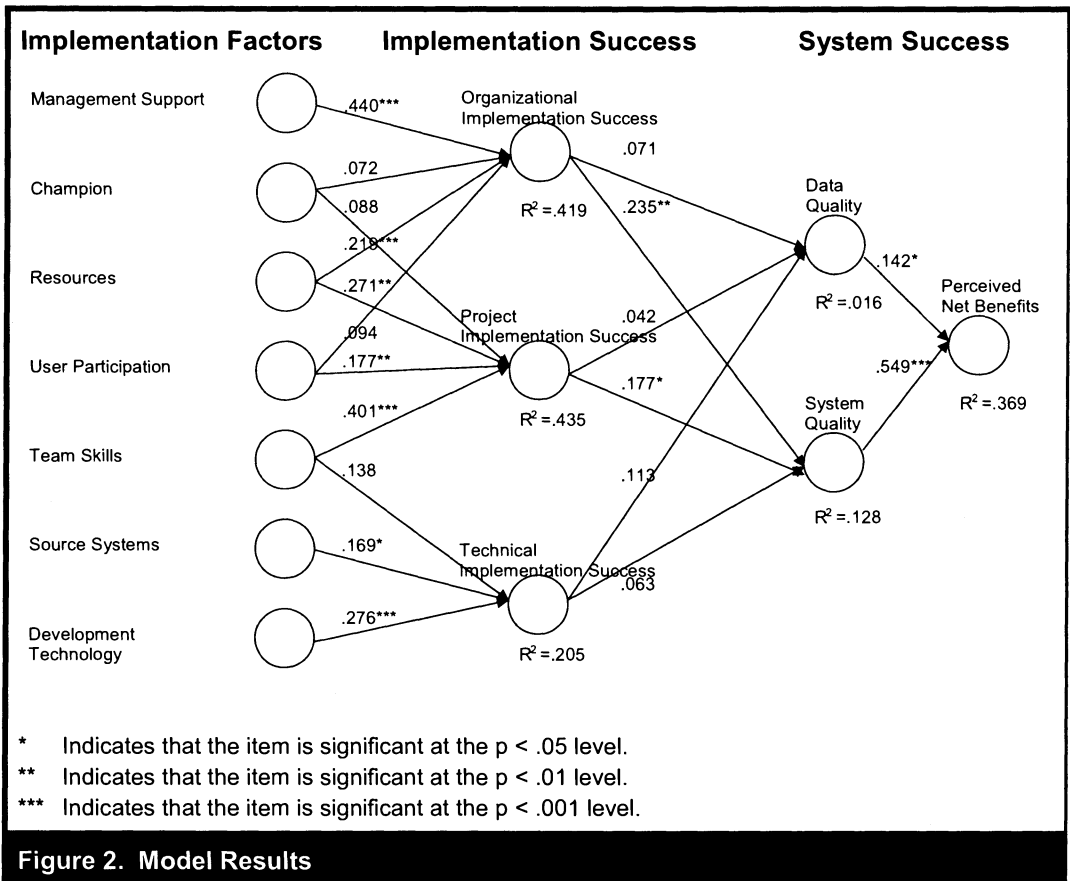
Table 3. Correlations of Latent Variables

	MANS	CHSM	RESO	USER	SKIL	SOUR	DEVT	ORGS	PROS	TECS	DATA	SYST	PNB
MANS	.79												
CHAM	0.423	.55											
RESO	0.411	0.285	.84										
USER	0.463	0.289	0.162	.76									
SKIL	0.357	0.218	0.350	0.224	.90								
SOUR	0.096	0.042	0.136	0.011	0.236	.64							
DEVT	0.293	0.235	0.385	0.122	0.525	0.323	.84						
ORGS	0.604	0.348	0.435	0.353	0.254	0.093	0.273	.88					
PROS	0.322	0.304	0.465	0.336	0.555	0.222	0.419	0.311	.80				
TECS	0.062	0.121	0.249	0.090	0.323	0.291	0.403	0.127	0.342	.84			
DATA	0.099	0.091	0.128	0.052	0.092	0.019	0.109	0.069	0.025	0.090	.75		
SYST	0.283	0.139	0.341	0.005	0.262	0.134	0.287	0.298	0.271	0.153	0.304	.78	
PNB	0.180	0.012	0.290	0.032	0.292	0.116	0.191	0.117	0.209	0.135	0.309	0.593	.84

Diagonal elements are the square root of Average Variance Extracted. These values should exceed the inter-construct correlations for adequate discriminant validity.

Legend:

- MANS = Management Support
- CHAM = Champion
- RESO = Resources
- USER = User Participation
- SKIL = Team Skills
- SOUR = Source Systems
- DEVT = Development Technology
- ORGS = Organizational Implementation Success
- PROS = Project Implementation Success
- TECS = Technical Implementation Success
- DATA = Data Quality
- SYST = System Quality
- PNB = Perceived Net Benefits



As hypothesized, perceived net benefits was associated with system quality and data quality, which together explained 37% of the dependent construct's variance. Both paths had positive effects, with path coefficients of .549 and .142, respectively. Hypotheses 1a and 1b were supported.

Against expectations, organizational, project, and technical implementation success had no effect on data quality, as shown by the three non-significant paths. Hypotheses 2a, 3a, and 4a were not supported. The R² value for data quality was .016, suggesting that factors not included in this model are more important in explaining the variance for data quality. Implementation success with organizational and project issues did have significant effects on system quality (paths = .235 and .177). Hypotheses 2b and 3b were supported.

The constructs explained 13% of the variance contained in system quality.

Management support and resources contributed to organizational implementation success, supporting hypotheses 5 and 7a. These factors had path coefficients of .440 and .219, and along with champion and user participation, they explained 42% of the variance. Consistent with hypotheses 7b, 8b, and 9b, resources, user participation, and team skills contributed to project implementation success, with path coefficients of .271, .177, and .401, respectively. When combined with the champion construct, they explained 44% of the variance for the dependent construct. As hypothesized in hypotheses 10 and 11, source systems and development technology contributed to technical implementation success, and they along with team skills explained 21% of the factor's variance.

Table 4. Hypothesis Results		
H1a	A high level of data quality will be associated with a high level of perceived net benefits.	Supported
H1b	A high level of system quality will be associated with a high level of perceived net benefits.	Supported
H2a	A high level of organizational implementation success is associated with a high level of data quality.	Not Supported
H2b	A high level of organizational implementation success is associated with a high level of system quality.	Supported
H3a	A high level of project implementation success is associated with a high level of data quality.	Not Supported
H3b	A high level of project implementation success is associated with a high level of system quality.	Supported
H4a	A high level of technical implementation success is associated with a high level of data quality.	Not Supported
H4b	A high level of technical implementation success is associated with a high level of system quality.	Not Supported
H5	A high level of management support is associated with a high level of organizational implementation success.	Supported
H6a	A strong champion presence is associated with a high level of organizational implementation success.	Not Supported
H6b	A strong champion presence is associated with a high level of project implementation success.	Not Supported
H7a	A high level of resources is associated with a high level of organizational implementation success.	Supported
H7b	A high level of resources is associated with a high level of project implementation success.	Supported
H8a	A high level of user participation is associated with organizational implementation success.	Supported
H8b	A high level of user participation is associated with project implementation success.	Supported
H9a	A high level of team skills is associated with project implementation success.	Supported
H9b	A high level of team skills is associated with technical implementation success.	Not Supported
H10	High-quality source systems are associated with technical implementation success.	Supported
H11	Better development technology is associated with technical implementation success.	Supported

The development technology had the greatest impact, with a path coefficient of .276, followed by the source systems with a path of .169. See Table 4 for a summary of the hypothesis test results.⁸

Discussion and Implications

This study examined the factors that affect data warehousing success by using a research model that was developed from the IT implementation and data warehousing literature, an exploratory survey, and structured interviews. Implementation success factors were used to help understand why the implementation factors affected the system success and ultimate success from the use of the system. The following sections present key observations regarding the major pieces of the model.

Perceived Net Benefits

Data quality and system quality had significant relationships with perceived net benefits and explained a good portion of the construct's variance. These results show that the quality of the data warehouse and the data that it provides are associated with the net benefits as perceived by the organization's data suppliers. In other words, a warehouse with good data quality and system quality improves the way data is provided to decision-support applications and decision makers. This supports the data warehousing literature that emphasizes that data warehouses must contain high-quality data, flexibly respond to users' requests for data, and integrate data in the ways that are required by users, all in order to create value for the organization.

⁸Because there is no generic model for IT implementation, two other alternative research models for data warehousing success were considered: the original model without implementation success factors and a model with direct relationships between the seven implementation factors and perceived net benefits. The primary research model was found to provide the greatest predictive power based on results from confirmatory factor analyses. Interested readers can obtain results for the alternative models by contacting the authors directly.

This study furthers the knowledge of IT success by supporting the use of multiple success dimensions and confirming other research findings that show the success dimensions (e.g., system quality, data quality, and perceived net benefits) to be interrelated. System quality and data quality do affect perceived net benefits in the context of data warehousing. More work is needed, however, to examine exactly how the dimensions of success interrelate. Theoretically, we need to understand why relationships exist, and practically, we need to explore how success measures can be applied most effectively. We also need to explore the role of other success dimensions, such as extent of implementation or use, in data warehousing.

Data Quality and System Quality in a Data Warehouse Context

Factors not included in the research model affect the data quality of the data warehouse. Further research is needed to understand warehouse data quality and the factors that affect it. For example, does poor data quality in source systems undermine the ability to provide high-quality data in a data warehouse? What role does the extraction, cleansing, and transformation process play in creating high-quality data? Do the data model and data storage format have any influence on the perception of the data's quality? Or, can a data warehouse even exist without data quality? The companies in this sample had at least somewhat successful warehouse implementations, and it may be possible that data quality is required before a warehouse project can be completed. Thus, does a relationship between implementation success and data quality not exist because organizations have to achieve an acceptable level of quality to roll out the warehouse to their users? There are many questions regarding data quality that remain unanswered.

In data warehousing, system quality depends on a number of factors, such as the selection of subject areas and data for the data store, the underlying data model that was created, and the warehouse architecture that was selected. Not all organizations have the vision and knowledge to properly include these considerations in the

design of their warehouses, which can lead to a future lack of flexibility and integration of the data. The findings of this study show that system quality was associated with implementation success with organizational and project issues. The reasons for these relationships are clear when one takes into account how much easier it is for a team to create a flexible and integrated data warehouse when organizational barriers are removed and a well-managed team is responsible for meeting the demands of the project.

Technical implementation success, however, was not significantly related to system quality. This finding may be because successful, operational data warehouses (as were the ones included in the study) have overcome the technical problems that were encountered. If they had not overcome the most serious problems, their warehouses would not be operational. In order to understand the relationship between system quality and technical implementation success, failed warehousing projects need to be studied.

It should be noted that the R squared value for system quality (.128) suggests that like data quality, other factors not included in the research model also affect the quality of the data warehouse. Thus, the integration and flexibility of the infrastructure that data warehousing creates is also influenced by factors other than those that were considered. For example, how important is the IT infrastructure already in place in the organization? If an organization does not have internal data warehousing expertise, how important is it to bring in external consultants? How do data planning and management practices influence the system quality for data warehousing? These questions still need to be addressed.

Implementation Factors for Data Warehousing

Management support, a champion, and resources are key ingredients to supporting the change management process in organizations. This finding is consistent with other IT implementation studies that substantiate the value of these

organizational factors. A data warehouse is an expensive, enterprise-wide endeavor with significant organizational impacts. Data warehousing creates changes that resonate throughout the entire organization, and it demands broad-based and lasting support. It requires the sponsorship and support of senior management, managers in the business units, and IT. There must be a substantial initial and ongoing commitment of financial and human resources. This commitment must be made while recognizing that the greatest benefits from data warehousing usually occur later rather than immediately. Together, all three organizational factors were found to be significant in the research model, and together they provide organizations with effective mechanisms for increasing widespread support for warehousing, addressing politics, and ensuring that the necessary resources are provided.

Interestingly, a champion for warehousing did not influence the project's ability to address organizational issues. Unlike decision support applications that may benefit from having a single proponent, the large scope and far-reaching impact of data warehousing appears to require broad-based support from multiple sources. A single warehouse champion may abandon the project at the first sign of trouble (Watson et al. 1999) and has limited influence and understanding outside his or her own area of the organization. Likewise, grass-roots support may not be sufficient for implementation success. Although studies have found that user participation can help manage user expectations, this may not be sufficient for the acceptance of a warehouse within the organization. All of these findings highlight some of the challenges that managers should expect when working with a warehouse initiative. An organization that has successfully rolled out applications in the past cannot assume that a data warehouse can be introduced with the same levels of sponsorship and resources.

According to the findings, having resources, appropriate people on the project team, and user participation have positive effects on the project's outcome. Unfortunately, companies sometimes experience problems in these areas. Warehousing

demands a large financial investment that can be difficult to sell to management without having guaranteed up-front tangible benefits. Currently, the demand for experienced warehousing personnel exceeds the supply. Many companies have little choice but to staff from within, independent of whether their staff have appropriate experience. As a result, the data warehousing staff may have little or no experience in how to plan for and manage a project of this type. User participation also can be challenging because the needs of many, diverse internal groups (e.g., marketing, production) must be understood and communicated to the project team. Much data warehousing literature advocates an *incremental approach* when building a warehouse, which means building a warehouse in three- to six-month increments that each deliver substantial value to the business. In this way, project teams can work toward goals that are more manageable in size, users can participate in only relevant parts of the project, and management can be satisfied that the project is delivering value. If management requires post-implementation assessments of its investments, the value that is created during beginning increments can be used as a foundation for a rigorous future cost-benefit analysis.

Technical factors also affect data warehousing implementations. The practitioner literature contains considerable debate over the merits of beginning a decision-support infrastructure with an enterprise-wide data warehouse versus a smaller-scale data mart. The data warehouse proponents argue that data marts can quickly grow into an unintegrated collection of *information silos* that counter the underlying purpose of data warehousing. Data mart supporters explain that data warehouses are more expensive and difficult to construct in a reasonable amount of time. Moreover, a data mart provides a *proof of concept*. This study indicates that more technical problems are related to warehouses that pull from diverse, unstandardized sources, undoubtedly due to the increased technical complexity. Organizations involved in building enterprise-wide data warehouses should prepare for technical obstacles that must be overcome. The development technology that is used also appears to affect the technical problems that may arise during

implementation. Data warehousing requires specialized software. The project team must learn how to use this software and how to fit it into the existing technical environment.

Although the source systems and development tools are related to the technical problems that occur during the development of a warehouse, the technical problems do not have long-lasting effects that ultimately affect the benefits from operational warehouses. Likely, project teams are ultimately able to address technical problems effectively, much more so than they are able to overcome organizational and project issues. Also, as was mentioned earlier, this sample includes operational warehouses and does not contain warehouses that failed. This study does *not* suggest that technical problems in data warehousing are easy to overcome.

Conclusions

There are few academic empirical studies on data warehousing. A valuable contribution of this study is the extension of the IS implementation literature through the investigation of data warehousing implementation factors. Both the IS implementation and data warehousing areas will benefit from the validation of current understandings and the development of new ideas.

The findings suggest that most of the traditional factors from the implementation literature (e.g., management support, resources) also affect the success of a data warehouse, thus providing further evidence of the existence of a common set of IT implementation factors. However, the study also shows that implementation success models cannot be used to investigate data warehousing without some modification. For example, other factors were needed to explain the data quality and system quality for the data warehouse.

Another contribution of this study is the way in which implementation success factors can be grouped together into organizational, project, and technical success to more clearly communicate the kinds of effects implementation factors can

have. This approach allowed us to tie implementation factors to system success and the benefits from the ultimate use of a system. The empirical evidence supported the idea that these connections are important to understand.

As noted previously, there has been little research on the success factors associated with infrastructure projects. Parr et al. (1999) investigated the success factors associated with ERP implementations, which can be viewed as infrastructure investments. Their list of success factors can be organized into three overarching implementation factors—organizational, project, and technical success—which is the same grouping used in this study. While the specific factors in the groups vary somewhat between data warehousing and ERP, it appears that there is a macro-level model for understanding the success factors associated with infrastructure projects that can be used in future research. It is likely that the organizational factors are the most generic (e.g., management support) to implementation success. Many of the project management success factors also are probably the same. The greatest differences are most likely with the technical success factors, because the technical issues vary with the nature of the infrastructure project.

More research is required to further develop our understanding of infrastructure and determine the differences between infrastructure and application-level IT phenomenon. This study presents data warehousing as a viable way of investigating such issues. This study also challenges the notion of applying IT implementation knowledge to an infrastructure context without giving careful thought to how changes should be made.

Acknowledgments

We would like to thank Dale Goodhue, Peter Todd, Wynne Chin, and Izak Benbasat for their helpful comments on this paper. We also are grateful to Ron Weber, Joe Valacich, and the reviewers whose comments have improved the quality of the paper substantially.

References

- Amason, A. C. "Distinguishing the Effects of Functional and Dysfunctional Conflict on Strategic Decision Making: Resolving a Paradox for Top Management Teams," *Academy of Management Journal* (39:1), 1996, pp. 123-148.
- Ancona, D. G., and Caldwell, D. F. "Bridging the Boundary," *Administrative Science Quarterly* (37:4), 1992, pp. 634-666.
- Banker, R. D., and Kauffman, R. J. "Reuse and Productivity in Integrated Computer-Aided Software," *MIS Quarterly* (15:3), 1991, pp. 375-402.
- Barclay, D., Higgins, C., and Thompson, R. "The Partial Least Squares Approach to Causal Modeling, Personal Computing Adoption and Use as an Illustration," *Technology Studies* (2:2), 1995, pp. 285-309.
- Barki, H., and Hartwick, J. "Measuring User Participation, User Involvement, and User Attitude," *MIS Quarterly* (18:1), 1994, pp. 59-82.
- Barquin, R. C., and Edelstein, H. *Planning and Designing the Data Warehouse*, Prentice Hall, Upper Saddle River, NJ, 1997.
- Beath, C. M. "Supporting the Information Technology Champion," *MIS Quarterly* (15:3), 1991, pp. 355-371.
- Beitler, S. S., and Leary, R. "Sears' EPIC Transformation: Converting from Mainframe Legacy Systems to On-Line Analytical Processing (OLAP)," *Journal of Data Warehousing* (2:2), 1997, pp. 5-16.
- Bergeron, F., and Raymond, L. "Managing EDI for Competitive Advantage: A Longitudinal Study," *Information & Management* (31), 1997, pp. 319-333.
- Bischoff, J., and Alexander, T. *Data Warehouse: Practical Advice from the Experts*, Prentice Hall, Upper Saddle River, NJ, 1997.
- Bollen, K. A. "Multiple Indicators: Internal Consistency or No Necessary Relationship?" *Quality and Quantity* (18), 1984, pp. 377-385.
- Brooks, F. P. *The Mythical Man-month: Essays on Software Engineering*, Addison Wesley, Reading, MA, 1975.
- Chau, P. Y., and Tam, K. Y. "Factors Affecting the Adoption of Open Systems: An Exploratory Study," *MIS Quarterly* (21:1), 1997, pp. 1-24.
- Chin, W. W. "The Partial Least Squares Approach to Structural Equation Modeling," in G. A. Marcoulides (ed.), *Modern Methods for Business Research*, Lawrence Erlbaum Associates, Mahwah, NJ, 1998, pp. 295-336.

- Chin, W. W., and Frye, T. A. *PLS Graph*, version 2.91.03.04, Department of Decision and Information Systems, University of Houston, 1996.
- Constantine, L. L. "Work Organizations: Paradigms for Project Management and Organization," *Communications of the ACM* (36:10), 1993, pp. 35-42.
- Converse, J. M., and Presser, S. *Survey Questions: Handcrafting the Standardized Questionnaire*, Sage Publications, Newbury Park, CA, 1986.
- Curtis, M. B., and Joshi, K. "Lessons Learned from the Implementation of a Data Warehouse," *Journal of Data Warehousing* (3:2), 1998, pp. 12-18.
- Davis, F. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), 1989, pp. 319-339.
- DeLone, W. H., and McLean, E. R. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), 1992, pp. 60-95.
- Devlin, B. *Data Warehouse: From Architecture to Implementation*, Addison Wesley Longman, Reading, MA, 1997.
- Dillman, D. A. *Mail and Telephone Surveys: The Total Design Method*, Wiley, New York, 1978.
- Duncan, N. B. "Capturing Flexibility of Information Technology Infrastructure: A Study of Resource Characteristics and Their Measure," *Journal of Management Information Systems* (12:2), 1995, pp. 37-57.
- Eckerson, W. W. "Post-Chasm Warehousing," *Journal of Data Warehousing* (3:3), 1998, pp. 38-45.
- Eckerson, W. W. *Evolution of Data Warehousing: The Trend Toward Analytical Applications*, The Patricia Seybold Group, April 28, 1999, pp. 1-8.
- Ein-Dor, P., and Segev, E. "Organizational Context and the Success of Management Information Systems," *Management Science* (24:10), 1978, pp. 1064-1077.
- Finlay, P. N., and Mitchell, A. C. "Perceptions of the Benefits from the Introduction of CASE: An Empirical Study," *MIS Quarterly* (18:4), 1994, pp. 353-371.
- Fornell, C., and Larcker, D. F. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18), 1981, pp. 39-50.
- Fox, R. J., Crask, M. R., and Kim, J. "Mail Survey Response Rate: A Meta-Analysis of Selected Techniques for Inducing Response," *Public Opinion Quarterly* (52), 1988, pp. 467-491.
- Fraser, S. G., and Salter, G. "A Motivational View of Information Systems Success: A Reinterpretation of DeLone and McLean's Model," working paper, Department of Accounting and Finance, The University of Melbourne, Australia, 1995.
- Garrity, E. J., and Sanders, G. L. *Information Success Measurement*, Idea Group Publishing, Hershey, PA, 1998.
- Goodhue, D. L., Quillard, J. A., and Rockart, J. F. "Managing the Data Resource: A Contingency Perspective," *MIS Quarterly* (12:3), 1988, pp. 373-392.
- Goodhue, D. L., Wybo, M. D., and Kirsch, L. J. "The Impact of Data Integration on the Costs and Benefits of Information Systems," *MIS Quarterly* (16:3), 1992, pp. 293-311.
- Graham, S. *The Foundations of Wisdom: A Study of the Financial Impact of Data Warehousing*, International Data Corporation, Toronto, 1996.
- Gray, P., and Watson, H. J. *Decision Support in the Data Warehouse*, Prentice Hall, Upper Saddle River, 1998.
- Grim, R., and Thornton, P. "A Customer for Life: The WarehouseMCI Approach," *Journal of Data Warehousing* (2:1), 1997, pp. 73-79.
- Guimares, T., Igbaria, M., and Lu, M. "The Determinants of DSS Success: An Integrated Model," *Decision Sciences* (23), 1992, pp. 409-430.
- Haley, B. J., Watson, H. J., and Goodhue, D. L. "The Benefits of Data Warehousing at Whirlpool," *Annals of Cases on Information Technology Applications and Management in Organizations* (1:1), 1999, pp. 14-25.
- Hartwick, J., and Barki, H. "Explaining the Role of User Participation in Information System Use," *Management Science* (40:4), 1994, pp. 440-465.
- Howell, J. M., and Higgins, C. A. "Champions of Technological Innovations," *Administrative Science Quarterly* (35:2), 1990, pp. 317-341.
- Hufnagel, E. M., and Conca, C. "User Response Data: The Potential for Errors and Biases," *Information Systems Research* (5:1), 1994, pp. 48-73.

- Igbaria, M., Zinatelli, N., Cragg, P., and Cavaye, A. L. "Personal Computing Acceptance Factors in Small Firms: A Structural Equation Model," *MIS Quarterly* (21:3), 1997, pp. 279-302.
- Karahanna, E., Straub, D. W., and Chervany, N. L. "Information Technology Adoption Across Time: Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs," *MIS Quarterly* (23:2), 1999, pp. 183-213.
- Kelly, S. *Data Warehousing in Action*, John Wiley & Sons, Chichester, 1997.
- Kraemer, K. L., Danzinger, J. N., Dunkle, D. E., and King, J. L. The Usefulness of Computer-Based Information to Public Managers," *MIS Quarterly* (17:2), 1993, pp. 129-148.
- Lawrence, M., and Low, G. "Exploring Individual User Satisfaction Within User-Led Development," *MIS Quarterly* (17:2), 1993, pp. 195-208.
- Leonard-Barton, D., and Sinha, D. K. "Developer-User Interaction and User Satisfaction in Internal Technology Transfer," *Academy of Management Journal* (36:5), 1993, pp. 1125-1139.
- Lewin, K. *Field Theory in Social Science: Selected Theoretical Papers*, Harper and Brothers, New York, 1951.
- Lohmoller, J.-B. "Predictive vs. Structural Modeling: PLS vs. ML," in *Latent Variable Path Modeling with Partial Least Squares*, Physica-Verlag, Heidelberg, 1989, pp. 212-55.
- Lucas, H. C. *Implementation: The Key to Successful Information Systems*, McGraw-Hill, New York, 1981.
- Lyon, J. "Customer Data Quality: Building the Foundation for a One-to-One Customer Relationship," *Journal of Data Warehousing*, (3:2), 1998, pp. 38-47.
- Maish, A. M. "A User's Behavior Toward His MIS," *MIS Quarterly* (3:1), 1979, pp. 39-52.
- Markus, M. L. "Power, Politics, and MIS Implementation," *Communications of the ACM* (26:6), 1983, pp. 430-444.
- Markus, M. L., and Robey, D. "Information Technology and Organizational Change: Causal Structure in Theory and Research," *Management Science* (34:5), 1988, pp. 583-598.
- Massetti, B., and Zmud, R. W. "Measuring the Extent of EDI Usage in Complex Organizations: Strategies and Illustrative Examples," *MIS Quarterly*, (20:3), 1996, pp. 331-345.
- McConnell, S. *Rapid Development* Microsoft Press, Redmond, WA, 1996.
- McFarlan, F. W. "Portfolio Approach to Information Systems," *Harvard Business Review* (59:5), 1981, pp. 142-159.
- Melone, N. "A Theoretical Assessment of the User-Satisfaction Construct in Information Systems Research," *Management Science* (36:1), 1990, pp. 76-91.
- META Group. "Industry Overview: New Insights in Data Warehousing Solutions," *Information Week*, 1996, pp. 1-27HP.
- Moore, G., and Benbasat, I. "Development of an Instrument to Measure the Perceptions of Adopting and Information Technology Innovation," *Information Systems Research* (2:3), 1991, pp. 192-222.
- Nunnally, J. C. *Psychometric Theory*, McGraw-Hill, New York, 1978.
- Parr, A., Shanks, G., and Darke, P. "Identification of Necessary Factors for Successful Implementation of ERP Systems," in *New Information Technologies in Organizational Process*, O. L. Ngwenyama, L. D. Introna, M. D. Myers, and J. I. DeCross (eds.), Kluwer Academic Publishers, Boston, 1999, pp. 99-119.
- Pitt, L., Watson, R. T., and Kavan, C. B. "Service Quality: A Measure of Information Systems Effectiveness," *MIS Quarterly* (19:2), 1995, pp. 173-185.
- Prescott, M. B., and Conger, S. A. "Information Technology Innovations: A Classification by IT Locus of Impact and Research Approach," *Data Base* (26: 2, 3), 1995, pp. 20-40.
- Ragowsky, A., Ahituv, N., and Neumann, S. "Identifying the Value and Importance of an Information System Application," *Information and Management* (31), 1996, pp. 89-102.
- Rainer, R. K., and Watson, H. J. "The Keys to Executive Information Systems Success," *Journal of Management Information Systems* (12:2), 1995, pp. 83-98.
- Reich, B. H., and Benbasat, I. "An Empirical Investigation of Factors Influencing the Success of Customer-Oriented Strategic Systems," *Information Systems Research* (1:3), 1990, pp. 325-347.
- Rist, R. "Challenges Faced by the Data Warehousing Pioneers," *Journal of Data Warehousing* (2:1), 1997, pp. 63-72.
- Ross, J. W., C. M. Beath, and Goodhue, D. L. "Develop Long-Term Competitiveness Through

- IT Assets," *Sloan Management Review* (38:1), 1996, pp. 31-42.
- Sakaguchi, T., and Frolick, M. N. "A Review of the Data Warehousing Literature," *Journal of Data Warehousing* (2:1), 1997, pp. 34-54.
- Seddon, P. "A Respecification and Extension of the DeLone and McLean Model of IS Success," *Information Systems Research* (8:3), 1997, pp. 240-253.
- Seddon, P. B., and Kiew, M-Y. "A Partial Test and Development of the DeLong and McLean Model of IS Success," in *Proceedings of the International Conference on Information Systems*, J. I. DeGross, S. L. Huff, and M. C. Munro (eds.), Vancouver, Canada, 1994, pp. 99-110.
- Seddon, P., Staples, S., and Patnayakuni, R. "Dimensions of Information Systems Success," *Communications of the AIS* (2:20), 1999.
- Shanks, G., and Darke, P. "A Framework for Understanding Data Quality," *Journal of Data Warehousing*, (3:3), 1998, pp. 46-51.
- Steinbart, P. J., and Nath, R. "Problems and Issues in the Management of International Data Communications Networks: The Experiences of American Companies," *MIS Quarterly* (16:1), 1992, pp. 55-76.
- Tait, P., and Vessey, I. "The Effect of User Involvement on System Success: A Contingency Approach," *MIS Quarterly* (12:1), 1988, pp. 91-108.
- Vandenbosch, B., and Huff, S. L. "Searching and Scanning: How Executives Obtain Information from Executive Information Systems," *MIS Quarterly* (21:1), 1997, pp. 81-107.
- Vatanasombut, B., and Gray, P. "Factors for Success in Data Warehousing: What the Literature Tells Us," *Journal of Data Warehousing* (4:3), 1999, pp. 25-33.
- Waldrop, J. H. "Project Management: Have We Applied All That We Know?" *Information & Management* (7:1), 1984, pp. 13-20.
- Wand, Y., and Wang, R. Y. "Anchoring Data Quality Dimensions in Ontological Foundations," *Communications of the ACM* (39:11), 1996, pp. 86-95.
- Wang, R. Y., and Strong, D. M. "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems* (12:4), 1996, pp. 5-34.
- Watson, H. J., Gerard, J. G., Gonzalez, L. E., Haywood, M. E., and Fenton, D. "Data Warehousing Failures: Case Studies and Findings," *Journal of Data Warehousing* (4:1), 1999, pp. 44-55.
- Watson, H. J., Haines, M., and Loiacono, E. T. "The Approval of Data Warehousing Projects: Findings from Ten Case Studies," *Journal of Data Warehousing* (3:3), 1998, pp. 29-37.
- Watson, H. J., and Haley, B. J. "Data Warehousing: A Framework and Survey of Current Practices," *Journal of Data Warehousing* (2:1), 1997, pp. 10-17.
- Wetherbe, J. C. "Executive Information Requirements: Getting It Right," *MIS Quarterly* (15:1), 1991, pp. 51-66.
- Wold, H., and Joreskog, K. *Systems Under Indirect Observation: Causality, Structure, Prediction*, Volume 2, North-Holland, Amsterdam, 1982.
- Wybo, M. D., and Goodhue, D. L. "Using Interdependence as a Predictor of Data Standards: Theoretical and Measurement Issues," *Information & Management* (29:6), 1995, pp. 317-330.
- Yoon, Y., Guimares, T., and O'Neal, Q. "Exploring the Factors Associated with Expert Systems Success," *MIS Quarterly* (19:1), 1995, pp. 83-106.

About the Authors

Barbara H. Wixom is an assistant professor of Commerce at the University of Virginia's McIntire School of Commerce. She received her Ph.D. in MIS from the University of Georgia. Dr. Wixom was made a Fellow of The Data Warehousing Institute for her research in data warehousing. She has published in journals that include *MIS Quarterly*, *Information Systems Research*, *Communications of the ACM*, and *Journal of Data Warehousing*. She has presented her work at national and international conferences.

Hugh J. Watson is professor of MIS and holds the C. Herman and Mary Virginia Terry Chair of Business Administration in the Terry College of Business at the University of Georgia. He specializes in the design of information systems to support decision making. Dr. Watson is the author of over 100 articles and 22 books, including *Decision Support in the Data Warehouse* (Prentice-Hall, 1998). He is the Senior Editor of the *Journal of Data Warehousing* and is a Fellow of The Data Warehousing Institute.