December 2002

IMPLEMENTING THE CORPORATE DATA WAREHOUSE TO SUPPORT CRM APPLICATIONS

Fay Payton
North Carolina State University

Debra Zahay
North Carolina State University

IMPLEMENTING THE CORPORATE DATA WAREHOUSE TO SUPPORT CRM APPLICATIONS

Fay Cobb Payton and Debra Zahay
College of Management
Department of Business Management
North Carolina State University
fay_payton@ncsu.edu   debra_zahay@ncsu.edu

Abstract

This study uses a previously validated diffusion model to explore the experiences of one large health care payer in its implementation of a corporate data warehouse to support the marketing function via CRM applications. The organization invested significant time and other valuable resources on its data warehouse applications, but perceived that the warehouse was underutilized by the marketing function. Current estimates indicate that as many as 70% of CRM implementations have failed and the figures for data warehouse applications are not much better (Nelson and Eisenfeld 2002). Using a case methodology as noted in Lee, et al. (1997) and five focus group sessions with twenty-one marketing, information systems and strategic managers, we documented and tracked the data warehouse implementation process. Qualitative data were, then summarized and coded using two independent coders. Our findings suggest that the ability of marketing to use the system for its functional needs, information quality and the internal training were factors that ranked highly in term of significance to implementation success and often failed to meet users' expectations. In particular, the data and analysis needs of the marketing function were not necessarily consistent with the concept of a corporate data warehouse. Marketing managers have special data needs, such as external data and data about prospects, not just current customers. This additional data is not always kept in the data warehouse. When data were available, marketing managers/users questioned its quality and found data warehousing data analysis tools difficult to use. While the users agreed that the data warehouse could have economic benefit, this company discovered, through this research, that data warehousing technologies implemented to meet strategic corporate and IT objectives can be less than successful. Our research suggests that data warehousing implementations that fail to adhere to marketing analyses (including data quality), business intelligence and functional support needs will not gain and sustain the benefits normally associated these technologies. In sum, the challenge remains for practitioners to migrate data warehousing applications from operational, financial, accounting and production functions to value-added marketing and strategic functions.

Introduction

The proliferation of data warehousing technologies and applications has been widely documented among information technology consultants and vendors. According to Gartner, organizations will be confronted with the challenges of managing 30 times more data by 2004 in a continual effort to meet the demands associated with electronic commerce and supply chain applications (Gartner, 2000). Defined as a well-defined central repository used for decision-support, data warehouses are subject oriented, time-variant and non-volatile. Data are collected over time and used to perform trend analyses, forecasting and comparative analyses. Typically, these data are not updated in real time. Rather, data refreshes on a periodic basis from operational systems. To enable these capabilities, data warehousing implementations can deploy multi, parallel and massive processing databases well as a series of other hardware and software technologies (Wixom and Watson, 2001).

Customer relationship management (CRM) applications as well as Sales Force Automation (SFA) and contact are often enabled by data warehousing technologies and have been viewed as an emerging research stream in the IS field (Kalakota and Robinson 1999, Zwass, 1996; Romano, 2001). Broadly defined, CRM has been defined as a process/application that permits organizations

to gather and analyze customer data rapidly while seeking to improve customer loyalty via targeted products and services (Riby, et. al., 2002). While the cases of data warehousing implementation supporting a myriad of applications have been documented (Beitler and Leary, 1997; Grim and Thorton, 1997), a typical implementation is costly with researchers and practitioners, alike, documenting expenditures well over $1-2 million in year one (Watson and Haley, 1997; Gartner, 2001). Despite widely reported implementation failures ranging in rates of 55 to 75 percent, data warehouses have been described as the linchpin to any effective marketing effort that intends to include CRM as part of its strategy. To this end, IDC (2000, in CIO Special Advertising Supplement, 2001) projected that the worldwide spending on CRM applications, including databases, would total $20 billion by 2004. In addition, the warehousing application, an appropriate, yet complicated alignment of IS and organizational infrastructure, processes and strategy is fundamental. Cooper, et al. (200), however, determined that these associations are often disregarded in the implementation and strategic planning processes. Moreover, prior IT implementation studies have offered the field both factors and process streams to investigate the change associated with technology adoption. More recently and in a study of 111 organizations, Wixon and Watson (2001) offered specific, yet empirical findings impacting data warehousing implementations. While their study expanded the existing IT implementation knowledge, these researchers have offered critical findings to "changes" associated with data warehouse. Few studies, however, have investigated the implementation of data warehouses for the purposes of supporting marketing applications, such as CRM. Moreover, few studies have pointed out the relevance of data quality to the data warehouse's users.

Consequently, when we were approached by this large healthcare payor with the question, "What factors impact and/or stand to impact implementing our corporate data warehousing to support CRM applications for the marketing function?", we drew from the existing data warehousing, CRM and IS implementation literatures. We suggested a series of focus group interviews with functional Marketing and IS managers at the organization which served as the basis for our case study. The model was based on an existing, validated, model that looked at the factors leading to successful IT implementations. This paper describes the model, the focus group process and presents the results of the study.

The Research Model

Figure 1 shows the research model adopted for this study. Based on the diffusion work of Cooper and Zmud (1990), Payton and Ginzberg (2001) used this model to explore the implementations of multiple health care information networks. Interactions with marketing, IT and strategy managers confirmed the significance of which factors applied to the organization's context as well as the implementation in question. Prior to deriving the model, those questions that guided the original model were shared with the management team and published in Payton and Ginzberg (2001).

Based on these sessions with management, a few changes to Figure 1 are noteworthy:

1) The government factor was eliminated in the push/pull cluster – given the intraorganizational focus of the data warehouse implementation
2) A prior composite factor, Quality of IT Management, was reduced to championship and top management support; thus, we eliminated the Association Support construct which was via in a cross-state and/or interstate applications as explored by Payton and Ginzberg (2001)

The dependent variable in Figure 1 is the success of the implementation effort. Others (Wixom and Watson, 2001) offered that implementation success impacts systems success which is the quality of data warehouse and the data that is extracted from the system. Our results imply that information quality is a central measure of success or failure of our data warehousing proposed to support CRM initiatives. While the work of Wixom and Watson (2001) offers an similar data warehousing implementation success model, our model in Figure rests on IT implementation, interorganizational systems (IOS) and economic literature. As a result, our work augments that of Wixom and Watson by including those variables that are germane to the push/pull climate (e.g., anticipated economic benefits and competitive pressures), noting political climates along with autonomy/control and appending quality of information sharing. Further, our work seeks to explore the implementation of a data warehouse that was not deployed initially and exclusively for CRM and other marketing applications.

Three Clusters of Factors in the Model

In Figure 1, three factor clusters are defined: push/pull factors, behavioral factors and shared systems topologies. *Push or pull factors* are elements that can persuade an organization's willingness to adopt a given technology, strategy, and/or change initiative.

Behavioral factors are aspects that stand to impact and/or influence stakeholders including end-user support. *Shared or integrated systems topologies* represent certain aspects of the infrastructure needed for a data warehouse. These factors include arrangements for cooperation and information sharing as well as for assuring information quality. Payton and Ginzberg (2002) provide a complete explanation for the Figure 1.

In our exploratory effort, the dependent variable in this model is *success of the implementation effort, or, in this case, use of the corporate data warehouse by marketing for its functional work.* It is interesting that we did not add any specifically marketing-oriented questions. As will be seen, questions themselves were broad enough to elicit responses to the larger question at hand, "What factors impact and/or stand to impact implementing our corporate data warehousing to support CRM applications for the marketing function"?
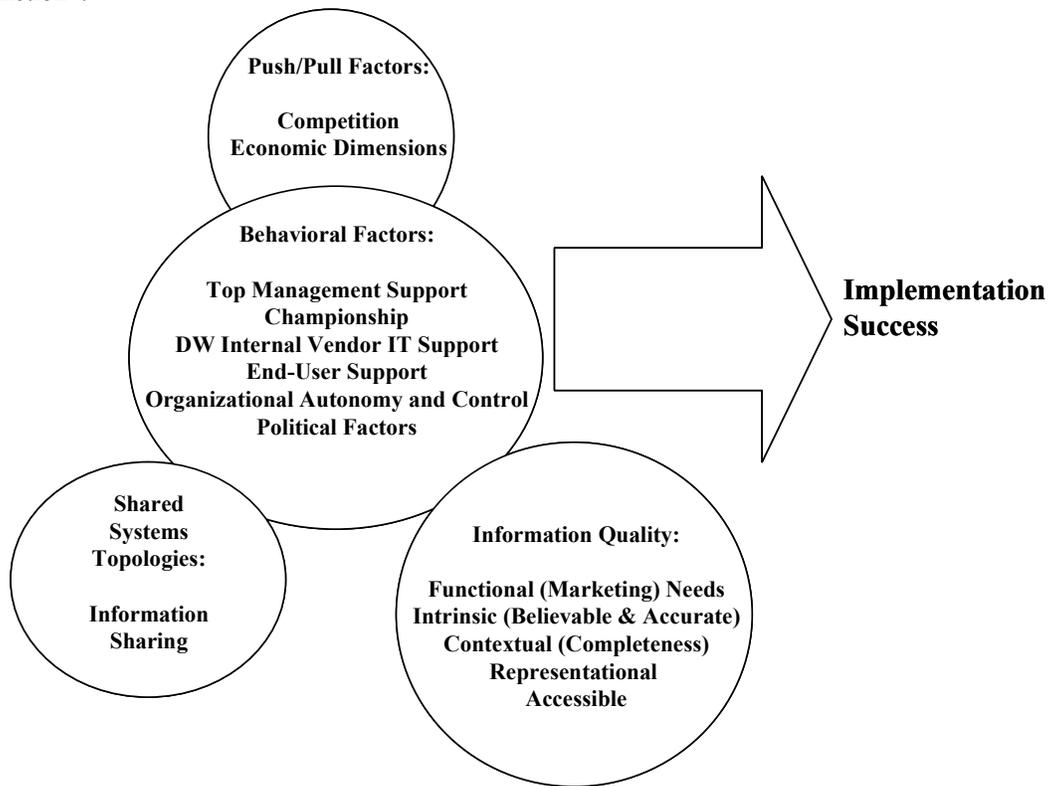

Figure 1. Corporate Data Warehouse (CDW) Implementation Model

Research Method

Data were gathered from five two-hour focus group sessions with a large health care payor group. Sessions were held almost exclusively with marketing and IT teams and included current, forthcoming and power users; data warehousing internal systems staff and middle management. Note that some power users represented the finance department. Appendix A lists the questions that were asked of all focus group participants, and these items had been adopted from the prior implementation work of Payton and Ginzberg (2001). Prior to using these questions in our focus group sessions, we pre-tested our research model via five one-hour interviews, participatory observation, meeting notes and other fact-finding sessions with managers from the health care organization. This pre-work enabled us to determine the validity of our constructs in the context of the organization.

All focus group sessions were recorded and transcribed by a professional writer on staff at the health care organization as well as a member of our research team within 48 hours. Names and titles were withheld from all transcribed documents. Top management agreed not to attend the focus group sessions to avoid biased responses to questions asked.

To analyze the focus group data, we applied analyses methods as prescribed by Currall, et al. (1999) and Boudreau, et al. (2001). Based on Payton and Ginzberg (2001), we developed coding dictionaries to capture the marketing and IS constructs. Our coding dictionaries can be found in Appendices A and B. As in the case of Currall, et al. (1999), we used a three-dimensional structure to code the participants' comments that includes individuals (total of twenty-one), topics and verbal contributions. A next phase is to add a third coder to allow us to determine inter-coder reliability using Scott's π.

Findings

The detailed results of the analysis of all five focus groups are reported in Appendices A and B and summarized in Table 1. As Table 1 indicates, the needs of marketing were most frequently mentioned as important to the participants as preventing them from using the Corporate Data Warehouse (CDW) application. The entire model was supported, but quality issues and needs of marketing dominated. Marketing needs that can be linked to data quality issues, marketing needs, overall data quality issues and training were the most important factors in preventing the widespread adoption of the CDW in this organization. These categories represent 66% of total statements coded by the two independent coders. Thus, the analysis will focus on describing these categories in more detail.

Table 1. Summary of All Coded Items, Ranked by Frequency of Mention

Frequency	Coded Elements	Model Category
99	Quality: Believability, completeness, timeliness, amount, ease of understanding and accessibility	Data Quality
76	Ease of use of interface (Quality Accessibility)	Marketing Needs
69	Top Management Support	Top Management Support:
64	Ease of use of CDW itself (Quality-Accessibility)	Marketing Needs
57	Marketing Analysis Needs	Marketing Needs
51	System provides accurate/right data	Data Accuracy
49	Standardization of data definitions & structures	Data Integration
38	Training should be based on the actual users roles, job tasks	Training
38	Need other data in CDW (aka Quality-Completeness)	Marketing Needs
38	Marketing Needs Not a Priority	Marketing Needs
34	Systems used prior to current applications	Legacy Systems
30	Lack of Trust in System (aka Quality-Believability)	Marketing Needs
29	The training potential of users should be assessed	Training
29	Legacy systems are vital data sources for CDW	Legacy Systems
27	Data Dictionary Needs (aka Quality-Ease of Understanding)	Marketing Needs
25	Determination of user requirements	Needs Assessment
23	External Data Needs (aka Quality-Completeness)	Marketing Needs
22	Communication between group members, users and managers	Needs Assessment
20	Power users have functional jobs and also are gurus for CDW	Power Users
18	Data that are fit to be used by data consumers	Data Quality
16	Perceived management support with willingness to commit resources	Needs Assessment
13	Tools used to educate users	Training
10	Users should be engaged in the training process from the beginning	Training
9	System supports functional elasticity	Flexibility
8	The ability of CDW to contribute to perceived economic benefits	Economic Impact
7	Strength of management's message re: CDW	Needs Assessment
6	Where & how the CDW must evolve (System Flexibility)	Flexibility
5	Links architecture, planning, HR skills	Flexibility
5	Implementing CDW is expensive but nevertheless useful	Economic Impact
3	Alternative System Needs	Marketing Needs
1	Definitions agreed upon over time	Data Integration

919 total statements, 462 per each of two coders

Overall, users supported the idea of the warehouse and the benefits derived from the technology. The focus groups, however, indicated that marketing needs included specialized analysis that could not be performed with the present combination of warehouse data and analysis tools. These analyses were fairly standard data mining/marketing types of applications (Shaw et. al. 2001), such as definition characteristics to target marketing programs, longitudinal analysis and the comparison of customers who had purchased and those who had not. The general category of marketing needs included several categories which, after our initial analysis, appeared to be strongly related to overall data quality dimensions (Wang and Strong, 1996). The users mentioned of the ease of use of the interface as well as overall ease of use of the warehouse, which could also be considered the Accessibility dimension of quality (Wang and Strong ,1996). Marketers expected the interface to resemble that experienced in training or to be similar to other packages (e.g., SPSS) that they used in their work by capturing "click and drag" features. Participants were less likely to use reporting tools (e.g., Business Objects) and quite likely to consult local "power users". These power users had become residential experts with the application and as a consequence were somewhat overworked in the organization. Moreover, our findings run counter to prior works (Wixom and Watson, 2001) and points out the significance of data quality to data warehouse users, particularly those external to the IT function. Further, more data oriented trends impacting the health care industry, as a whole, can have bearing on this finding – as health care payors and providers are seeking more CRM related data to meet competitive market pressures.

Users also mentioned that they needed other data that were not in the warehouse, which relates to the Completeness dimension of Data Quality. These needs were for external, demographic and descriptive data for the consumer market and Dun & Bradstreet company descriptive data for the commercial (B2B) applications which were missing from the data warehouse. Other missing data included information on former customers and prospective customers, neither of which are in the CDW. In addition, users expressed interest in a data dictionary, which relates to the perceived Ease of Understanding of the data. Users did not have an exact idea what data items meant and consequently could not use them. Perhaps most importantly, users and potential users continue to distrust data extracted from the warehouse, which relates to the dimension of quality known as Believability (Strong and Wang, 1996). Focus group participants reported "reconciling" mainframe reports to the information from the data warehouse. In addition, overall quality issues were mentioned frequently, such as overall accuracy of the information and other dimensions of data quality. See Figure 1 for bold highlighted constructs added to the research model.

Finally, the issue of training was one of the most frequently mentioned, with users reporting that training should be based on actual job roles and responsibilities, not on imagined data sets not related to their work. Also, the focus groups indicated that users should be involved in the training from the beginning and that potential users should be assessed before entering internal IS CDW training sessions. In fact, it seems as if training and data quality could be related, since users might be more comfortable with the data if they understood the content of the warehouse.

Conclusions

This study indicates that marketing applications must be considered carefully before the data is developed for a Corporate Data Warehouse if marketing is going to use the data. While this paper complements prior research which focused on strategic issues (Cooper et al., 2000), our work suggests that data quality is a critical factor in determining data warehousing implementation among users peripheral to the IT function. The contribution of this research is in further identifying the components of an overarching model which can facilitate success in the field. Finally, in addition to receiving training to use the system, why are the dimensions of quality so important in the implementation of a data warehouse by the marketing function? One possible explanation for the overriding importance of data quality issues in the adoption of the Corporate Data Warehouse comes from social relations theory. The adoption of an IT system is analogous to a commercial relationship in which trust is a critical element in willingness to enter into and continue in a relationship (Morgan and Hunt, 1994). Quality appears to signal trust in a system and a willingness to move forward in the relationship, or, in this case, the implementation the data warehouse. Lastly, data quality can also be an indication of competitive pressures in health care and other data intensive organizations – as they strive to strategic use corporate data to obtain new and retain existing customers.

References

Beitler, S.S., and R. Leary. "Sears' EPIC Transformation: Converting from Mainframe Legacy Systems to OLAP", *Journal of Data Warehousing* (2:2), 1997, pp 5-16.

Boudreau, M.C., D. Gefen and D. Straub. "Validation in Information Systems Research: A State-of-the-Art Assessment", *MIS Quarterly* (25:1), 2001, pp 1-16.

Burton, B., 2001, Gartner Research Note: DF-13-5981, Avoiding the Data Warehousing 'Money Pit'

Cooper, B.L., H.. Watson, B. Wixom and D.L. Goodhue. "Data Warehousing Supports Corporate Strategy at First American Corporation", *MIS Quarterly* (24:4), 2000, pp 547-567.

Cooper, R.B. and R.W. Zmud. "Information Technology Implementation Research: A Technological Diffusion Approach", *Management Science*, 1990, pp 404-420.

Gartner (author: Arthur Hochberg). "Changing IT Priorities in the Year 2000" as noted at http://www.knightsbridge.com/big_data3.html#sources).

Grim, R. and P. Thorton. "A Customer for Life: The WarehouseMCI Approach", *Journal of Data Warehousing* (2:1), 1997, pp 73-79.

Kalakota, R., and Robinson, M. *e-Business, Roadmap for Success*, Addison-Wesley, Reading, Massachusetts, 1999, pp. 67-69.

Lee, A.S., Liebenau, J. & DeGross J.I. (Eds.). 1997. *Information Systems and Qualitative Research*, New York: Kluwer/Chapman and Hall.

Morgan, R. M., and Hunt, S. D. "The Commitment-Trust Theory of Relationship Marketing," *Journal of Marketing* (58:3), July 1994, pp. 20-38.

Nelson, S., and B. Eisenfeld, "Salvaging a Failed CRM Initiative," Gartner Report SPS-15-4007, 2002.

Payton, F.C. and M.J. Ginzberg. "Interorganizational Health Care Systems Implementations: An Exploratory Study of Early Electronic Commerce Initiatives", *Health Care Management Review* (26:2), 2001, pp 20-32.

Rigby, D.K., F.F. Reichheld and P. Schefter. "Avoid the Four Perils of CRM", *Harvard Business Review* (80:2), pp 101-109.

Romano, N.C. "Customer Relationship Management Research: An Assessment of Sub Field Development and Maturity", *Proceedings of the 34th Hawaii International Conference on System Sciences*, 2001.

Shaw, M. J., Subramaniam, C., Tan, G. W., and Welge, M. E. "Knowledge management and Data Mining for Marketing," *Decision Support Systems* (31:1), May 2001, pp. 127-137.

Wang, R. Y., and Strong, D. M. "What Data Quality Means to Data Consumers", *Journal of Management Information Systems* (12:4), Spring 1996, pp. 5-34.

Watson, H. J. and B. J. Haley. "Data Warehousing: A Framework and Survey of Current Practices", *Journal of Data Warehousing* (2:1), 1997, pp 10-17.

Wixom, B. and H. Watson. "An Empirical Investigation of the Factors Affecting Data Warehousing Success", *MIS Quarterly* (25:1), 2001, pp 17-41.

Zwass, V. "Electronic Commerce: Structures and Issues", *International Journal of Electronic Commerce* (1:1), 1996, pp 3-23.

Appendix A
Coded Items Relating to Marketing Needs,
Ranked by Frequency of Mention and
Sub-Topic Within Major Topic

Frequency	Average	Coded item
Ease of use of interface		
15	7.5	Can't meet time frame using CDW currently
14	7	Takes too long for analysis
12	6	Business Objects not accessible to all analysts
10	5	Takes too long for requests for changes
10	5	Business Objects are easier to use
7	3.5	Executive area can't get answers to marketing questions easily
5	2.5	SQL not accessible to all analysts
3	1.5	Ad hoc queries only, not production of marketing applications
76	38	Total-Ease of use of interface
Ease of use of CDW itself		
54	27	Expect to be able to access data easily
10	5	CDW can't be used to answer basic questions
64	32	Total- Ease of use of CDW itself
Marketing Analysis Needs		
		Types of analysis that CDW can't perform:
15	7.5	Group characteristics
9	4.5	Contains information only on enrolled groups and members
7	3.5	How many members in each group
4	2	Identifying profitable customers
4	2	Link leads historical data (longitudinal)
4	2	Can't monitor prospects because they are in CDW
3	1.5	Segmentation of offers
3	1.5	Inability to track things longitudinally and across products
3	1.5	Should identify type of business producer (sales person) has been selling over time
3	1.5	Need to manage customers more effectively and efficiently (analysis)
1	0.5	Differences between Customers who purchased and did not
1	0.5	Others
57	28.5	Total-Marketing Analysis Needs
Need other data in CDW		
38	19	Data not there (Need to gather information)
38	19	Total-Need other Data in CDW
Marketing Needs Not a Priority		
21	10.5	Marketing Needs not given top priority
10	5	Technical people don't know marketing terms so they can't help us
7	3.5	Marketing lost influence over the warehouse
38	19	Total-Marketing Needs not a Priority
Lack of Trust in System (Quality)		
17	8.5	CDW reports are "reconciled" to mainframe reports
13	6.5	Mainframe reports are still used, believed in more than CDW
30	15	Total-Lack of Trust in System (Quality)

Frequency	Average	Coded item
		Data Dictionary Needs
16	8	Have common definitions
8	4	Need "data definition sheets"
3	1.5	Expect all data to be in one place
27	13.5	Total-Data Dictionary Needs
		External Data Needs
11	5.5	Need Demographic data in CDW
9	4.5	Need a marketing "data universe"
3	1.5	Need Dun & Bradstreet data in CDW
23	11.5	Total-External Data Needs
		Alternative System Needs
		Other systems that might meet needs:
2	1	Opportunity Management (being considered)
1	0.5	Marketing System outside of CDW
3	1.5	Total-Alternative System Needs

924 total statements coded, 462 per each of two coders

Appendix B
Coded Items Relating to IS Needs, Ranked by Frequency
of Mention and   ub-topic   ithin Major Topic

Frequency	Average	Coded item
		Data Integration
49	24.5	Standardization of data definitions & structures
1	0.5	Definitions agreed upon over time
50	25	Total-Data Integration
		Flexibility
9	4.5	System supports functional elasticity
6	3	Where & how the CDW must evolve (System Flexibility)
10	5	Links architecture, planning, HR skills
25	12.5	Total-Flexibility
		Training
38	19	Training should be based on the actual users roles, job tasks
29	14.5	The training potential of users should be assessed
13	6.5	Tools used to educate users
10	5	Users should be engaged in the training process from the beginning
90	45	Total-Training
		Data Accuracy
51	25.5	System provides accurate/right data (as dimension of quality)
		Data Quality
18	9	Data that are fit to be used by data consumers a
99	49.5	Other Quality dimensions: Believability, completeness, timeliness, amount, ease of understanding and accessibility
117	58.5	Total-Data Quality
		Legacy Systems
34	17	Systems used prior to current applications
29	14.5	Legacy systems are vital data sources for CDW
63	31.5	Total-Legacy Systems

Frequency	Average	Coded item
Needs Assessment		
25	12.5	Determination of user requirements
22	11	Communication between work group members, users and managers
16	8	Perceived management support with willingness to commit resources
7	3.5	Strength of management's message re: CDW
70	3.5	Total-Needs Assessment
Power User Roles		
20	10	Power users have functional jobs and also are gurus for CDW
20	10	Total-Power Users
Economic Impact		
8	4	The ability of the application to contribute to perceived economic benefits
5	2.5	Implementing CDW is expensive but nevertheless useful
13	6.5	Total-Economic Impact
Top Management Support:		
13	6.5	Message is not communicated throughout the company that CDW is important
13	6.5	Top managers don't have time to learn about warehouse
11	5.5	Top Managers need to find out what is going on in Warehouse
10	5	Top management role not visible
9	4.5	Managers are key to making CDW a success, vendor evaluation
7	3.5	Top Management role not clear
3	1.5	Managers came to training session
2	1	Should have realistic expectation out of warehouse, have better communication
1	0.5	Managers left the training session early; questions importance of CDW
69	34.5	Total-Top Management Support

924 total statements coded, 462 per coder