Conditions for the Detection of Data Errors in Organizational Settings: Preliminary Results from a Field Study

Barbara D. Klein
University of Michigan-Dearborn

Dale L. Goodhue and Gordon B. Davis
University of Minnesota

Abstract

There is strong evidence that data stored in organizational databases have a significant rate of errors. As computerized databases continue to proliferate and as organizations become increasingly dependent upon these databases to support business processes and decision making, the number of errors in stored data and the organizational impact of these errors are likely to increase. Two main approaches to this problem are validating data as they are input to or stored in databases and (2) depending on users to detect and correct errors. A research program examining the efficacy of the second approach is underway. This paper discusses the results of a field study conducted as part of this research stream. The overall objective of the field study is to increase our understanding of how users in different professional domains deal with errors in data. This objective was met through field interviews in three professional domains that were selected based on expected variation along two theoretically derived dimensions: the base rate of errors in the domain and perceived payoffs for error detection in the domain. The findings of the field study suggest that the materiality of data errors and expectations about the base rate of errors are related to error detection performance. Although we did not find a straight-forward relationship between different levels of incentives to detect errors and error detection performance, it appears that users who believe that strong incentives to detect data errors are present in their organizations are more likely to detect data errors than those without this belief.

1.0 Introduction

There is strong evidence (e.g., Laudon, 1986; Morey, 1982; Redman, 1992) that data stored in organizational databases have a significant rate of errors. Between one and ten percent of data items in critical organizational databases are estimated to be inaccurate (Laudon, 1986; Madnick and Wang, 1992; Morey, 1982; Redman, 1992). This estimate is based on the findings of several studies such as those by Laudon (1986) and Morey (1982). Inaccurate data have been reported in a student loan database maintained by the U.S. Department of Education (Knight, 1992), in records maintained by the U.S. Department of Agriculture ("Dead farmer," 1992), and in records maintained by credit reporting bureaus ("Consumer enemy," 1991).

Technical developments that have occurred during the past decade have increased the need to understand the detection of errors in data. For example, the proliferation of end-user computing has increased the potential for data errors in computer applications (Boockholdt, 1989). As end users develop applications, it is possible that fewer data validation methods such as logic tests and control totals will be in place and it is likely that less rigorous testing will occur before applications are used in production (Corman, 1988; Davis, 1984; Davis et al., 1983). End-user computing may also increase data errors in organizations because data that are processed and stored using information systems developed and maintained by end users may not be kept consistent with data that are stored in centralized organizational databases containing information referring to the same entities (Maxwell, 1989; Nesbit, 1985). Thus, problems related to the currency of data may develop.

As computerized databases continue to proliferate and as organizations become increasingly dependent upon these databases to support business processes and decision making, the number of errors in stored data and the organizational impact of these errors are likely to increase. Indeed, Mason (1986) argues that the scope of this problem is such that data quality will become an important issue facing MIS managers. Redman (1992) argues that inaccurate and incomplete data may adversely affect the competitive success of an organization. For example, strategies such as total quality management may be difficult to implement if the data needed to support the decisions required by the strategy are not of adequate quality (Fox et al., 1993; Madnick and Wang, 1992; Redman, 1995). Errors in data can have a significant financial impact on organizations. For example, Dun & Bradstreet paid $350,000 to a construction company after they incorrectly reported that the company was bankrupt because a Dun & Bradstreet employee had entered inaccurate data into their credit reporting system (Percy, 1986).

Two main approaches to this problem are validating data as they are input to or stored in databases (e.g., Morey) and (2) depending on users to detect and correct errors. A research program examining the efficacy of the second approach is underway. To data, several studies have been completed in this research stream. The first study was a field study in one business domain (actuarial science) showing that at least some users of information systems detect errors in data in organizational settings (Klein, 1996). As the next step in the research stream, two laboratory experiments were conducted to examine the impact of base rate expectations, incentive structures, and error detection goals on performance in the detection of errors. The conclusion drawn from the laboratory experiments is that incentive structures and error detection goals affect error detection performance. No main effect for base rate expectations was found, possibly because this factor was not successfully manipulated (Klein, 1995b; Klein, 1995c).

This paper discusses the results of a field study conducted to link the findings of one of these laboratory experiments to practice in organizations. The overall objective of the field interview study is to increase our understanding of how users in different professional domains deal with errors in data. The field study described in this paper and the laboratory experiment conducted as part of the overall research stream are complementary. While the experiment allows us to make valid measurements of performance outcomes, the findings of laboratory experiments are often criticized as not being generalizable to organizational settings. While it is not possible to develop valid measures of performance outcomes using the data collected in the field interviews, the findings from the analysis of the interviews increase our understanding of the applicability of the laboratory findings to organizational practices.

A second objective of the field study described here is to examine whether users working in other professional domains attend to errors in data in a manner similar or dissimilar to that of actuaries. Actuaries were selected as the subjects of the first field study in the research stream because it was thought that they would be quite likely to detect errors in data. One possibility is that users in other domains also detect many errors in the data they use. On the other hand, it may be that users from domains that vary systematically from actuarial science infrequently detect errors in data.

The remaining sections of this paper present (1) a review of prior research bearing directly on the question of the conditions under which individuals detect errors in data, (2) a theory of error detection, (3) the design of the field study, (4) the results of the field study, and (5) conclusions and suggestions for further research.

2.0 Background

In a broad sense, this investigation falls in the literature on data quality. Several general conclusions can be drawn from the existing research on data quality. First, while no single definition of data quality has been accepted by researchers working in this area, there is agreement that data accuracy, currency, and completeness are important areas of concern (Agmon and Ahituv, 1987; Davis and Olson, 1985; Fox et al., 1993; Huh et al., 1990; Madnick and Wang, 1992; Wand and Wang, 1994; Zmud, 1978). Second, while it is difficult to compare error rates across studies, rates substantially greater than zero have been found in all of the studies addressing the extent to which data errors exist in databases (Ham et al., 1985; Johnson et al., 1981; Knight, 1992; Laudon, 1986; Stone and Bublitz, 1984). Third, there is disagreement about the extent to which efforts to purge all errors from databases should be attempted. Some researchers propose methods designed to completely rid databases of errors (Janson, 1988; Misvanks, 1988; Naus, 1975; Parsaye and Chignell, 1993), while others propose tools for determining how to best allocate limited resources to controlling the level of data errors (Ballou and Pazer, 1987; Ballou and Tayi, 1989; Ballou et al., 1987; Bowen, 1992; Paradice and Fuerst, 1991). Fourth, many researchers argue that users need not discard data containing errors. A variety of approaches for using imperfect data have been suggested (Ballou and Pazer, 1985; Ballou and Pazer, 1995; Ballou and Pazer, 1987; Bansal, 1993; Gaba and Winkler, 1992; Garfinkel et al., 1986; O'Leary, 1993; O'Neill and Vizine-Goetz, 1988).

This study fits into the literature on data quality that describes ways in which users of information systems might modify their use of data if they are aware that errors exist in data. In general, the data quality literature argues that users are not very capable of finding errors in data and then altering the way in which they use the data. More specifically, there is considerable evidence of poor user performance in detecting data errors. Davis et al. (1967) conducted a field experiment in which individuals were mailed banking confirmation statements with imbedded errors. The individuals were asked to verify their account information, and approximately half failed to detect important errors. Laudon (1986) found that users of criminal information systems rarely detected errors in these records even though information provided to police departments by the FBI is accompanied by a warning stating that the user should verify that the information is accurate. Ricketts (1990) conducted a laboratory experiment in which over ninety percent of the subjects failed to detect a substantial data error in production planning reports. The failure of humans to detect errors in data is also assumed in much of the literature on data quality in which it is argued that resources should be devoted to the up-front improvement of the quality of data in organizational databases (e.g., Redman, 1992; Redman, 1995).

Guided primarily by a strong sense that the above research misrepresented organizational behavior in handling data, one of the authors conducted a pilot study to investigate whether actuaries typically detect errors in data (Klein, 1996). Actuaries were studied because it was expected that (1) they would be highly motivated to detect errors since they use data of uncertain accuracy generated outside of their organizations and (2) they would develop effective methods for working with inaccurate data because errors frequently occur in this domain.

A series of semi-structured interviews was conducted with ten actuaries. All of the actuaries reported instances in which they successfully detected errors in data. The degree to which data items were reviewed for errors prior to use was affected by the actuaries' expectations about the likelihood of errors in a given dataset. The actuaries did not attempt to detect all errors, apparently responding to a perceived tradeoff between the time and effort required to find additional errors versus the potential impact of a more accurate dataset. The pilot study results suggested the value of additional research.

Conflicting findings between prior studies and the study of the actuaries motivated the study discussed in this paper. While prior findings (e.g., Davis et al., 1967; Laudon, 1986; Ricketts, 1990) suggest that error detection is not a frequently occurring phenomenon, the findings of the field study of actuaries suggest that users in at least one professional domain detect data errors. This raised the question of whether error detection in organizational settings is a rare phenomenon occurring only under a limited set of conditions or a more frequently occurring phenomenon. This question was addressed through field interviews in three additional professional domains that were selected based on expected variation along two theoretically derived dimensions: the actual base rate of errors in the domain and perceived payoffs for error detection in the domain.

3.0 Theoretical Framework

A theory of individual task performance and theories of effort and accuracy in decision making underlie this research.

3.1 Theories of Individual Task Performance

Theories of individual task performance provide some general guidance for identifying conditions under which users detect errors in data. For example, experience seems to affect performance in general and may affect performance in error detection. One theory we could use is Campbell's (1990; Campbell and Pritchard, 1976) theory of individual task performance (depicted in Figure 1). This suggests that experience (e.g., Weber et al., 1993), knowledge, and effort (e.g., Payne, 1982; Payne et al., 1988) all affect error detection. This suggestion is consistent with the findings of the interviews with the actuaries.

Performance = f (declarative knowledge X
procedural knowledge and skills X
choice to expend effort X
choice of degree of effort to expend X
choice to persist)

Figure 1 Determinants of Individual Task Performance

Campbell's (1990) theory argues that performance on a particular component of a job is a function of an individual's declarative knowledge, procedural knowledge and skill, and motivation. Declarative knowledge is defined as knowledge of the facts required to complete a task. Procedural knowledge refers to skill-based knowledge about how to perform a task. Declarative knowledge and procedural knowledge are said to be partially a function of education, training, and experience; and motivation is said to be a function of three choices: the choice to expend effort, the choice of the degree of effort to expend, and the choice to persist in task performance. The theory suggests that experience and motivational influences can only affect job performance through changes in declarative knowledge, procedural knowledge and skill, or the three choices related to effort.

Error detection is viewed here as a very specific component of some jobs that is influenced by these determinants, and performance is viewed in this research as the successful or unsuccessful detection of an error in data. We argue here that variation in declarative knowledge and procedural knowledge affect error detection performance and that differences in expectations about the base rate of errors in data and assessments of the payoffs of error detection affect the rate at which errors are detected through the choices related to effort.

3.1.1 Experience and Knowledge.

Studies of expert performance suggest that significant amounts of experience are necessary for the development of expertise (e.g., Ericsson and Chase, 1982; Johnson et al., 1992a; Johnson et al., 1992b). This suggests that the actual number of errors that users of data encounter will influence performance if they recognize the problem and try to detect the errors. A high base rate of errors has the potential, when adequate feedback occurs, to facilitate the development of declarative knowledge about the number and types of errors in data. Users working with data containing many errors also have more opportunities to develop the procedural knowledge and skills needed to detect errors than users working with data with a low base rate of errors. Thus, users in domains with a high base rate of errors may develop effective strategies for error detection, and performance in the task of error detection may be enhanced.

3.1.2 Effort.

Effort expended to detect errors may be a function, at least in part, of expectations about the base rate of errors in data and of user assessments of the payoffs of error detection. Campbell's (1990) theory of performance suggests that choices about the degree of effort to expend in the detection of errors will influence performance. Analysis of the data collected in the study of the actuaries suggests that there are several factors that influence these choices when users work with imperfect data. Factors influencing choices related to effort in error detection are discussed below.

Expectations about the Base Rate of Errors in Data. As users expect more errors in data, greater effort may be devoted to this task. Compared to users who expect a low base rate of errors in data, users who expect a high base rate of errors may expend more effort to detect errors simply because they expect to detect more errors at any level of expended effort. There is evidence from the study of actuaries that expectations about the base rate of errors in a source of data influence effort. For example, one user reported that she considers the base rate of errors in published mortality tables to be low and that she does not attempt to find errors in these tables. There is also evidence from the work of Weber et al. (1993) that at least some decision makers are sensitive to base rates in the generation of hypotheses in diagnostic tasks.

Payoffs of Error Detection. A specific task described by a subject in the study of actuaries will be used to illustrate the impact of assessments of payoffs on error detection performance. In this incident an actuary was using data provided by a client. The objective was to determine whether an organization's financial reserves for its pension fund were sufficient. This judgment typically depends in part on the pay rate and the number of years of organizational service of each employee in the organization. The data provided by the client contained this information along with other personnel information for each employee as of the end of the year. Imagine a specific case in which this data (as of the end of 1995) contains a record holding information about an accountant in a position requiring a CPA certificate in which the value of the Date of Birth field is "December 31, 1970" and the value of the Number of Years of Service field is "10".

A user working with this dataset might or might not suspect that the data in one of these fields is inaccurate (i.e., it is unlikely that a firm would hire an accountant at the age of 15). An actuary analyzing a pension fund might be likely to detect this error because it is quite material to the judgment about the sufficiency of the firm's pension reserves. On the other hand, a payroll manager reviewing the same dataset might be unlikely to find the error because errors in the Date of Birth and Number of Years of Service fields are not material to a firm's payroll.

Materiality. Thus, beliefs about the materiality of an error may influence the degree of effort expended to detect the error. Users may expend more effort to detect errors that they believe will have a significant impact on their calculations or decisions. There is evidence from the pilot study that the impact of data errors on the work being performed using the data is explicitly considered in the determination of the level of effort to expend in error detection. For example, one actuary stated that there are some types of errors that he does not try to detect when pricing insurance because the errors would not have a significant impact on his calculations.

Incentives. Organizational incentives may also play an important role in users' assessments of the payoffs of error detection. For example, an error in data that is successfully detected may generate additional work for the detector; and an incentive system that discourages the use of time to investigate and correct errors may create an environment in which many errors in data go unnoticed.

Ease of Verification and Correction. The ease with which an error can be corrected may also influence the degree of effort expended to detect the error. For example, it is possible that individuals won't try very hard to detect an error if (a) it is difficult to confirm that a suspected error is actually an error, or (b) if a confirmed error cannot be corrected. Perceived payoffs of error detection may be quite low if detected errors cannot be corrected.

3.2 Theories of Effort and Accuracy in Decision Making

An underlying assumption of theories of effort and accuracy in decision making is that humans will devote no more mental resources or effort to a task than what is demanded by task requirements. This suggests that performance in the task of error detection may be sensitive to the specific performance requirements implicit in payoffs for error detection. Payne (1982; Payne et al., 1988; Johnson and Payne, 1985) has demonstrated that task demands influence the selection of information processing strategies. Cryer et al. (1990) built on Payne's research to examine the impact of incentive schemes on information use and on task performance. Cryer et al. (1990) found that an incentive scheme rewarding accuracy would lead to more normative information processing and higher levels of task performance while an incentive scheme rewarding the minimization of effort would lead to the use of heuristic processing and lower levels of task performance. This finding supports the contention that error detection performance may be sensitive to variation in payoffs.

4.0 Design of the Field Study

This section begins by discussing the procedure used to select professional domains for inclusion in the field study. Next, the methodology of the study is outlined.

4.1 Selection of the Professional Domains

Field interviews were conducted in three professional domains. Domains were selected that were expected to vary with respect to (1) the base rate of errors in the data and (2) the perceived payoffs of error detection. Theories of effort and accuracy in decision making and Campbell's theory of individual performance suggest that these domain differences will affect error detection. The choice of domains for further study follows the advice of Eisenhardt (1989) and Yin (1989) with respect to the selection of theoretically interesting cases. The three additional domains studied are (1) consumer product management, (2) inventory management, and (3) municipal bond analysis. An initial investigation consisting of one interview with a domain expert in each area was conducted. This investigation suggested that the domains differ with respect to the base rate of errors and perceived payoffs of error detection. The expectations coming out of the initial investigation are summarized below.

Consumer Product Managers. Product managers analyze data that are collected using scanners located at checkout counters in retail organizations. An initial interview with an information systems manager of an organization that sells this data to consumer product companies indicated that the base rate of errors in this data is high and that the motivation of product managers to detect errors is low. The interview suggested that errors in scanning, errors in aggregation, and errors of classification (e.g., classification of the data by sales territory) occur in this data. It was suggested that effort to detect errors may be low in this domain because there is little feedback about the successful and unsuccessful detection of errors. A consumer product manager using scanner data to forecast sales of a product may never know if there is an undetected error because forecasts are not expected to exactly equal future sales. Ease of correction may also contribute to low motivation in this domain because it is difficult to verify that a suspected error in scanning is actually an error and it is difficult to correct a scanning error if it is verified. The initial investigation also suggested that the materiality of errors in this domain may be low because some errors in scanner data are random in nature (e.g., some errors in scanning), and the impact of these errors on analyses made using the data may not be significant.

Inventory Managers. The second domain that was selected for the field interviews is inventory management. Observations in two organizations suggested that the base rate of errors is relatively low in this domain while effort to detect errors is high. The value of the inventory held in both of these organizations was relatively high, and we acknowledge that the base rate of errors may be higher in organizations holding less valuable inventories. The specific users studied in this domain were material managers using inventory control data. The base rate of errors in inventory control data is relatively low because its accuracy is periodically checked by internal and external auditors. Some errors do occur, however, as discrepancies between the number of items held in physical inventory and the number of items listed in computerized inventory records develop between audits. There are several reasons that we expected effort to detect those errors that do exist to be high. First, errors in inventory data may materially impact the competitive success of a firm. If inventory records overstate the number of goods in physical inventory, stockouts may occur and customer orders may not be filled. On the other hand, if inventory records understate the number of goods in physical inventory, unnecessary inventory carrying costs are incurred. It is also more likely that theft of inventory will occur if it is believed that inventory records do not correspond to physical inventories. A second reason we expected effort to detect errors to be high in this domain is that it is relatively easy to confirm and correct errors by making a physical count of the goods about which the accuracy of inventory records is questioned. Formal auditing procedures also encourage high detection motivation in this domain in two ways. First, errors that have not been detected during normal business operations may be detected during an audit. This provides regular performance feedback to users of these data. Second, organizational incentives discourage an outcome in which so many errors are detected during an audit that the auditors conclude that the inventory system is not properly controlled.

Municipal Bond Analysts. Municipal bond analysts work with data presented in the financial statements of not-for-profit organizations and state and local governments. An initial interview with a municipal bond analyst at a large investment bank suggested that the base rate of errors in the data with which these professionals work is low and that effort to detect errors is also low. The informant in the initial interview acknowledged that there are some errors in this data. However, he said that he does not actively look for errors when working with the data. One factor that could influence effort in this domain is that some of the data with which bond analysts work is audited by public accountants before it is used. The analysts may therefore assume that the most significant errors in this data have already been detected. It is possible that effort to detect errors among this group of professionals is low because any problems stemming from undetected errors could be attributed to the failure of the external auditors to detect the errors. The initial investigation also suggested that ease of correction is low in this domain.

Figure 2 depicts our *a priori* expectations about the theoretical variance among the groups of users studied in the field interviews. These groups of users were selected to get as much variance as possible in order to detect differences in the detection of data errors across the two theoretical dimensions shown in Figure 2.

		Base Rate of Errors in the Domain	
		Low	High
Perceived Payoff of Error Detection in the Domain	Low	Municipal Bond Analysis	Consumer Product Management
	High	Inventory Management	Actuarial Science

Figure 2. ***A Priori* Expectations About Domains**

4.2 Methodology

Five professionals were selected for each of the three additional sets of field interviews. To control for selection bias, potential interviewees were asked to participate in a study of the use of data in their work. The terms "error detection" and "data quality" were not used when recruiting subjects. Data were collected using a semi-structured interview. Several of the questions in the interview protocol are a variation on the critical incidents methodology developed by Flanagan (1954). These questions were designed to elicit descriptions of incidents in which the interviewees successfully detected errors in data and incidents in which errors were missed.

The semi-structured interviews were recorded and transcribed. An analysis of the interview transcripts was performed using methodologies outlined by Miles and Huberman (1994) and King (1994). A coding scheme based on the theoretical framework was developed, and the transcripts were coded using this scheme.

Two researchers independently coded two of the interview transcripts. Overall, the level of agreement was 92 percent. The two coders were in complete agreement about the presence or absence of evidence reflecting incidents of error detection and incidents in which errors in data were missed. One coder scored the remaining thirteen transcripts.

5.0 Results

This section begins with a domain-level analysis of the field interview data. Next, an analysis at the level of the individual respondent is presented.

5.1 Domain-Level Analysis

Figure 3 shows the number of interviewees in each domain reporting actual incidents of error detection.

		Base Rate of Errors in the Domain: Low	Base Rate of Errors in the Domain: High
Perceived Payoff of Error Detection in the Domain	Low	Municipal Bond Analysis (5 out of 5)	Consumer Product Management (2 out of 5)
	High	Inventory Management (2 out of 5)	Actuarial Science (10 out of 10)

Figure 3. Number of Interviewees Reporting Error Detection Incidents

As expected, the consumer product managers and the inventory managers were less likely to report specific incidents in which they had detected errors in data than were the actuaries in the pilot study. However, in contrast to our prediction, all of the municipal bond analysts reported incidents of error detection. Our *a priori* expectation was that the municipal bond analysts would be unlikely to report incidents of error detection because of a low base rate of errors and a low payoff for error detection in this domain. This expectation was based on the initial investigation into the domain that focused exclusively on the use of audited financial statements by municipal bond analysts. Interestingly, only one of the municipal bond analysts reported an error detection incident involving an audited financial statement. However, it became clear while conducting the five field interviews with the municipal bond analysts that they also use other types of data (e.g., hospital utilization reports, reports on loan portfolios) that have much higher base rates of errors and higher perceived payoffs for error detection compared to audited financial statements. It is quite likely that the municipal bond analysts were misclassified in the framework shown in Figure 2 on the basis of the initial investigation into the domain and that none of the domains studied are actually characterized by both a low base rate of errors and low payoffs for error detection.

A detailed analysis of the field interviews within each of the professional domains as well as an analysis of base rates and payoffs across the three professional domains is presented in Klein (1995a). This analysis shows that our *a priori* expectations about base rates and payoffs may have been flawed. First, both the perceived payoffs and the perceived base rate of errors are higher for the municipal bond analysts than for the inventory managers or the consumer product managers. Second, the perceived rate of data errors within the consumer product management domain is inconsistent with our belief about the actual base rate of errors. The perceived rate of errors for the consumer product managers is not higher than the perceived rate of errors for the inventory managers. There is evidence in the field interviews that perceptions of the base rate of errors in data do not necessarily reflect the actual error rate. Even so, the interview data do not strongly support our expectation that the rate of errors is higher for consumer product managers than for inventory managers. Finally, there is a considerable amount of within-domain variation with respect to perceived payoffs and the perceived base rate of errors.

5.2 Individual-Level Analysis

The likely misclassification of the municipal bond analysts and the conflicts between our *a prior* expectations about base rates and payoffs and the actual evidence about base rates and payoffs by domain make additional domain-level analyses of the effects of base rates and payoffs on error detection quite difficult. Because of these problems, an individual-level analysis of the effects of these factors on performance was conducted.

A detailed analysis of error detection performance, perceptions about the base rate of errors, and perceptions about payoffs of error detection within each professional domain is presented in Klein (1995a). The summary data from this analysis is presented below in Table 1 for all three domains. Perceptions about the payoffs of error detection are divided into three categories: incentives for error detection, the materiality of data errors, and the ease of verifying and correcting data errors. Each interviewee was assigned a judgment of High, Moderate, or Low for each of these three categories when the interview transcripts were coded. A detailed explanation of the procedure used to assign the judgments of High, Moderate, or Low for the perceived base rate of errors is presented in section 5.2.2.

			Payoffs		
Interviewee	Error Detected	Perceived Base Rate of Errors	Incentives	Materiality	Ease of Verification and Correction
CPM #1	Yes	High	Low	Low	Moderate
CPM #2	Yes	Moderate	Low	Low	High
CPM #3	No	Low	Low	Low	Moderate
CPM #4	No	Moderate	Low	Moderate	High
CPM #5	No	Low	Low	Low	Low
IM #1	No	Low	Moderate	Low	Moderate
IM #2	No	Moderate	Low	Low	High
IM #3	Yes	High	Low	High	High
IM #4	No	Moderate	Moderate	Low	High
IM #5	Yes	Moderate	High	High	High
MBA #1	Yes	High	High	High	High
MBA #2	Yes	High	High	High	Moderate
MBA #3	Yes	High	Moderate	Low	High
MBA #4	Yes	Low	High	High	Moderate
MBA #5	Yes	High	High	High	High

CPM = Consumer Product Manager
IM = Inventory Manager
MBA = Municipal Bond Analyst

Table 1. Performance, Base Rates, and Payoffs

The data presented in Table 1 were used to conduct an individual-level analysis of the relationship between perceived payoffs and self-reports of error detection as well as between perceived base rates and self-reports of error detection across the three professional domains.

5.2.1 Payoffs and Error Detection

Incentives. Table 2 shows the number of interviewees reporting High, Moderate, and Low incentives to detect errors who did and did not report detecting data errors. Notice that none of the interviewees who did not report error detection incidents reported that strong incentives to detect data errors exist in their organizations. In contrast, five of the interviewees who did report error detection incidents also reported strong incentives to detect data errors.

	Incentives		
Performance	High	Moderate	Low
Error Detected	5	1	3
Error Not Detected	0	2	4

Table 2. Relationship Between Incentives and Performance

A chi-square test does not support the hypothesis that there is a relationship between incentives and self-reports of error detection at a level of significance of .05 ($\chi_2^2 = 4.74$; $p < .10$). Because this test is not quite statistically significant, an exploratory analysis was performed to determine whether the effect of a high level of incentives is different than the effect of moderate or low incentives. Table 3 pools the interviewees reporting moderate or low incentives to detect errors.

	Incentives	
Performance	High	Moderate/Low
Error Detected	5	4
Error Not Detected	0	6

Table 3. Relationship Between High Levels of Incentives and Performance

A chi-square test for this contingency table suggests that the effect of a high level of incentives on error detection performance is different than the effect of all lower levels of incentives to detect errors ($\chi_1^2 = 5.0$; $p < .05$).

Materiality. Table 4 shows the number of interviewees reporting High, Moderate, and Low materiality of data errors who did and did not report detecting data errors. Notice that none of the interviewees who did not report error detection incidents reported that errors in data are highly material to business outcomes. In contrast, six of the interviewees who did report error detection incidents also reported that errors in data are highly material to business outcomes.

	Materiality		
Performance	High	Moderate	Low
Error Detected	6	0	3
Error Not Detected	0	1	5

Table 4. Relationship Between Materiality and Performance

A chi-square test supports the hypothesis that there is a relationship between the materiality of data errors and self-reports of error detection ($\chi_2^2 = 7.19$; $p < .05$).

Ease of Verification and Correction. Table 5 shows the number of interviewees reporting High, Moderate, and Low ease of verifying and correcting data errors who did and did not report detecting data errors. Half (three out of six) of the interviewees who did not report error detection incidents reported that the ease of verifying and correcting suspected data errors is high. This proportion is not quite as high as for those interviewees who did report error detection incidents. Two-thirds (six out of nine) of the interviewees who did report error detection incidents also reported that it is easy to verify and correct suspected data errors.

	Ease of Verifying and Correcting Errors		
Performance	High	Moderate	Low
Error Detected	6	3	0
Error Not Detected	3	2	1

Table 5. Relationship Between Verifying and Correcting Errors and Performance

A chi-square test does not support the hypothesis that there is a relationship between the ease of verifying and correcting data errors and error detection ($\chi_2^2 = 1.67$; $p > .10$). Even when interviewees reporting moderate and low ease of correction are pooled, the effect on error detection is not statistically significant ($\chi_1^2 = .42$; $p > .10$).

5.2.2 Base Rate of Errors and Error Detection

Table 6 presents the estimates of the base rate of errors for all fifteen interviewees. All of the base rate estimates collected in this study refer to the percentage of reports believed to contain at least one data error. Some of the interviewees provided base rate estimates for several types of data and some of the interviewees provided a range when asked for a base rate estimate for a particular type of data. For these interviewees, the numbers presented in Table 6 are the average base rate estimate given. The third column of Table 6 shows that all of the interviewees who did not report error detection incidents estimated the rate of serious errors to be less than two percent. In contrast, only three of the nine interviewees who did report an error detection incident gave an estimate this low for serious data errors. The other six estimated the rate of serious data errors to be between 3.4 and 7.5 percent. Four of the six interviewees who did not report error detection incidents estimated the rate of trivial errors to be less than two percent. Only one of the interviewees who reported an error detection incident gave an estimate of trivial data errors this low. Only seven of the interviewees who reported detecting data errors were able to provide an estimate of the base rate of trivial errors. Six of the seven gave an estimate of at least ten percent.

Professional Domain	Detected an Error	Base Rate of Serious Errors	Base Rate of Trivial Errors
Consumer Product Mgmt.	YES	4%	50%
Consumer Product Mgmt.	YES	1%	*
Inventory Management	YES	7.5%	15%
Inventory Management	YES	.5%	*
Municipal Bond Analysis	YES	5.5%	11%
Municipal Bond Analysis	YES	5.3%	19.5%
Municipal Bond Analysis	YES	3.4%	18.8%
Municipal Bond Analysis	YES	0%**	0%**
Municipal Bond Analysis	YES	4.8%	10%
Consumer Product Mgmt.	NO	0%	5%
Consumer Product Mgmt.	NO	.5%	33%
Consumer Product Mgmt.	NO	0%	0%
Inventory Management	NO	0%	1%
Inventory Management	NO	.5%	.5%
Inventory Management	NO	2%	1.5%

* = This interviewee did not provide an estimate.

** = This interviewee gave an estimate of "almost zero" for serious and trivial errors.

Table 6. Error Detection Performance and Base Rate Estimates

Table 7 shows the number of interviewees reporting High, Moderate, and Low rates of serious data errors who did and did not report detecting data errors. For perceived base rates, estimates of the rate of serious errors were used to assign a judgment of High, Moderate, or Low to each interviewee. An average estimate was used for interviewees who provided base rate estimates for several types of data and for interviewees who provided a range when asked for a base rate estimate. The judgment Low was assigned to the four interviewees who gave zero as a base rate estimate. Because we had no *a prior* definition of a moderate and high rate of errors, the median of the remaining eleven estimates (two percent) was used as the threshold between Moderate and High.

	Perceived Base Rates		
Performance	High	Moderate	Low
Error Detected	6	2	1
Error Not Detected	0	3	3

Table 7. Relationship Between Perceived Base Rates and Performance

A chi-square test supports the hypothesis that there is a relationship between the perceived base rate of errors and self-reports of error detection ($\chi_2^2 = 6.27$; $p < .05$).

6.0 Discussion of Results and Conclusion

The findings of the field interview study suggest that the materiality of data errors and expectations about the base rate of errors are related to error detection performance. Although we did not find a straight-forward relationship between different levels of incentives to detect errors and error detection performance, it appears that users who believe that strong incentives to detect data errors are present in their organizations are more likely to detect errors than those without this belief. It is not possible to understand the directionality of these relationships from the data collected in the field interviews. It is entirely possible that the experience of detecting data errors causes users to increase their estimates of organizational incentives to detect errors, the materiality of data errors, and the base rate of errors in data.

Even so, an examination of selected excepts from the interview transcripts gives additional support to the observed relationships between incentives and error detection performance and between base rate perceptions and error detection performance. The extent to which incentives may influence performance in organizational settings can be seen by examining the contrasts between the comments made by the municipal bond analysts (all of whom reported detecting errors) and the consumer product managers (only two of whom reported detecting errors) when they were asked about the extent to which organizational incentives encourage users to detect errors in data. The municipal bond analysts answered this question with comments similar to the two that follow.

Well, the incentive is obviously monetary. To the extent that you're keeping people out of trouble or making them money, it's to your benefit to spot and correct errors...I think there's a high incentive for people to spot errors and try to get the correct information before they buy bonds.

There are very strong incentives to check for accuracy in data. Because a lot is riding on it and I have to rely on that data to be able to identify value in the market. To keep us out of more risky situations in the market...Accuracy is probably the most critical component of my job responsibilities.

In contrast, the consumer product managers gave responses similar to the following two statements.

You aren't encouraged to look for errors.

No. Nobody has enough time to figure out if it's (the data) right or not.

A similar contrast was found between the municipal bond analysts and the consumer product managers for perceptions about the error rate. The consumer product managers tend to assume that the data they use are correct.

I'll tell you, 99.9% of the time it's not wrong.

There aren't a lot of errors...It doesn't happen that often...I may have had one experience (involving a data error) in four years.

I think most people are going under the assumption that it's good data.

In contrast, the municipal bond analysts provide substantially higher estimates of the error rate, and they appear to assume that unaudited data may contain errors.

I always go into it assuming that there might be an error.

You want to make sure that the system is accurate, because we do a lot of work off it, and we need to make sure that it is correct.

These contrasting comments suggest that incentives and expectations about the base rate of errors influence error detection performance not only in controlled laboratory environments, but also in organizational settings.

Limitations of the field study. There are a number of limitations inherent in the methodology used in the field interview study. One limitation is that it is very difficult to know the base rate of errors in a given professional domain with certainty. Indeed, the findings of the field interviews suggest that we did not successfully classify each professional domain along this dimension.

Second, we were not able to gather objective performance measures using the interview methodology. There are two limitations on measuring performance in this study through self-reported data. First, we cannot assume that users can accurately report task outcomes (i.e., whether or not an error actually existed in data). Thus, we cannot classify reported behavior into four categories (1. the user found an error that existed, 2. the user failed to find an error that existed, 3. the user believed that an error existed when an error did not exist, and 4. the user did not believe that an error existed when an error did not exist). Without this classification, it is not possible to characterize users in one domain as more or less successful than users in another domain. The second limitation of self-reported performance is potential biases in accounts of error detection. For example, interviewees may be completely unaware of or forget incidents in which they failed to detect errors in data. Although six of the professionals interviewed in the field study did provide accounts of incidents in which they learned that an error had been missed, it is not possible to know the extent to which the interviewees failed to detect other data errors.

Directions for future research. Additional studies addressing the detection of errors in data by users of information systems could take several paths. The results of the present study suggest that new studies built on the premise that users of information systems can detect errors in data are worthwhile. Four avenues will be suggested here: (1) development of a taxonomy of types of errors that occur in data including an assessment of the types of errors likely to be found by users, (2) studies to develop and test prescriptive interventions to improve error detection, (3) studies identifying individual differences that affect error detection performance, and (4) analysis of the process through which errors in data are detected using process tracing methods.

REFERENCES

Agmon, N., & Ahituv, N. (1987). Assessing data reliability in an information system. **Journal of Management Information Systems, 4**(2), 34-44.

Ballou, D. P., & Pazer, H. L. (1985). Modeling data and process quality in multi-input, multi-output information systems. **Management Science, 31**, 150-162.

Ballou, D. P., & Pazer, H. L. (1987). Cost/quality tradeoffs for control procedures in information systems. **OMEGA: International Journal of Management Science, 15**, 509-521.

Ballou, D. P., & Pazer, H. L. (1995). Designing information systems to optimize the accuracy-timeliness tradeoff. **Information Systems Research.**

Ballou, D. P., Pazer, H. L., Belardo, S., & Klein, B. (1987). Implications of data quality for spreadsheet analysis. **Data Base, 18**(3), 13-19.

Ballou, D. P, & Tayi, G. K. (1989). Methodology for allocating resources for data quality enhancement. **Communications of the ACM, 32**, 320-329.

Bansal, A., Kauffman, R. J., & Weitz, R. R. (1993). Comparing the modeling performance of regression and neural networks as data quality varies: A business value approach. **Journal of Management Information Systems, 10**, 11-32.

Boockholdt, J. L. (1989). Implementing security and integrity in micro-mainframe networks. **MIS Quarterly, 13**, 135-144.

Bowen, P. L. (1992). Managing data quality in accounting information systems: A stochastic clearing system approach. **Dissertation Abstracts International**, (University Microfilms No. 9319179).

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette and L. M. Hough (Eds.), **Handbook of Industrial and Organizational Psychology** (2nd ed., Vol. 1, pp. 687-732). Palo Alto, CA: Consulting Psychologists Press, Inc.

Campbell, J. P., & Pritchard, R. D. (1976). Motivation theory in industrial and organizational psychology. In M. D. Dunnette (Ed.), **Handbook of Industrial and Organizational Psychology** (pp. 63-130). Chicago: Rand McNally College Publishing Company.

Consumer enemy no. 1. (1991, October 28). **Newsweek**, pp. 42, 47.

Corman, L. S. (1988). Data integrity and security of the corporate data base: The dilemma of end user computing. **Data Base, 19**, 1-5.

Cryer, E. H., Bettman, J. R., & Payne, J. W. (1990). The impact of accuracy and effort feedback and goals on adaptive decision behavior. **Journal of Behavioral Decision Making, 3**, 1-16.

Davis, G. B. (1984). Caution: User developed systems can be dangerous to your organization. MISRC Working Paper 82-04, MIS Research Center, University of Minnesota.

Davis, G. B., Adams, D. L., & Schaller, C. A. (1983). **Auditing & EDP**. New York: American Institute of Certified Public Accountants, Inc.

Davis, G. B., Neter, J., & Palmer, R. R. (1967). An experimental study of audit confirmation. **Journal of Accountancy, 123**(6), 36-44.

Davis, G. B., & Olson, M. H. (1985). **Management information systems: Conceptual foundations, structure, and development**. New York: McGraw-Hill Book Company.

Dead farmer syndrome haunts efforts to trim USDA offices. (1992, April 19). **Minneapolis Star Tribune**, p. 5A.

Eisenhardt, K. M. (1989). Building theories from case study research. **Academy of Management Review, 4**, 532-550.

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. **American Scientist, 70**, 607-614.

Flanagan, J. C. (1954). The critical incident technique. **Psychological Bulletin, 51**, 327-358.

Fox, C., Levitin, A., & Redman, T. (1993). The notion of data and its quality dimensions. **Information Processing & Management, 30**, 9-19.

Gaba, A., & Winkler, R. L. (1992). Implications of errors in survey data: A Bayesian model. **Management Science, 38**, 913-925.

Garfinkel, R. S., Kunnathur, A. S., & Liepins, G. E. (1986). Optimal imputation of erroneous data: Categorical data, general edits. **Operations Research, 34**, 744-751.

Ham, J., Losell, D., & Smieliauskas, W. (1985). An empirical study of error characteristics in accounting populations. **The Accounting Review, 60**, 387-406.

Huh, Y. U., Keller, F. R., Redman, T. C., & Watkins, A. R. (1990). Data quality. **Information and Software Technology, 32**, 559-565.

Janson, M. (1988). Data quality: The Achilles heel of end-user computing. **OMEGA: International Journal of Management Science, 16**, 491-502.

Johnson, E. J., & Payne, J. W. (1985). Effort and accuracy in choice. **Management Science, 31**, 395-414.

Johnson, J. R., Leitch, R. A., & Neter, J. (1981). Characteristics of errors in account receivable and inventory audits. **The Accounting Review, 56**, 270-293.

Johnson, P. E., Grazioli, S., & Jamal, K. (1992a). Fraud detection: Sources of error in a low base-rate world. Paper presented at the 10th European Conference on Artificial Intelligence Research Workshop on Expert Judgment, Human Error, and Intelligent Systems, Vienna, Austria, August 1992.

Johnson, P. E., Grazioli, S., Jamal, K., & Zualkernan, I. A. (1992b). Success and failure in expert reasoning. **Organizational Behavior and Human Decision Processes, 53**, 173-203.

King, N. (1994). The qualitative research interview. In C. Cassell and G. Symon (Eds.), **Qualitative Methods in Organizational Research** (p. 14-36). Thousand Oaks, CA: Sage Publications.

Klein, B. D. (1995a). Base rates and payoffs in the detection of errors in data. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

Klein, B. D. (1995b). Base rates and payoffs in the detection of errors in data. In M.K. Ahuja, D. F. Galletta, H. J. Watson (Eds.), **Proceedings of the First Americas Conference on Information Systems** (pp. 381-383). Pittsburgh, Pennsylvania.

Klein, B. D. (1995c). End user detection of errors in data: Preliminary findings and an experimental study. **Proceedings of the Northeast Decision Sciences Institute 1995 Annual Meeting**, March 1994.

Klein, B. D. (1996). Error detection and correction in actuarial data. In M. Khosrowpour (Ed.), **Proceedings of the 1996 Information Resources Management Association International Conference** (pp. 73-79). Washington D.C., Idea Group Publishing.

Knight, B. (1992). The data pollution problem. **Computerworld, 26**(39), 81-83.

Laudon, K. C. (1986). Data quality and due process in large interorganizational record systems. **Communications of the ACM, 29**, 4-11.

Madnick, S. E., & Wang, R. Y. (1992). Introduction to the TDQM research program. Total Data Quality Management Research Program Working Paper #92-01.

Mason, R. O. (1986). Four ethical issues of the information age. **MIS Quarterly, 10**, 5-12.

Maxwell, B. S. (1989). Beyond data validity: Improving the quality of HRIS data. **Personnel, 66**(4), 48-58.

Miles, M. B., & Huberman, A. M. (1994). **Qualitative data analysis**. Thousand Oaks, CA: Sage Publications.

Misvanks. (1988). Integrity analysis. **Information and Software Technology, 30**, 595-605.

Morey, R. C. (1982). Estimating and improving the quality of information in a MIS. **Communications of the ACM, 25**, 337-342.

Naus, J. I. (1975). **Data quality control and editing**. New York: Marcel Dekker, Inc.

Nesbit, I. S. (1985). On thin ice: Micros and data integrity. **Datamation, 31**(21), 80-84.

O'Leary, D. E. (1993). The impact of data accuracy on system learning. **Journal of Management Information Systems, 9**, 83-98.

O'Neill, E. T., & Vizine-Goetz, D. (1988). Quality control in online databases. In M. E. Williams (Ed.), **Annual Review of Information Science and Technology**, (pp. 125-156). Amsterdam: Elsevier Science Publishers.

Paradice, D. B., & Fuerst, W. L. (1991). An MIS data quality methodology based on optimal error detection. **Journal of Information Systems, 5**(1), 48-66.

Parsaye, K., & Chignell, M. (1993). Data quality control with SMART databases. **AI Expert, 8**(5), 23-27.

Payne, J. W. (1982). Contingent decision behavior. **Psychological Bulletin, 92**, 382-402.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, **14**, 534-552.

Percy, T. (1986). My data, right or wrong. **Datamation, 32**(11), 123-128.

Redman, T. C. (1992). **Data quality: Management and technology**. New York: Bantam Books.

Redman, T. C. (1995). Improve data quality for competitive advantage. **Sloan Management Review, 36**(2), 99-107.

Ricketts, J. A. (1990). Powers-of-ten information biases. **MIS Quarterly, 14**, 63-77.

Stone, M., & Bublitz, B. (1984). An analysis of the reliability of the FASB data bank of changing price and pension information. **The Accounting Review**, **59**, 469-473.

Wand, Y., & Wang, R. Y. (1994). Anchoring data quality dimensions in ontological foundations. Total Data Quality Management Research Program Working Paper #94-03.

Weber, E. U., Bockenholt, U., Hilton, D. J., & Wallace, B. (1993). Determinants of diagnostic hypothesis generation: Effects of information, base rates, and experience. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, **19**, 1151-1164.

Yin, R. (1989). **Case study research**. Beverly Hills, CA: Sage Publications.

Zmud, R. W. (1978). An empirical investigation of the dimensionality of the concept of information. **Decision Sciences, 9**, 187-195.