IMPACT OF APPRAISAL EXPECTATION AND INFORMATION ABOUT FACTORS AFFECTING PERFORMANCE ON SUPERVISORY INFORMATION SEARCH AND EVALUATION ACCURACY

Mary Alison Waterhouse

A Thesis

in

The Faculty

of

Commerce and Administration

Presented in Partial Fulfilment of the Requirements

for the Degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

June 1996

ISBN 0-612-18456-0

Abstract

Impact of Appraisal Expectation and Information about Factors Affecting Performance on Supervisory Information Search and Evaluation Accuracy

Mary A. Waterhouse

Computerized Performance Monitoring Systems (CPM) are the software that capture performance data from employees. There are two distinct uses of information provided by computerized performance monitors. The first use is examining the employees' work as it happens, so that any drop in performance can be identified and corrected. The seccnd use is in employee performance evaluations. What happens when CPM systems are used by supervisors for both purposes? How will the supervisors' knowledge that they must rate performance affect their ability to accurately spot temporary performance declines? In addition, temporary declines may be caused by factors internal to the employee such as dissatisfaction, or factors external to the employee, such as hardware problems. How does supervisors' knowledge of these factors affect their monitoring performance?

Using a computerized experiment, where 127 subjects looked at performance data of three simulated employees, we studied the effect on the decline spotting accuracy and search patterns of the supervisors who knew they had to do performance appraisals, and either thought that the performance declines were caused by hardware and software problems or employee dissatisfaction.

The results show that the subjects spotted the performance declines more accurately if they did not know they had to give performance appraisals, and they thought that hardware was the cause of the performance decline.

Acknowledgements

I would like to thank Dr. Deborah Ballou for her supervision and constructive critisism during the preparation of this thesis. Many thanks to my family and friends for their support and encouragement while I was writing this thesis.

Table Of Contents

List Of Tables

List Of Figures

Chapter 1: Introduction

Computerized Performance Monitoring Systems are the software that capture performance data from employees. Their use has become common in industries where employees spend much or all of their time using computers. In 1987 the U.S Government estimated that computerized performance monitoring affects from 25 to 35% of all U.S. clerical workers (U.S. Congress, OTA 1987). By 1991 the percentage of workers electronically monitored was 40% (Bylinsky, 1991; Halpern, 1992). In 1985, 25,000 companies in the U.S. used computer technology to monitor their employees. In 1991, the number was over 70,000. The total amount spent on software used to monitor employees is expected to exceed one billion (U.S.) by 1996 (Aiello, 1993). There is a large amount of literature which looks at many aspects of computer monitoring. Some research considers the relationship between management and employees, and the effects of monitoring on the relationship (Garson 1988, Nussbaum 1984, Shaiken 1984). Other authors have looked at the issues of privacy, and the ethical use of monitoring systems (U.S. Congress, OTA 1987; Danaan 1990, Marx & Sherizen 1987, Westin et al. 1985). Other research examines changes in motivation, satisfaction and the quality of life in the workplace caused by computerized performance monitoring (Olson & Lucas 1982, Irving et al. 1986, Smith et al. 1986, Walton & Vittori 1983). A fourth group looks at the impact of monitoring on the work itself and its control (Attewell 1987, Eisenman 1986. Grant et al. 1988, Griffith, 1993, Kraut et al. 1989, Tamuz 1987, Grant & Higgins 1991). Another group examines the supervisors who must incorporate this new source of information into assessments of employee performance (Fenner et al. 1993, Kulik & Ambrose, 1993). There are, however, several important issues that have yet to be addressed by this literature.

There are two distinct uses of information provided by computerized performance monitors. The first use of computerized performance monitors is examining the employees' work as it happens, so that any drop in performance can be identified and corrected. Grant and Higgins (1991) refer to this control system as a thermostat (Lawler & Rhode, 1976). As in a thermostat, a 'sensor', which is a control device or evaluation system, measures the level of performance. A 'discriminator' (in this case a supervisor) compares the measurements to a standard set by the organization. The 'discriminator' reports any lack of performance to an 'effector' (employee). The effector then increases performance. Most computerized performance monitors are designed to report performance data in detail, covering relatively short time spans, so that corrective action can begin immediately if a decline in performance is noted (OTA, 1987). The second use is in employee performance evaluations. What happens when computerized performance monitors are used by supervisors for both purposes? How will the supervisors' knowledge that they must rate performance affect their ability to accurately spot temporary performance declines?

As most or all of the work that is being monitored is computerized, the decline in performance could be caused by problems with hardware or software. When supervisors are examining data for performance declines, will they be able to make allowances for the system being down? If the computerized performance systems collect data over several months so that they can be used in performance appraisals, do they contain information about the "down-time" of the system? Without such information, will the supervisors make a more severe judgment on the performance of the employees? The research literature has not yet addressed this problem (Ballou, 1996).

This research studies the effect on the decline spotting accuracy and search patterns of supervisors who either knew or did not know that they had to do performance appraisals, and either thought that the performance declines were caused by hardware problems or employee dissatisfaction.

The study consisted of a computerized experiment, where 127 subjects (undergraduate and graduate students) looked at performance data of three simulated employees. The experiment was conducted at Carnegie Mellon University by Drs. Ballou, Lerch and Kulik. The experiment contained four weeks of performance data for these employees, who were said to be entering time card information into a computer. The data was given both in tabular and graph form for each employee. It was also given in daily and hourly formats. Before the experiment began, the supervisors were given different information. They were told that the performance declines of the employees were either caused by hardware problems, or they were told that the performance declines were caused by dissatisfaction with the task. The supervisors were either told that they would be required to appraise the employees after they had viewed the data, or they were not told about being required to perform a subsequent performance appraisal.

The following chapters contain a review of the literature on computer performance monitoring, performance appraisals and information search; the methodology and the results of the experiment as well as a discussion of the results.

Chapter 2: Computerized Performance Monitoring Literature

There is little empirical research on computerized performance monitoring (CPM), although it is a controversial subject in information technology (Ambrose & Kulik, 1994). Monitoring opponents believe it violates worker's privacy, makes jobs more boring, increases employee stress, produces less job satisfaction and morale, and makes modern offices comparable to sweat-shops (Austin & Drake, 1985; Smith, Carayon, Saunders, Lim & LeGrande, 1992; Ambrose & Kulik, 1994; Aiello, 1993). They also claim that computer monitoring reduces employees' human contact with supervisors and co-workers (Aiello, 1993). Monitoring proponents believe it provides a reliable, unbiased, more accurate (less subjective) evaluation of employee performance that allows for more timely feedback. This can both motivate the employee, and lead to increases in levels of performance (Angel, 1989; Henriques, 1986; Aiello, 1993). Aiello states that in his experience with organizations that have installed computerized performance monitors, the reality of the situation is somewhere between the two extremes (Aiello, 1993). He began his research in the area of computerized performance monitoring in 1985, with a field study which investigated employee stress in two large insurance companies. During his work he interviewed employees who worked on video display terminals (VDT), some of whom were monitored by computer. These employees described various symptoms of stress. Many of these employees stated that an important factor in the way they feel was the lack of human contact with workers and supervisors. This lowered their levels of perceived social support, and caused higher levels of stress.

There are few empirical studies which examine how supervisors use the information gathered from the computerized performance monitoring systems (Ambrose & Kulik, 1994). Grant et al. (1988) reported that while monitored workers perceive that quality of work is less important that quantity of work in their supervisors' appraisals, the supervisors reported that both quantity and quality were of importance. Chalykoff and Kochan (1989) reported that there has been a dramatic increase in the work-load of supervisors caused by the introduction of computerized performance monitors into their organization. Many supervisors resent this time-consuming part of their work. These findings occurred during a study in which Chalykoff and Kochan interviewed supervisors and employees of the Internal Revenue Service. Supervisors reported spending approximately 25 hours per week in monitoring and related activities. Aiello's (1993) experience was also consistent with the results of Chalykoff and Kochan. He observed that in many organizations which had introduced monitoring systems, an increased pressure on supervisors to give more feedback to the employees about their performance occurred. Employees expect more feedback with the introduction of computerized performance monitoring because of the increased amount of data available to the supervisors (Aiello, 1993).

Kulik and Ambrose (1993) used theories of person perception that differentiated between category-based and feature-based information processing strategies as described by Fiske and Pavelchak (1986) to determine how raters integrate performance information from both visual and computerized sources. They define feature-based processing as the method by which perceivers are highly motivated to process all of the available information carefully and accurately. Category-based information processing is the other extreme where perceivers use less effortful methods of processing. In this processing method, perceivers may rely heavily on heuristics (Maheswaran & Chaiken, 1991), or on social categories (Fiske & Pavelchak, 1986). In Kulik and Ambrose's research, subjects received both positive and negative computerized data and viewed positive or negative instances of a secretary's performance. They found that visual data was most influential in determining the processing strategies of raters, which suggests that raters used category-based strategies when viewing positive performance and used feature-based strategies when viewing negative performance, despite the evaluative implications of the computerized data. These results contradict the documented concerns of employees that are monitored by computerized performance monitoring systems (Grant et al., 1988), who feel that the objective character of the computerized data may lead supervisors to put more weight on the quantitative data when appraising performance. However, Kulik and Ambrose found that raters effectively evaluated performance data, when computerized data was applicable.

Fenner, Lerch, and Kulik (1993) studied the impact of computerized performance monitoring and prior performance knowledge on performance appraisals. They used a laboratory experiment to assess the impact of prior performance level on requests for computerized performance data, and how this requested information and the prior performance level affected the performance appraisal. The results indicated that prior performance levels and employee performance during the monitoring period independently affected both current and future ratings. Also, when the performance of the simulated employees differed from the prior performance, the subjects: 1) requested more information about the employees' performance; 2) were less certain about their ratings of the employee's present and future performance; and 3) rated the employee's present performance as more changeable than when the employee performed at a level similar to that of their prior performance. These results suggest that there are several ways that the computerized performance monitoring systems could be designed to combine features to help supervisors with the performance appraisal process. The authors suggest that it would be helpful to design computerized performance monitors that directly indicate changes in performance over time, but it is important that the systems do not highlight small or temporary variations that do not indicate a performance trend (Fenner et al., 1993).

Ambrose and Kulik (1994) examined how supervisors actually process the data produced by computerized monitoring systems, and use the data in performance appraisals. Their study examines performance pattern and information format. In a computer simulation, the subjects evaluated the performance of a typist monitored by computer. Each subject was assigned one of three format conditions, a periodic, delayed or summarized format. The performance pattern of the typist also varied between improved, stable, or worsened. The results indicate that the pattern affected the subjects' ratings of overall performance, quality of performance and consistency of performance. The factors also affected ratings of future performance and recall of precise performance information. These results have significant implications for the use of performance monitoring systems. While advocates of computerized performance monitors have maintained that these systems will increase the objectivity and accuracy of the performance evaluation, Ambrose and Kulik's (1994) results clearly show that the presentation of performance data on its own will not ensure that performance evaluations are accurate. The monitoring systems, however, must be made to fit the specific purpose of the appraisal. According to Aiello (1993) :

> "Exactly what gets measured, for what periods of time, and toward what type of comparison with established standard(s) in the computer-based work monitored environment will probably be the crucial components of any monitoring system." p 504

Ambrose and Kulik's (1994) research shows how the information is presented to the supervisor is also an important consideration.

Another related issue that has not been explored by researchers is whether supervisors are able to discount declines in performance that are caused by hardware or software problems. As computerized performance monitoring systems are normally used when employees conduct part or all of their work on computers, their work performance is affected by the performance of the hardware and software the employee is working on. Will supervisors be able to factor in the effects of hardware or software problems on the employee's perfo .nance (Ballou, 1996)?

To summarize, Fenner, Lerch and Kulik (1993) looked at the impact of computerized performance monitoring and prior performance knowledge in performance appraisals. Kulik and Ambrose (1993) looked at integrating visual performance information and computerized performance information. Ambrose and Kulik (1994) looked at how supervisors actually process data produced by computerized performance monitors and how they use that information in performance appraisals.

As the literature on the effect of computerized performance monitoring on performance appraisals is scarce, studying the performance appraisal literature may give some insight into how the knowledge of performance appraisal affects computerized performance monitoring and if the supervisors will be able to discount declines in performance that are caused by hardware or software problems.

Chapter 3: Performance Appraisal Literature

Performance appraisal refers to the process by which the job performance of an employee is rated by an observer, usually a supervisor or peer. These appraisals are often conducted annually or semi-annually. The product of the appraisal has important consequences in enhancing organizational effectiveness (Landy & Farr, 1980; DeNisi et al., 1984). Practical utility, however, is restrained by their demonstrable susceptibility to bias, which is caused by a number of personal, contextual, and psychometric factors (Borman, 1977; Cooper, 1981; Landy & Farr, 1980; DeNisi et al., 1984). This results in the reduction of accuracy of subjective appraisals. Accuracy is defined as the degree to which the rank ordering of the ratings for a group of employees approximates the rank ordering of their objective performance. This implies an existence of a "true performance" score (DeNisi et al., 1984). There is a growing amount of literature that shows that both computerized monitored workers and their supervisors believe that computerized performance monitoring may result in more accurate performance evaluations by providing more accurate data about performance (Fenner et al., 1993; Eisenman, 1986; Grant & Higgins, 1989; Irving et al., 1986). Computerized performance monitors may considerably increase the accuracy of performance evaluations by aiding the supervisors to avoid systematic biases which occur in the performance appraisal process (Fenner et al., 1993).

There are two different types of research concerning the issue of employee evaluation, one body of research is concerned with the instrument, the other is concerned with the user. The first issue concentrates on the development of formal instruments for evaluation, and the susceptibility of these instruments to both random error and various systematic biases (e.g., halo effect, leniency/stringency, central tendency). The other research area concentrates on social psychology and looks at attribution and stereotyping and how they influence evaluation (Feldman, 1981). Little attention, however, has been paid to the way raters search for and obtain information for performance appraisals (DeNisi et al., 1984). The way humans make decisions cannot be understood simply by studying final decisions. The perceptual, intellectual, and emotional process which ultimately leads to a choice of a decision alternative must also be studied if we want to gain adequate understanding of human decision making (Svenson, 1979). Performance appraisals are a type of decision in which the rater has to evaluate the performance of an employee. Crocker (1981) proposed an active information searching role for raters. She suggested that raters deliberately determine the types of information relevant to their decisions and then methodically sample from the area of relevant behaviour. This idea of sampling seems to be apt for performance appraisals as most supervisors deal with many subordinates, and are only able to use a sample of work-related behavior for any one employee (DeNisi et al., 1984).

Causal attribution also appears to be relevant to the problem of obtaining information in performance appraisal. We often seek to determine the cause of an action or outcome in terms of the part played by the actor, the situation, or some combination of the two. In the context of performance appraisals, a rater seeks to determine whether an employee's poor performance was due to the worker, the task, or some interaction of the two. To find a cause the rater may use three types of information: 1) Consensus information which is information about how others performed the same task; 2) Distinctiveness information which is how the employee performed on all job related tasks; 3) Consistency information which is how the employee performed on this task (Kelley, 1967). If the rater feels that other workers have also performed unsatisfactorily on the task (high consensus), this employee performs well on other tasks (high distinctiveness), and this employee has had difficulty with the task in the past (high consistency), the rater is likely to credit the poor performance to some attribute of the task (an external attribution), rather than to some special lack in the employee (an internal attribution). Conversely, an employee who always performs a particular task well that others usually perform poorly, and also performs well on all other job-related tasks, would lead a rater to credit the good performance of the employee to positive attributes of the worker rather than the task (an internal attribution) (Kelley, 1967; DeNisi et al., 1984).

Lalljee, Lamb, Furnham and Jaspars, (1984), however, reported results of three studies that claim that people usually have hypotheses about why an event occurred, and that rather than seek consensus, consistency and distinctiveness information as described by Kelley (1967, 1972, 1973), they instead seek out information that enables them to 'disambiguate' their hypotheses. Lalljee et al. cite the works of Garland et al., (1975), Hanson (1980) and Major (1980) who all indicate that the pattern of information search for different events is different and is related to the type of attribution generally made about the event. Their results support the hypothesis-testing model (Lalljee, et al., 1984).

DeNisi et al. (1984) cite the work of Major (1980) who found that subjects were willing to make attribution decisions based on less than 25% of the available information. This suggests that they tended to forego consensus information, and the small amount of information they sought was generally distinctiveness or consistency information. The research has important implications for performance appraisals because they suggest that raters might depend upon internal standards of what constitutes "good" performance, rather than comparing the performance of several workers and making relative judgments. How a rater searches for information will determine what behavior is observed by the rater, but the factors that determine what information a rater searches for need to be examined.

There are several authors who have proposed different models that consider the way the rater searches for information. DeNisi and Williams (1988) explain four of these models, three of which will be discussed below:

3.1 Feldman's Model

Feldman (1981) developed a model which is "an applied person perception model" (DeNisi & Williams, 1988 p.112). This model recognizes that there are many demands on a supervisor's time in addition to the appraisal-related tasks. Feldman (1981) said "In this uncertain, informationally 'noisy' environment, the supervisor must perform several cognitive tasks before performance appraisals are possible" (p.128). These tasks are: 1) The rater must recognize and attend to relevant information about employees. 2) The information must be organized so that it can be stored and retrieved when necessary. New information must be linked with previously gathered data. 3) When judgments are required - recall this information in a timely and organized fashion. 4) Information must be summarized and integrated into the appraisal form. The major feature of Feldman's model is the notion of automatic and controlled processing. (DeNisi & Williams, 1988). When automatic processing occurs, a rater encounters a stimulus with which he or she is familiar, or has had considerable prior experience. The rater continues the rating process without being consciously aware of it, and it does not interfere with ongoing mental activities (Posner & Snyder, 1975). The rater would see how a ratee was performing a particular task, the rater would automatically process this behavior as being relevant or irrelevant to the appraisal, and would automatically rate the performance as being good or poor. The automatic processing would also take place if the rater come across a piece of information about the particular employee which was consistent with the rater's preconceived idea about the employee. On the other hand, controlled processing is always under the conscious, intentional control of the rater. Controlled processing takes place when the rater encounters information about the ratee which is inconsistent with the view the rater holds of the ratee, or the rater does not feel that the information is relevant to the appraisal decision. Controlled processing requires active concentration on the part of the rater. DeNisi and Williams (1988) cited Feldman who proposed that reliance upon automatic processing results in categorization of performance incidents and ratees. The incidents and ratees are assigned to a category such as: good or poor, lazy, ambitious, male or female. Any future processing of information about these ratees is based upon the attributes of the category. Any other information about the ratee is perceived to be the same as the characteristics of the category to which the ratee has been assigned. The assigned category will also screen out any pieces of information as irrelevant if they do not fit the stereotype of a person in that particular category. It is only when controlled processing occurs, in the case of the ratee exhibiting some kind of behaviour that does not fit into the category, that the rater will possibly reassign the ratee to a new category. The category assigned will determine much of what happens in the appraisal process.

3.2 Ilgen and Feldman's Model

Ilgen and Feldman (1983) also discuss categorization. They cite works that say that categorization is necessary to cognitive economy, as it reduces the amount of information that must be stored and processed (Smith, Adams & Schorr, 1978; Behling, Gifford & Tolliver, 1980. Ilgen and Feldman (1983) define categorization and perception to be the same. They expand upon automatic processes and controlled process to include causal and trait or disposition attribution. They also highlight several biases that Ross (1977) has discussed that can have an effect on performance appraisals.

The first bias is "fundamental error", which is a tendency to underestimate the importance of situational factors, and overestimate the dispositional factors as causes of the observed behavior. In other words people are seen as causes too often, and situations too infrequently. Ilgen and Feldman (1983) cite several authors, who have noted that situational prototypes and scripts of behaviour settings have a tendency to describe the behaviour and dispositions of the people in the situation (Bem & Funder, 1978; Cantor, 1979; Shank & Abelson, 1977). In the supervisory monitoring case, prototyping and scripts may influence the supervisors' perception of the behaviour and disposition of the employees.

The second bias is that actors and observers differ in causal attributions: Situational factors are emphasized by actors, and observers emphasize the actor's personal dispositions. In the computerized performance monitoring situation, the raters will emphasize the ratees personal dispositions, and the ratees will emphasize the hardware or software declines as the cause of any performance decline.

The third bias is the tendency to see as causal the most salient features of the environment.

The fourth bias is the tendency to see actions having positive or negative consequences for the observer as being more dispositional than other actions.

The fifth bias is the tendency for people to pay insufficient attention to base rate information (the degree to which a behaviour is common), and the tendency of people to use their own behaviour as a basis for judging the behaviour of others. This results in a tendency to end the search for causes when the first prototype providing a satisfactory explanation for observed behaviour is found (Ilgen & Feldman, 1983).

3.3 DeNisi, Cafferty and Meglino

Four potential determinants were suggested in a model by DeNisi et al. (1984). They are as follows: 1) Preconceived notions the rater has concerning the ratee; 2) The purpose of the appraisal; 3) The nature of the rating instrument; and 4) The time pressures of the rater.

3.3.1 Preconceived Notions

Preconceived notions the rater has concerning the ratee, are one element of the types of information sought. DeNisi et al. cited two reviews (Tajfel, 1969; Slovic, Fischoff & Lichenstein, 1977) which conclude that these preconceived notions could be manipulated by providing observers with certain "background information" prior to the raters observing the ratee behavior. The notion of priming (Wyer & Srull, 1981), which is the prior release of information in a social perception task that would bias the subject to select one of a number of information sources with which to interpret subsequent information, may also be related to the preconceived notion idea. Preconceived notions can also be obtained from earlier interaction with the ratee in other contexts, prior evaluations, or information about the ratee from others.

Preconceived notions are a part of a schemata[1] entrenched by the rater to provide a framework of reference for interpreting incoming stimuli (DeNisi et al., 1984). In the context of computerized performance monitoring, will the daily or weekly rating of the ratee by the rater result in preconceived notions that will affect the performance appraisal?

The activation of a schema can affect how the rater searches the information. A rater who has categorized an employee in terms of a "good employee" schema, based upon some previous salient observations may feel that he or she already knows that the employee is performing a good job, and does not feel that it is necessary to collect further information about that employee. Also, any other information sought may be distorted to conform with the rater's preconceived notions (DeNisi et al., 1984). There have been several studies that have found that people tend to search for information that confirms their initial expectations, rather than searching for information that contradicts that information (Snyder, 1981; Synder & Cantor, 1979; Ebbesen, 1981; Swann, Stephenson & Pittman, 1981, Wong & Weiner, 1981). On the other hand, there have been studies that found that raters tended to seek diagnostic rather than corroborative information. They tended to collect information that would help the rater to classify target persons into "good" or "poor" categories, for example, even when this

---

1 A "cognitive structure that consists in part of the representation of some defined stimulus domain. The schema contains general knowledge about that domain, including a specification of the relationships among its attributes, as well as specific examples of instances of the stimulus domain." (Taylor & Crocker, 1981, p 91)

information would contradict their initial expectations. This is consistent with authors who have discovered that raters tend to rely more upon negative information when making decisions. (Trope & Basok, 1982; Hamilton & Huffman, 1971). Whichever way the preconceived notion effects the search, the evidence suggests that a raters preconceived notion does in fact affect the type of information the rater seeks. The effect would be caused by the activation of some schema followed by the rater searching for some information that would prove or disprove the validity of the schema. The way the search is carried out would then have an effect on the behavior that the rater actually observes (DeNisi et al., 1984).

3.3.2 Purpose of the Appraisal

To explain why the purpose of the appraisal has been included in the model, DeNisi et al. (1984) cited the results of several studies that suggest that the purpose of the search for information affects what types of information are sought and how that information is later stored (Chaiken, 1980; Cohen, 1981; Crockett, Mahood & Press, 1975; Jeffrey & Mischel, 1979; Press et al., 1975; Ross, Lepper, Strack, & Steinmetz, 1977; Wyer, Srull, Gordon, & Hartwick, 1982). Williams, Blencoe, DeNisi and Cafferty (1983) found that when raters were making comparative decisions, raters sought information that would allow them to judge the performance of the ratee relative to the performance of others. The purpose of the appraisal would appear to play a greater role in the appraisal process than has been previously suggested (DeNisi et al., 1984).

3.3.3 Time Pressures

The time pressure on the rater is hypothesized to affect what behaviours are observed. DeNisi et al. (1984) cite Staelin and Payne's study (1976) which found that when facing time pressures and distraction, shoppers tried to reduce search time by collecting fewer pieces of information, and often searching for negative information. As the supervisors using computerized performance monitors are under severe time pressures as shown by Aiello (1993) and Chalykoff and Kochan (1989), will they compensate for this by searching for negative information?

3.3.4 Nature of the Instrument

According to DeNisi et al. (1984), the rating scale performs a directive function by guiding the rater to look for certain measurements or aspects of ratee performance and not for others. The measurements required for a rating determine those measurements for which information is sought; the rating of traits requires information about ratee traits, the rating of behaviours requires information about ratee behaviours.

Much of the work on DeNisi et al.'s (1984) model focuses on purposeful acquisition of information by the rater (DeNisi & Williams, 1988). This research stresses rater activities in situations where the rater intentionally searches for ratee performance information in order to make a judgment. Studies that looked at these search processes found evidence of systematic search strategies, but also demonstrated that these strategies were sensitive to constraints on the amount of information a rater could acquire (representing time constraints) and the purpose for which the appraisal was conducted (Cafferty, DeNisi & Williams, 1986; DeNisi, Cafferty, Williams, Blencoe, & Meglino, 1983; Williams, DeNisi, Blencoe & Cafferty, 1985; DeNisi & Williams, 1988).

An issue relating to the role of memory in appraisal decisions in concerned with retrieval from memory (DeNisi & Williams, 1988). This could have implications in the use of computerized performance monitors. Recall can be affected by general impressions about ratees. Murphy, Gannett, Herr and Chen (1986) suggested that the effect on ratings they found for subsequent performance information was caused by biased recall. Accuracy in ratings is related to the accuracy of the information recalled from memory. Williams and his colleagues (Williams, 1984; Williams, DeNisi, Meglino & Cafferty, 1986) have shown that the study of retrieval is further complicated by the fact that information acquired and encoded for one purpose must often be retrieved and used for a different purpose. This has been referred to as the problem of reprocessing objectives (Anderson & Pichart, 1978), and suggests that information originally acquired for one purpose must be reprocessed before it is used for a different purpose. The studies mentioned above found that it was the original purpose that established how the information was organized. This would mean that reprocessing would be difficult and could cause inaccuracy in ratings, if one assumes that any information acquired by means other than by a deliberate search (and even some acquired through a deliberate search) must be reprocessed before it can be used for performance appraisals (Williams, 1984; Williams et al., 1986).

In this chapter we examined different models that consider the way raters search for information and the determining factors. To have insight into how the raters searched the computer performance monitoring data, and why they searched the data the way they did, we need to know about the factors that could influence the way they searched. The first model we looked as was Feldman's (1981) model. The main point of Feldman's model was the notion of automatic and controlled processing. When the data obtained from the computerized performance monitoring experiment is examined, we need to consider if the way the subjects search the data can be explained by automatic and controlled processing. When the subjects were primed with the information that either hardware or dissatisfaction was the cause of the performance decline, would this cause the subjects to have preconceived notions about the employees? Would whichever preconceived notion the subjects had cause the subjects to search the data differently? The second model was Ilgen and Feldman's (1983) model. In this model Ilgen and Feldman discuss how Ross's (1977) different biases affect performance appraisals. The first bias of "fundamental error", which is a tendency to underestimate the importance of situational factors, and overestimate the dispositional factors as causes of the observed behavior could be a factor in the way the raters search the data. As hardware is a situational factor and dissatisfaction is a dispositional factor, the subjects searching the data differently depending upon what they thought was the cause of decline in performance could be attributable to fundamental error. The third was DeNisi et al.'s model which discusses the following factors which may cause the raters to search the data differently:

1) Preconceived notions the rater has concerning the ratee. These preconceived notions have been primed in our experiment. We want to discover if these preconceived notions have an effect on the way the raters search the information to try and ascertain if the supervisors can distinguish between hardware or software factors causing performance decline and decline in performance caused by employee dissatisfaction; and

2) The purpose of the appraisal. In our experiment, the purpose of the raters searching the data is different. They either have one objective in searching the data, which is that of monitoring the data to find declines in performance, or they have a dual objective. Those in the dual objective sections think they are looking at the data both to find declines in performance and to do performance appraisals on the employees. DeNisi et al. said that the purpose of the appraisal is a factor in how the raters search for information. We want to discover if this will be the case in this experiment.

Another stream of literature we looked at that discusses different ways people search for information is that of Lalljee et al. (1984), Major (1980), Hanson (1980) and Garland et al.(1975) who all indicate that the pattern of information search for different events is different and is related to the type of attribution generally made about the event. This is referred to as the hypotheses-testing model. These authors imply that there should be a close relationship between the information sought about an event and the explanations provided to interpret that event.

To be able to discover if the knowledge of performance appraisal affects the way supervisors search the data, and if supervisors are able to discount hardware problems, we want to see if the way the raters search the data is different depending upon what they believe to be the purpose of the search and the preconceived notions they have concerning the employees declines in performance. If there is a difference in the way the search is carried out, would that have an effect on the behavior that the rater actually observes?

The following chapter has further information about the way people search data and how the use of graphs and tables helps them in their search.

Chapter 4: Information Search Literature

4.1 Factors affecting rater's acquisition of ratee performance information

Kozlowski and Ford (1991) performed studies examining the dynamic effects of the rater information acquisition process. One of these studies examined the effects of prior ratee knowledge (familiarity) and ratee performance level on subsequent acquisition of ratee performance information under varying levels of control. This study was carried out using a computer controlled information board procedure which gave the subjects additional information concerning the ratees. Before the experiment started the subjects were given prior ratee information. The results indicated that the amount of prior ratee information had an impact on the search as raters searched for more information on the ratees for whom there was less prior information. The raters searched for more information for the ratees who had a 'poor' performance rating, as opposed to those who had a 'good' performance rating.

4.2 Comparison of Graphs and Tabular Data

There are many circumstances that affect a supervisor's decisions; one of the most important is the way the data is presented to the supervisor. There are several conflicting studies on the impact of graphical and tabular displays on managerial decision making (Remus, 1987). Powers, Lashley, Sanchez and Shneiderman (1984) hypothesized that more "usable" information can be conveyed using a combination of graphical and tabular data than by using either form by itself. They felt that when a user is given an option as to which type to use, the user would get a better overall "view" of the information and would understand it better. The combination of graphic and tabular data also allows users to extract information from the presentation form they are most comfortable with, and then further enhance their understanding with the other presentation mode. This hypothesis was only partially supported in the study. They found that if the goal was to speed up performance then the combination of tabular and graphic data should be avoided (their results showed that tabular data would be most appropriate), but if the goal was to increase the accuracy of performance, then the combination of graphics and tabular data appears to be most helpful to that goal. One of the recommendations given was that a menu selection process could be implemented so that the same data could be presented in both forms. Users could then select the forms they most readily understood, and with which they feel most comfortable. The authors cautioned that both tables and graphs can distort results (Powers et al., 1984). Jarvenpaa and Dickson (1988) stated that graphics have been recommended as more effective than tables for portraying time series data, and for allowing comparisons to be made more easily.

In this chapter Kozlowski and Ford's (1991) discussion about factors affecting raters' acquisition of ratee performance information was outlined. They found that raters search for more information for ratees with 'poor' performance rating. In our experiment would the subjects search for more information for the employees that they thought were the poorer (dissatisfied) employees?

As the data was presented to the subjects in both graphical and tabular form, a brief discussion about how subjects search graphs and tables is relevant as we are looking at how subjects search data.

The following chapter discusses the current research and gives hypotheses about what we were expecting to find considering the literature studied.

Chapter 5: Current Research

The first thing that was looked at is the accuracy with which the declines in performance were spotted. There were two types of decline periods, ones that spanned a day or more, and ones that were just for a few hours in the day. The decline periods were easier to spot if they were for a few days, so we referred to them as easy accuracy decline periods. The decline periods that were just for a few hours were referred to as difficult accuracy. Total accuracy is the sum of the other two types of accuracy.

After studying the literature the following hypotheses were proposed:

5.1 Accuracy Hypotheses

H1a: If subjects know that they have to give performance appraisals, they will be <u>more</u> accurate in spotting <u>easy decline periods</u>.

In DeNisi et al.'s (1984) model that considers the way the rater searches for information, the purpose of the appraisal has been included as one of the determinants. In this case will the knowledge that the subject has to give a performance appraisal after searching the data affect the way the subject searches the data? If the data has been searched differently, will that effect the accuracy with which the rater spots the decline periods? Lalljee et al. also said that the pattern of information search for different events is different. If the subjects think they will be required to do a performance appraisal, they will search the data differently than they would if they had no appraisal knowledge. The knowledge of performance appraisal should cause the rater to be more accurate in spotting performance declines.

H1b: *If subjects think that hardware problems are responsible for the decline in performance they will be <u>less</u> accurate in spotting <u>easy decline periods</u>.*

In Ilgen and Feldman's (1983) model they highlight several biases that can have an effect on performance appraisals. The first bias is "fundamental error", a tendency to underestimate the importance of situational factors and overestimate the dispositional factors as causes of the observed behaviour. Will the subjects search the data differently if they think the cause of the declines is hardware related? Will the preconceived notions the rater has concerning the ratee effect the type of information sought (De Nisi et al., 1984)?

H1c: *If subjects know that they have to give performance appraisals they will be <u>less</u> accurate in spotting <u>difficult decline periods</u>.*

This hypothesis is linked to H1a, but the difficult accuracy decline periods are the hourly declines in performance. Will the subjects who have appraisal knowledge search the data differently, so that they disregard the hourly data and concentrate on the daily data that gives a better picture of the employees' overall performance?

H1d: *If subjects think that hardware problems are responsible for the decline in performance they will be <u>less</u> accurate in spotting <u>difficult decline periods</u>.*

This hypothesis is linked to H1b. If the subjects think that dissatisfaction is the cause of the decline in performance, will they be searching for signs of hourly declines in performance? Will the subject be using automatic processing when they are looking at the hourly performance data? This automatic processing would allow the rater to disregard hourly fluctuations in performance as being part of the rater's preconceived idea about the employee.

Total accuracy is the sum of easy accuracy and difficult, therefore the next two hypotheses follow from the first four hypotheses.

***H1e**: If subjects know that they have to give performance appraisals, they will be <u>more</u> accurate in spotting <u>total decline periods</u>.*

***H1f**: If subjects think that hardware problems are responsible for the decline in performance they will be <u>less</u> accurate in spotting <u>total decline periods</u>.*

For hypotheses H1a, H1c and H1e which refer to knowledge of performance appraisals and easy, difficult and total accuracy respectively, it was predicted that the subjects would be more accurate in spotting the declines in performance if they have appraisal knowledge for easy accuracy, and total accuracy and less accurate in spotting difficult declines. The rationale behind these predictions was that DeNisi at al.'s model has the purpose of appraisal as a determinant. According to DeNisi et al. the purpose of the appraisal affects the way the subject searches the data. Williams et al. (1983) found that when raters were making comparative decisions, they sought information that would allow them to judge the performance of the ratee relative to the performance of others. The subjects in the appraisal knowledge sections could be trying to make comparative decisions. Recall that accuracy in terms of performance appraisal is defined as the degree to which the rank ordering of the ratings for a group of employees approximates the rank ordering of their objective performance. This implies an existence of a "true performance" score (DeNisi et al., 1984). With appraisal knowledge the subjects should be attempting to rank the employees, so that they can appraise them correctly. To be able to do this they would have to make a more thorough search of the data, and hence be more accurate in spotting the easy accuracy and total accuracy performance declines. The reason why the subjects can reasonably be expected to be less accurate in the difficult accuracy is that the subjects who had appraisal knowledge would tend to disregard the hourly declines and concentrate on the daily data that gives them a better overall picture of the employees' performance. The small or temporary variations do not indicate a performance trend (Fenner et al., 1993), therefore, the hourly declines are classed as temporary declines, and they should be disregarded when doing performance appraisals. There are only two hourly performance declines, so even if the subjects are less accurate in spotting the hourly declines, they will still be more accurate in spotting the total declines, as there are six decline periods in all. According to Kelley (1967), a rater seeks to determine whether an employee's poor performance was due to the worker, the task, or some interaction of the two when doing performance appraisals. The knowledge of having to do a performance appraisal could cause the subjects to attempt to find consensus, distinctiveness, or consistency information to determine the performance of the subjects. This again would suggest a more thorough search on the part of those subjects who had performance appraisal knowledge.

For hypotheses H1b, H1d and H1f, which are the hypotheses concerned with reason, and easy, difficult and total accuracy respectively. The following theory points to subjects being more accurate if they thought dissatisfaction was the cause of the performance declines. The first bias of fundamental error (Ross, 1977) as discussed by Ilgen and Feldman (1983), says that subjects have a tendency to underestimate the importance of situational factors, and overestimate the importance of dispositional factors. As the subjects have been primed with the notion that the subjects are dissatisfied employees, and dissatisfaction is a dispositional factor, the subjects will search more data if they think dissatisfaction is the cause.

Kozlowski and Ford (1991) said that raters search for more information for ratees with 'poor' performance ratings. If the subjects think that the employees are dissatisfied, they will be likely to search for more information thinking that they are dealing with poor employees. If the subjects search more data they will spot the declines in performance more accurately.

DeNisi et al. (1984) said that the evidence suggests that a rater's preconceived notion does in fact affect the type of information the rater seeks. The effect would be caused by the activation of some schema followed by the rater searching for some information that would prove or disprove the validity of the schema. The way the search is carried out would then have an effect on what the rater actually observes. If the above hypotheses 1a-1f are accepted, it would show that there has been an effect on what the rater has actually observed. We expected to find that there would be a difference in the accuracy in which the raters spotted the performance declines. Upon acceptance of any of the above hypotheses, the data will be examined further to see if there is any difference in the way the raters have searched the information.

5.2 Breadth of Search Hypotheses

Breadth of Search is measured by the number of screens the rater accesses. If the rater only accesses 50% of the screens, they would only have access to 50% of the available data. The following hypotheses address the breath of search:

H2a: If subjects know that they have to do a performance appraisal, the breadth of search will be greater than that of those subjects who did not know that they would have to appraise performance.

The literature has shown that the purpose of the appraisal affects what information is sought, and how that information is later stored (DeNisi et al., 1984). Does it follow that if the subjects had no appraisal knowledge would they search fewer screens than those subjects who had appraisal knowledge? Will the subjects who have appraisal knowledge access more screens in an attempt to sample more work-related behaviour of an employee (Crocker, 1981)? The above literature points to the breadth of search being greater if subjects had appraisal knowledge. The literature discussed under the accuracy hypotheses for subjects with appraisal knowledge, namely those referring to searching for data to make comparisons, such as the work by Williams et al. (1983) and Kelley (1967) also points to the subjects being more accurate if they had appraisal knowledge

H2b: If subjects think that decline in performance is caused by hardware problems, the breadth of search will less than that of those subjects who think that task dissatisfaction is the reason for the decline in performance.

Again we refer to the bias of "fundamental error", in Ilgen and Feldman (1983)'s model, which is a tendency to underestimate the importance of situational factors, and overestimate the dispositional factors as causes of the observed behavior. Will the subjects be able to differentiate between hardware problems causing performance declines, or task dissatisfaction causing the decline? Will the subjects search more screens if they think dissatisfaction is the cause of the performance declines? Or will they be willing to make their decision based on a small amount of the available information (Major, 1980)? Kowslowski and Ford's (1991) contention that raters search for more information for ratees with 'poor' performance rating would suggest that the subjects would search more screens if they thought that dissatisfaction was the cause of the performance decline.

5.3 Hypotheses Regarding Average Time

Averages refer to the average time the subject spends on the various types of screens, and the average time the subject spends upon each screen as a percentage of the average time the subject spends on all screens. Will the preconceived notions and the reason for the search affect the average time the subject spends on each screen? This leads to the following hypotheses:

H3a: If subjects think that declines in performance are caused by hardware problems, the average time spent on each screen will be less than those subjects who think that task dissatisfaction is the reason for the declines in performance.

Again this is related to fundamental error (Ilgen & Feldman, 1983), a tendency to underestimate the importance of situational factors and overestimate the dispositional factors as causes of the observed behaviour. In this case will underestimating the situational factors cause the rater to spend more time on each screen? Those subjects who think that dissatisfaction is the cause of the performance decline will spend more time on each screen. They will underestimate the situational factors and not consider that hardware could be a cause of the performance declines, and hence spend more time on each screen.

H3b: If subjects know that they have to do a performance appraisal, the average time spent on each screen will be greater than that of subjects who have no appraisal knowledge.

Does the purpose for which the raters think they are searching for information affect the way they search the information? Will the subjects spend more time on each screen if they know that they have to do a performance appraisal? In this case we expect that the subjects will spend more time on each screen if they know have to do a performance appraisal. They are looking at the data for two purposes. Those who think they have to do a performance appraisal will have to spend time trying to compare the employees so that they can accurately access the employees' performance. The subjects who have appraisal knowledge also have to spend more effort in processing the information so they can store it in memory in an organized fashion so that they can recall it for the appraisal process. This will also cause them to spend more time on each screen.

5.4 Times and Frequencies Hypotheses

Times and frequencies refer to how long the subjects spend on the various screens, also how many times they access the different types of screens. The way the subjects use the different types of screens could give us an insight into how the data is used differently depending upon the preconceived notions and knowledge of the appraisal. The following hypotheses follow:

***H4a:** If subjects think that declines in performance were caused by hardware problems, the time spent on the various screens, and the frequency with which they access them will be less than that of the subjects who think that task dissatisfaction was the reason for the decline in performance.*

Do the subjects look at tables and graphs differently if they think hardware is the cause of performance declines? According to Powers at al. (1984) accuracy of performance can be increased by the use of graphs and tables. Do the subjects spend more time on graph screens rather than table screens? Does the use of graphs and tables help the subjects search the data better and hence to more accurately spot the performance declines? Do they spend more time on the data and access the screens more frequently trying to find out if there is any performance decline in common which could be caused by a common breakdown in hardware? If the subjects think that dissatisfaction is the cause of performance decline we predict that they will spend more time on the data, and access the screens more frequently. This again is a result of Kozlowski and Ford's (1991) raters search for more information for ratees with a poor performance rating.

***H4b**: If subjects know that they have to do a performance appraisal, the time spent on the various screens, and the frequency in which they access them will be greater than that of the subjects who have no appraisal knowledge.*

Do the subjects spend more time on the screens and access them more frequently in an attempt to find consensus information? If the subjects know that they have to give performance appraisals we expect that they will spend more time on the screens and access them more frequently. Spending more time on the screens and accessing them more frequently suggests that the subjects are attempting to make comparisons. This again points to Williams et al. (1983) research which found that when raters were making comparative decisions, they sought information that would allow them to judge the performance of the ratee relative to the performance of others.

Chapter 6: Method

6.1 Subjects

The study was conducted at Carnegie Mellon University by Drs. Deborah J. Ballou, F. Javier Lerch, and Carol T. Kulik. Data was obtained from 127 graduate and undergraduate students, who each received $10 for their participation in the study.

6.2 Design and Stimuli

In this research we only looked at decline spotting accuracy. We did not look at the performance ratings, only at the effect of the performance appraisal knowledge.

The study consisted of a 2 x 2 x 3 factorial design. The two between subject factors were the objective of monitoring (single -- only identifying decline periods, dual -- identifying decline periods and rating performance) and reason for performance declines (machine, task dissatisfaction). The within subject factor was the number of performance declines (one, two or three). Subjects were assigned at random to one of the four following experimental cells: (1) dual monitoring objectives, dissatisfaction caused performance declines; (2) dual monitoring objectives, machine caused performance declines; (3) single monitoring objective, dissatisfaction caused performance declines; (4) single monitoring objective, machine caused performance declines (See table 1).

All subjects monitored and evaluated the performance data of three simulated employees who had three levels of performance declines (one, two or three). Performance data for each simulated employee was available in both graph and table form for two weekly periods, in daily and hourly format. Thus, four weeks of data was available.

Table 1: Experimental Cells

Cell Number	Appraisal Knowledge	Reason for Performance Decline
100 (n=36)	Yes	Hardware
200 (n=30)	No	Hardware
300 (n=31)	Yes	Dissatisfaction
400 (n=30)	No	Dissatisfaction

Subjects in the 'dual objective monitoring' section were told on three separate occasions that they would be required "both to identify performance decline periods and to rate the performance of the three employees." This occurred in the computerized introductions to the experiment and in the data search section. Subjects in the 'single objective monitoring' section were told on the same three occasions that they would be required "to identify performance decline periods."

Subjects in the 'task dissatisfaction caused declines' section were told on three separate occasions that, "the declines in performance are frequently caused by employees' dissatisfaction with their assignments." The three occasions were in the computerized introductions to the experiment, in the first practice session and in the data search session. Likewise, subjects in the 'machine caused performance declines' section were also told on the same three occasions that, "The declines in performance are frequently caused by problems in the computer hardware." To ensure that the subjects had been aware of the performance decline reasons given to them, they were asked in the post study questionnaire if they could "recall WHY clerks had temporary performance declines at Shane Data, Inc." Fifty-six percent of the 127 subjects or 72 subjects, were able to accurately identify the reason for the decline in the questionnaire. We then looked at the responses of the other 44 percent who were not able to accurately identify the reason for the decline in performance on the questionnaire. Eight percent of the hardware group said dissatisfaction was the cause, but no one in the dissatisfaction caused performance declines group suggested that the declines could be caused by hardware problems.

Each simulated employee had a 160-member performance array created for them. The members of these arrays were values representing the number of time cards each employee had processed every hour during twenty working days of eight hours each. The array for the employee who had the lowest number of performance declines consisted of numbers between 31 and 60. This array had 7955 time cards in total, with a mean of 50 time cards per hour, and a standard deviation of 6.4. The array contained one 16 hour (two day) decline period occurring on days 16 and 17. The average number of time cards processed during this time was 42.5 cards per hour. The array for the employee with the medium number of performance consisted of numbers between 32 and 60. This array had a total of 7841 time cards, with a mean of 49 time cards per hour, and a standard deviation of 6.6. The array contained two decline periods, one four hour decline on day 3 and one 16 hour (two day) decline period occurring on days 16 and 17, giving a total of 20 hours of decline periods. The average number of time cards processed during these decline periods was 40 cards per hour. The array for the employee who had the highest number of performance declines consisted of numbers between 30 and 60. This array had 7584 time cards in total, with a mean of 47 time cards per hour, and a standard deviation of 7. The array contained one 16 hour (two day) decline period on days 8 and 9, one 4 hour decline period on day 12 and one 16 hour (two day) decline period occurring on days 16 and 17, giving a total of 36 decline hours in total (3 periods). The average number of time cards processed during this time was 40 time cards per hour.

Each simulated employee was differentiated by the number of decline periods they had (1,2 and 3), the duration of the decline (16 hours, 20 hours, 36 hours) and the total number of time cards processed during the twenty day period (7955, 7841, 7584). There was, however, one decline period in common, which occurred on days 16 and 17. The subjects were asked on the post study questionnaire if they had noticed "that all three employees had a common two day decline period," they were then asked to identify the period. This was in order to verify the accuracy of the subjects' observation. Sixty-eight percent of the one hundred and twenty-seven subjects (eighty-six subjects) were able to identify the common decline period.

The subjects had two types of tables and graphs available to them. The first type of table summarized daily performance for ten days in a two-week period. The second type of table summarized the hourly performance for the eight hours of each day in a two-week period. The first type of graph summarized the daily performance for the ten days in a two-week period in a line graph. The second type of graph summarized the hourly performance for the eight hours of each day in a two-week period again in a line graph. The hourly and daily data for each simulated employee was available in both graph and table format. There were 24 screens available: two daily tables for each employee, one for each week for each employee giving six screens; two daily graphs for each employee, again one for each week for each employee (six screens); two hourly tables for each employee, one for each week (six screens) and two hourly graphs for each employee, one for each week (six screens).

6.3 Procedure

The experiment that was designed using the cT language[2] was entirely computerized. A file was created for each subject that collected information about how they did the search. When the subject went from screen to screen a time was recorded. Also when the subject recorded a decline period the time was recorded. At the beginning of the practice session subjects were told by the experimenter that the study was designed to explore how supervisors search through performance information to spot periods when the employees were performing poorly. It was explained that computers are being used more often to collect performance data and supervisors are frequently being asked to monitor employees' performance using this data. Thus, the issue has important consequences for companies that are using this data.

Using a computer set up for demonstration purposes, the experimenter showed the subjects' sample search menus and data screens in both table and graph format. The subjects were told that they could search through the data in any way they liked. When the subjects had identified a decline period, they had to click on a button. They were shown how to do this by the experimenter. They also had to fill in the decline period manually on a form. Subjects were told that they would get more information about what constituted a decline period in the instructions on the computer, and there

---

2 Copyright 1989 by Bruce Arne Sherwood of Carnegie Mellon University and Judith N. Sherwood of Sonoma State University. All rights reserved.

would be a practice session to orientate themselves with the experiment and to help them to become familiar with the decline periods. The experiment took place in private cubicles. Before the subjects entered into the cubicle they were given a manila envelope which contained two decline statements for the two practice sessions and six decline statements that had to be completed during the experiment. These decline statements had spaces in which the subject had to identify the employee, the day/s and/or hours of the decline period.

The subjects were presented with two screens of introductory information. They were allowed to look at these screens for as long as they liked. The subjects could go back and forth between the two screens at the click of a button. The first screen told the subjects that they would be searching through performance data of three employees of Shane Data Inc., a data processing company whose main business was processing payroll information. These employees were responsible for transferring the information provided by Shane Data's customers into the computer system. The time cards were randomly assigned to the employees as they arrived, and they were all different. The subjects were told that they would be using a computerized performance monitoring system developed by Shane Data Inc. that records the number of time cards that the employees processed every hour. The introduction also stated that one of the main objectives of the performance monitoring system was to help the company's supervisors identify temporary declines in the employees' performance, and in this session the subject would be responsible for identifying these performance declines. At this point the subjects who were in the 'machines caused performance decline' were told that the performance declines were usually due to problems in the computer hardware, whereas the subjects in the 'dissatisfaction caused performance decline' were told that the performance declines were usually caused by employees' dissatisfaction with their tasks. Finally, subjects in the appraisal knowledge condition were told that they would be asked to rate the performance of the three employees after the decline periods had been identified.

On the second screen the subjects were informed that the work day consisted of eight hours, and that the monitoring system would show how many time cards would be processed both daily and hourly. Subjects were also advised that the performance declines would follow two different patterns: The first pattern would be a few hours within a day (e.g., hours of performance decline in day 8); the second pattern would be a few days within a week (e.g., 2 days of performance decline within a week). They were then told that their task was to search through the data for the three employees and to identify these performance declines. Subjects in the 'appraisal knowledge' section were also reminded that they would be expected to rate the performance of the three employees. Subjects were told that they would have a practice session in which they would be able to familiarize themselves with the monitoring system and how to identify decline periods.

In the two practice sessions, subjects searched through the data for a sample employee to find an example of each of the main types of decline period, namely a few hours within the day and a couple of days within a week. They were then told in the computerized instructions and on the sheet in which the decline period was recorded exactly when the decline period occurred. The subjects had only to search through the data to familiarize themselves with how the decline periods were displayed in both the graph and table formats and in the hourly and daily data. When they discovered a decline period they had to click on a black box on the screen and fill in the day and time if applicable on the form provided. It was emphasized that they must click on the black box each time they entered the decline period. Subjects were told that they should look at both the table and the graph in the practice session, so that they would be able to spot the decline periods in the actual study. In addition subjects in 'the machines caused reason for decline' were also told the declines may have been caused by problems with the computer hardware, subjects in the 'dissatisfaction caused performance decline' were again told that the declines were caused by employees' dissatisfaction with the task. The subjects who had appraisal knowledge were told that they would be required to rate the performance of the employee after the decline periods had been identified.

In the practice sessions, the monitoring system that the subjects encountered was identical to what they would use in the following data search except for the fact that there was only one employee listed, whereas there are three employees on the search menu. The search menu screen showed the name of the employee whose data could be accessed. It also displayed the following categories: daily performance data and hourly performance data. Directly below the categories, the weeks for which data were available was listed. These weeks were: weeks 1 and 2, and weeks 3 and 4. Subjects could click in boxes to reach the appropriate category and week in order to access a screen containing performance data. The words "tables and graphs" were displayed at the top of the screen, with a box beside each word. The subject was required to click the box in order to access the format in which they wanted the data to appear.

There was a total of eight screens that the subject could access, one for each format/aggregation/time period combination. Each screen contained two boxes. One of these boxes was used to return the subject to the search menu, so that they could search other screens. The other button was used by the subjects to signal that they had found a decline period on the screen, and they had recorded it on the decline sheet. The subjects were allowed to search the screens in any order they wished. For example, a subject could look at the data for weeks 1 and 2 in a graph format, then they could look at the table for the same two weeks. They were allowed to search a screen as many times as they wanted.

In each of the two practice sessions the subjects filled in the decline form as soon as they had located the decline period and clicked on the black box. In order to fill in the decline form the subjects had to enter the numbers of the day and the hours in which the decline had occurred or in the case of a multiple decline, the numbers of the days on which the decline occurred. For the practice session only the decline statements were preprinted with the name of the employee and identified the periods during which the declines occurred.

At the end of each practice session subjects were shown screens which displayed the exact hours and/or days in which the decline periods occurred in the same format as they should have appeared on the decline form. The subjects were told that they could return to the search menu if they wanted to verify the declines or look at the values again.

After the subjects completed the two practice sessions they were ready to do the actual search. The first screen they saw was an introductory screen on which they were told that they were now going to search through the data for three employees, Ann Landy, Betty Miller, and Jill White. They would be required to find the periods of decline in performance for these three employees. Subjects were informed that there was a total of six decline periods distributed among the three employees, and they were only given six decline forms to complete. It was also explained to the subjects that each employee could have more than one decline period. Again the subjects in the 'machine reason for decline in performance' were told that the reason for the decline was frequently caused by problems with the computer hardware. The subjects in the 'dissatisfaction caused decline in performance' section were told that task dissatisfaction was often the reason for decline in performance. Those in the 'appraisal knowledge' section were told that they would be required to do a performance appraisal when they had finished finding the decline periods. Finally, all subjects were told again that they were required to click the black box after filling in **each** decline on the decline form.

The subjects then saw a search menu similar to the one they had seen in the practice session, but this time there were the names of three employees instead of one. As before the subjects had to click the appropriate box to specify if they wanted to see a table or graph, which weeks' data they wanted to see, and if they wanted to see days or hours. The only difference being that they were also required to click the name of the employee this time. A total of 24 screens could be accessed, one for each employee for each graph, table, time period, and days or hourly data. As before these screens contained two boxes, one to be used to return the subject to the search menu, and one to click when the decline form had been completed. Subjects had to enter the name of the employee on the decline form and then fill out the decline details as they had on the practice sheet. Also, as in the practice session the subjects were allowed to access the screens in any order they wished, and return to any screen so desired. When the subjects were satisfied that they had filled in all six decline periods, they clicked a box on the search menu which allowed them to exit the data search session. The next screen gave them instructions to place their completed decline statements in the manila envelope provided, and give it to the experimenter before continuing with the experiment.

The next segment of the experiment consisted of a practice evaluation session. The subjects were told that they would practice rating an employee's performance by answering questions about the performance of the employees whose data they had examined. They were told to think carefully about their ratings as they would only have one chance to fill in the ratings sheet. They were then advised that the practice session was only meant to familiarize themselves with the type of questions that they would be required to answer in the performance evaluation. The subjects then rated the three employees during the four weeks in which the data had been collected. They also rated how important skill or ability, level of effort, difficulty of the task and luck (random chance) were in determining the employee's performance. Finally they were asked to predict the employee's future performance.

Next, the introductory screen of the main evaluation session was displayed. Subjects were informed that they would now be evaluating the performance of the three employees whose data they had examined in the first part of the experiment. Again they were instructed to try as hard as they could do to remember the performance data. After they had viewed this screen subjects began the rating task. They answered the entire series of questions for each employee in turn. Again this series was composed of ratings of performance and consistency during the four weeks for which there was data, a prediction of future performance, and ratings regarding the importance of skill or ability, level of effort, difficulty of the task, and luck (random chance) in determining the employee's current performance. When subjects reached the end of this evaluation session, they were told that the computerized portion of the experiment was over and were instructed to fill out the post-study questionnaire before notifying the experimenter that they had finished.

On the post study questionnaire, subjects identified their majors and year in college. They provided ratings regarding the acceptability of using computer systems to monitor employee's work, and whether they would like their work to be monitored by a computer system. They also answered questions which asked them if they could recall why employees had temporary performance declines, if they had any strategy for finding the decline periods. If they answered yes to the previous question, they were asked to describe the strategy. Finally, they were asked if they had noticed that all three employees had a common two day decline period, and, if they had, were asked when it was. When the subjects finished filling out the post-study questionnaire they notified the experimenter. They were then able to ask any questions about the experiment, sign up on a mailing list in order to receive a future written description of the study, and fill out a questionnaire about their experience during the experiment. At the end of the experiment the participants were paid $10, thanked for their participation and asked not to discuss the experiment with their friends.

Chapter 7: Analysis

To test the hypotheses, both the decline statements and the computerized files that traced the subjects' search process were analyzed. Whenever the subject accessed or left a particular screen, the time was recorded. When the subjects recorded a decline period, by clicking the black box the time was also recorded. All the data was stored on a computer file for each subject. The data from the files for each subject was imported into a spreadsheet. The number of times the graph and table screens were visited in both the hourly and daily screens was aggregated. The amount of time each subject spent on each type of screen was also totaled (see table 5). When a screen was visited it was noted, which gave a value for the breadth of search, i.e., a value was calculated for the number of screens visited (see table 3). The average time the subject spent on each type of screen was also calculated (see table 4). When all calculations were completed, the file was analyzed using univariate ANOVAs. The computer printouts are in Appendices A - D.

7.1 Independent variables.

The independent variables investigated were appraisal knowledge (APPRKNOW) and reason (REASON) for decline. The subjects were each given a number. The subjects with numbers in the 100's had appraisal knowledge and they were given 'hardware problems' as the reason for the decline. The subjects in the

200's had no appraisal knowledge and 'hardware problems' for reason for decline. The subjects in the 300's had appraisal knowledge and were given 'dissatisfaction with the task' as a reason for decline. The subjects in the 400's had no appraisal knowledge and were given 'dissatisfaction with the task' as a reason for decline (see table 1). There were 36 subjects in cell 100, 30 in cell 200, 31 in cell 300, and 30 in cell 400. The independent variables were indicator variables with 1 = appraisal knowledge, 0 = no appraisal knowledge; 0 = hardware problems for reason of decline and 1 = dissatisfaction with task.

7.2 Dependent variables

7.2.1 Accuracy Variables

The first dependent variables considered in this study were the accuracy with which the subjects spotted the decline periods. These were the decline periods the subjects entered on the decline statements. In all there were six decline periods possible, two of which were classified as "difficult accuracy" (DIFFACC) - these were the hourly decline periods. The other four decline periods were classified as "easy accuracy" (EASYACC) i.e., the decline periods spanned a day or more. The total accuracy was also considered. This was the number of decline periods spotted correctly out of the six decline periods. The way the experiment was set up was that employee Ann Landy had high performance with one 'easy accuracy' decline period on days 16 and 17. Employee Jill White had medium performance with two decline periods one 'easy accuracy' on days 16 and 17 and one 'difficult accuracy' on day three, of 4 hours (hours 5-8 inclusively). Betty Miller was considered the poor employee. She had three decline periods. These were two 'easy accuracy' decline periods, one on days 16 and 17, another on days eight and nine and a 'difficult accuracy' decline period on day 12 (hours 3-6) (see table 2).

Table 2: Description of Accuracy Variables

VARIABLE	DESCRIPTION	Hypotheses
EASYACC (Easy Accuracy)	The number of decline periods spotted that span a day or more	H1a, H1b
DIFFACC (Difficult Accuracy)	The number of decline periods spotted that spanned a few hours.	H1c, H1d
TOTACC (Total Accuracy)	The total number of decline periods spotted.	H1e, H1f

The other dependent variables were obtained from the computer generated files which showed ways in which the subjects searched the data. For a complete list of these variables see tables 3-5. All these variables were analyzed as dependent variables, with appraisal knowledge, reason for decline and the interaction between appraisal knowledge and reason as independent variables.

7.2.2 Breadth of Search Variables.

There are two different breadth of search variables, one that denotes the number of times each unique screen type is accessed out of the total number of times possible, and the percentage of the number of unique screens accessed out of the number of screens of that type available. For example, there were 6 day table screens available. Two for each of the three employees. Every time a subject accessed a different day table screen it was counted. If the subject looked at all the screens they would have scored 6 for breadth of search day tables (BSDAYTAB), and their score for breadth of search day tables percentages (BSDTPER) would have been 100%. See table 3.

7.2.3 Averages

There are two different average variables, one was calculated by the amount of time spent on the various types of screens, divided by the number of times the screens were visited (see table 4a). For example, the average hour variable (AVHOUR), which is the average time spent on the hour screens, was calculated by taking the amount of time spent on the hour screens by the subject and dividing it by the number of times the hour screens were accessed by the subject. The other variable types are the overall averages (see table 4b). These were calculated by taking the average found above and calculating it as a percentage of the overall average time the subject spent on all the screens. These percentages will be over 100% in many cases. In fact a figure of 100% is the average. If a subject spent an average of 20 seconds on all screens, and she spent 20 seconds on the table day screens (AVTDPER), her overall score would be 100% for average table day percent. Thus the variable AVTDPER is calculated by

Table 3: Breadth of Search Variables

VARIABLE	DESCRIPTION	Hypothesis
SCRSRCH (Screen Search)	The number of unique screens accessed out of the twenty four screens available.	H2
BRSRCH (Breadth of Search)	The percentage of the unique screens accessed out of the possible screens.	H2
BSDAYTAB (Breadth Search Day Tables)	The number of unique day table screens accessed.	H2
BSDTPER (Breadth Search Day Tables Percentage)	The number of unique day table screens accessed as a percentage of the number of times possible (6).	H2
BSTAB (Breadth of Search Tables)	The number of unique table screens accessed.	H2
BSTABPER (Breadth of Search Tables Percentage)	The number of unique table screens accessed as a percentage of the total number of times possible (12).	H2
BSHRTAB (Breadth of Search Hour Tables)	The number of unique hour table screens accessed.	H2
BSHTPER (Breadth Search Hour Table Percentage)	The number of unique hour table screens accessed as a percentage of the total number of times possible (6).	H2
BSDAYS (Breadth of Search Days)	The number of unique day screens accessed.	H2
BSDAYPER (Breadth Search Days Percentage)	The number of unique day screens accessed as a percentage of the total number of times possible (12).	H2
BSHRS (Breadth of Search Hours)	The number of unique hour screens accessed.	H2
BSHRSPER (Breadth of Search Hours Percentage)	The number of unique hour screens accessed as a percentage of the total number of times possible (12).	H2
BSDAYGR (Breadth Search Day Graph)	The number of unique day graph screens accessed .	H2

Table 3: Breadth of Search Variables (Continued)

VARIABLE	DESCRIPTION	Hypothesis
BSDGPER (Breadth of Search Day Graph Percentage)	The number of unique day graph screens accessed as a percentage of the total number of times possible (6).	H2
BSGRAPHS (Breadth Search Graphs)	The number of unique graph screens accessed.	H2
BSGRPPER (Breadth Search Graphs Percentage)	The number of unique graph screens accessed as a percentage of the total number of times possible (12).	H2
BSHRGRP (Breadth of Search Hour Graphs)	The number of unique hour graph screens accessed .	H2
BSHGPER (Breadth of Search Hour Graphs Percentage)	The number of unique hour graph screens accessed as a percentage of the total number of times possible (6).	H2

taking the average time the subject spent on the table day screens divided by the average time the subject spent on all screens and multiplying it by one-hundred.

Table 4a: Averages Variables

VARIABLE	DESCRIPTION	Hypothesis
AVTABDAY (Average Table Day)	The average time spent on the day tables. This was amount of time spent on the table day screens divided by the number of times the screens were accessed.	H3
AVGRDAY (Average Graph Day)	The average time spent on the day graphs. This was the amount of time spent on the graph day screens divided by the number of times the screens were accessed.	H3
AVTABHR (Average Table Hour)	The average time spent on the hour tables. This was the amount of time spent on the table hour screens divided by the number of times the screens were accessed.	H3
AVGRHR (Average Graph Hour)	The average time spent on the hour graphs. This was the amount of time spent on the graph hour screens divided by the number of times the screens were accessed.	H3
AVDAY (Average Day)	The average time spent on the day screens. This was the amount of time spent on the day screens divided by the number of times the screens were accessed.	H3
AVHOUR (Average Hour)	The average time spent on the hour screens. This was the amount of time spent on the hour screens divided by the number of times the screens were accessed.	H3
AVTABLE (Average Table)	The average time spent on the day graph screens. This was the amount of time spent on the table screens divided by the number of times the screens were accessed.	H3
AVGRAPH (Average Graph)	The average time spent on the day graphs. This was the amount of time spent on the graph screens divided by the number of times the screens were accessed.	H3
AVSCR (Average Screen)	The average time spent on each screen. This was the total time spent on the screens divided by the number of times the screens were accessed.	H3

Table 4b: Overall Averages Variables

VARIABLE	DESCRIPTION	Hypothesis
AVTDPER (Average Table Day Percentage)	The average time spent on the table day screens as a percentage of the average time per screen. This was the average amount of time spent on the table day screens divided by the average time spent on all screens, multiplied by one-hundred.	H3
AVGDPER (Average Graph Day Percentage)	The average time spent on the graph day screens as a percentage of the average time per screen. This was the average amount of time spent on the graph day screens divided by the average time spent on all screens, multiplied by one-hundred.	H3
AVTHPER (Average Table Hour Percentage)	The average time spent on the table hour screens as a percentage of the average time per screen. This was the average amount of time spent on the table hour screens divided by the average time spent on all screens, multiplied by one-hundred.	H3
AVGHPER (Average Graph Hour Percentage)	The average time spent on the graph hour screens as a percentage of the average time per screen. This was the average amount of time spent on the graph hour screens divided by the average time spent on all screens, multiplied by one-hundred.	H3
AVDAYPER (Average Day Percentage)	The average time spent on the day screens as a percentage of the average time per screen. This was the average amount of time spent on the day screens divided by the average time spent on all screens, multiplied by one-hundred.	H3
AVHRPER (Average Hour Percentage)	The average time spent on the hour screens as a percentage of the average time per screen. This was the average amount of time spent on the hour screens divided by the average time spent on all screens multiplied by one-hundred.	H3
AVTBPER (Average Table Percentage)	The average time spent on the table screens as a percentage of the average time per screen. This was the average amount of time spent on the table screens divided by the average time spent on all screens, multiplied by one-hundred.	H3
AVGRPER (Average Graph Percentage)	The average time spent on the graph screens as a percentage of the average time per screen. This was the average amount of time spent on the graph screens divided by the average time spent on all screens, multiplied by one-hundred.	H3

7.2.4 Times and Frequencies

There are four different variables in this section. There are two types of frequency variables, one of which is the number of times each screen type is accessed. This is different from the variable, breadth of search. To calculate this variable we counted every time the screen type was accessed. Thus, if each of the six table day screens were accessed four times the subject would score 24 on the table day (TABDAY) variable. The other type of frequency variable was calculated by taking the number of times the screen was accessed and dividing it by the total number of screens accessed. Thus, table day percentage (TDPER) is the percentage of table days screens accessed by the subject as a percentage of the total number of screens accessed by the same subject (see table 5a). The other types of variables are the times variables. There are also two of these variable types. The first one is the amount of time the subjects spent on the various screens. If the subject spent a total of 100 seconds on the table day screens they would score 100 on the table day time (TDTIME) variable. The second time variable is the amount of time spent on a certain screen divided by the total amount of time spent on the experiment, calculated from the time they entered the main menu screen to them handing in the decline sheet. Thus, the table day time (TDTPER) percentage variable was calculated by taking the time spent on the table day screens as a percentage of the total time (TOTTIME) (see table 5a).

Table 5a: Frequency Variables

VARIABLE	DESCRIPTION	Hypotheses
TOTTIME (Total time)	Total time spent on the actual experiment. This time was taken as the time from when the subjects accessed the first search menu until the time the experiment was handed in.	H4
TABLES	The number of times the subjects looked at the tables screens.	H4
GRAPHS	The number of times the subjects looked at the graphs screens.	H4
DAYS	The number of times the subjects looked at the days screens.	H4
HOURS	The number of times the subjects looked at the hour screens.	H4
TABDAY (Table days)	The number of times the subjects looked at the days tables.	H4
TABHOUR (Table Hours)	The number of times the subject looked at the hourly data in the table format.	H4
GRPHDAY (Graph Days)	The number of times the subject looked at the daily information screens with the information in the graph format.	H4
GRPHHOUR (Graph Hours)	The number of times the subject looked at the screens with hourly data in the graphical format.	H4
TOTSCR (Total Screens)	The total number of screens the subject accessed.	H4
TDPER (Table Day Percentage)	The percentage of table days screens accessed as a percentage of the total number of screens accessed.	H4
THPER (Table Hour Percentage)	The number of table hour screens accessed as a percentage of the total number of screens accessed.	H4
GDPER (Graph Day Percentage)	The number of graph day screens accessed as a percentage of the total number of screens accessed.	H4
GHPER (Graph Hour Percentage	The number of graph hour screens accessed as a percentage of the total number of screens accessed.	H4

Table 5a: Time Variables

VARIABLE	DESCRIPTION	Hypothesis
TABTIME (Table Time)	The total time spent on the table screens.	H4
GRPHTIME (Graph Time)	The total time spent on the graph screens.	H4
DAYTIME (Day Time)	The total time spent on the day screens.	H4
HOURTIME (Hour Time)	The total time spent on the hour screens.	H4
TDTIME(Table Day Time)	The total time spent on the table day screens.	H4
THTIME (Table Hour Time)	The total time spent on the table hour screens.	H4
GDTIME (Graph Day Time)	The total time spent on the graph day screens.	H4
GHTIME (Graph Hour Time)	The total time spent on the graph hour screens.	H4
TDTPER (Table Day Time Percentage)	The time spent on the table day screens as a percentage of the total time.	H4
THTPER (Table Hour Time Percentage)	The total time spent on the table hour screens as a percentage of the total time.	H4
GDTPER (Graph Day Time Percentage)	The total time spent on the graph day screens as a percentage of the total time.	H4
GHTPER (Graph Hour Screens)	The total time spent on the graph hour screens as a percentage of the total time.	H4

Chapter 8: Results

8.1 Accuracy

To test the first group of hypotheses the variables Easy Accuracy (EASYACC), Difficult Accuracy (DIFFACC) and Total Accuracy (TOTACC) were examined (see Appendix A for complete results). These are the variables that measure the accuracy in which the subjects spotted the decline periods. Easy Accuracy is the number of decline periods spotted which lasted more than a day. Difficult Accuracy is the number of decline periods spotted which lasted only a few hours. Total Accuracy is the total number of decline periods spotted. Recall that there were four easy accuracy decline periods, and two difficult accuracy decline periods, giving six total accuracy decline periods.

Hypothesis H1a predicted that there would be a difference in spotting the easy accuracy decline periods if the subjects knew that they had to give performance appraisals. The null hypothesis is not rejected as the F Value for appraisal knowledge = 0.29 (NS). There is no evidence to show that there is a difference in spotting the easy accuracy decline periods if the subjects knew that they would have to give performance appraisals (see table 6).

Hypothesis H1b predicted that there would be a difference in spotting the easy accuracy decline periods if the subjects thought that hardware problems were responsible for declines in performance. The null hypothesis is rejected as the F Value for reason was 5.99 ($p < .05$). Examining the means (see table 7) we see that the subjects were more accurate in spotting the decline period if they thought that hardware

Table 6: F Values for Accuracy

VARIABLE	Appraisal Knowledge	Reason	Interaction Appraisal Knowledge * Reason
Easy Accuracy (EASYACC)	0.29	5.99*	0.48
Difficult Accuracy (DIFFACC)	5.02*	6.13*	5.66*
Total Accuracy (TOTACC)	0.13	7.79**	0.08

*$p < .05$ ** $p < .01$

was responsible for the decline in performance. If the subjects thought that hardware was the problem, they spotted an average of 3.03 decline periods, if the subjects thought employee dissatisfaction was the cause of the decline periods they only spotted an average of 2.34 decline periods.

Hypothesis H1c predicted that there would be a difference in spotting the difficult accuracy decline periods if the subjects knew that they had to give performance appraisals. The null hypothesis is rejected as the F value is 5.02 ($p < .05$) (see table 6). Looking at the means, we see that there was greater accuracy if the subjects have no knowledge of the appraisal (see table 8). The subjects spotted an average of 0.66 decline periods if they did not have appraisal knowledge, but they only spotted an average of 0.38 if they had appraisal knowledge.

Hypothesis H1d predicted that there would be a difference in spotting the difficult accuracy declines periods if the subjects thought that hardware problems were responsible for decline in performance. The null hypothesis is rejected as the F Value is 6.13 ($p<.05$) (see table 6). Looking at the means we see that there is greater accuracy if the subjects thought that hardware problems caused a decline in performance (see table 7). If the subjects thought hardware was responsible for the performance decline they spotted an average of 0.65 decline periods, but if they thought employee dissatisfaction was the cause they only spotted an average of 0.36 decline periods.

Table 7: Means for Accuracy (Reason)

REASON	Hardware	Dissatisfaction
Easy Accuracy (EASYACC)	3.03 (1.44)	2.34 (1.67)
Difficult Accuracy (DIFFACC)	0.65 (0.79)	0.36 (0.63)
Total Accuracy (TOTACC)	3.68 (1.92)	2.71 (2.03)

Standard deviations in brackets

Table 8: Significant Means for Accuracy (Appraisal Knowledge)

APPRAISAL KNOWLEDGE	No Appraisal Knowledge	Appraisal Knowledge
Difficult Accuracy (DIFFACC)	0.66 (0.77)	0.38 (0.67)

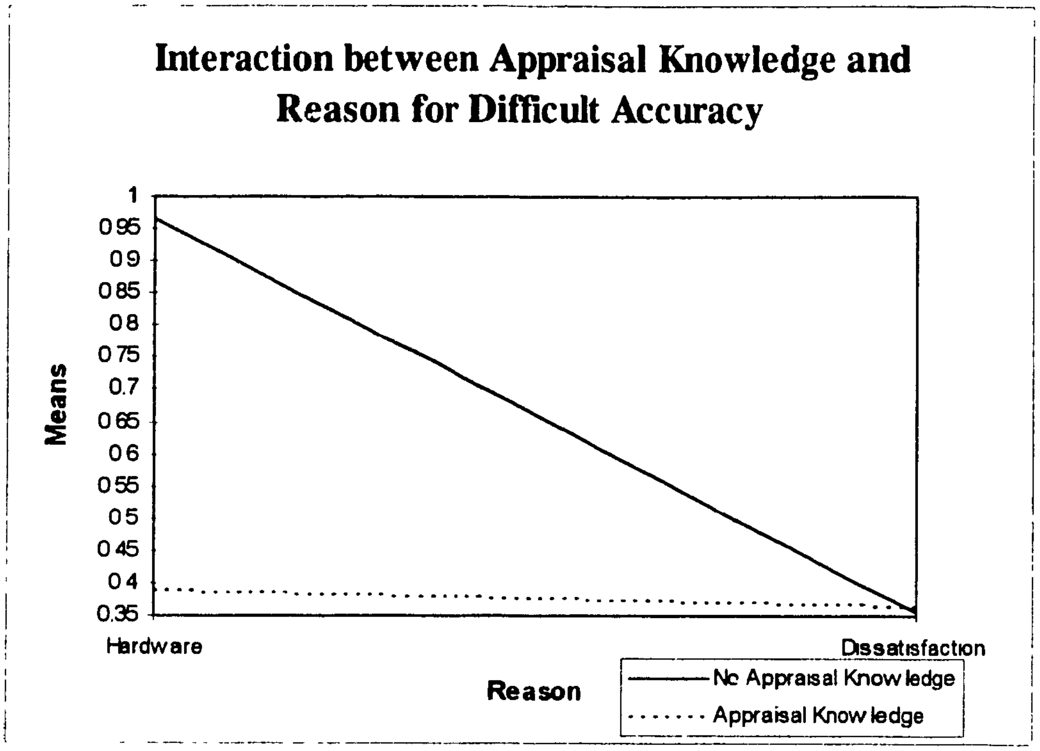
Standard deviations in brackets

There is, however, a very large interaction between appraisal knowledge and reason. These interactions were not predicted in the hypotheses, so they may be chance occurrences. When we look at the interaction we see that when the subjects did not have appraisal knowledge and they thought hardware was the problem, they spotted an average of 0.97 of the hourly declines (see table 9). The subjects in the other three groups only spotted an average of .35 to .39 decline periods. Thus, the subjects who thought hardware was the cause of the decline in performance and had no appraisal knowledge were nearly three times more accurate than the other three groups. This interaction effect seems to drive the two main effects. Figure 1 shows a graph of the interaction between appraisal knowledge and reason for difficult accuracy.

Table 9: Means for Accuracy (Interaction)

	No Appraisal Knowledge		Appraisal Knowledge	
VARIABLE	Hardware	Dissatisfaction	Hardware	Dissatisfaction
Difficult Accuracy (DIFFACC)	0.97 (1.44)	0.35 (0.85)	0.39 (0.65)	0.37 (0.72)

Standard deviations in brackets



(Note: this graph and all graphs following do not cross at the origin)

Fig 1: Interaction between Appraisal Knowledge and Reason for Difficult Accuracy

Hypothesis H1e predicted that there would be a difference in spotting all decline periods if the subjects knew that they had to give performance appraisals. The null hypothesis is not rejected as the F Value for appraisal knowledge = 0.13 (NS). There is no evidence to show that there is a difference in spotting all decline periods if the subjects knew that they had to give performance appraisals (see table 6)

Hypothesis H1f predicted that there would be a difference in spotting all decline periods if the subjects thought that hardware problems were responsible for declines in performance. The null hypothesis is rejected as the F Value is 7.79 ($p < .01$). Looking at the means we see that there was greater accuracy if the subjects thought that hardware problems caused a decline in performance (see table 7). If the subjects thought that hardware problems were the cause of performance decline they spotted an average of 3.68 decline periods, if they thought employee dissatisfaction caused the performance declines, they only spotted an average of 2.71 decline periods.

8.2 Breadth of Search

The value for Breadth of Search was calculated by the number of unique screens accessed out of the total number of screens available. For example, there were 24 screens in total (comprised of both graph and table screens), which had data in hourly or daily format. If a subject had accessed 19 out of the possible 24 screens, their score was given as 19 or 79.17%. These values were calculated for all the different types of screens so we could understand how the data had been searched. For a complete copy of the results of this section see Appendix B.

Hypothesis H2a predicted that if the subjects knew they had to do a performance appraisal, the breadth of search would be different from those subjects who had no appraisal knowledge. The null hypothesis is rejected as the F value is 4.3 ($p < .05$) see table 10. Looking t the means, we see that more screens were searched if the subjects had no knowledge of the appraisal (see table 11). However, there was only a mean of 66.53% of total screens searched. Only 59.47% of screens were searched if the subject had appraisal knowledge. The results for breadth of search hour tables (BSHRTAB and BSHTPER) and breadth of search hours (BSHRS & BSHRSPER) also show that there was a difference in the way the subjects searched if they had appraisal knowledge. In each case more screens were searched when the subjects <u>did not</u> have appraisal knowledge. For the breadth of search hour tables, which was the number of unique tables accessed where the data was in an hourly format, the subject accessed 66.67% of the available tables if they had no appraisal knowledge, as opposed to 52.27% if they subjects had appraisal knowledge. There is, however, a strong interaction effect for this variable. For the breadth of search hours, which was the number of unique screens accessed that had information in an hourly format, 63.11% of the screens were accessed if the subjects had no appraisal knowledge, as opposed to only 50.63% in the appraisal knowledge group.

Table 10: Significant F values for Breadth of Search

VARIABLE	Appraisal Knowledge	Reason	Appraisal Knowledge * Reason
Breadth of Search (SCRSRCH & BRSRCH)	4.3*		
Breadth of Search Day Table (BSDAYTAB & BSDTPER)			4.96*
Breadth of Search Table (BSTAB & BSTABPER)			7.78**
Breadth of Search Hour Table (BSHRTAB & BSHTPER)	6.57*		6.43*
Breadth of Search Hours (BSHRS & BSHRSPER)	8.28**		
Breadth of Search Graphs (BSGRAPHS & BSGRPPER)		4.88*	
Breadth of Search Hour Graphs (BSHRGRP & BSHGPER)		4.08*	

*p<.05 ** p<.01

Hypothesis H2b predicted that if the subjects thought that declines in performance were caused by hardware problems, the breadth of search would be different from those subjects who thought that task dissatisfaction was the reason for the declines in performance. The null hypothesis was rejected for the variables breadth of search graph screens (BSGRAPHS and BSGRPPER), F-value 4.88 (p <

.05) and breath of search hour graphs screens (BSHRGR and BSHGPER), F-value 4.08 ($p < .05$). They all searched more screens if they thought that hardware was the cause of the performance decline (see table 12).

Table 11: Means for Breadth of Search (Appraisal Knowledge)

APPRAISAL KNOWLEDGE	No Appraisal Knowledge	Appraisal Knowledge
Breadth of Search (%) (BRSRCH)	66.5 (19.57)	59.47 (19.25)
Breadth of Search Hour Table (%) (BSHTPER)	66.67 (31.47)	52.27 (32.82)
Breadth of Search Hours (%) (BSHRSPER)	63.11 (23.98)	50.63 (25.33)

Standard deviations in brackets

The percentage of unique graph screens accessed was 71.84% if the subjects thought that hardware was the cause of the performance declines. 60.93% of unique graph screens were accessed if the subjects thought that dissatisfaction was the cause of the performance declines. The subjects searched 59.85% of graph screens with data in an hourly format if they thought hardware was the cause of performance declines. They searched 47.81% of the hour graph screen if they thought dissatisfaction was the cause of performance declines.

There were significant interactions between appraisal knowledge and reason, These interactions were not predicted in the hypotheses, so they may be chance occurrences. ɔr breadth search tables, however, the interactions were highly significant ($p. < .01$). Looking at the interaction between appraisal knowledge and reason, the variables breadth search tables (BSTAB and BSTABPER), breadth of search day tables (BSDAYTAB and BSDTPER) and breadth of search hour tables

Table 12: Means for Breadth of Search (Reason)

REASON	Hardware	Dissatisfaction
Breadth Search Graphs Percentage(BSGRPPER)	71.84 (27.69)	60.93 (29.37)
Breadth Search Hour Graphs Percentage (BSHGPER)	59.85 (34.59)	47.81 (36.19)

Standard deviations in brackets

(BSHRTAB and BSHTPER) showed that the subjects searched more screens if they had no appraisal knowledge and they thought hardware was the cause of the performance declines. The group that searched the second highest number of screens was the group that had appraisal knowledge, but thought that dissatisfaction was the cause of performance declines (see table 13 and figures 2, 3 and 4 which are graphs of the interaction). For the breadth of search tables, the subjects accessed 71.67% of the available tables if they had no appraisal knowledge and they thought hardware was the cause of the performance decline.

Table 13: Means for Breadth of Search (Interaction)

	No Appraisal Knowledge		Appraisal Knowledge	
VARIABLE	Hardware	Dissatisfaction	Hardware.	Dissatisfaction
Breadth Search Day Tables Percentage (BSDTPER)	68.33 (38.49)	54.30 (42.16)	49.54 (37.9)	66.11 (35.69)
Breadth Search Tables Percentage (BSTABPER)	71.67 (26.41)	56.45 (34.74)	48.15 (29.28)	62.5 (27.92)
Breadth Search Hour Tables Percentage (BSHTPER)	75.00 (25.05)	58.6 (35.19)	46.76 (33.29)	58.89 (31.48)

Standard deviations in brackets

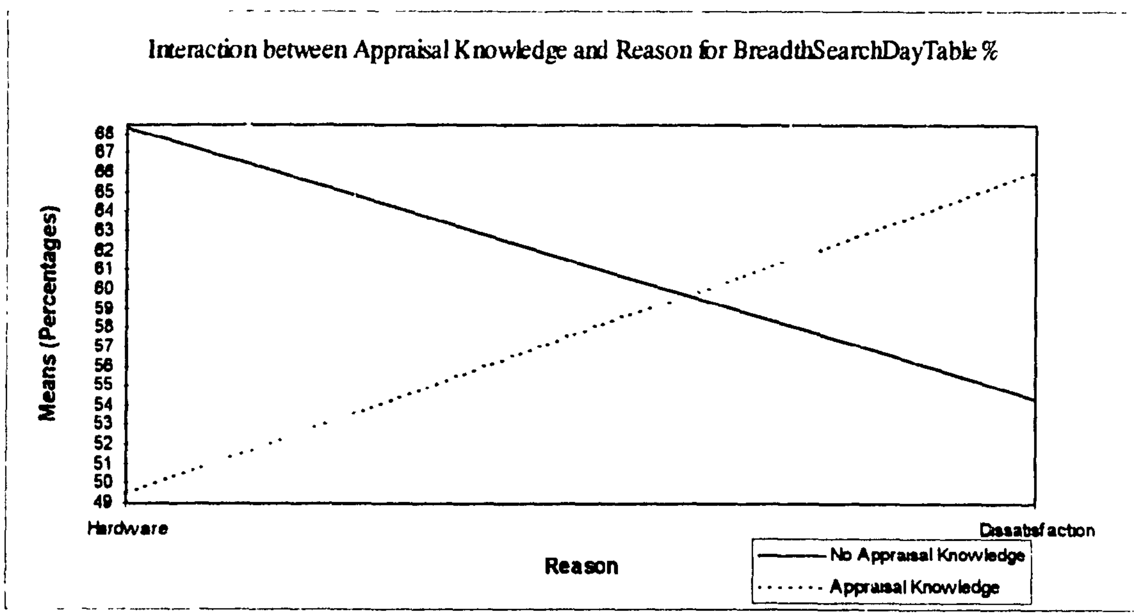


Fig 2: Interaction between Appraisal Knowledge and Reason for Breadth of Search Day Table Percentage

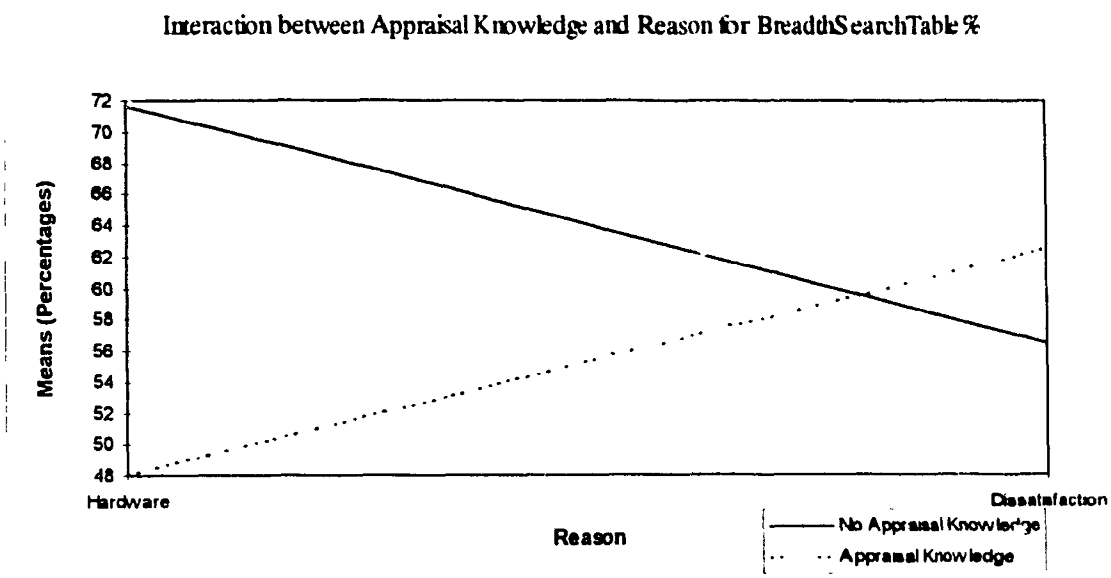


Fig 3: Interaction between Appraisal Knowledge and Reason for Breadth of Search Table Percentage

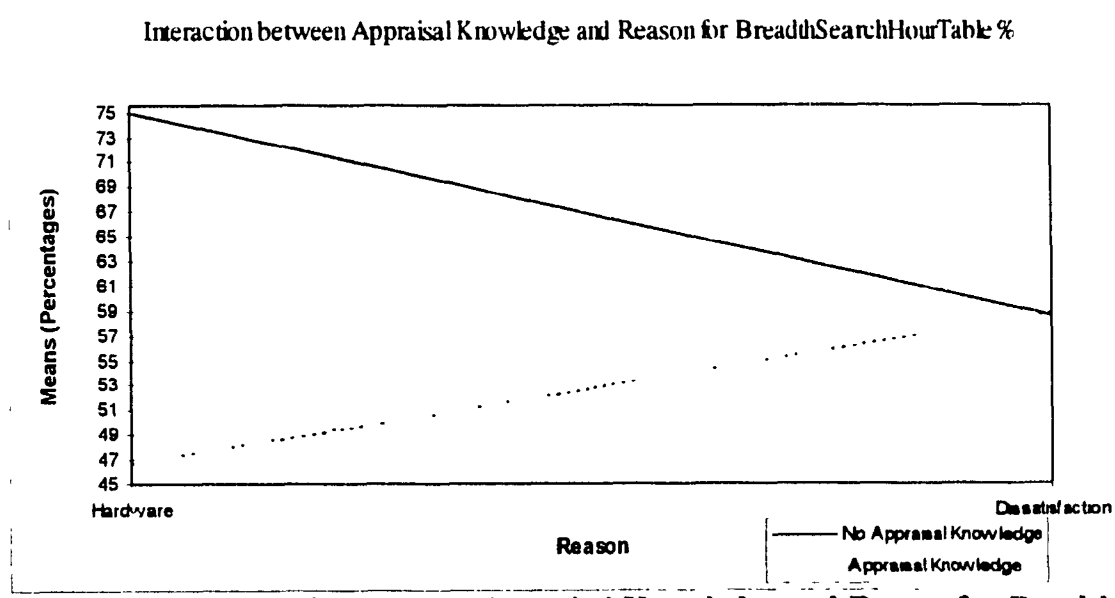


Fig 4: Interaction between Appraisal Knowledge and Reason for Breadth of Search Hour Table Percentage

8.3 Averages

There were two different types of averages examined. The first type was the average amount of time spent on the different screens. This was calculated by taking the total amount of time spent on a particular screen, and dividing it by the number of times the screen was accessed. For example, the average graph screen was the amount of time spent on the graph screen, divided by the amount of times the graph screens were visited. If a subject had spent a total of 100 seconds on the graph screens, and visited 5 different graph screens, the value for average graph for this subject would be 20. The second type of average was the average amount of time spent on a particular type of screen as a percentage of the average time spent on all screens. Thus a subject who spent an average of 30 seconds on all screens, but only spent an average of 20 seconds on the table screens would have a value of 66.66% for average table percentage. Any amount over 100% means that they have spent more time than the average on a particular screen.

Hypothesis H3a predicted that if the subjects thought that declines in performance were caused by hardware problems, the average time spent on each screen would be different than that of those subjects who thought that task dissatisfaction was the reason for the declines in performance. All of the variables in tables 4a and 4b were tested, none of the F-values were significant for any of the variables for reason. There is no direct evidence to reject the null hypothesis. For a full copy of results for this section see Appendix C.

Hypothesis H3b predicted that if the subjects knew that they had to do a performance appraisal, the average time spent on each screen would be different for those subjects who had no appraisal knowledge. The null hypothesis was rejected for average graph (AVGRAPH) screens F-value = 8.41 ($p < .01$) (see table 14). Average graph was the amount of time spent on all the graph screens divided by the number of times the screens were accessed. The average time the subjects spent on the graph screens was 12.47 seconds if they had appraisal knowledge, as opposed to 18.11 seconds if they had appraisal knowledge (see table 15). The subjects spent longer on the screens if they had appraisal knowledge.

Table 14: Significant F Values for Average Times

VARIABLE	Appraisal Knowledge	Reason	Appraisal Knowledge * Reason
Average Graph (AVGRAPH)	8.41**		
Average Table Day Percentage (AVTDPER)			7.04**
Average Table Hour Percentage (AVTHPER)			6.88**
Average Table Percentage (AVTBPER)			6.85*

*p<.05 ** p<.01

Table 15: Means for Average Times (Appraisal Knowledge)

APPRAISAL KNOWLEDGE	No Appraisal Knowledge	Appraisal Knowledge
Average Graph (AVGRAPH)	12.47 (7.59)	18.11 (13.31)

Standard deviations in brackets

Although there was no direct evidence to show that there was any reason to reject the null hypothesis in hypothesis 3a. There was, however, a significant interaction between appraisal knowledge and reason for the following variables: Average table day percent (AVTDPER), average table percent (AVTBPER) and average table hour percent (AVTHPER). These interactions were not predicted in the hypotheses, so they may be chance occurrences. For average table day, and average table hour, however, they are highly significant. These interactions show that the percentage of time the subjects spent on the screens was higher if they thought that

Table 16: Means for Average Times (Interaction)

	No Appraisal Knowledge		Appraisal Knowledge	
VARIABLE	Hardware	Dissatisfaction	Hardware.	Dissatisfaction
Average Table Day Percentage (AVTDPER)	57.88 (37.66)	33.58 (28.33)	43.5 (34.54)	49.67 (27.03)
Average Table Hour Percentage (AVTHPER)	137.58 (57.58)	104.30 (65.70)	105.52 (72.96)	133.67 (64.53)
Average Table Percentage (AVTBPER)	111.35 (44.04)	84.10 (48.14)	84.46 (42.65)	99.84 (48.4)

Standard deviations in brackets

hardware was the problem, and they did not have appraisal knowledge (see table 16 and figures 5, 6 and 7 which show the interaction graphically). These variables measured the amount of time the subjects spent on the table day, table, and table hour screens respectively, as a percentage of the average amount of time spent on all screens. For the table hour screens the values were all over 100%. This shows that the subjects spent much more than the average on the table hour screens, no matter to which group they belonged.

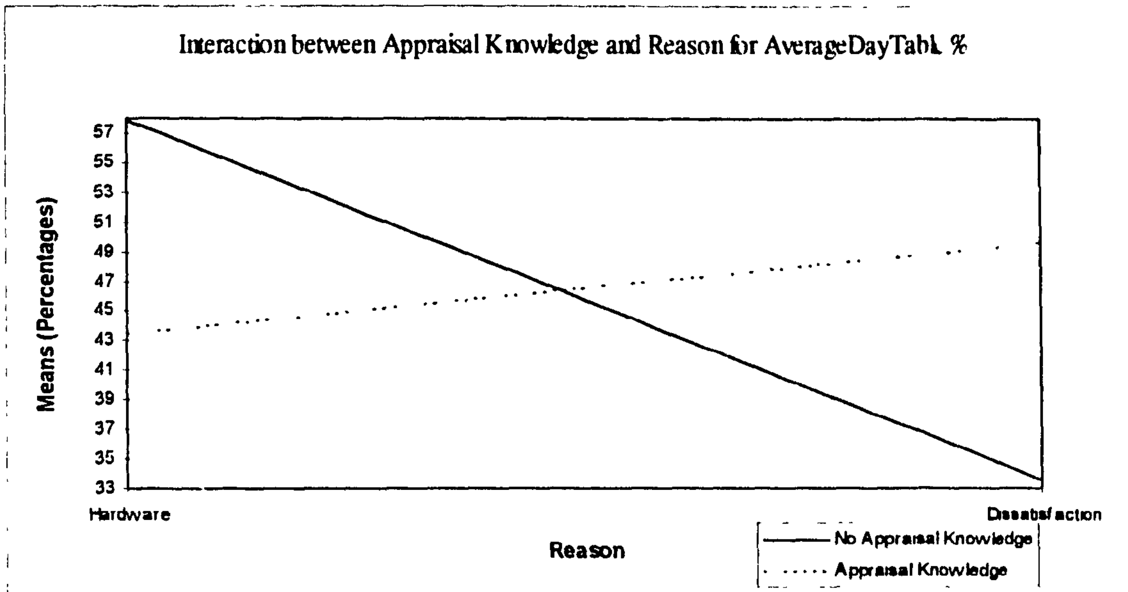


Fig 5: Interaction between Appraisal Knowledge and Reason for Average Day Table Percentage

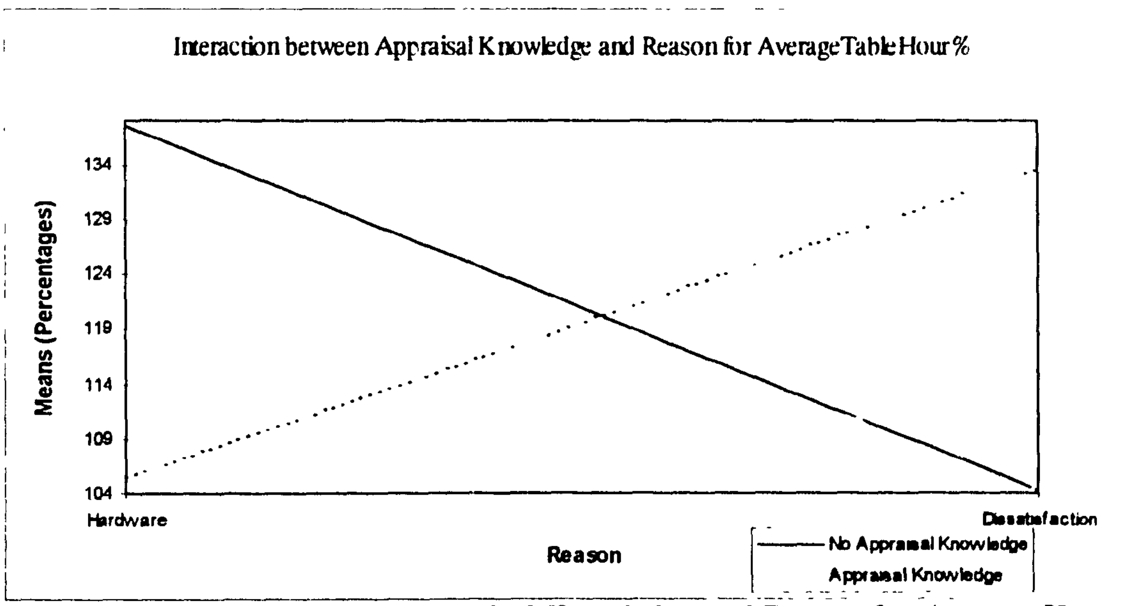


Fig 6: Interaction between Appraisal Knowledge and Reason for Average Hour Table Percentage

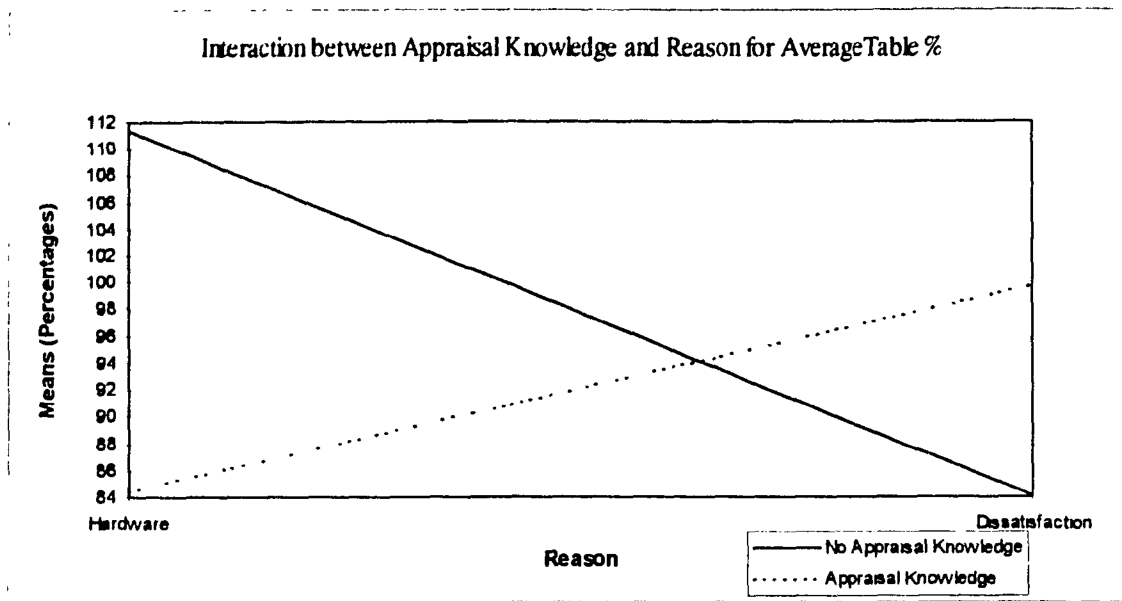


Fig 7: Interaction between Appraisal Knowledge and Reason for Average Table Percentage

8.4 Times and Frequencies

There are four different types of variables in this section. The first type is the number of times the subjects looked at the different graphs and table screens. These variables are different from the breadth of search variables. A subject can look at each screen any number of times. For example, if a subject looked at one graph screen 10 times, and another graph screen five times, and never looked at the other graph screens she would score 15 for graphs. (For the breadth of search she would have scored two because she only looked at two unique screens). The second type of variable is the number of different types of screens accessed as a percentage of the total number of screens accessed. For example, table day percentage is the number of table screens with data aggregated in a daily format, as a percentage of the total number of screens accessed. If the subject had visited 12 table day screens, and accessed a total of 50 screens in all he would have a value of .24 for table day (TDPER). The third type of variable is the total time the subject spent on each type of screen. The fourth type of variable is the amount of time the subject spent on the screen type divided by the total time spent on the experiment. For example, if the subject spent 100 seconds on graph day screens, and 900 seconds on the experiment in total, she would have a score of .11 for graph day time percentage (GDTPER). For a complete copy of the results see Appendix D.

Hypothesis H4a predicted that if the subjects thought that declines in performance were caused by hardware problems, the time spent on the different screens, and the frequency with which they accessed them would be different from that of the subjects who thought that task dissatisfaction was the reason for the declines in performance. All the variables shown on table 5a were tested, none of them had any significant F-values. There is, therefore, no evidence to reject the null hypothesis.

Hypothesis H4b predicted that if the subjects knew that they had to do a performance appraisal, the time spent on the different screens, and the frequency with which they accessed them, would be different from that of the subjects who had no appraisal knowledge. The null hypothesis was rejected in the case of table hour time percentage (THTPER) as the F-value is 4.42 ($p < .05$). There is, however, a highly significant interaction of appraisal knowledge and reason for this variable (see table 18b). The amount of time the subjects spent on the table hour time screen as a percentage of total time was greater if the subjects had appraisal knowledge (see table 17).

Table 17: Means for Times (Appraisal Knowledge)

APPRAISAL KNOWLEDGE	No Appraisal Knowledge	Appraisal Knowledge
Table Hours Time Percentage (THTPER)	0.3 (0.17)	0.24 (0.16)

Standard deviations in brackets

There was a significant interaction between appraisal knowledge and reason for declines in the following frequency variables (see table 18a), tables (TABLES), table

Table 18a: Significant F Values for Frequencies

VARIABLE	Appraisal Knowledge	Reason	Appraisal Knowledge * Reason
Tables (TABLES)			8.07**
Tables Day (TABDAY)			6.78*
Tables Hour (TABHOUR)			5.54*
Table Day Percentage (TDPER)			5.02*
Table Hour Percentage (THPER)			5.38*
Graph Day Percentage (GDPER)			7.68**

*p<.05 ** p<.01

day (TABDAY), table hour (TABHOUR), table day percent (TDPER), table hour percent (THPER), and the following time variables (see table 18b), table time (TABTIME), table day time (TDTIME), table hour time (THTIME), table day time percent (TDTPER), table hour time percent (THTPER). These interactions were not predicted in the hypotheses, so they may be chance occurrences. They were highly significant (p. < .01), however, for tables, graph day percentage, table time, table day time, table day time percentage, table hour time percentage and graphs day time percentage.

Table 18b: Significant F Values for Times

VARIABLE	Appraisal Knowledge	Reason	Appraisal Knowledge * Reason
Table Time (TABTIME)			9.31**
Graph Time (GRPHTIME)			5.37*
Table Day Time (TDTIME)			6.[illegible]9**
Table Hour Time (THTIME)			5.87*
Graph Day Time (GDTIME)			4.57*
Graph Hour Time (GHTIME)			4.27*
Table Day Time Percentage (TDTPER)			7.00**
Table Hour Time Percentage (THTPER)	4.42*		7.01**
Graphs Day Time Percentage (GDTPER)			9.05**
Graphs Hour Time Percentage (GHTPER)			6.05*

*p < .05 ** p < .01

Table 19a: Means for Frequencies (Interaction)

	No Appraisal	Knowledge	Appraisal	Knowledge
VARIABLE	**Hardware**	**Dissatisfaction**	**Hardware**	**Dissatisfaction**
Tables (TABLES)	14.80 (8.81)	10.64 (7.25)	8.72 (6.37)	12.67 (9.55)
Table Day (TABDAY)	7.40 (5.65)	5.13 (4.58)	4.28 (3.47)	6.63 (6.13)
Table Hour (TABHOUR)	7.40 (3.86)	5.52 (3.90)	4.44 (4.17)	6.03 (4.60)
Table Day Percentage (TDPER)	0.25 (0.17)	0.18 (0.16)	0.20 (0.21)	0.29 (0.20)
Table Hour Percentage (THPER)	0.24 (0.09)	0.21 (0.15)	0.16 (0.11)	0.23 (0.14)
Graphs Day Percentage (GDPER)	0.31 (0.17)	0.41 (0.23)	0.44 (0.16)	0.34 (0.22)

Standard deviations in brackets

Table 19b: Means for Times (Interaction)

	No Appraisal	Knowledge	Appraisal	Knowledge
VARIABLE	**Hardware**	**Dissatisfaction**	**Hardware**	**Dissatisfaction**
Table Time (TABTIME)	355.19 (200.64)	283.93 (303.89)	197.83 (130.2)	381.19 (277.29)
Graph Time (GRPHTIME)	165.35 (126.95)	222.41 (136.50)	286.24 (271.33)	192.37 (133.59)
Table Day Time (TDTIME)	129.78 (137.08)	81.36 (118.15)	75.31 (84.28)	138.72 (137.67)
Table Hour Time (THTIME)	225.41 (126.51)	202.57 (215.94)	122.53 (103.83)	242.47 (197.71)
Graph Day Time (GDTIME)	95.20 (74.15)	130.83 (69.76)	153.3 (117.67)	120.64 (83.06)
Graph Hour Time (GHTIME)	70.15 (64.18)	91.58 (90.07)	132.94 (167.93)	71.73 (83.83)
Table Day Time Percentage (TDTPER)	0.19 (0.17)	0.09 (0.10)	0.13 (0.17)	0.17 (0.13)
Table Hour Time Percentage (THTPER)	0.329 (0.127)	0.26 (0.20)	0.20 (0.15)	0.28 (0.16)
Graphs Day Time Percentage (GDTPER)	0.148 (0.094)	0.23 (0.15)	0.24 (0.13)	0.18 (0.14)
Graphs Hour Time Percentage (GHTPER)	0.11 (0.09)	0.14 (0.14)	0.17 (0.17)	0.09 (0.12)

The subjects spent more time on the table screens, or accessed the table screens more frequently if they did not have appraisal knowledge, and they thought hardware was the cause of performance decline, or if they had appraisal knowledge and they thought task dissatisfaction was the reason for the decline in performance (see table 19a).

Conversely, graph day percent (GDPER), graph time (GRPHTIME, graph day time (GDTIME), graph hour time (GHTIME), graphs day time percentage (GDTPER) and graph hour time percentage (GHTPER) were greater if the subjects had appraisal knowledge and they thought hardware was the cause of performance declines, or if they did not have appraisal knowledge but they thought that task dissatisfaction was the cause of the decline in performance (see tables 19a and 19b). These results are graphed in figures 8,9,10, and 11 following and in figures 12-22 in Appendix D.

The graphs that follow (Fig 8, 9, 10 and 11) are typical of the times and frequencies with which the subjects looked at the graphs and table screens. They show that the subjects looked at the graphs and tables screens differently depending upon appraisal knowledge, and whether they thought hardware or task dissatisfaction was the cause of performance decline.

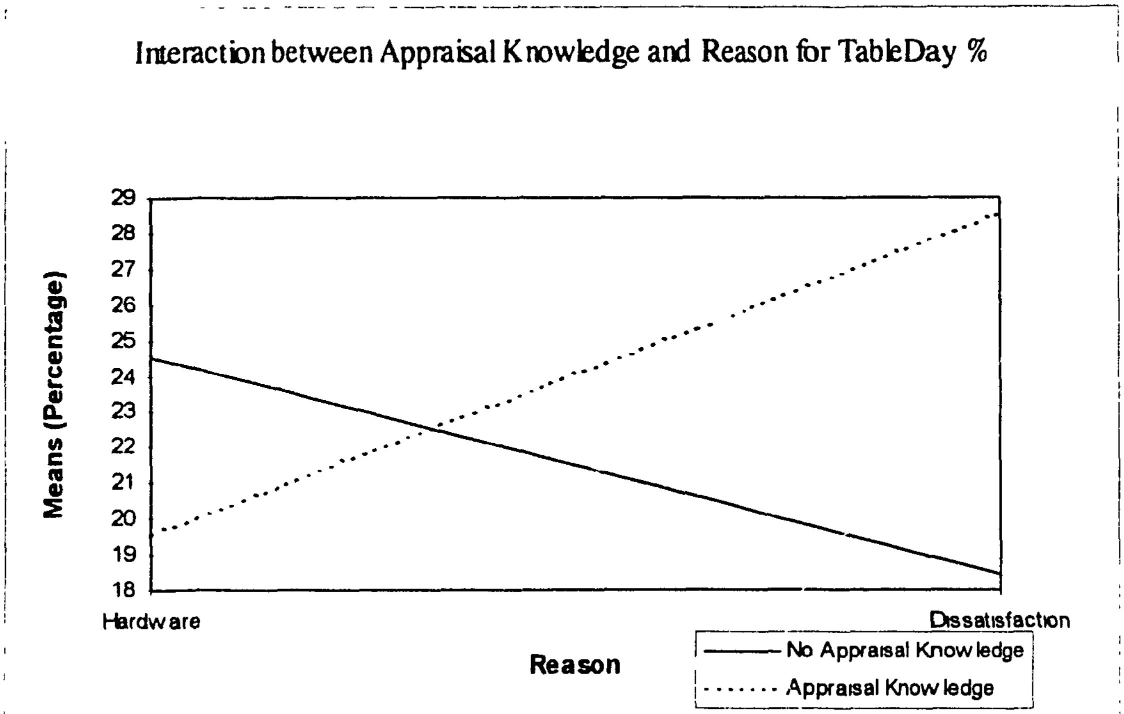


Fig 8: Interaction between Appraisal Knowledge and Reason for Table Day Percentage

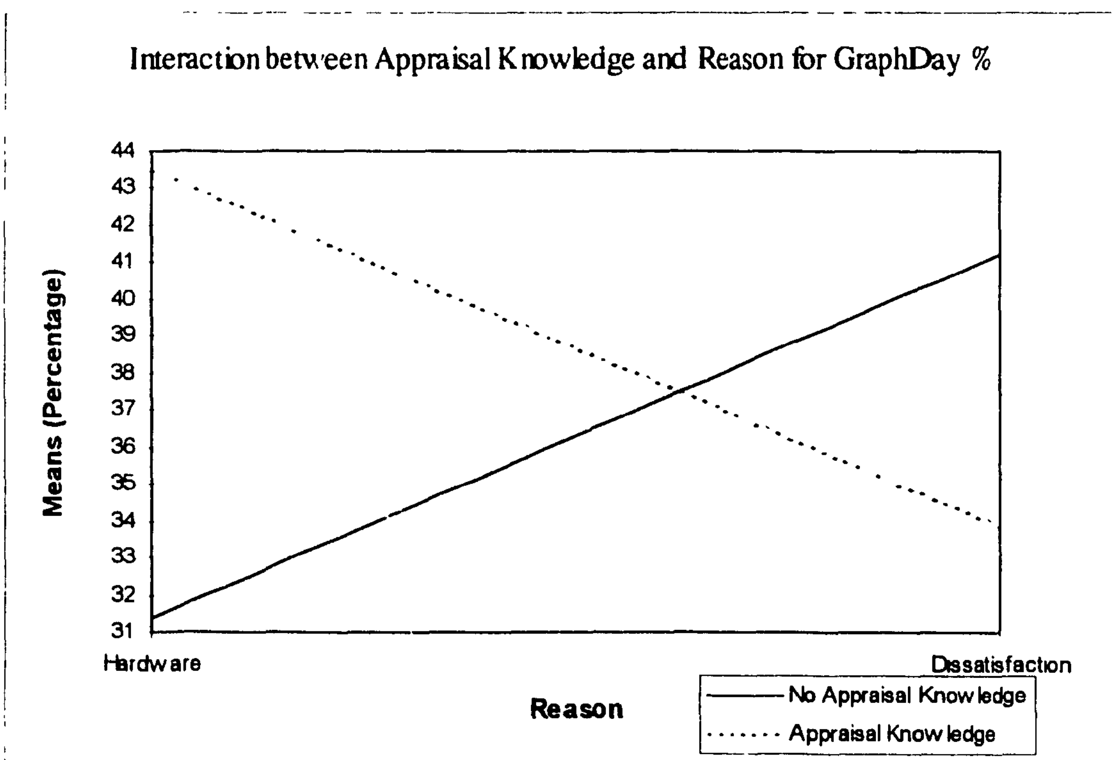


Fig 9: Interaction between Appraisal Knowledge and Reason for Graph Day %

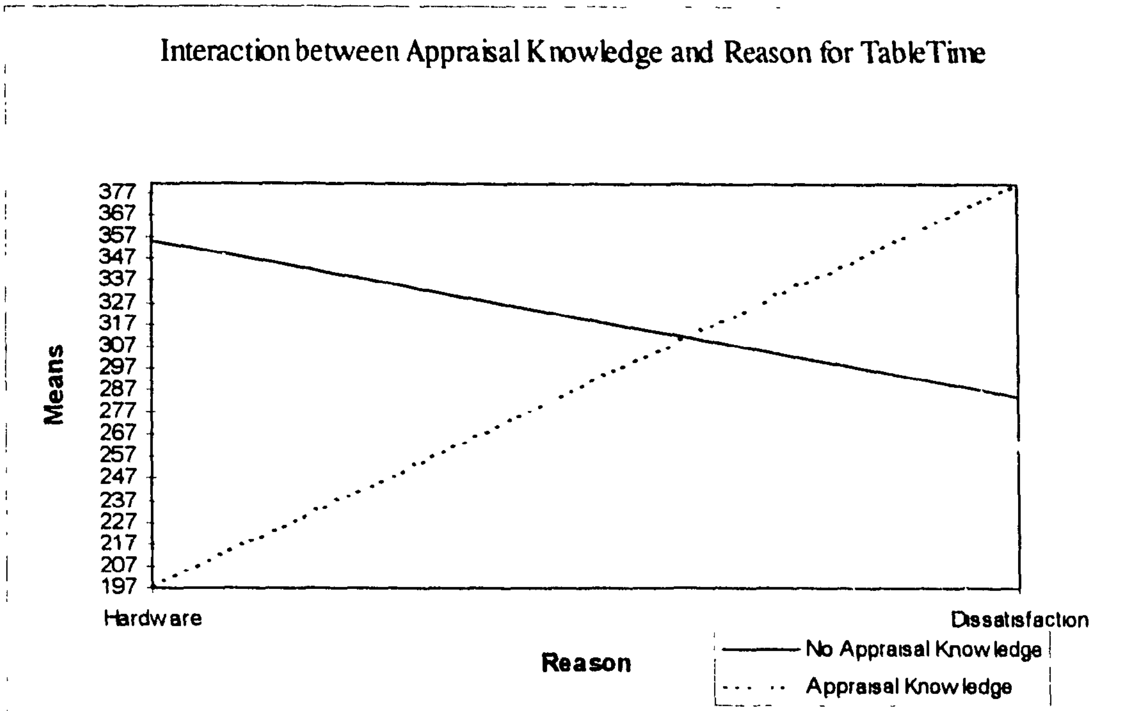


Fig 10: Interaction between Appraisal Knowledge and Reason for Table Time

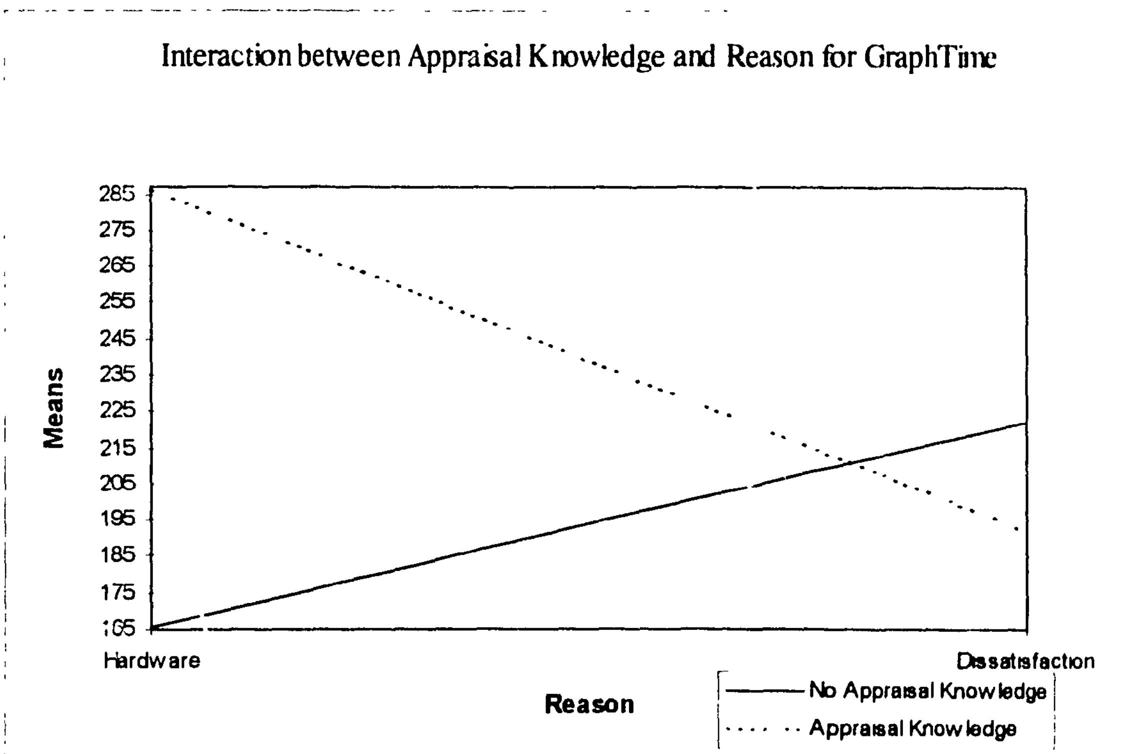


Fig 11: Interaction between Appraisal Knowledge and Reason for Graph Time

Chapter 9: Discussion of Results

9.1 Accuracy

The results show that when the subjects thought that hardware was the cause of the performance declines, they were more accurate in spotting the performance declines. This is true for all types of accuracy: easy, difficult and total accuracy. The knowledge that the employees were dissatisfied caused the subjects to be less accurate in spotting both types of declines, hourly and daily. In addition, for difficult accuracy, there was also a significant main effect for appraisal knowledge -- subjects who did not know they were expected to appraise performance were more accurate. For difficult accuracy, there was a strong interaction effect. This interaction effect showed that the subjects who had no appraisal knowledge and thought that the declines were caused by hardware were much more accurate in spotting the performance declines. This interaction seems to drive the two significant main effects. The knowledge of appraisal or employee dissatisfaction seems to interfere with objective decline spotting. These results are not what we expected to find. We expected that the subjects would be more accurate in spotting declines if they had appraisal knowledge for easy accuracy and total accuracy, but less accurate in spotting the hourly declines. The reason why we thought that they would be less accurate in spotting the hourly declines was because it was felt that they would disregard the hourly declines and try to concentrate on daily data that gave a better overall picture of the data. This is what could have happened here. The subjects in the dual purpose of appraisal condition have two objectives in searching the data. This could have caused them to be so preoccupied with the thought of having to do performance appraisals that they could not accurately spot the declines in performance. When people only have one thing on their minds at any particular time they are likely to perform better. Looking at the breadth of search, the subjects with appraisal knowledge searched less of the hourly screens. The overall time the subjects spent on the experiment was no different for all groups, so the appraisal knowledge groups were spending their time on the daily screens. Recall that Feldman (1981) said that supervisors must perform several subjective tasks before performance appraisals are possible. He said that the rater must recognize and attend to relevant information about employees and the information must be organized so that it can be stored and retrieved when necessary. Both these tasks must be done before the information is retrieved. The subjects in the dual monitoring condition could be spending time on organizing the data in memory, which prevented them from spending more time accessing all the screens. Or they could have just assumed that hourly data is less relevant to an assessment of performance. The work of Williams and his colleagues (Williams, 1984: Williams, et al. 1986) and Anderson and Pichart (1978) said that they found that it was the original purpose of the rating that established how the information was organized. The subjects in the dual monitoring section had a more 'noisy' environment in which to work, which could have caused them to be less accurate in spotting the performance declines.

The subjects who thought hardware was the cause of the performance decline could have done a better job searching the data and hence been more accurate because they did not want to penalize employees who declined in performance through no fault of their own. The activation of a schema can affect how the rater searches the information. Recall that DeNisi et al. (1984) said that a rater who has categorized an employee in terms of a "good employee" schema, based upon some previous salient observations may feel that he or she already knows that the employee is doing a good job, and does not feel that it is necessary to collect further information about that employee. The subjects in the dissatisfaction caused performance could have categorized the employees in terms of a "bad employee" schema, based upon the information they were primed with. This could have caused them to feel that it was not necessary to collect further information about that employee. They already 'knew' that the employees were doing a bad job, because they were dissatisfied employees. Those in the hardware caused performance declines would not have invoked such a schema. If we put these two effects together we can explain why the subjects were more accurate in spotting the daily and total declines in performance when they thought hardware was the cause of the performance declines, and why they were much more accurate in spotting the hourly declines when they thought hardware was the cause of the performance decline and they did not have appraisal knowledge. As the subjects were told that there would be six decline periods, we have no way of measuring if the subjects who thought that dissatisfaction was the cause of the performance declines would have spotted more declines if they were given an unlimited amount of decline sheets to complete.

9.2 Breadth of Search

When the subjects thought that hardware was the cause of performance declines, the subjects made a much broader search of the graph screens and more of the hour graph screens. In the no appraisal knowledge condition, the subjects searched more screens overall, more hour tables and more hour screens. In the no appraisal knowledge, hardware caused performance declines group, the subjects made a much more broad examination of the hour tables. This group searched 75% of the available screens, as opposed to the appraisal knowledge, dissatisfaction group who searched 58.89% of the available screens. This would account for their increased accuracy in spotting the hourly declines (difficult accuracy). The breadth of search was different than expected. In the hypothesis section we stated that the breadth of search would have been greater if the subjects had appraisal knowledge and they thought that dissatisfaction was the cause of the performance decline. This can be explained by the subjects spending more time on the other screens trying to store the information in memory if they knew that they would have to do a performance appraisal after monitoring the data. The search by the subjects in the hardware caused performance decline, and those who did not have appraisal knowledge was more controlled. They seemed to know what they were looking for, and their search strategy was more efficient. They were not wasting time trying to find consensus information, or trying to find negative information about the employees to be used on the performance appraisal. Recall what Feldman (1981) said about controlled processing. Controlled processing takes place when the rater encounters information about the ratee which is inconsistent with the view the rater holds of the ratee, or the rater does not feel that the information is relevant to the appraisal decision. Controlled processing requires active concentration on the part of the rater. The subjects who thought that hardware was a cause of decline in performance could have discounted hardware information as not being relevant to the appraisal decision, thus invoking controlled processing. They would then search in a more controlled fashion, and thus more efficiently and effectively.

9.3 Averages

The average time the subjects spent on the graph screens was longer if they had appraisal knowledge. It seems that subjects with appraisal knowledge looked longer at fewer screens. This did not serve then well. The reason they spent longer on the screens could have been because they were attempting to store the information thinking that they would have to retrieve it to perform a performance appraisal after the experiment was over. The subjects who had no appraisal knowledge and thought that hardware caused declines spent much more time than average on the table hour screens. These are the screens where the subjects could accurately spot the hourly performance declines. However, subjects with appraisal knowledge who thought dissatisfaction was the cause of the performance declines also spent much more than average time on the table hour screens. But the hourly declines were not spotted as accurately by this group. This again could be caused by the subjects attempt to store relevant information, thinking that they would have to retrieve it when they were required to do a performance appraisal. These results were consistent with what we expected at the beginning of the experiment. We expected subjects who had appraisal knowledge to spend more time on the screens due to the subjects spending more time in trying to store the information if they knew they had to do a performance appraisal. If we look back to the breadth of search section, the number of unique hour tables accessed was greater for the no appraisal knowledge, hardware group. This shows that the breadth of search was the most important factor in being able to spot the decline periods more accurately.

9. 4 Times and Frequencies

The people who scrutinized table days and table hour screens more than they did other screens were in the hardware, no appraisal knowledge group. The percentage of time they spent on the table day and table hour screens as a percentage of the total time was greater than the other groups. The frequency that the hardware, no appraisal knowledge group accessed the table hour screens, as a percentage of total screens accessed was greater than other groups. The time they spent on the table hour screens was much more than the time they spent on the table day screens. The people who had appraisal knowledge and thought dissatisfaction caused performance declines spent the most time on the table hour screens. However, their difficult decline appraisal accuracy was not as good as that of subjects who were in the no appraisal knowledge, hardware group. Again it is not the amount of time that they spent upon the screens, but the number of unique hourly table screens accessed that is the main factor in the accuracy of spotting the performance declines.

It was interesting to note that the people who spent the most time on the graphs, both in graph day time and graph hour time were in the appraisal knowledge, hardware group. This group was not significantly better at accurately spotting performance declines. In the case of graph hour time, the groups that spent more time on these screens were the ones that did not spot the hourly declines accurately. This would suggest that the use of graphs does not help the supervisors in this case. Using graphs however, did help with spotting easy accuracy decline periods. This would suggest that a combination of graphs for easy accuracy and tables for difficult accuracy would be best. The combination of graphic and tabular data also allows users to extract information from the presentation form they are most comfortable with, and then further enhance their understanding with the other presentation mode. This is partially consistent with the findings of Powers, Lashley, Sanchez and Shneiderman (1984) who found that if the goal was to speed up performance then the combination of tabular and graphic data should be avoided (their results showed that tabular data would be most appropriate), but they found if the goal was to increase the accuracy of performance, then the combination of graphics and tabular data appears to be most helpful to that goal.

The dissatisfaction-caused declines, appraisal knowledge group spent the second largest amount of time on, and made the second largest amount of screen accesses of hour table screens yet they did not perform nearly as accurately as the hardware-caused declines, no appraisal group. This could have been because they were trying to find consensus information. If they thought they had to do performance appraisals on the employees, they could have been searching more hourly graphs, in an attempt to make comparisons (Jarvenpaa & Dickson, 1988). However, in doing so they were not able to spot the hourly declines in performance correctly. The subjects in the appraisal knowledge/ dissatisfaction group could also have been trying to collect information that would help them classify target persons into "good" or "poor" categories, for example, even when this information would contradict their initial expectations. This is consistent with authors who have discovered that raters tend to rely more upon negative information when decision making. (Trope & Basok, 1982; Hamilton & Huffman, 1971). Because they thought that they would have to do a performance appraisal after rating the employees, they could have been trying to find some information that would help them classify the employees. This does not explain, however, their poor performance in spotting the hour decline periods.

Chapter 10: Limitations of the Study

Although the results of this study show that appraisal expectation and information about factors affecting performance do have an impact on supervisory information search and evaluation accuracy, there are some limitations that should be noted. This research consisted of a laboratory experiment carried out using graduate and undergraduate students. The effects of computerized performance monitoring systems on supervisors in an organization may be very different than was found here. There was no information given about how much work experience the subjects had. If the students had very little work experience, and hence even less supervisory experience, which is highly likely considering they were mainly undergraduate students, the external validity of the study would be affected. Supervisors in an organization might look at computerized performance monitoring data differently. They might be able to discount the slight hourly declines in performance as being normal. They also might be more aware of the decline in performance caused by badly written time cards and other situational problems.

Another possible effect on the external validity of the experiment is that the subjects knew in advance about the number of decline periods, and the two patterns of decline. This kind of foreknowledge does not occur in the real world. However, supervisors might know from experience to look for two different patterns of decline periods. They might also have historic data to look back on if they were relatively new to the task, so they might be able to estimate how many declines to expect. The reason why the subjects were told in advance about the number of decline periods was in order to make the experimental task manageable. Because of the subjects foreknowledge of these decline periods, we have no way of knowing if the subjects in the dissatisfaction caused performance decline cells would have spotted more decline periods than those who thought hardware was the cause of the performance declines.

When the subjects were questioned on the post experiment questionnaire, 56% of the 127 subjects were able to accurately identify the reason for the decline in performance. This figure is very low. There could be several reasons for this. One of the reasons could have been because they were tired after doing the experiment, and the questionnaire answers are suspect. The results show that there were differences in the way the subjects in the various cells looked at the data, even if they could not consciously remember the reason for the performance decline. If we had obtained no significant results there would be reason to believe that the subjects were not sufficiently primed with the information. However, the results clearly show that there was a difference in the way the data was searched in the various cases. If the subjects had been informed of the reason for the possible decline in performance more often, the results would possibly have been even more significant. When the subjects were asked on the questionnaire if they had noticed a decline period in common, seventy-five percent of the subjects who said that they had not noticed a decline period in common had, either fully or partially, correctly identified the decline period in common on the decline sheet. This shows that the questionnaire data may have been filled in by subjects who were tired. The low percentage of subjects who were accurately able to identify the reason for the decline they were primed with should have been higher When Feldman (1981) talked about automatic and controlled processing, he said that people learn to attend to certain stimulus features without consciously monitoring the process. The subjects in this study could have attended to the stimuli of the employees being dissatisfied without being fully conscious of it, but they reacted to it anyway.

The fact that there was one decline period in common could suggest to the subjects that an external factor such as hardware was really the problem and not a given individual's dissatisfaction. This could have had an impact on internal validity, but when the replies on the questionnaire were examined, not one of the subjects in the dissatisfaction caused performance decline group said that they thought hardware was a cause for decline in performance. When the subjects in the dissatisfaction caused performance decline cell were asked on the questionnaire if they had noticed the decline period in common over 60% said they had, but none attributed it to hardware problems. The number of subjects who accurately spotted the decline period in common was higher than those who said they had noticed the common decline period. This shows that the decline period in common did not have any effect on the internal validity of the study.

Kulik and Ambrose (1993) found in their study that visual data influenced the way the raters rated the ratees, and that data from performance monitors plays a limited role in appraisals. In a work setting the supervisors will have more chance of interaction with the employees to obtain visual data, and will have more knowledge of their prior performance.

Prior performance knowledge has been shown by Fenner et al. (1993) to be an important factor in performance appraisals, but they also cautioned that a laboratory study manipulating the prior performance information would likely produce less persistent performance expectations than interactions with and observations of subordinates in a work setting. Supervisors in the workplace have to deal with competing tasks and consider a broader range of data than was available in the study (Fenner et al., 1993). This is also true in the study discussed in this thesis. Supervisors may be aware of hardware breakdowns, and may not attribute all declines in performance as being caused by employees' dissatisfaction.

Supervisors also have more time contraints than the subjects in this study. Supervisors have already been shown to be spending more time on evaluating employees monitored by computers than they had been prior to the introduction of the monitors (Aiello, 1993; Chalykoff & Kochan, 1989). The subjects in this experiment had no such time constraints, which could cause them to look at the data differently.

As in any laboratory study, the effects of computerized performance monitoring systems on supervisors in an organization may be very different than was found here. Supervisors in most cases have more chance of interaction with their employees. There are, however, some employees that are monitored by supervisors at remote sites who do not have much contact with the supervisors (Aiello, 1993). These are the employees where the supervisors may have incorrect information about the employees, which could affect the way they look at computerized performance monitoring data. Supervisors will have had more experience in looking over the data, however, and they should be more aware of hardware and software problems, and other external attributes that can cause declines in performance.

Chapter 11: Summary and Conclusions

This research shows that there is a difference in the way supervisors search computerized performance monitoring data depending upon what prior information they have and whether they thought they would have to do a performance appraisal or not. The level of accuracy in spotting the performance declines is different depending upon the rater's preconceived notions. They were much more accurate if they thought hardware was the cause of the performance decline.

The more we understand about computerized performance monitoring systems and the way supervisors search the data, the better we will be able to design computerized performance monitoring systems to eliminate effects of preconceived notions and other negative effects. These results show that there is a difference in the way raters search computerized performance monitoring data depending upon what prior information they had and if they know that they may have to do a performance appraisal using the data. Computerized performance monitors should be designed so that they record when there has been a "down-time" in the system. The instructions given to every supervisor who monitors computerized performance monitoring systems data should make them aware of the possibility of computer failure.

Irving, Higgins, and Safayeni (1986) said:

> "A computerized performance monitoring system may provide an increase in the amount of quantitative feedback available to all levels of the organization. This quantitative information may underlie the general perception that these systems increase the fairness of the performance evaluation system and increase the organization's ability to judge performance. It seems that increasing the quantitative information available to management tends to minimize the emphasis given to subjective information in performance evaluations.

Reducing the emphasis on subjective measures may improve performance evaluations only to the extent that job performance can realistically be measured using numerical counts. On the other hand, many aspects of an individual's work may not lend themselves well to numerical measurement" (p. 800).

If the supervisors are not aware of hardware failure, they may look at the data from computerized performance monitoring systems less objectively, than they would if they thought that hardware was the cause of the decline in performance. One of the strengths of computerized performance monitoring systems is that they minimize the emphasis given to subjective information in evaluating performance. Without knowledge of hardware problems, the data from computerized performance monitoring systems is looked at more subjectively.

Kulik and Ambrose (1993)'s research showed that supervisors are able to factor in both visual and computerized performance monitoring systems data when evaluating performance despite the evaluative implications of the computerized data. Kulik and Ambrose (1993)'s results contradict the documented concerns of employees that are monitored by computerized performance monitoring systems (Grant et al., 1988), who feel that the objective character of the computerized data may lead supervisors to put more weight on the quantitative data when appraising performance. This research shows that if supervisors are not made aware of the possibility of hardware breakdowns they may make more subjective judgments, especially if they think the employees are dissatisfied.

The knowledge that they will have to do performance appraisals using the same data they are monitoring does not serve the supervisors well. Supervisors should look at computerized performance monitoring data for one purpose only at any one time. In this experiment we had subjects looking at the monitoring data with a dual purpose, one of which was monitoring the data as it happens and the other was the knowledge that they would have to do a performance appraisal after monitoring the data. This should never be allowed to happen. The subjects who had appraisal knowledge, were less accurate in using the data for the main purpose in which it was intended. One of the uses of computerized performance monitoring systems is to spot temporary declines in performance, and correct these performance declines as they occur. They should be used for the main purpose for which they are designed. The subjects we used here were not supervisors, so the effect of using actual supervisors may be different. But the subjects in this experiment were not looking at the data under severe time pressures, which the supervisors are under, according to Chalykoff and Kochan (1989). There is documented work that says that people under severe time pressures and distraction tend to look for more negative information (DeNisi et al., 1984; Staelin and Payne, 1976). The subjects here were looking for negative information, and they were not under the same time pressures that actual supervisors would be under.

As stated before, this research was not looking at the performance appraisal accuracy, just the effect on the accuracy of the day-to-day monitoring of the employees if the supervisors have appraisal knowledge. The effect on the performance appraisal may be completely different if supervisors are looking at computerized performance monitoring systems data with the sole purpose of doing a performance appraisal. Kulik and Ambrose (1993) and Fenner, Lerch, and Kulik (1993) found that there are ways in which the data from computerized performance monitoring systems can be used to help in the performance appraisal process This research just shows that the performance appraisal should not take place at the same time as the monitoring of the system to find temporary declines.

Murphy, Garcia, Kerkar, Martin and Balzer, (1982) showed in their study that their results suggest that accuracy in observation is related to accuracy in performance evaluations, but they said

> "It is hazardous to make generalizations from a laboratory study such as ours to performance appraisals in real organizations. Our data provide information on the strength and nature of the relationship between accuracy in observing behavior and accuracy in evaluating performance, but a number of questions remain unanswered, ..." (p. 324).

This work shows that the supervisors should be made aware of the possibility of problems with the computer system, and that they should not be made to use the data from computerized performance monitoring systems for more than one purpose at a time. However, as Murphy at al. said, this was a laboratory study, and the results should be treated as such. What might be found in real organizations may be different, given the limitations of the study.

There is still much more research needed in the field of computerized performance monitoring systems, and how they impact performance appraisals. The findings of this research provide a step in understanding the impact of appraisal expectation on supervisory information search and evaluation accuracy.

BIBLIOGRAPHY

Aiello, J.R, (1993) Computer-Based Work Monitoring: Electronic Surveillance and Its Effects. Journal of Applied Social Psychology, 23, 499-507.

Aiello, J.R., & Svec, C.M., (1993). Computer Monitoring of Work Performance: Extending the Social Facilitation Framework to Electronic Presence. Journal of Applied Social Psychology, 23, 537-548.

Ambrose, M.L., Kulik, C.T., (1994). The Effect of Information Format and Performance Pattern on Performance Appraisal Judgments in a Computerized Performance Monitoring Context. Journal of Applied Social Psychology, 24, 801-823.

Anderson, N.H., (1974). Cognitive Algebra: Integration Theory Applied to Social Attribution. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Vol. 7. New York: Academic Press.

Anderson, N.H., (1976). How Functional Measurement Can Yield Validated Interval Scales of Mental Quantities. Journal of Applied Psychology, 61, 677-692.

Anderson R.C., & Pichart, J.W., (1978). Recall of Previously Unrecallable Information Following a Shift in Perspective. Journal of Verbal Learning and Verbal Behavior, 2, 379-391.

Angel, R.N. (1989). Evaluating Employees by Computer. Personnel Administrator, 34, 67-72.

Attewell, P., (Spring, 1987). Big Brother and the Sweat-Shop: Computer Surveillance in the Automated Office. Sociological Theory, 5, 87-99.

Austin, W., & Drake, L. C., Jr. (1985). Office Automation. Occupational Outlook Quarterly, 29, 16-19.

Ballou, D.J., (1996). Integrating Electronic Monitoring Information into Performance Appraisals: How are Evaluative Impressions Formed and Attributions Made? Working Document.

Behling, V.W., Gifford, W.E., & Tolliver, J.M., (1980). Effects of Grouping Information on Decision Making Under Risk. Decision Sciences, 11, 272-283.

Bem, D.J. & Funder, D.C. (1978). Predicting More of the People More of the Time: Assessing the Personality of Situations. Psychology Review, 85, 485-501.

Berry, D.C., & Broadbent, D.E., (1986). Human Search Procedures and the Use of Expert Systems. Current Psychological Research & Reviews, 130-147

Bettman, J.R., Johnson E.J., & Payne, J.W., (1990). A Componential Analysis of Cognitive Effort in Choice. Organizational Behavior and Human Decision Processes, 45, 111-139.

Borman, W.C., (1977). Consistency of Rating Accuracy and Rating Errors in the Judgment of Human Performance. Organizational Behavior and Human Performance, 20, 238-252.

Bruner, J.S., (1958). Social Psychology and Perception. In E.E. Maccoby, T.M. Newcomb, & E.L. Hartley (Eds.) Readings In Social Psychology. New York: Holt, Rinehart, & Winston.

Bylinsky, G., (1991, November). How Companies Spy on Employees. Fortune, 131-140.

Cafferty, T.P., DeNisi, A.S., & Williams, K.J., (1986). Search and Retrieval Patterns for Performance Information: Effects of Evaluations of Multiple Targets. Journal of Personality and Social Psychology, 50, 676-683.

Cantor, J. H., Individual Needs and Salient Constructs in Interpersonal Perception. (1976). Journal of Personality and Social Psychology, 35, 38-48.

Cantor N. & Mischel, W., (1979). Prototypes in Person Perception. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Vol. 12. New York: Academic Press.

Chaiken, S., (1980). Heuristic vs. Systematic Information Processing and the Use of Source vs. Message Cues in Persuasion. Journal of Personality and Social Psychology, 39, 752-766.

Chalykoff, J., & Kochan, T. A. (1989). Computer-Aided Monitoring: Its Influence on Employee Satisfaction and Turnover. Personnel Psychology, 40, 807-834.

Cohen, C.E., (1979). Prototypes: Some Properties of Person Categories. Paper Presented at the Symposium on Categorical Processing and Representation of Person Information, Annual Convention of the American Psychological Association, New York, September, 1979.

Cohen, C. E., (1981). Person Categories and Social Perception: Testing Some Boundaries of the Processing Effects of Prior Knowledge. Journal of Personality and Social Psychology, 40, 441-452.

Cooper, W.H., (1981). Ubiquitous Halo. Psychological Bulletin, 90, 218-244.

Crocker, J., (1981). Judgment of Covariation by Social Perceivers. Psychological Bulletin, 90, 272-292.

Crockett, W.H., Mahood, S., & Press, A.N., (1975). Impressions of a Speaker as a Function of Variations in the Cognitive Characteristics of the Perceiver and the Message. Journal of Personality, 43, 168-178

Danaan, S., Stories of Mistrust and Manipulation: The Electronic Monitoring of the American Workforce, 9 to 5 Working Women Education Fund, Cleveland, 1990.

DeNisi, A.S., Cafferty, T.P., & Meglino B.M., (1984). A Cognitive View of the Performance Appraisal Process: A Model and Research Propositions. Organizational Behavior and Human Performance, 33, 360-396.

DeNisi, A.S., Cafferty, T.P., Williams, K.J., Blencoe, A.G., & Meglino, B.M., (1983). Rater Information Acquisition Strategies: Two Preliminary Experiments. Proceeding of the Academy of Management, 169-172

DeNisi, A.S., & Williams, K.J., (1988). Cognitive Approaches to Performance Appraisal. Personnel and Human Resources Management, 6, 109-155.

Ebbesen, E.B., (1981). Cognitive Processes in Inferences about a Person's Personality. In F. Higgins, C. Herman, & M. Zanna (Eds.), Social Cognition:The Ontario Symposium (Vol. 1), Hillsdale, NJ, Erlbaum.

Eisenman, E., (1986). Employee Perceptions and Supervisory Behaviors in Clerical VDT Work on Systems that Allow Electronic Monitoring. OTA-CIT-333, U.S.

Government Printing Office, Washington, D.C., Office of Technology Assessment. (NTIS No. PB88-156351).

Fenner, D.J., Lerch, F.J., & Kulik, C.T., (1993). The Impact of Computerized Performance Monitoring and Prior Performance Knowledge on Performance Evaluation. Journal of Applied Social Psychology, 23, 573-601.

Feldman, J. M., (1981). Beyond Attribution Theory: Cognitive Processes in Performance Appraisal. Journal of Applied Psychology, 66, 127-148.

Fiske, S.T., & Pavelchak, M.A., (1986). Category-based versus piece-meal-based affective responses: Developments in schema-triggered affect. In R.M. Sorrentino & E.T. Higgins (Eds.), Handbook of Motivation and Cognition; Foundations of Social Behavior (pp. 167-203). New York: Guilford Press.

Garland, H., Hardy, A. & Stephenson, L. (1975). Information Search as Affected by Attribution Type and Response Category. Personality and Social Psychology Bulletin, 4, 612-615

Garson, B., (1988) The Electronic Sweatshop: How Computers are Transforming the Office of the Future Into the Factory of the Past. Simon & Schuster, New York.

Gerlach, J.H., Kuo, F-Y, (1991). Understanding Human- Computer Interaction for Information Systems Design. MIS Quarterly, 527-549.

Grant, R.A., & Higgins, C.A., (1991). The Impact of Computerized Performance Monitoring on Service Work: Testing a Causal Model. Information Systems Research, 2, 116-142.

Grant, R.A., & Higgins, C.A., & Irving, R.H., (1988 Spring). Computerized Performance Monitoring Systems: Are they Costing You Customers? Sloan Management Review,29, 39-45.

Griffith, T.L., (1993). Monitoring and Performance: A Comparison of Computer and Supervisor Monitoring. Journal of Applied Social Psychology, 23, 549-572.

Halpern, S. (1992, May). Big Boss is Watching You. Details 18-23.

Hansen, R.D., (1980). Commonsense Attribution. Journal of Personality and Social Psychology, 39, 996-1009.

Hamilton, D.L., & Huffman, L.J., (1971). Generality of Impression Formation for Evaluative and Non-evaluative Judgments. Journal of Personality and Social Psychology, 20, 200-207.

Harvey, R.J., (1982). The Future of Partial Correlation as a Means to Reduce Halo in Performance Ratings. Journal of Applied Psychology, 67, 171-176.

Henriques, V.E. (1986, July 14). Computer Monitoring: Boon to Employee and Manager? ComputerWorld, 20, 17.

Hulin, C.L., (1982). Some Reflections on General Performance Dimensions and Halo Rating Error. Journal of Applied Psychology,67, 165-170.

Ilgen, D.R., & Feldman, J.M. (1983). Performance Appraisal: A Process Focus. Research In Organizational Behavior, 5, 141-197.

Irving, R.H., Higgins, C.A., & Safayeni, F.R. (1986). Computerized Performance Monitoring Systems: Use and Abuse. Communications of the ACM, 29, 794-801.

Jarvenpaa, S.L., & Dickson, G.W., (1988). Graphics and Managerial Decision Making: Research Based Guidelines. Communications of the ACM, 31, 764-773.

Jeffrey K.M., & Mischel, W., (1979). Effects of Purpose on the Organization and Recall of Information in Person Perception. Journal of Personality, 47, 297-419.

Johnson, E.J., Payne, J.W., & Bettman, J.R., (1988). Information Displays and Preference Reversals. Organizational Behavior and Human Decision Processes, 42, 1-21.

Kelley, H.H., (1967). Attribution Theory in Social Psychology. In D. Levine (Ed.), Nebraska Symposium on Motivation. Lincoln: Univ. of Nebraska Press.

Kelley, H.H., (1972). Causal Schemata and the Attribution Process. In E.E. Jones, D.E. Kanouse., H.H. Kelley, R.E. Nisbett, S. Valins & B. Weiner (eds). Attribution: Perceiving the Causes of Behavior. Morristown. NJ: General Learning Press.

Kelley, H.H., (1973). The Process of Causal Attribution. American Psychologist, 23. 107-128.

Kerstholt, J.H., (1992). Information Search and Choice Accuracy as a Function of Task Complexity and Task Structure. Acta Psychologica, 80, 185-197.

Kimble, G.A., & Perlmuter, L.C., 1970. The Problem of Volition. Psychological Review, 77, 361-384.

Kozlowski, S.W.J., & Ford, J. K., (1991). Rater Information Acquisition Processes: Tracing the Effects of Prior Knowledge, Performance Level, Search Constraint, and Memory Demand. Organizational Behavior and Human Decision Processes, 49, 282-301.

Kraut, R., Dumais, S., & Koch, S., (1989). Computerization, Productivity, and Quality of Work-Life. Communications of the ACM, 32, 220-238.

Kulik, C.T., & Ambrose, M.L., (1993). Category-Based and Feature-Based Processes in Performance Appraisal: Integrating Visual and Computerized Sources of Performance Data. Journal of Applied Psychology, 78, 821-830.

Lalljee, M., Lamb, R., Furnham, A., & Jaspars, J., (1984). Explanations and Information Search: Inductive and Hypothesis-Testing Approaches to Arriving at an Explanation. Journal of Social Psychology, 23, 201-212.

Langer, E.J., (1978). Rethinking the Role of Thought in Social Interaction. In J.H. Harvey et al. (Eds.), New Directions in Attribution Research, Vol.2. Hillsdale, N.J.: Erlbaum.

Landy, F.J., & Farr, J.L.,(1980). Performance Rating. Psychological Bulletin, 87, 72-107.

Larson, J.R., (1984). The Performance Feedback Process: A Preliminary Model. Organizational Behavior and Human Decision Processes, 33, 42-76.

Lawler, E. III., & Rhode, J., (1976). Information and Controls in Organizations. Goodyear Publishing Co., Pacific Palisades, CA.

Lord, C.G., (1980). Schemas and Images as Memory Aids: Two Modes of Processing Social Information. Journal of Personality and Social Psychology, 38, 257-269.

Maheswaren, D., & Chaiken, S., (1991). Promoting Systematic Processing in Low-motivation Settings: Effect of Incongruent Information on Processing and Judgment. Journal of Personality and Social Psychology, 61, 13-25.

Major, B., (1980). Information Acquisition and Attribution Processes. Journal of Personality and Social Psychology, 39, 1010-1023.

Marcus, H., (1977). Self-schemata and Processing Information About the Self. Journal of Personality and Social Psychology, 35, 63-78.

Marx, G. & Sherizen, S. (1987). Social Aspects of Changes in Worker Monitoring and Computer/Communications Privacy and Security Practices, (Revised)(1986) in U.S. Congress, Office and Technology Assessment, The Electronic Supervisor: New Technology, New Tensions - Contractor Documents. Vol 2 (PB88-156369), U.S. Department of Commerce, NTIS, Springfield, V.A. December, 1987, 120-310.

Murphy, K.R., (1982). Difficulties in the Statistical Control of Halo. Journal of Applied Psychology, 67, 161-164.

Murphy, K.R., Garcia, M., Kerkar, S., Martin, C., & Balzer, W. K., (1982). Relationship Between Observational Accuracy and Accuracy in Evaluating Performance. Journal of Applied Psychology, 67, 320-325.

Murphy, K.R., Gannett, B.A., Herr, B.M., & Chen, J.A., (1986). Effects of Subsequent Performance on Evaluations of Previous Performance. Journal of Applied Psychology, 71, 427-431.

Nebeker, D.M., Tatum, B.C., (1993). The Effects of Computer Monitoring, Standards, and Rewards on Work Performance, Job Satisfaction, and Stress. Journal of Applied Social Psychology, 23, 508-536.

Nisbett, R.E., & Wilson, T.D., (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. Psychological Review, 84, 231-259.

Nussbaum, K. (1984). The 9 to 5 National Survey of Women and Stress: Office Automation Addendum, 9 to 5 National Association of Women, Cleveland, 1984.

Olson, M. & Lucas, H. Jr., (1982). The Impact of Office Automation on the Organization: Some Implications for Research & Practice. Communications of the ACM, 25, 838-847.

Payne, J.W., (1976). Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. Organizational Behavior and Human Performance, 16, 366-387.

Payne, J.W., Braunstein, M.L., & Carroll, J. S., (1978). Exploring Predecisional Behavior: An Alternative Approach to Decision Research. Organizational Behavior and Human Performance, 22, 17-44.

Press A.N., Crockett, W.H., & Delia , J.G., (1975). Effects of Cognitive Complexity and Perceiver's Set upon the Organization of Impressions. Journal of Personality and Social Psychology, 32, 865-872.

Posner, M.I., & Snyder, C.R., (1975). Attention and Cognitive Control. In R.L. Solso (ed.), Information Processing and Cognition: The Loyola Symposium. Hillsdale. NJ: Erlbaum.

Powers, M., Lashley, C., Sanchez, P., & Shneiderman, B., (1984), An Experimental Comparison of Tabular and Graphic Data Presentation. International Journal of Man-Machine Studies, 20, 545-566.

Remus, W. (1987). A Study of Graphical and Tabular Displays and Their Interaction with Environmental Complexity. Management Science, 33, 1200-1205

Remus, W. (1984). An Empirical Investigation of the Impact of Graphical and Tabular Data Presentations on Decision Making. Management Science, 30, 533-542.

Rosch, E., (1977). Human Categorization. In N. Warren (Ed.), Studies in Cross-Cultural Psychology, Vol. 1. New York: Academic Press.

Rosch, E., Mervis, C.G., Gray, W.D., Johnson D.M., & Boyes-Braem, P., (1976). Basic Objects in Natural Categories. Cognitive Psychology, 8, 382-439.

Ross, L., (1977). The Intuitive Psychologist and his Shortcomings: Distortions in the Attribution Process. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Vol. 10. New York: Academic Press.

Ross, L. Lepper, M. R., Strack, F., & Steinmetz J. (1977). Social Expectation: Effects of Real and Hypothetical Explanations on Subjective Likelihood. Journal of Personality and Social Psychology, 35, 817-829.

Schneider, W., & Shiffrin R.M., (1977). Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention. Psychological Review, 84, 1-66.

Shaiken, H., (1984) Work Transformed: Automation and Labor in the Computer Age, Holt, Rinehart and Winston, New York.

Shank, R., & Ahelson, R.P., (1977). Scripts, Plans, Goals, and Understanding. Hillsdale, N.J.: Erlbaum.

Shevrin, H., & Dickman, S., (1980). The Psychological Unconscious: A Necessary Assumption for all Psychological Theory? American Psychologist, 35, 421-434.

Shiffrin R.M., & Schneider, W., (1977). Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory. Psychological Review, 84, 127-190.

Simon, H.A., (1957). Models of Man. New York: Wiley.

Slovic, P., Fischhoff B., & Lichenstein, S., (1977). Behavioral Decision Theory. Annual Review of Psychology, 28, 1-39.

Smith, E.E., Adams, N., & Shorr, D. (1978) Fact Retrieval and the Paradox of Interference. Cognitive Psychology, 10, 438-464.

Smith, M.J., Carayon, P., Miezio, K., (1986). Motivational, Behavioral, and Psychological Implications of Electronic Monitoring of Worker Performance. U.S. Congress, Office of Technology Assessment The Electronic Supervisor: New Technology, New Tensions. Vol. 2 (. B88-156369), U.S. Department of Commerce, NTIS, Springfield, V.A. December, 1987, 1-119.

Snyder, M., (1981). Seek and Ye Shall Find: Testing Hypotheses about Other People: In F. Higgins, C. Herman, & M. Zanna (Eds.), Social Cognition:The Ontario Symposium (Vol. 1), Hillsdale, NJ, Erlbaum.

Synder M., & Cantor N., (1979). Testing Hypotheses about Other People: The Use of Historical Knowledge. In F. Higgins, C. Herman, & M. Zanna (Eds.), Social Cognition:The Ontario Symposium (Vol. 1), Hillsdale, NJ, Erlbaum.

Srull, T.K., & Wyer, R.S. Jr., (1979). The Role and Category Accessibility in the Interpretation of Information About: Some Determinants and Implications. Journal of Personality and Social Psychology, 10, 1660-1672.

Staelin, R., and Payne, J.W., (1976). Studies of the Information Seeking Behavior of Consumers. In J.S. Carroll & J.W. Payne (Eds.). Cognition and Social Behavior. Hillsdale. NJ: Erlbaum.

Stokmans, M., (1992). Analyzing Information Search Patterns to Test the Use of a Two-Phased Decision Strategy. Acta Psychologica, 80, 213-227.

Svenson, O., (1979). Process Descriptions of Decision Making. Organizational Behavior and Human Performance, 23, 86-112.

Swann, W.B. Jr., Stephenson, B., & Pittman, T.S. (1981). Curiosity and Control: On the Determinants of the Search for Social Knowledge. Journal of Personality and Social Psychology, 40, 635-642.

Tajfel, H., (1969). Social Perception. In G. Lindzey & E. Aronson (Eds.) Handbook of Social Psychology (Vol. 3) Reading MA: Addison-Wesley.

Tamuz, M. (1987). The Impact of Computer Surveillance on Air Safety Reporting. Columbia Journal of World Business, 22, 69-77.

Taylor, S.E., & Crocker, J. (1981). Schematic Bases of Social Information Processing. In F. Higgins, C. Herman, & M. Zanna (Eds.), Social Cognition: The Ontario Symposium (Vol. 1), Hillsdale, NJ, Erlbaum.

Taylor S.E., & Fiske, S.T., (1978). Salience, Attention, and Attributions: Top of the Head Phenomena. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Vol. 11. New York: Academic Press.

Thorngate, W., (1976). Must We Always Think Before We Act? Personality and Social Psychology Bulletin, 2, 31-35.

Trope, Y., Bassok, M., (1982). Confirmatory and Diagnosing Strategies in Social Information Gathering. Journal of Personality and Social Psychology, 23, 22-34.

U.S. Congress, Office o Technology Assessment (OTA, September 1987), The Electronic Supervisor: New Technology, New Tensions, OTA-CIT-333, U.S. Government Printing Office, Washington, D.C., Walton, R.E. & Vittori, W., (1983). New Information Technology: Organizational Problem or Opportunity? Office Technology & People , 1, 249-273.

Westin, A.F., Schweder, H.A., Baker, M.A. & Lehman, S. (1985). The Changing Workplace. Knowledge Industry Publications, Inc., White Plains, NY.

Williams, K.J., Blencoe, A.G., DeNisi A.S., & Cafferty T.P., (1983). Does Appraisal Purpose and Outcome Affect Rater Information Search? Unpublished Manuscript. Department of Psychology. University of South Carolina.

Williams, K.J., DeNisi A.S., Blencoe, A.G., & Cafferty T.P., (1985). The Role of Appraisal Purpose: Effects of Purpose on Information Acquisition and Utilization. Organizational Behavior and Human Decision Processes, 36, 314-339

Wong, P.T.P., Weiner, B., (1981). When People Ask "Why" Questions, and the Heuristics of Attributional Search. Journal of Personality and Social Psychology, 40, 650-663.

Wyer, R.S. & Srull, T.K., (1981). Category Accessibility: Some theoretical and Empirical Issues Concerning the Processing of Social Stimuli Information. In F. Higgins, C. Herman, & M. Zanna (Eds.), Social Cognition:The Ontario Symposium (Vol. 1), Hillsdale, NJ, Erlbaum.

Wyer, R.S., Srull, T.K., Gordon, S.E., & Hartwick, J., (1982). Effects of Processing Objectives on the Recall of Prose Material. Journal of Personality and Social Psychology, 43, 674-688.

Zadny, J., & Gerard, H.B., (1974). Attributed Intentions and Information Selectivity. Journal of Experimental Social Psychology, 10, 34-52.

Zedeck, S., & Cascio, W.F., (1982). Performance Appraisal Decisions as a Function of Rater Training and Purpose of the Appraisal. Journal of Applied Psychology, 67, 752-758.

Appendix A

Analysis of Variance

Accuracy

General Linear Models Procedure

Class Level Information

Class	Levels	Values
APPRKNOW	2	0 1
REASON	2	0 1

Number of observations in data set = 127

Dependent Variable: EASYACC

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	16.50841588	5.50280529	2.26	0.0853
Error	123	300.12150538	2.44001224		
Corrected Total	126	316.62992126			

R-Square	C.V.	Root MSE	EASYACC Mean
0.052138	57.83698	1.562054	2.700787

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.71114332	0.71114332	0.29	0.5903
REASON	1	14.61462896	14.61462896	5.99	0.0158
APPRKNOW*REASON	1	1.18264361	1.18264361	0.48	0.4876

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.35774969	0.35774969	0.15	0.7024
REASON	1	14.30400146	14.30400146	5.86	0.0169
APPFRKNOW*REASON	1	1.18264361	1.18264361	0.48	0.4876

Level of APPRKNOW	N	EASYACC Mean	EASYACC SD
0	61	2.62295082	1.63466545
1	66	2.77272727	1.54716072

Level of REASON	N	EASYACC Mean	EASYACC SD
0	66	3.03030303	1.43548113
1	61	2.34426230	1.67217668

Level of APPRKNOW	Level of REASON	N	EASYACC Mean	EASYACC SD
0	0	30	2.86666667	1.59164485
0	1	31	2.38709677	1.66688171
1	0	36	3.16666667	1.29835060
1	1	30	2.30000000	1.70496233

Dependent Variable: DIFFACC

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	8.14662038	2.71554013	5.61	0.0012
Error	123	59.58566308	0.48443629		
Corrected Total	126	67.73228346			

R-Square	C.V.	Root MSE	DIFFACC Mean
0.120277	135.9905	0.696015	0.511811

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	2.43148863	2.43148863	5.02	0.0269
REASON	1	2.97142190	2.97142190	6.13	0.0146
APPRKNOW*REASON	1	2.74370985	2.74370985	5.66	0.0189

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	2.52796257	2.52796257	5.22	0.0241
REASON	1	3.17294135	3.17294135	6.55	0.0117
APPRKNOW*REASON	1	2.74370985	2.74370985	5.66	0.0189

Level of APPRKNOW	N	DIFFACC Mean	DIFFACC SD
0	61	0.65573770	0.77212361
1	66	0.37878788	0.67402697

Level of REASON	N	DIFFACC Mean	DIFFACC SD
0	66	0.65151515	0.79406305
1	61	0.36065574	0.63331895

Level of APPRKNOW	Level of REASON	N	DIFFACC Mean	DIFFACC SD
0	0	30	0.96666667	0.85028731
0	1	31	0.35483871	0.55065943
1	0	36	0.38888889	0.64488217
1	1	30	0.36666667	0.71839540

Dependent Variable: TOTACC

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	31.60213643	10.53404548	2.67	0.0507
Error	123	485.65770609	3.94843663		
Corrected Total	126	517.25984252			

R-Square	C.V.	Root MSE	TOTACC Mean
0.061095	61.85234	1.987067	3.212598

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.51269895	0.51269895	0.13	0.7192
REASON	1	30.76576686	30.76576686	7.79	0.0061
APPRKNOW*REASON	1	0.32367062	0.32367062	0.08	0.7751

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.98373662	0.98373662	0.25	0.6186
REASON	1	30.95073332	30.95073332	7.84	0.0059
APPRKNOW*REASON	1	0.32367062	0.32367062	0.08	0.7751

Level of APPRKNOW	N	TOTACC Mean	TOTACC SD
0	61	3.27868852	2.16895634
1	66	3.15151515	1.89933125

Level of REASON	N	TOTACC Mean	TOTACC SD
0	66	3.68181818	1.92281117
1	61	2.70491803	2.02767734

Level of APPRKNOW	Level of REASON	N	TOTACC Mean	TOTACC SD
0	0	30	3.83333333	2.24504563
0	1	31	2.74193548	1.98272106
1	0	36	3.55555556	1.62910044
1	1	30	2.66666667	2.10636692

Appendix B

Analysis of Variance

Breadth of Search

Dependent Variable: SCRSRCH

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	196.1730505	65.3910168	3.09	0.0298
Error	123	2605.8741935	21.1859691		
Corrected Total	126	2802.0472441			

R-Square	C.V.	Root MSE	SCRSRCH Mean
0.070011	30.50931	4.602822	15.08661

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	91.02190877	91.02190877	4.30	0.0403
REASON	1	56.24641029	56.24641029	2.65	0.1058
APPRKNOW*REASON	1	48.90473149	48.90473149	2.31	0.1312

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	93.61873432	93.61873432	4.42	0.0376
REASON	1	59.94405256	59.94405256	2.83	0.0951
APPRKNOW*REASON	1	48.90473149	48.90473149	2.31	0.1312

Level of	-----------SCRSRCH-----------		
APPRKNOW	N	Mean	SD
0	61	15.9672131	4.69740075
1	66	14.2727273	4.61950865

Level of	-----------SCRSRCH-----------		
REASON	N	Mean	SD
0	66	15.6818182	4.92141034
1	61	14.4426230	4.43292451

Level of	Level of	-----------SCRSRCH-----------		
APPRKNOW	REASON	N	Mean	SD
0	0	30	17.3000000	4.25197723
0	1	31	14.6774194	4.81239439
1	0	36	14.3333333	5.08780054
1	1	30	14.2000000	4.07176993

Dependent Variable: BRSRCH

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	3405.782128	1135.260709	3.09	0.0298
Error	123	45240.871416	367.811963		
Corrected Total	126	48646.653544			

R-Square	C.V.	Root MSE	BRSRCH Mean
0.070011	30.50931	19.17842	62.86089

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1580.241472	1580.241472	4.30	0.0403
REASON	1	976.500178	976.500178	2.65	0.1058
APPRKNOW*REASON	1	849.040477	849.040477	2.31	0.1312

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1625.325249	1625.325249	4.42	0.0376
REASON	1	1040.695357	1040.695357	2.83	0.0951
APPRKNOW*REASON	1	849.040477	849.040477	2.31	0.1312

Level of APPRKNOW	N	BRSRCH Mean	BRSRCH SD
0	61	66.5300546	19.5725031
1	66	59.4696970	19.2479527

Level of REASON	N	BRSRCH Mean	BRSRCH SD
0	66	65.3409091	20.5058764
1	61	60.1775956	18.4705188

Level of APPRKNOW	Level of REASON	N	BRSRCH Mean	BRSRCH SD
0	0	30	72.0833333	17.7165718
0	1	31	61.1559140	20.0516433
1	0	36	59.7222222	21.1991689
1	1	30	59.1666667	16.9657080

Dependent Variable: BSDAYTAB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	28.93743826	9.64581275	1.80	0.1514
Error	123	660.57437276	5.37052336		
Corrected Total	126	689.51181102			

R-Square	C.V.	Root MSE	BSDAYTAB Mean
0.041968	65.40328	2.317439	3.543307

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1.94797595	1.94797595	0.36	0.5481
REASON	1	0.37364177	0.37364177	0.07	0.7924
APPRKNOW*REASON	1	26.61582055	26.61582055	4.96	0.0278

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1.38677764	1.38677764	0.26	0.6123
REASON	1	0.18357161	0.18357161	0.03	0.8536
APPRKNOW*REASON	1	26.61582055	26.61582055	4.96	0.0278

Level of		----------BSDAYTAB----------	
APPRKNOW	N	Mean	SD
0	61	3.67213115	2.44077386
1	66	3.42424242	2.25361662

Level of		----------BSDAYTAB----------	
REASON	N	Mean	SD
0	66	3.48484848	2.34187580
1	61	3.60655738	2.35427759

Level of	Level of		----------BSDAYTAB----------	
APPRKNOW	REASON	N	Mean	SD
0	0	30	4.10000000	2.30964992
0	1	31	3.25806452	2.52939706
1	0	36	2.97222222	2.27390301
1	1	30	3.96666667	2.14127473

Dependent Variable: BSDTPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	8038.177295	2679.392432	1.80	0.1514
Error	123	183492.881321	1491.812043		
Corrected Total	126	191531.058615			

R-Square	C.V.	Root MSE	BSDTPER Mean
0.041968	65.40328	38.62398	59.05512

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	541.104431	541.104431	0.36	0.5481
REASON	1	103.789380	103.789380	0.07	0.7924
APPRKNOW*REASON	1	7393.283484	7393.283484	4.96	0.0278

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	385.216011	385.216011	0.26	0.6123
REASON	1	50.992113	50.992113	0.03	0.8536
APPRKNOW*REASON	1	7393.283484	7393.283484	4.96	0.0278

Level of APPRKNOW	N	BSDTPER Mean	BSDTPER SD
0	61	61.2021858	40.6795644
1	66	57.0707071	37.5602770

Level of REASON	N	BSDTPER Mean	BSDTPER SD
0	66	58.0808081	39.0312634
1	61	60.1092896	39.2379598

Level of APPRKNOW	Level of REASON	N	BSDTPER Mean	BSDTPER SD
0	0	30	68.3333333	38.4941654
0	1	31	54.3010753	42.1566176
1	0	36	49.5370370	37.8983836
1	1	30	66.1111111	35.6879122

Dependent Variable: BSTAB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	138.5245647	46.1748549	3.62	0.0152
Error	123	1570.3415771	12.7670047		
Corrected Total	126	1708.8661417			

R-Square	C.V.	Root MSE	BSTAB Mean
0.081062	50.36437	3.573095	7.094488

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	39.16594302	39.16594302	3.07	0.0824
REASON	1	0.00378192	0.00378192	0.00	0.9863
APPRKNOW*REASON	1	99.35483972	99.35483972	7.78	0.0061

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	34.68723011	34.68723011	2.72	0.1018
REASON	1	0.08468414	0.08468414	0.01	0.9352
APPRKNOW*REASON	1	99.35483972	99.35483972	7.78	0.0061

Level of APPRKNOW	N	BSTAB Mean	BSTAB SD
0	61	7.67213115	3.79350898
1	66	6.56060606	3.52192599

Level of REASON	N	BSTAB Mean	BSTAB SD
0	66	7.06060606	3.62418707
1	61	7.13114754	3.77480865

Level of APPRKNOW	Level of REASON	N	BSTAB Mean	BSTAB SD
0	0	30	8.60000000	3.16881355
0	1	31	6.77419355	4.16901009
1	0	36	5.77777778	3.51414376
1	1	30	7.50000000	3.35024446

Dependent Variable: BSTABPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	9619.761436	3206.587145	3.62	0.0152
Error	123	109051.498408	886.597548		
Corrected Total	126	118671.259843			

R-Square	C.V.	Root MSE	BSTABPER Mean
0.081062	50.36437	29.77579	59.12073

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	2719.857154	2719.857154	3.07	0.0824
REASON	1	0.262634	0.262634	0.00	0.9863
APPRKNOW*REASON	1	6899.641648	6899.641648	7.78	0.0061

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	2408.835424	2408.835424	2.72	0.1018
REASON	1	5.880843	5.880843	0.01	0.9352
APPRKNOW*REASON	1	6899.641648	6899.641648	7.78	0.0061

Level of APPRKNOW	N	-----------BSTABPER---------- Mean	SD
0	61	63.9344262	31.6125748
1	66	54.6717172	29.3493833

Level of REASON	N	-----------BSTABPER---------- Mean	SD
0	66	58.8383838	30.2015589
1	61	59.4262295	31.4567387

Level of APPRKNOW	Level of REASON	N	-----------BSTABPER---------- Mean	SD
0	0	30	71.6666667	26.4067796
0	1	31	56.4516129	34.7417508
1	0	36	48.1481481	29.2845313
1	1	30	62.5000000	27.9187038

Dependent Variable: E3HRTAB

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	47.06983180	15.68994393	4.36	0.0059
Error	123	442.34749104	3.59632107		
Corrected Total	126	489.41732283			

R-Square	C.V.	Root MSE	BSHRTAB Mean
0.096175	53.40186	1.896397	3.551181

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	23.64459556	23.64459556	6.57	0.0115
REASON	1	0.30224165	0.30224165	0.08	0.7724
APPRKNOW*REASON	1	23.12299458	23.12299458	6.43	0.0125

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	22.20267398	22.20267398	6.17	0.0143
REASON	1	0.51761976	0.51761976	0.14	0.7051
APPRKNOW*REASON	1	23.12299458	23.12299458	6.43	0.0125

Level of APPRKNOW	N	-----------BSHRTAB----------- Mean	SD
0	61	4.00000000	1.88856206
1	66	3.13636364	1.96810228

Level of REASON	N	BSHRTAB----------- Mean	SD
0	66	3.57575758	1.96946386
1	61	3.52459016	1.98835407

Level of APPRKNOW	Level of REASON	N	-----------BSHRTAB----------- Mean	SD
0	0	30	4.50000000	1.50287082
0	1	31	3.51612903	2.11141292
1	0	36	2.80555556	1.99741897
1	1	30	3.53333333	1.88886635

Dependent Variable: BSHTPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	13074.95328	4358.31776	4.36	0.0059
Error	123	122874.30306	998.97807		
Corrected Total	126	135949.25634			

R-Square	C.V.	Root MSE	BSHTPER Mean
0.096175	53.40186	31.60661	59.18635

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	6567.943212	6567.943212	6.57	0.0115
REASON	1	83.956015	83.956015	0.08	0.7724
APPRKNOW*REASON	1	6423.054050	6423.054050	6.43	0.0125

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	6167.409438	6167.409438	6.17	0.0143
REASON	1	143.783266	143.783266	0.14	0.7051
APPRKNOW*REASON	1	6423.054050	6423.054050	6.43	0.0125

Level of APPRKNOW	N	BSHTPER Mean	BSHTPER SD
0	61	66.6666667	31.4760344
1	66	52.2727273	32.8017046

Level of REASON	N	BSHTPER Mean	BSHTPER SD
0	66	59.5959596	32.8243976
1	61	58.7431694	33.1392345

Level of APPRKNOW	Level of REASON	N	BSHTPER Mean	BSHTPER SD
0	0	30	75.0000000	25.0478469
0	1	31	58.6021505	35.1902153
1	0	36	46.7592593	33.2903162
1	1	30	58.8888889	31.4811058

Dependent Variable: BSHRS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	111.7522550	37.2507517	4.34	0.0061
Error	123	1056.9249104	8.5928855		
Corrected Total	126	1168.6771654			

R-Square	C.V.	Root MSE	BSHRS Mean
0.095623	43.13824	2.931362	6.795276

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	71.13792045	71.13792045	8.28	0.0047
REASON	1	23.16461030	23.16461030	2.70	0.1032
APPRKNOW*REASON	1	17.44972421	17.44972421	2.03	0.1567

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	72.71422209	72.71422209	8.46	0.0043
REASON	1	24.57843708	24.57843708	2.86	0.0933
APPRKNOW*REASON	1	17.44972421	17.44972421	2.03	0.1567

Level of APPRKNOW	N	BSHRS Mean	BSHRS SD
0	61	7.57377049	2.87784072
1	66	6.07575758	3.03979051

Level of REASON	N	BSHRS Mean	BSHRS SD
0	66	7.16666667	3.16997129
1	61	6.39344262	2.87679618

Level of APPRKNOW	Level of REASON	N	BSHRS Mean	BSHRS SD
0	0	30	8.40000000	2.72409428
0	1	31	6.77419355	2.83677842
1	0	36	6.13888889	3.18166872
1	1	30	6.00000000	2.91251757

Dependent Variable: BSHRSPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	7760.573261	2586.857754	4.34	0.0061
Error	123	73397.563221	596.728156		
Corrected Total	126	81158.136482			

R-Square	C.V.	Root MSE	BSHRSPER Mean
0.095623	43.13824	24.42802	56.62730

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	4940.133364	4940.133364	8.28	0.0047
REASON	1	1608.653493	1608.653493	2.70	0.1032
APPRKNOW*REASON	1	1211.786404	1211.786404	2.03	0.1567

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	5049.598756	5049.598756	8.46	0.0043
REASON	1	1706.835908	1706.835908	2.86	0.0933
APPRKNOW*REASON	1	1211.786404	1211.786404	2.03	0.1567

Level of APPRKNOW	N	BSHRSPER Mean	BSHRSPER SD
0	61	63.1147541	23.9820060
1	66	50.6313131	25.3315876

Level of REASON	N	BSHRSPER Mean	BSHRSPER SD
0	66	59.7222222	26.4164274
1	61	53.2786885	23.9733015

Level of APPRKNOW	Level of REASON	N	BSHRSPER Mean	BSHRSPER SD
0	0	30	70.0000000	22.7007857
0	1	31	56.7516129	23.6398202
1	0	36	51.1574074	26.5139060
1	1	30	50.0000000	24.2709798

Dependent Variable: BSGRAPHS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	76.79355968	25.59785323	2.19	0.0931
Error	123	1440.19856631	11.70893143		
Corrected Total	126	1516.99212598			

R-Square	C.V.	Root MSE	BSGRAPHS Mean
0.050622	42.81504	3.421832	7.992126

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	10.77329836	10.77329836	0.92	0.3393
REASON	1	57.17262300	57.17262300	4.88	0.0290
APPRKNOW*REASON	1	8.84763831	8.84763831	0.76	0.3864

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	14.33448414	14.33448414	1.22	0.2707
REASON	1	55.52260295	55.52260295	4.74	0.0313
APPRKNOW*REASON	1	8.84763831	8.84763831	0.76	0.3864

Level of APPRKNOW	N	-----------BSGRAPHS---------- Mean	SD
0	61	8.29508197	3.45612626
1	66	7.71212121	3.48520014

Level of REASON	N	-----------BSGRAPHS---------- Mean	SD
0	66	8.62121212	3.32248818
1	61	7.31147541	3.52392293

Level of APPRKNOW	Level of REASON	N	-----------BSGRAPHS---------- Mean	SD
0	0	30	8.70000000	3.55401667
0	1	31	7.90322581	3.37001324
1	0	36	8.55555556	3.16629070
1	1	30	6.70000000	3.63080688

Dependent Variable: BSGRPPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5332.886089	1777.628696	2.19	0.0931
Error	123	100013.789327	813.120238		
Corrected Total	126	105346.675415			

R-Square	C.V.	Root MSE	BSGRPPER Mean
0.050622	42.81504	28.51526	66.60105

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	748.145720	748.145720	0.92	0.3393
REASON	1	3970.321041	3970.321041	4.88	0.0290
APPRKNOW*REASON	1	614.419327	614.419327	0.76	0.3864

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	995.450288	995.450288	1.22	0.2707
REASON	1	3855.736316	3855.736316	4.74	0.0313
APPRKNOW*REASON	1	614.419327	614.419327	0.76	0.3864

Level of APPRKNOW	N	BSGRPPER Mean	SD
0	61	69.1256831	28.8010522
1	66	64.2676768	29.0433345

Level of REASON	N	BSGRPPER Mean	SD
0	66	71.8434343	27.6874015
1	61	60.9289617	29.3660244

Level of APPRKNOW	Level of REASON	N	BSGRPPER Mean	SD
0	0	30	72.5000000	29.6168056
0	1	31	65.8602151	28.0834437
1	0	36	71.2962963	26.3857559
1	1	30	55.8333333	30.2567240

Dependent Variable: BSHRGRP

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	31.33092033	10.44364011	2.34	0.0763
Error	123	548.10215054	4.45611505		
Corrected Total	126	579.43307087			

R-Square	C.V.	Root MSE	BSHRGRP Mean
0.054072	65.07058	2.110951	3.244094

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	12.75746232	12.75746232	2.86	0.0932
REASON	1	18.17485418	18.17485418	4.08	0.0456
APPRKNOW*REASON	1	0.39860382	0.39860382	0.09	0.7654

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	14.55645389	14.55645389	3.27	0.0731
REASON	1	17.96239449	17.96239449	4.03	0.0469
APPRKNOW*REASON	1	0.39860382	0.39860382	0.09	0.7654

Level of APPRKNOW	N	BSHRGRP Mean	BSHRGRP SD
0	61	3.57377049	2.08533783
1	66	2.93939394	2.16886208

Level of REASON	N	BSHRGRP Mean	BSHRGRP SD
0	66	3.59090909	2.07532961
1	61	2.86885246	2.17160010

Level of APPRKNOW	Level of REASON	N	BSHRGRP Mean	BSHRGRP SD
0	0	30	3.90000000	2.02314197
0	1	31	3.25806452	2.12865751
1	0	36	3.33333333	2.11119465
1	1	30	2.46666667	2.17720690

Dependent Variable: BSHGPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	8703.033424	2901.011141	2.34	0.0768
Error	123	152250.597369	1237.809735		
Corrected Total	126	160953.630794			

R-Square	C.V.	Root MSE	BSHGPER Mean
0.054072	65.07058	35.18252	54.06824

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	3543.739534	3543.739534	2.86	0.0932
REASON	1	5048.570606	5048.570606	4.08	0.0456
APPRKNOW*REASON	1	110.723284	110.723284	0.09	0.7654

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	4043.459415	4043.459415	3.27	0.0731
REASON	1	4989.554024	4989.554024	4.03	0.0469
APPRKNOW*REASON	1	110.723284	110.723284	0.09	0.7654

Level of APPRKNOW	N	BSHGPER Mean	BSHGPER SD
0	61	59.5628415	34.7556306
1	66	48.9898990	36.1477013

Level of REASON	N	BSHGPER Mean	BSHGPER SD
0	66	59.8484848	34.5888269
1	61	47.8142077	36.1933350

Level of APPRKNOW	Level of REASON	N	BSHGPER Mean	BSHGPER SD
0	0	30	65.0000000	33.7190329
0	1	31	54.3010753	35.4776252
1	0	36	55.5555556	35.1865775
1	1	30	41.1111111	36.2867817

Appendix C:

Analysis of Variance

Averages

Dependent Variable: AVGRAPH

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1234.780075	411.593358	3.43	0.0192
Error	123	14746.736320	119.892165		
Corrected Total	126	15981.516394			

R-Square	C.V.	Root MSE	AVGRAPH Mean
0.077263	71.11795	10.94953	15.39629

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1008.580978	1008.580978	8.41	0.0044
REASON	1	167.905218	167.905218	1.40	0.2389
APPRKNOW*REASON	1	58.293879	58.293879	0.49	0.4869

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1031.526093	1031.526093	8.60	0.0040
REASON	1	174.760501	174.760501	1.46	0.2296
APPRKNOW*REASON	1	58.293879	58.293879	0.49	0.4869

Level of APPRKNOW	N	AVGRAPH Mean	AVGRAPH SD
0	61	12.4649903	7.5942060
1	66	18.1055261	13.3085383

Level of REASON	N	AVGRAPH Mean	AVGRAPH SD
0	66	14.4378288	10.9501433
1	61	16.4333185	11.5912453

Level of APPRKNOW	Level of REASON	N	AVGRAPH Mean	AVGRAPH SD
0	0	30	10.5787399	6.5394699
0	1	31	14.2903939	8.1855753
1	0	36	17.6537362	12.8002232
1	1	30	18.6476740	14.0955504

Dependent Variable: AVTDPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	9647.851337	3215.950446	3.09	0.0298
Error	123	128166.617105	1042.005017		
Corrected Total	126	137814.468442			

R-Square	C.V.	Root MSE	AVTDPER Mean
0.070006	70.27353	32.28010	45.93494

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	18.798855	18.798855	0.02	0.8934
REASON	1	2296.903671	2296.903671	2.20	0.1402
APPRKNOW*REASON	1	7332.148812	7332.148812	7.04	0.0090

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	22.896145	22.896145	0.02	0.8824
REASON	1	2593.244994	2593.244994	2.49	0.1172
APPRKNOW*REASON	1	7332.148812	7332.148812	7.04	0.0090

Level of	-----------AVTDPER-----------		
APPRKNOW	N	Mean	SD
0	61	45.5347434	35.1729849
1	66	46.3048143	31.2723367

Level of	-----------AVTDPER-----------		
REASON	N	Mean	SD
0	66	50.0373792	36.4326890
1	61	41.4962305	28.6428302

Level of	Level of	-----------AVTDPER-----------		
APPRKNOW	REASON	N	Mean	SD
0	0	30	57.8854601	37.6635539
0	1	31	33.5824370	28.3347749
1	0	36	43.4973119	34.5385159
1	1	30	49.6738172	27.0320753

Dependent Variable: AVTHPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	30029.39569	10009.79856	2.31	0.0795
Error	123	532715.97101	4331.02415		
Corrected Total	126	562745.36670			

R-Square	C.V.	Root MSE	AVTHPER Mean
0.053362	55.09539	65.81052	119.4483

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	175.25251	175.25251	0.04	0.8409
REASON	1	68.99577	68.99577	0.02	0.8998
APPRKNOW*REASON	1	29785.14742	29785.14742	6.88	0.0098

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	57.05768	57.05768	0.01	0.9088
REASON	1	207.47097	207.47097	0.05	0.8271
APPRKNOW*REASON	1	29785.14742	29785.14742	6.88	0.0098

Level of APPRKNOW	N	-----------AVTHPER----------- Mean	SD
0	61	120.670231	63.5756812
1	66	118.318990	70.1710086

Level of REASON	N	-----------AVTHPER----------- Mean	SD
0	66	120.095317	67.8573491
1	61	118.748303	66.2555103

Level of APPRKNOW	Level of REASON	N	-----------AVTHPER----------- Mean	SD
0	0	30	137.582710	57.5807269
0	1	31	104.303316	65.6953221
1	0	36	105.522491	72.9638143
1	1	30	133.674790	64.5298409

Dependent Variable: AVTBPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	16343.84076	5447.94692	2.60	0.0550
Error	123	257377.26521	2092.49809		
Corrected Total	126	273721.10597			

R-Square	C.V.	Root MSE	AVTBPER Mean
0.059710	48.47756	45.74383	94.36084

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	1158.47955	1158.47955	0.55	0.4583
REASON	1	848.59237	848.59237	0.41	0.5254
APPRKNOW*REASON	1	14336.76884	14336.76884	6.85	0.0100

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	979.32253	979.32253	0.47	0.4952
REASON	1	1111.97879	1111.97879	0.53	0.4674
APPRKNOW*REASON	1	14336.76884	14336.76884	6.85	0.0100

Level of APPRKNOW	N	AVTBPER Mean	AVTBPER SD
0	61	97.5024293	47.7998146
1	66	91.4572504	45.6530881

Level of REASON	N	AVTBPER Mean	AVTBPER SD
0	66	96.6865614	45.0209583
1	61	91.8444862	48.5163588

Level of APPRKNOW	Level of REASON	N	AVTBPER Mean	AVTBPER SD
0	0	30	111.348300	44.0440493
0	1	31	84.103199	48.1372697
1	0	36	84.468446	42.6507418
1	1	30	99.843816	48.4009443

Appendix D:

Graphs & Analysis of Variance

Times and Frequencies

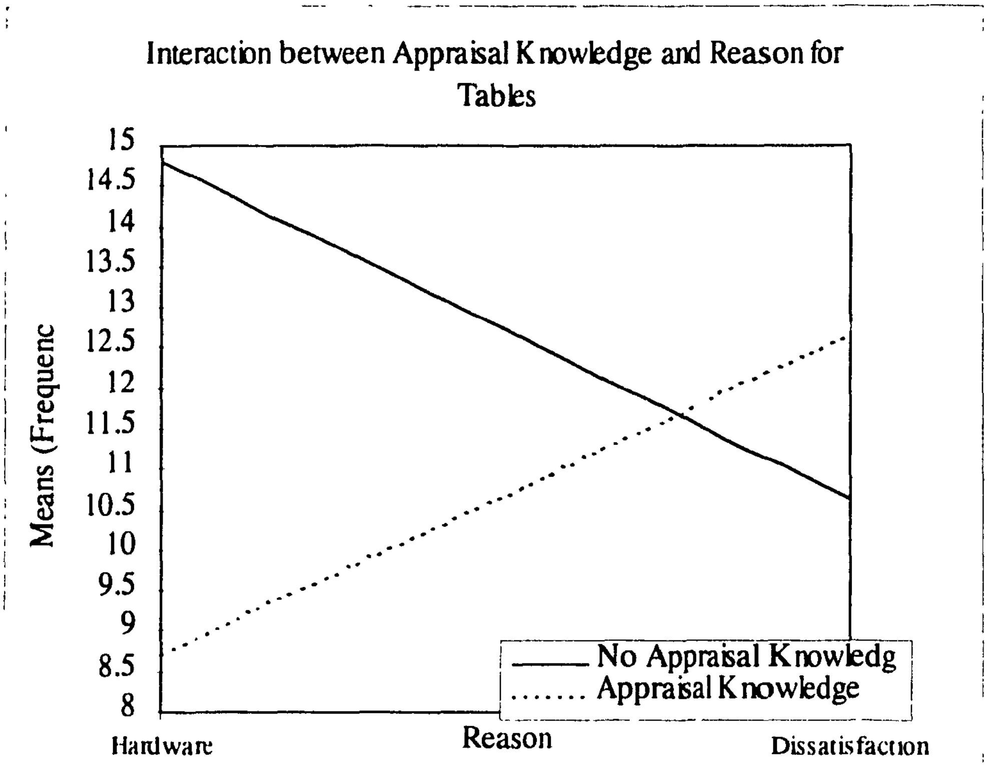


Fig 12: Interaction between Appraisal Knowledge and Reason for Tables

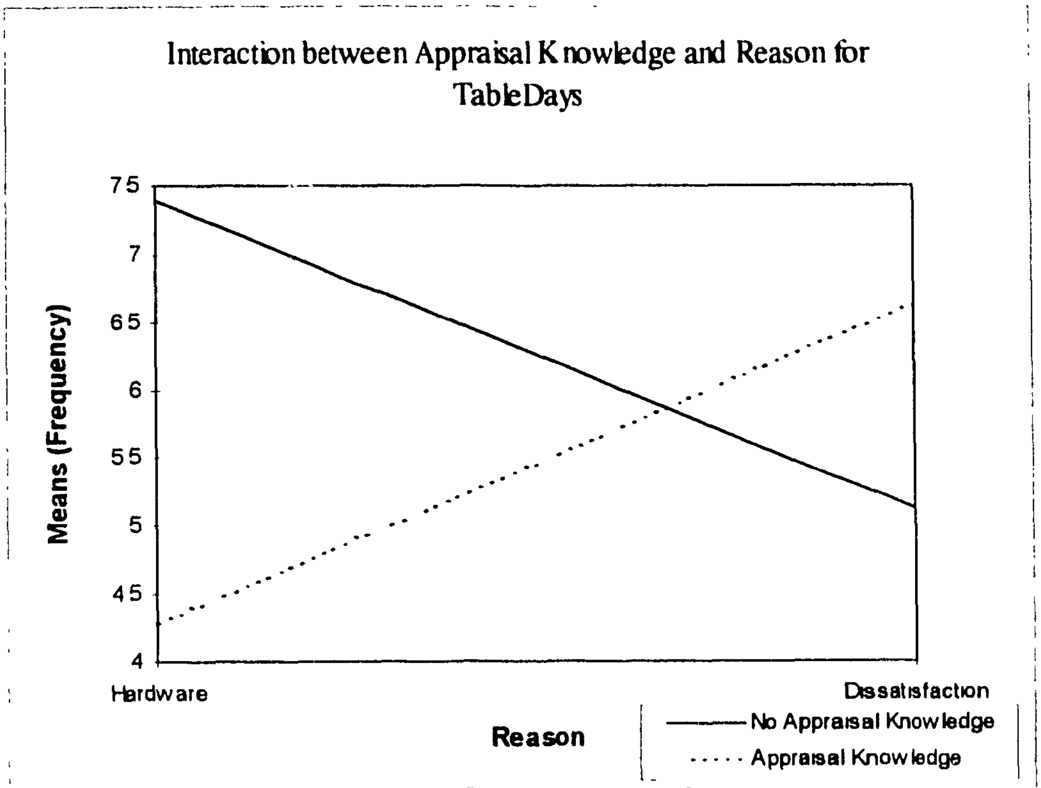


Fig 13: Interaction between Appraisal Knowledge and Reason for Table Days

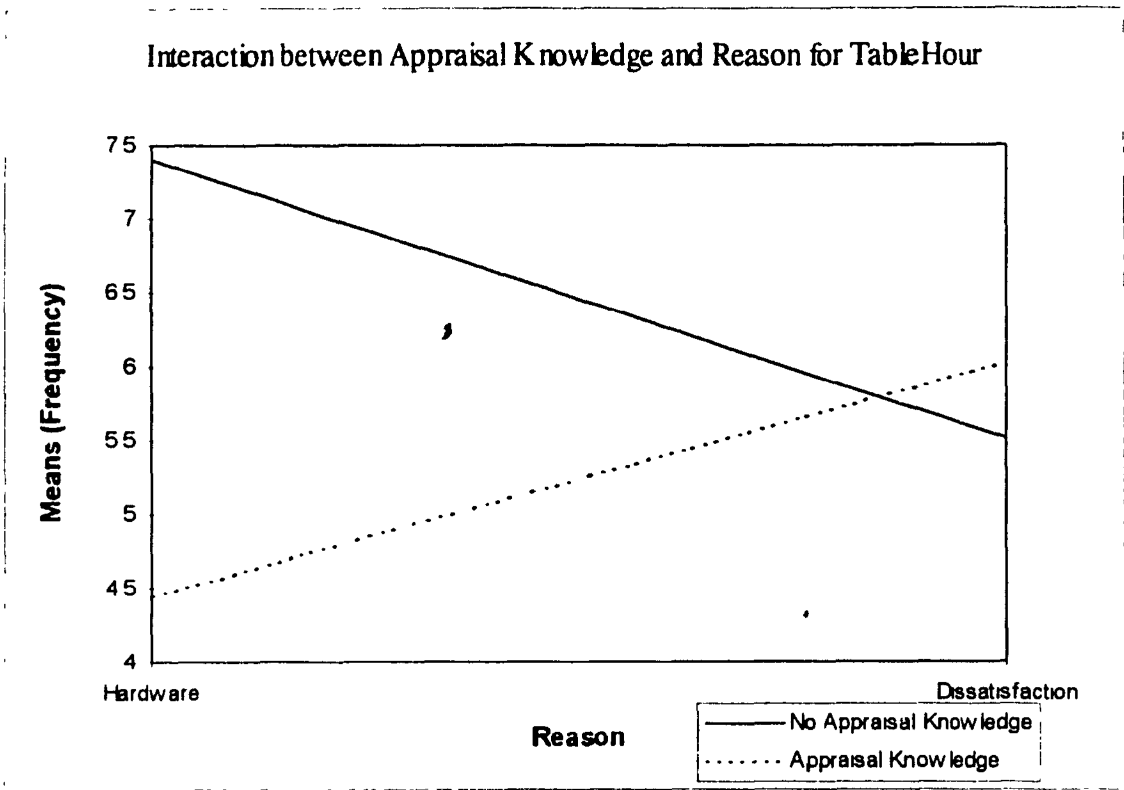


Fig 14: Interaction between Appraisal Knowledge and Reason for Table Hour

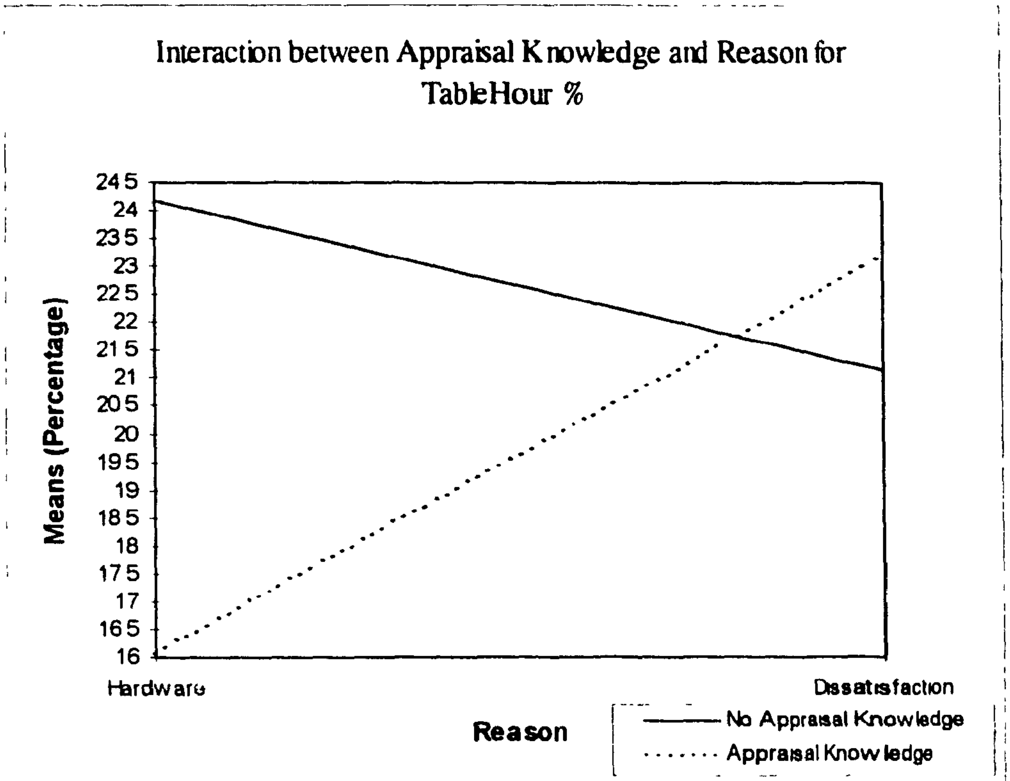


Fig 15: Interaction between Appraisal Knowledge and Reason for Table Hour

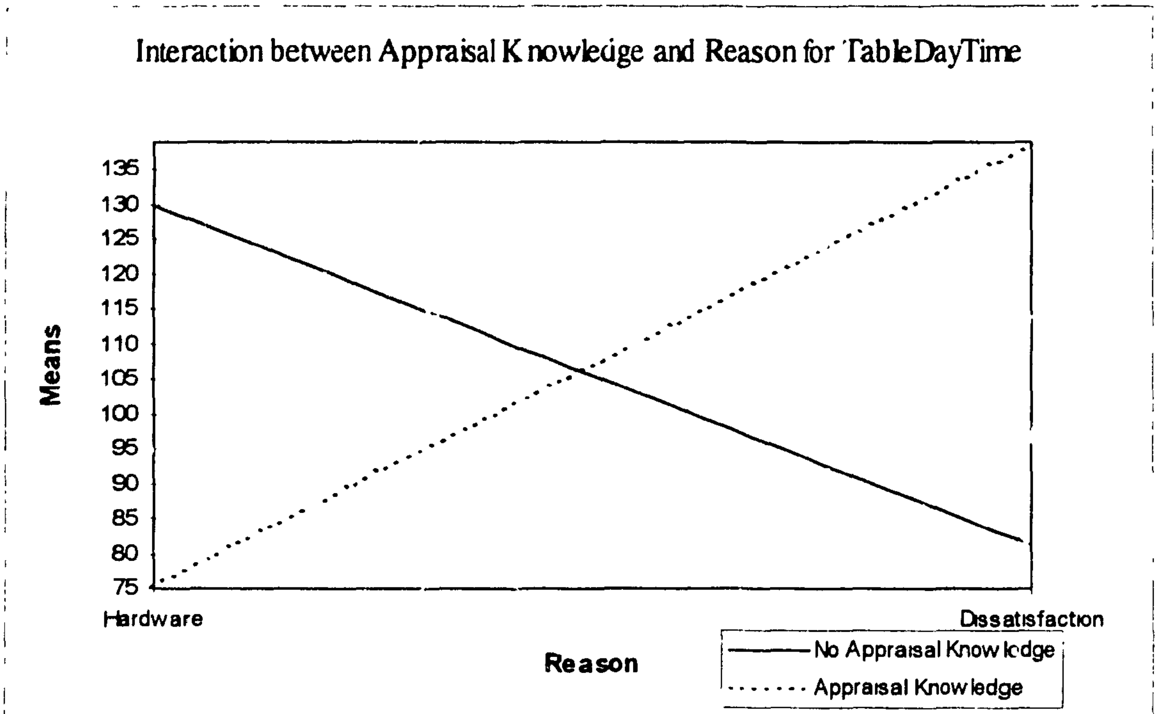


Fig 16: Interaction between Appraisal Knowledge and Reason for Table Day Time

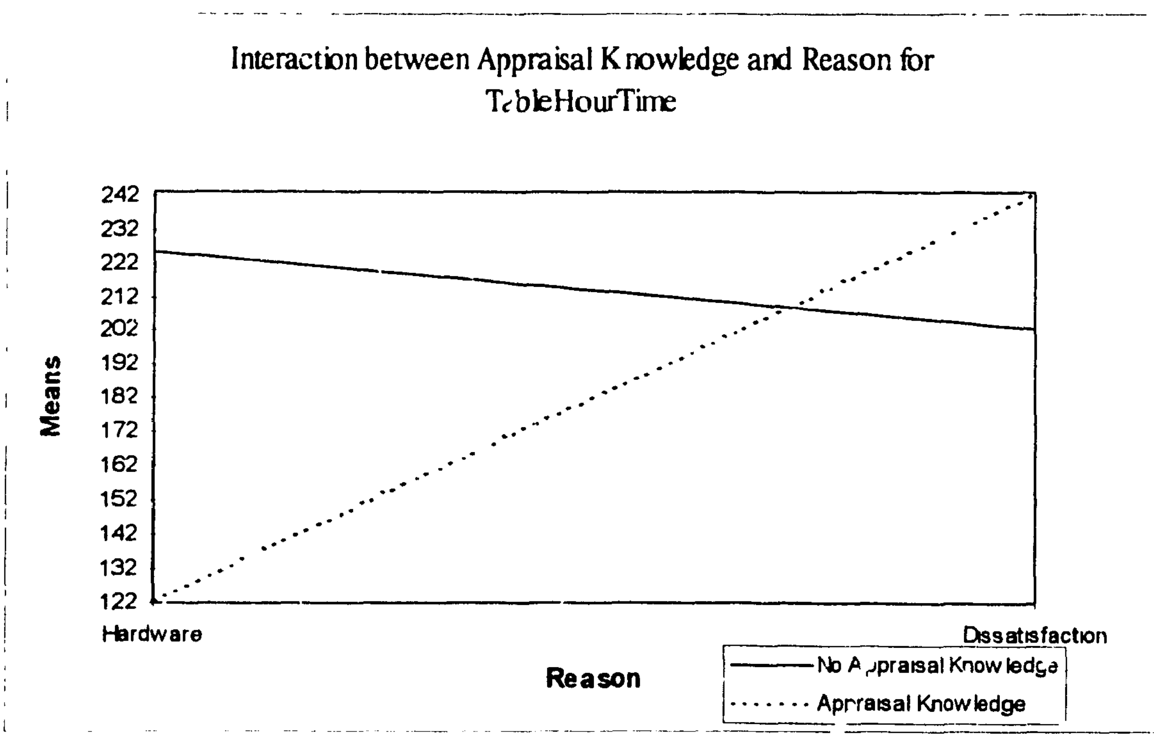


Fig 17: Interaction between Appraisal Knowledge and Reason for Table Hour Time

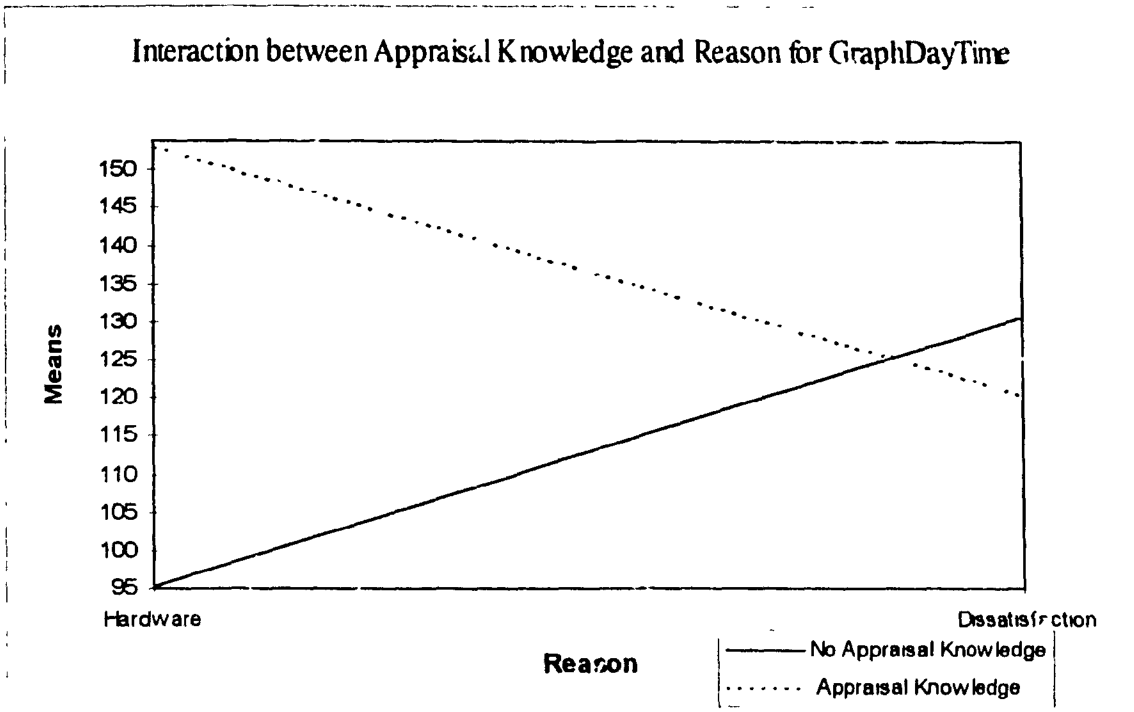


Fig 18: Interaction between Appraisal Knowledge and Reason for Graph Day Time

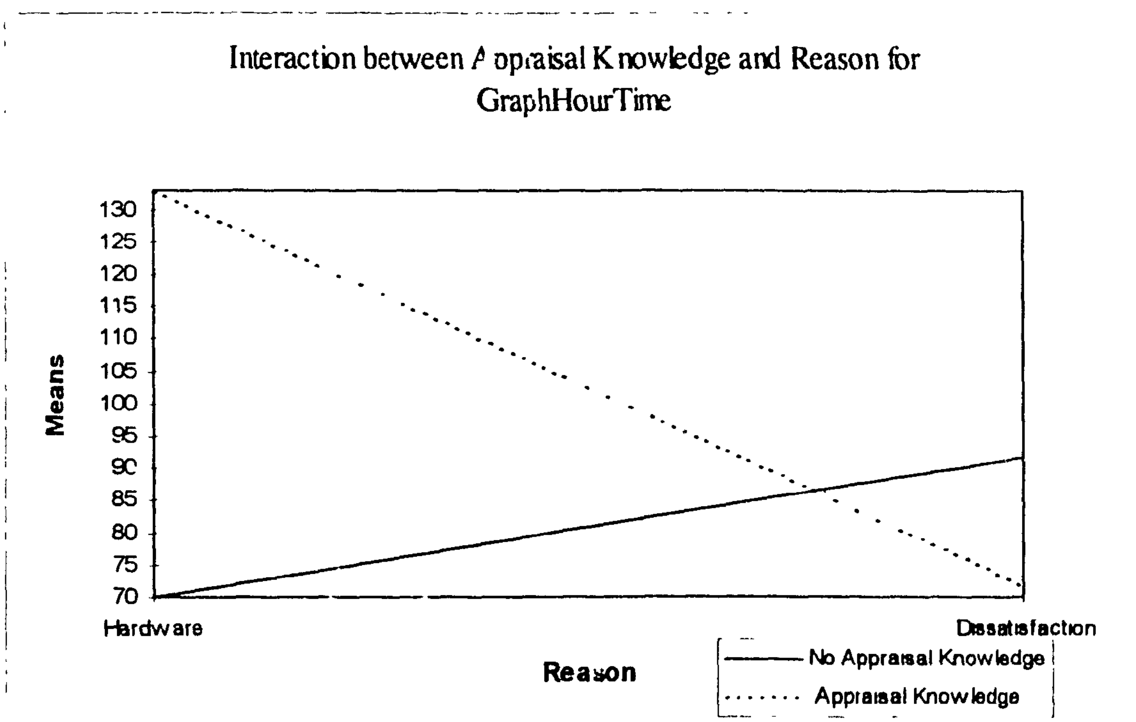


Fig 18: Interaction between Appraisal Knowledge and Reason for Graph Hour Time

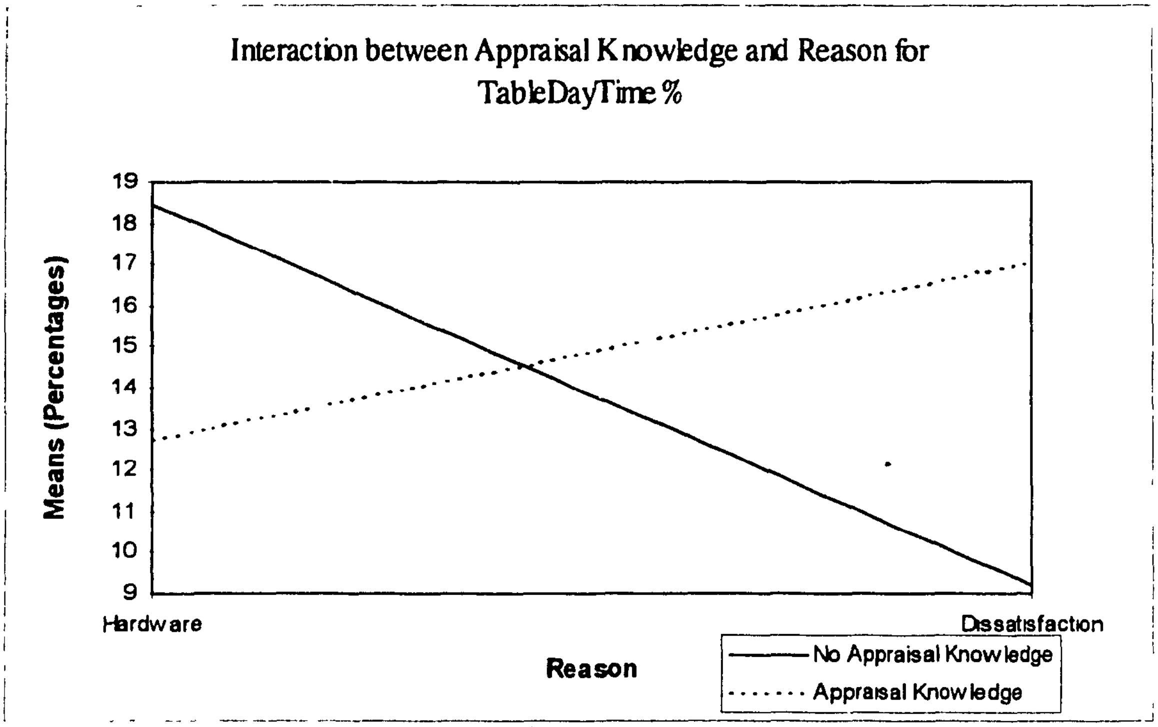


Fig 19: Interaction between Appraisal Knowledge and Reason for Table Day Time%

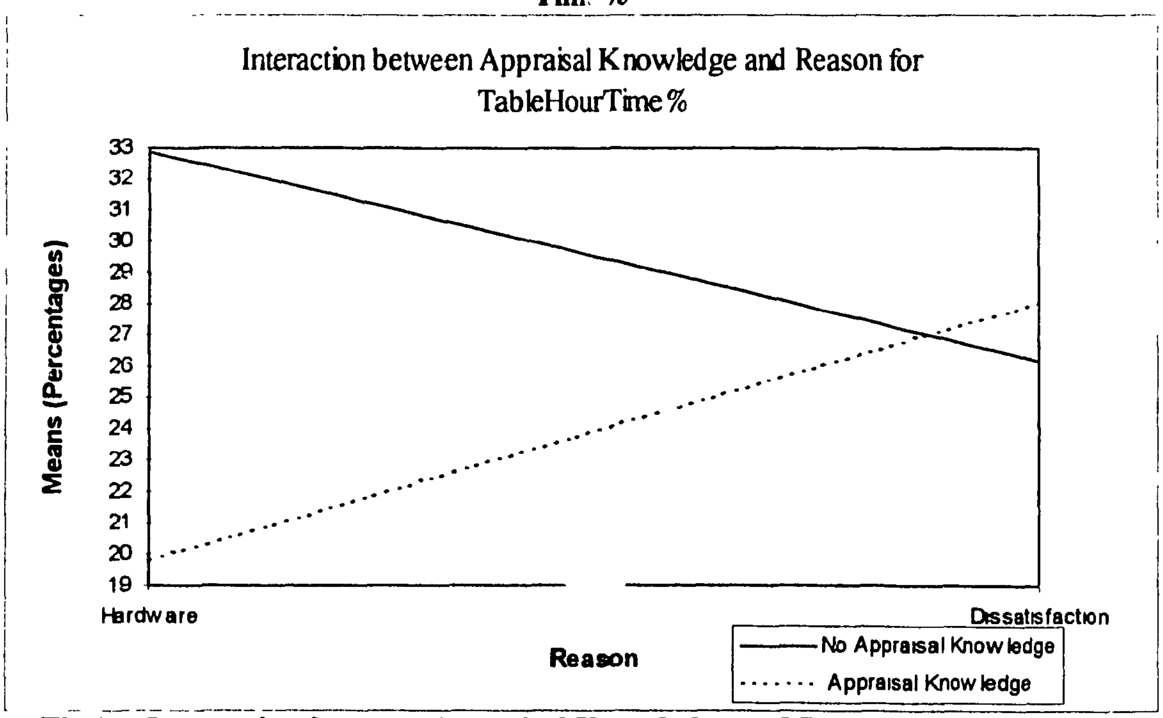


Fig 20: Interaction between Appraisal Knowledge and Reason for Table Hour Time%

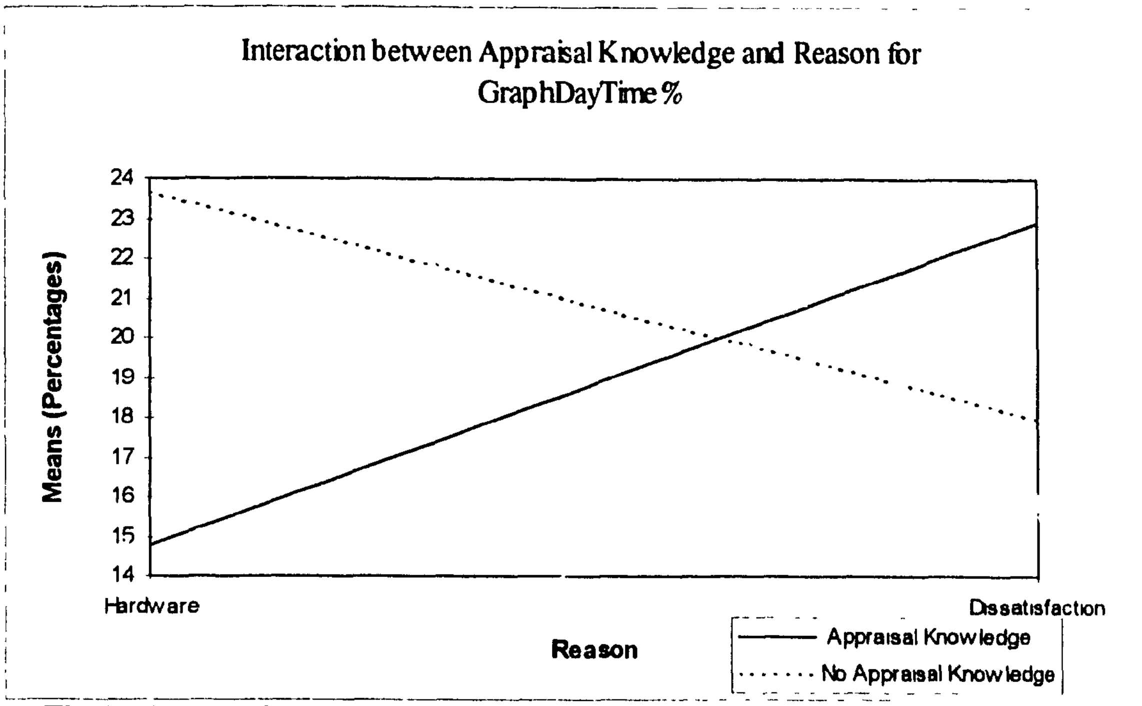


Fig 21: Interaction between Appraisal Knowledge and Reason for Graph Day Time%

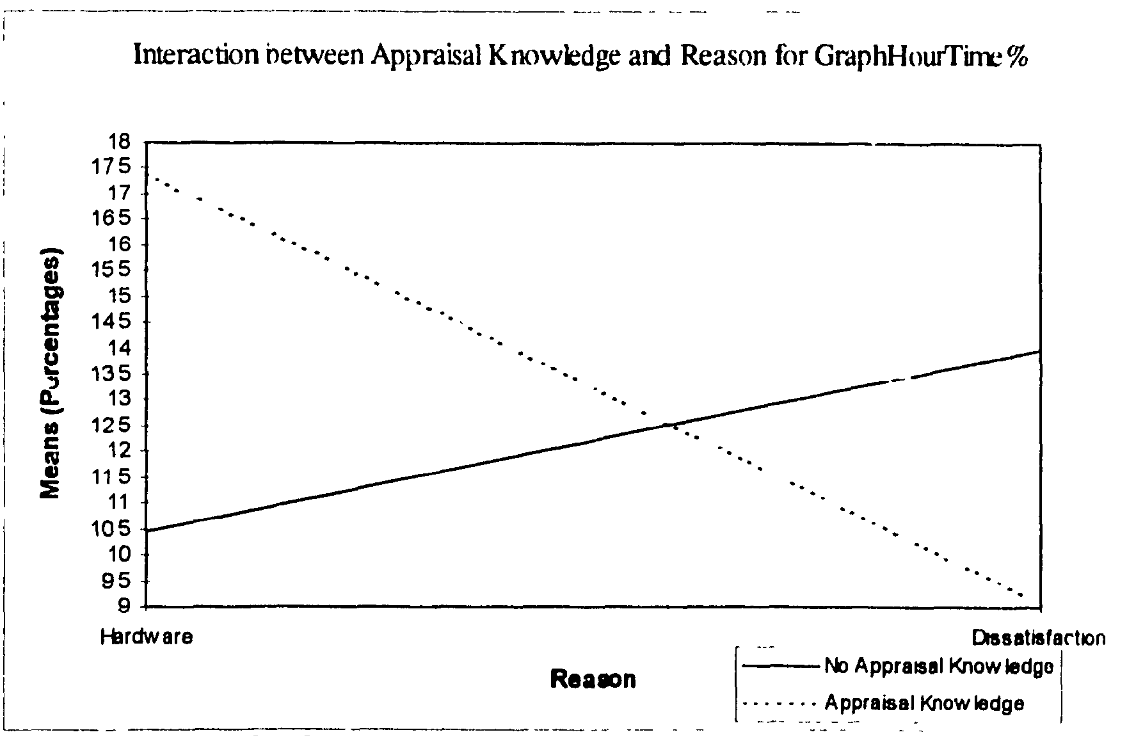


Fig 22: Interaction between Appraisal Knowledge and Reason for Graph Hour Time %

Dependent Variable: TABLES

Source	DF	Sum of Squares	Mean Square	F Value	Pr> F
Model	3	667.5214235	222.5071412	3.4	0.0184
Error	123	7893.7856631	64.1771192		
Corrected Total	126	8561.3070866			

R-Square	C.V.	Root MSE	TABLES Mean
0.077970	69.30551	8.011062	11.55906

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	149.7402709	149.7402709	2.30	.1292
REASON	1	0.0456464	0.0456464	0.00	0.9788
APPRKNOW*REASON	1	517.7355062	517.7355062	8.07	0.0053

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	129.8580899	129.8580899	2.02	0.1574
REASON	1	0.[illegible]93676	0.3493676	0.01	0.9413

APPRKNOW*REASON	1	517.7355062	517.7355062	8.07	0.0053

Level of APPRKNOW	N	TABLES Mean	TABLES SD
0	61	12.6885246	8.25740271
1	66	10.5151515	8.15285212

Level of REASON	N	TABLES Mean	TABLES SD
0	66	11.4848485	8.11123139
1	61	11.6393443	8.44991674

Level of APPRKNOW	Level of REASON	N	TABLES Mean	TABLES SD
0	0	30	14.8000000	8.81378231
0	1	31	10.6451613	7.24591097
1	0	36	8.7222222	6.36782587
1	1	30	12.6666667	9.55323860

Dependent Variable: TABDAY

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	194.9540118	64.9846706	2.61	0.0548
Error	123	3066.8727599	24.9339249		
Corrected Total	126	3261.8267717			

R-Square	C.V.	Root MSE	TABDAY Mean
0.059768	86.39786	4.993388	5.779528

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	25.5304478	25.5304478	1.02	0.3136
REASON	1	0.4867529	0.4867529	0.02	0.8891
APPRKNOW*REASON	1	168.9368111	168.9368111	6.78	0.0104

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	20.6599606	20.6599606	0.83	0.3645
REASON	1	0.0564716	0.0564716	0.00	0.9621
APPRKNOW*REASON	1	168.9368111	168.9368111	6.78	0.0104

Level of APPRKNOW	N	TABDAY Mean	TABDAY SD
0	61	6.24590164	5.21745704
1	66	5.34848485	4.96601504

Level of REASON	N	TABDAY Mean	TABDAY SD
0	66	5.69696970	4.81334586
1	61	5.86885246	5.40825113

Level of APPRKNOW	Level of REASON	N	TABDAY Mean	TABDAY SD
0	0	30	7.40000000	5.64831376
0	1	31	5.12903226	4.58069817
1	0	36	4.27777778	3.46913831
1	1	30	6.63333333	6.13347701

Dependent Variable: TABHOUR

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	147.0292806	49.0097602	2.85	0.0401
Error	123	2112.7974910	17.1772154		
Corrected Total	126	2259.8267717			

R-Square	C.V.	Root MSE	TABHOUR Mean
0.065062	71.71071	4.144540	5.779528

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	51.61092466	51.61092466	3.00	0.0855
REASON	1	0.23428170	0.23428170	0.01	0.9072
APPRKNOW*REASON	1	95.18407426	95.18407426	5.54	0.0202

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	46.92532839	46.92532839	2.73	0.1009
REASON	1	0.68676167	0.68676167	0.04	0.8418
APPRKNOW*REASON	1	95.18407426	95.18407426	5.54	0.0202

Level of APPRKNOW	N	TABHOUR Mean	TABHOUR SD
0	61	6.44262295	3.96032234
1	66	5.16666667	4.41529974

Level of REASON	N	TABHOUR Mean	TABHOUR SD
0	66	5.78787879	4.27340668
1	61	5.77049180	4.22844906

Level of APPRKNOW	Level of REASON	N	TABHOUR Mean	TABHOUR SD
0	0	30	7.40000000	3.85602976
0	1	31	5.51612903	3.89761438
1	0	36	4.44444444	4.18462795
1	1	30	6.03333333	4.59747557

Dependent Variable: TDPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.20410362	0.06803454	1.90	0.1327
Error	123	4.39856143	0.03576066		
Corrected Total	126	4.60266505			

R-Square	C.V.	Root MSE	TDPER Mean
0.044345	83.69780	0.189105	0.225938

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.01517615	0.01517615	0.42	0.5160
REASON	1	0.00927064	0.00927064	0.26	0.6116
APPRKNOW*REASON	1	0.17965683	0.17965683	5.02	0.0268

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.02047407	0.02047407	0.57	0.4507
REASON	1	0.00659929	0.00659929	0.18	0.6683
APPRKNOW*REASON	1	0.17965683	0.17965683	5.02	0.0268

Level of APPRKNOW	N	TDPER Mean	TDPER SD
0	61	0.21456705	0.16998488
1	66	0.23644697	0.20953419

Level of REASON	N	TDPER Mean	TDPER SD
0	66	0.21829955	0.19506367
1	61	0.23420197	0.18803408

Level of APPRKNOW	Level of REASON	N	TDPER Mean	TDPER SD
0	0	30	0.24555633	0.17495597
0	1	31	0.18457742	0.16220590
1	0	36	0.19558556	0.21008161
1	1	30	0.28548067	0.20144827

Dependent Variable: THPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.13254091	0.04418030	2.91	0.0373
Error	123	1.86733406	0.01518158		
Corrected Total	126	1.99987497			

R-Square	C.V.	Root MSE	THPER Mean
0.066275	58.93884	0.123214	0.209054

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.03513338	0.03513338	2.31	0.1308
REASON	1	0.01570856	0.01570856	1.03	0.3117
APPRKNOW*REASON	1	0.08169897	0.08169897	5.38	0.0220

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.02897646	0.02897646	1.91	0.1696
REASON	1	0.01325912	0.01325912	0.87	0.3519
APPRKNOW*REASON	1	0.08169897	0.08169897	5.38	0.0220

Level of APPRKNOW	N	THPER Mean	THPER SD
0	61	0.22635398	0.12300455
1	66	0.19306313	0.12751679

Level of REASON	N	THPER Mean	THPER SD
0	66	0.19751873	0.10826286
1	61	0.22153317	0.14257970

Level of APPRKNOW	Level of REASON	N	THPER Mean	THPER SD
0	0	30	0.24179168	0.08587044
0	1	31	0.21141428	0.15054347
1	0	36	0.16062460	0.11214935
1	1	30	0.23198936	0.13561915

Dependent Variable: GDPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.32293866	0.10764622	2.77	0.0446
Error	123	4.78229749	0.03888047		
Corrected Total	126	5.10523615			

R-Square	C.V.	Root MSE	GDPER Mean
0.063256	52.12877	0.197181	0.378258

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.02402903	0.02402903	0.62	0.4333
REASON	1	0.00019498	0.00019498	0.01	0.9437
APPRKNOW*REASON	1	0.29871465	0.29871465	7.68	0.0064

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.01812199	0.01812199	0.47	0.4961
REASON	1	0.00002885	0.00002885	0.00	0.9783
APPRKNOW*REASON	1	0.29871465	0.29871465	7.68	0.0064

Level of APPRKNOW	N	GDPER Mean	GDPER SD
0	61	0.36395029	0.20843763
1	66	0.39148199	0.19511071

Level of REASON	N	GDPER Mean	GDPER SD
0	66	0.38015708	0.17334578
1	61	0.37620348	0.22918593

Level of APPRKNOW	Level of REASON	N	GDPER Mean	GDPER SD
0	0	30	0.31403080	0.16995458
0	1	31	0.41225948	0.23252361
1	0	36	0.43526230	0.15811564
1	1	30	0.33894562	0.22342236

Dependent Variable: TABTIME

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	672852 1972	224284.0657	4.08	0.0084
Error	123	6761146.9115	54968.6741		
Corrected Total	126	7433999.1087			

R-Square	C.V.	Root MSE	TABTIME Mean
0.090510	78.32534	234.4540	299.3335

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	45287.4922	45287.4922	0.82	0.3658
REASON	1	1.5894.8595	115894.8595	2.11	0.1490
APPRKNOW*REASON	1	511669.8456	511669.8456	9.31	0.0028

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	28500.8825	28500.8825	0.52	0.4729
REASON	1	99178.7356	99178.7356	1.80	0.1817
APPRKNOW*REASON	1	511669.8456	511669.8456	9.31	0.0028

Level of APPRKNOW	N	TABTIME Mean	SD
0	61	318.975902	258.697036
1	66	281.179182	227.807687

Level of REASON	N	TABTIME Mean	SD
0	66	269.359561	182.541957
1	61	331.764344	292.820854

Level of APPRKNOW	Level of REASON	N	TABTIME Mean	SD
0	0	30	355.189500	200.640083
0	1	31	283.930484	303.896702
1	0	36	197.834611	130.196787
1	1	30	381.192667	277.291194

Dependent Variable: GRPHTIME

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	270601.1304	90200.3768	2.69	0.0491
Error	123	4120588.0026	33500.7155		
Corrected Total	126	4391189.1330			

R Square	C.V.	Root MSE	GRPHTIME Mean
0.061624	83.22375	183.0320	219.9276

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	76804.5190	76804.5190	2.29	0.1326
REASON	1	14039.1725	14039.1725	0.42	0.5186
APPRKNOW*REASON	1	179757.4389	179757.4389	5.37	0.0222

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	65136.6912	65136.6912	1.94	0.1657
REASON	1	10695.8537	10695.8537	0.32	0.5731
APPRKNOW*REASON	1	179757.4389	179757.4389	5.37	0.0222

Level of APPRKNOW	N	GRPHTIME Mean	SD
0	61	194.347738	133.912746
1	66	243.569636	223.208326

Level of REASON	N	GRPHTIME Mean	SD
0	66	231.284409	224.745006
1	61	207.034541	134.801098

Level of APPRKNOW	Level of REASON	N	GRPHTIME Mean	SD
0	0	30	165.354267	126.951876
0	1	31	222.405935	136.497503
1	0	36	286.235361	271.328728
1	1	30	192.370767	133.594810

Dependent Variable: TDTIME

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	101574.1934	33858.0645	2.36	0.0744
Error	123	1761986.1051	14325.0903		
Corrected Total	126	1863560.2985			

R-Square	C.V	Root MSE	TDTIME Mean
0.054505	114.3882	119.6875	104.6327

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	34.63989	34.63989	0.00	0.9609
REASON	1	2837.62420	2837.62420	0.20	0.6571
APPRKNOW*REASON	1	98701.92927	98701.92927	6.89	0.0098

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	65.48019	65.48019	0.00	0.9462
REASON	1	1774.66941	1774.66941	0.12	0.7255
APPRKNOW*REASON	1	98701.92927	98701.92927	6.89	0.0098

Level of APPRKNOW	N	Mean TDTIME	SD
0	61	105.[illegible]	129.06[illegible]
1	66	104.1306[illegible]	115.[illegible]

Level of REASON	N	Mean TDTIME	SD
0	66	100.068000	113.8[illegible]
1	61	109.571574	130.2[illegible]

Level of APPRKNOW	Level of REASON	N	Mean TDTIME	SD
0	0	30	129.781333	137.082011
0	1	31	81.364290	118.15449[illegible]
1	0	36	75.306889	84.277528
1	1	30	138.719100	137.66[illegible]

Dependent Variable: THTIME

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	286195.4986	95398.4995	3.48	0.0181
Error	123	3373932.3922	27430.3447		
Corrected Total	126	3660127.8908			

R Square	C.V.	Root MSE	THTIME Mean
0.078193	85.06441	165.6211	194.7008

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	42817.1340	42817.1340	1.56	0.2139
REASON	1	82463.1546	82463.1546	3.01	0.0854
APPRKNOW*REASON	1	160915.2101	160915.2101	5.87	0.0169

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	31298.5729	31298.5729	1.14	0.2875
REASON	1	74419.6931	74419.6931	2.71	0.1021
APPRKNOW*REASON	1	160915.2101	160915.2101	5.87	0.0169

Level of APPRKNOW	N	THTIME Mean	THTIME SD
0	61	213.799951	176.590449
1	66	177.048561	163.907082

Level of REASON	N	THTIME Mean	THTIME SD
0	66	169.291561	124.938583
1	61	222.192770	206.429359

Level of APPRKNOW	Level of REASON	N	THTIME Mean	THTIME SD
0	0	30	225.408167	126.509841
0	1	31	202.566194	215.944369
1	0	36	122.527722	103.825502
1	1	30	242.473567	197.706253

Dependent Variable: GDTIME

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	56847.90773	18949.30258	2.35	0.0753
Error	123	990146.72478	8049.97337		
Corrected Total	126	1046994.63251			

R-Square	C.V.	Root MSE	GDTIME Mean
0.054296	70.99557	89.72164	126.3764

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	20049.83629	20049.83629	2.49	0.1171
REASON	1	2.45634	2.45634	0.00	0.9861
APPRKNOW*REASON	1	36795.61510	36795.61510	4.57	0.0345

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	18118.81952	18118.81952	2.25	0.1361
REASON	1	69.70984	69.70984	0.01	0.9260
APPRKNOW*REASON	1	36795.61510	36795.61510	4.57	0.0345

Level of APPRKNOW	N	GDTIME Mean	GDTIME SD
0	61	113.306852	73.574944
1	66	138.455833	103.934036

Level of REASON	N	GDTIME Mean	GDTIME SD
0	66	126.890970	103.723014
1	61	125.819656	76.120312

Level of APPRKNOW	Level of REASON	N	GDTIME Mean	GDTIME SD
0	0	30	95.201967	74.147099
0	1	31	130.827710	69.763976
1	0	36	153.298472	117.671887
1	1	30	120.644667	83.059353

Dependent Variable: GHTIME

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	86680.34639	28893.44880	2.29	0.0819
Error	123	1553642.44659	12631.23940		
Corrected Total	126	1640322.79299			

R-Square	C.V.	Root MSE	GHTIME Mean
0.052843	120.1361	112.3888	93.55122

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	18370.77655	18370.77655	1.45	0.2301
REASON	1	14413 03163	14413.03163	1.14	0.2875
APPRKNOW*REASON	1	53896.53821	53896.53821	4.27	0.0410

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	14547.43400	14547.43400	1.15	0.2853
REASON	1	12492.53344	12492.53344	0.99	0.3219
APPRKNOW*REASON	1	53896.53821	53896.53821	4.27	0.0410

Level of APPRKNOW	N	GHTIME Mean	GHTIME SD
0	61	81.040885	78.509643
1	66	105.113803	138.792938

Level of REASON	N	GHTIME Mean	GHTIME SD
0	66	104.398439	134.220198
1	61	81.814885	86.907600

Level of APPRKNOW	Level of REASON	N	GHTIME Mean	GHTIME SD
0	0	30	70.152300	64.182359
0	1	31	91.578226	90.067571
1	0	36	132.936889	167.929681
1	1	30	71.726100	83.831176

Dependent Variable: TDTPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.16559898	0.05519966	2.62	0.0535
Error	123	2.58684793	0.02103128		
Corrected Total	126	2.75244692			

R-Square	C.V.	Root MSE	TDTPER Mean
0.060164	101.7424	0.145022	0.142538

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.00278405	0.00278405	0.13	0.7166
REASON	1	0.01567451	0.01567451	0.75	0.3897
APPRKNOW*REASON	1	0.14714042	0.14714042	7.00	0.0092

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.00354100	0.00354100	0.17	0.6823
REASON	1	0.01923313	0.01923313	0.91	0.3408
APPRKNOW*REASON	1	0.14714042	0.14714042	7.00	0.0092

Level of APPRKNOW	N	TDTPER Mean	SD
0	61	0.13766794	0.14352656
1	66	0.14703932	0.15260156

Level of REASON	N	TDTPER Mean	SD
0	66	0.15344463	0.16658703
1	61	0.13073760	0.12465092

Level of APPRKNOW	Level of REASON	N	TDTPER Mean	SD
0	0	30	0.18490594	0.16497748
0	1	31	0.09195374	0.10246246
1	0	36	0.12722688	0.16562274
1	1	30	0.17081426	0.13422838

Dependent Variable: THTPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.28977040	0.09659013	3.85	0.0112
Error	123	3.08225763	0.02505901		
Corrected Total	126	3.37202803			

R-Square	C.V.	Root MSE	THTPER Mean
0.085934	59.98396	0.158300	0.263905

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.11081361	0.11081361	4.42	0.0375
REASON	1	0.00341732	0.00341732	0.14	0.7126
APPRKNOW*REASON	1	0.17553947	0.17553947	7.01	0.0092

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.09875298	0.09875298	3.94	0.0494
REASON	1	0.00190149	0.00190149	0.08	0.7834
APPRKNOW*REASON	1	0.17553947	0.17553947	7.01	0.0092

Level of APPRKNOW	N	THTPER Mean	THTPER SD
0	61	0.29463020	0.16722633
1	66	0.23550651	0.15607374

Level of REASON	N	THTPER Mean	THTPER SD
0	66	0.25740116	0.15150677
1	61	0.27094090	0.17673836

Level of APPRKNOW	Level of REASON	N	THTPER Mean	THTPER SD
0	0	30	0.32858116	0.12659428
0	1	31	0.26177443	0.19536947
1	0	36	0.19808450	0.14624927
1	1	30	0.28041292	0.15798447

Dependent Variable: GDTPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.16790569	0.05596856	3.38	0.0205
Error	123	2.03747956	0.01656487		
Corrected Total	126	2.20538525			

R-Square	C.V.	Root MSE	GDTPER Mean
0.076134	64.16140	0.128705	0.200595

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.01480215	0.01480215	0.89	0.3464
REASON	1	0.00313327	0.00313327	0.19	0.6644
APPRKNOW*REASON	1	0.14997027	0.14997027	9.05	0.0032

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.01225901	0.01225901	0.74	0.3913
REASON	1	0.00484895	0.00484895	0.29	0.5895
APPRKNOW*REASON	1	0.14997027	0.14997027	9.05	0.0032

Level of APPRKNOW	N	GDTPER Mean	GDTPER SD
0	61	0.18936538	0.13124850
1	66	0.21097402	0.13341739

Level of REASON	N	GDTPER Mean	GDTPER SD
0	66	0.19638361	0.12066228
1	61	0.20515173	0.14471742

Level of APPRKNOW	Level of REASON	N	GDTPER Mean	GDTPER SD
0	0	30	0.14804063	0.09425953
0	1	31	0.22935707	0.15001086
1	0	36	0.23666942	0.12650938
1	1	30	0.18013954	0.13703248

Dependent Variable: GHTPER

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.13715132	0.04571711	2.51	0.0616
Error	123	2.23672726	0.01818477		
Corrected Total	126	2.37387858			

R-Square	C.V.	Root MSE	GHTPER Mean
0.057775	103.9636	0.134851	0.129710

Source	DF	Type I SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.00603412	0.00603412	0.33	0.5656
REASON	1	0.02102812	0.02102812	1.16	0.2843
APPRKNOW*REASON	1	0.11008908	0.11008908	6.05	0.0153

Source	DF	Type III SS	Mean Square	F Value	Pr > F
APPRKNOW	1	0.00336534	0.00336534	0.19	0.6678
REASON	1	0.01773891	0.01773891	0.98	0.3253
APPRKNOW*REASON	1	0.11008908	0.11008908	6.05	0.0153

Level of APPRKNOW	N	GHTPER Mean	GHTPER SD
0	61	0.12253990	0.11869971
1	66	0.13633650	0.15304435

Level of REASON	N	GHTPER Mean	GHTPER SD
0	66	0.14241812	0.14132876
1	61	0.11595979	0.13250085

Level of APPRKNOW	Level of REASON	N	GHTPER Mean	GHTPER SD
0	0	30	0.10457628	0.08774212
0	1	31	0.13992405	0.14178227
1	0	36	0.17395298	0.16871584
1	1	30	0.09119672	0.11949938

Appendix E:

Glossary

GLOSSARY

"Attribution	The process by which perceivers assign specific causes to observed events. *Causal* attribution is the assignment of causal responsibility for an event to some aspect of an actor or to an environment feature; *trait* attribution is a categorization of the causal agent in terms of groups of agents sharing similar characteristics (e.g., *hostile* people or *difficult* tasks).
Automatic process	A cognitive or behavioral process occurring without conscious monitoring or awareness (Kimble & Perlmuter, 1970; Langer, 1978; Nisbett & Wilson, 1977; Shevrin & Dickman, 1980; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Thorngate, 1976).
Category	A "fuzzy set" of objects considered equivalent, in which membership in the set is defined by family resemblance rather than by possession of all attributes considered to be critical (Cantor & Mischel, 1979; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976). Categorization is considered basic to perception, information storage, and organization (Bruner, 1958).
Controlled process	A cognitive or behavioral process that proceeds under conscious control, about which the person is continuously aware.
Implicit personality theory/Personal constructs	The set of categories and their associated prototypes/schemata used by an individual to represent, evaluate, and predict the behavior of others.
Information integration	The combination of two or more separate items of information into a summary judgment or evaluation (Anderson, 1974; 1976).
Prototype	An abstract *analog*, or *image*, summarizing "central tendencies" or resemblances among category members (Cohen, 1979; Rosch et al., 1976, Rosch, 1977). A prototype is the cognitive representation of a typical category member, and examplifies the category in memory (see also *Schema*)

GLOSSARY (Continued)

Salience	*Prominence* of a stimulus in the perception field. In memory it is the availability of a cognition of a memory trace (Taylor & Fiske, 1978).
Schema	Verbal or propositional memory structures representing categories of a more complex nature than prototypes but serving the same function in a perceptual organization and memory. Schemata are used to represent the self (Markus, 1977) and well-known other people as well as familiar situations and objects.
Stereotype	A subtype of implicit personality theory, specifically one based on racial, sexual, ethnic, or occupational categories rather than trait categories" (Ilgen & Feldman, 1983. p152).