

1989

BUILDING AND TESTING A CAUSAL MODEL OF AN INFORMATION TECHNOLOGY'S IMPACT

Rebecca A. Grant
University of Cincinnati

Follow this and additional works at: <http://aisel.aisnet.org/icis1989>

Recommended Citation

Grant, Rebecca A., "BUILDING AND TESTING A CAUSAL MODEL OF AN INFORMATION TECHNOLOGY'S IMPACT" (1989). *ICIS 1989 Proceedings*. 44.
<http://aisel.aisnet.org/icis1989/44>

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 1989 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

BUILDING AND TESTING A CAUSAL MODEL OF AN INFORMATION TECHNOLOGY'S IMPACT

Rebecca A. Grant

Department of Quantitative Analysis and Information Systems
University of Cincinnati

ABSTRACT

An increasing number of firms are adopting Computerized Performance Monitoring and Control Systems (CPMCS) in an effort to improve the productivity of employees in labor-intensive service industries. The service sector has not historically used volume of output or similar quantitative measures of performance when evaluating employees. Thus, monitoring often represents a new evaluation method and a new application of information technology. It is an application prone to controversy: Proponents claim it improves measurement accuracy, fairness and consistency, while opponents argue that it degrades the quality of work life, increases stress and undermines customer service.

Despite the need to understand the impact of CPMCS, there have been few attempts to predict what effects can be anticipated and explain how these effects arise. The methodology described in this paper was used to integrate existing anecdotal work with literature from reference disciplines to build a conceptual model of CPMCS impact on role definition. The three-phased research then used an intensive case study to build a theory of impact and generate testable, causal research hypotheses. Subsequent to this theory-building stage, surveys from 1500 service workers provided data to test the causal model.

The research produced three outcomes. First, it combined theory building and theory testing in a study of information technology impact to give structure and direction to a field characterized by anecdotal research. Second, it provided two, tested causal models of CPMCS impact with good explanatory and predictive power. These models explained the influence of monitor design on attitudes toward production and customer service. Third, it demonstrated the use of a holdout technique to increase the amount of knowledge gained in the hypothesis testing stage of empirical research.

1. INTRODUCTION

An increasing number of firms are adopting Computerized Performance Monitoring and Control Systems (CPMCS) in an effort to improve the productivity of employees in labor-intensive service industries. These systems use a wide variety of computer software, hardware, and surveillance technology to sense and record information about computer-mediated employee performance. Simple systems store summary counts of completed transactions, while more complex systems continuously track and report on activity.

The service sector has not historically used automated systems to track employee performance. Thus, monitoring often represents a new evaluation method and a new application of information technology. It is an application prone to controversy: Supporters claim it improves measurement accuracy, fairness and consistency (Smith, Carayon and Miezio 1987), while opponents argue that it degrades the quality of work life, increases stress and undermines customer service (Garson 1988; Gregory and Nussbaum 1982).

Despite the need to understand the impact of CPMCS, there have been few attempts to predict what effects can be anticipated and to explain how these effects arise. Most empirical research has focused on health and privacy issues (Nussbaum 1984; U.S. Congress, OTA 1987; Westin 1987). The few works examining the impact of monitoring on performance or attitudes toward work have relied primarily on case studies or exploratory designs (Eisenman 1987; Irving, Higgins and Safayeni 1986; Walton and Vittori 1983). The exploratory studies suggested links between monitoring and certain work attitudes, but did not use designs that tested causal relationships. It is not clear how far one can generalize their findings.

Theory building and theory testing are continuous phases of the research process (Zaltman, LeMasters and Heffring 1982). The inductive process of integrating new observations with existing theory produces new theory. Subsequent deductive reasoning leads to hypotheses, which can be tested empirically and which then provide input to further theory development. It is common in Information Systems (IS) research to see each phase of the larger process carried out in separate studies, often by different

researchers. Our research, however, developed *and* tested a theory of CPMCS impact.

This paper describes how reference disciplines, exploratory research and a case study were used to build causal theory. It discusses the results of a field survey of 1500 employees, which tested the theory, and concludes by examining the contributions of this research to studying the impact of CPMCS and other new information technologies.

2. BUILDING A CONCEPTUAL MODEL

The first stage of this research asked the general question "How does a CPMCS affect the productivity and customer service performance of service workers?" Both the popular press (Oreskovich 1985; Koepf 1986) and the academic literature (Gregory and Nussbaum 1982; Walton and Vittori 1983; Irving, Higgins and Safayeni 1986) contend that productivity improvements resulting from monitor use came at the expense of customer service. These contentions had never been tested. Furthermore, no prior research proposed a model to explain how this might happen and how it might be avoided.

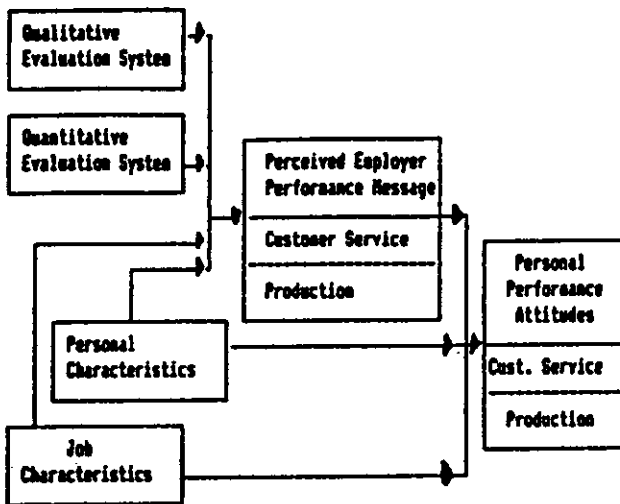


Figure 1. Conceptual Model

Theoretical and empirical precedents in organizational behavior and control systems (Ilgen, Fisher and Taylor 1979; Lawler 1976; Lawler and Rhode 1976; Smith, Carayon and Miezio 1987; Taylor, Fisher and Ilgen 1984) suggested that an employee's attitude toward the importance of productivity and customer service would be a critical factor in determining actual performance in these areas. Our study called these "Personal Performance Attitudes" (see Figure 1). They would evolve from three elements: (1) the importance employers seem to attach to production and service ("Perceived Employer Performance Message"); (2) personal characteristics of the employee ("Personal Characteristics"); and (3) the characteristics of the job ("Job Characteristics").

The "Perceived Employer Performance Message," in turn, seemed to be a function of the design and use of the qualitative and quantitative evaluation systems, as well as of characteristics of the employee and of the job. The "Quantitative Evaluation System" refers to systems which use quantitative measures of performance to evaluate the employee. In firms which monitor, this would include the CPMCS and its use. In other firms, a system of manual measures of quantified performance might comprise the system. The "Qualitative Evaluation System" incorporates the subjective or nonquantified evaluation mechanisms. These might include direct observation, comments from co-workers and customers, and random review of work quality.

This literature from areas outside computer monitoring research suggested a broad, conceptual framework. We were specifically concerned with the impact of CPMCS design and use in this framework. This led us to focus on the role of the "Quantitative Evaluation System." Four dimensions of the design and use of evaluation systems seemed important: the choice of tasks to monitor (or evaluate), the frequency with which data were collected, the audience for the data, and the object (individual or work group) of measurement. We posited that they would be the key independent variables of the final research model. We then used Lawler's (1976) adaptation of thermostat control models to translate these static elements of CPMCS design into the dynamic processes of control needed for a causal study of impact.

Finally, the literature suggested that the use of a *computer*, rather than a human supervisor, to collect data about performance could be an important feature of the quantitative evaluation systems (Turkle 1984; Whisler and Shultz 1960; Zuboff 1982). Such a deterministic view was the subject of conflicting research results. However, the measuring device is often the single element which distinguishes CPMCSs from manual quantitative systems. Thus, we chose to include the perceived credibility and appropriateness of computers as aspects of CPMCS design which could moderate impact, much as one would include the credibility and authority of the human supervisor in studies of qualitative evaluation (Ilgen, Fisher and Taylor 1979).

3. THE CASE STUDY

This review produced the conceptual model of Figure 1, as well as indications of lower level constructs apt to explain the impact of CPMCS. There was little CPMCS research to explicitly support the choice of the lower level constructs. The constructs of monitor design and use were essentially new, since they had never been tested within the context of CPMCS impact. Furthermore, few of the other constructs had been tested via instruments with demonstrated reliability and validity. This convinced us to conduct an initial exploratory study. Such work was needed to ensure that the research model captured the

significant constructs and relationships and helped to develop a reliable survey instrument for hypothesis testing.

We wanted to establish whether and how unobtrusive monitoring effects differed from the effects of manual quantitative systems. This objective addressed concerns about research and anecdotes which attributed worker attitudes to monitoring. Prior research had not compared monitored workers to those working under a manual, but still quantitative, evaluation system. It was entirely possible that the findings actually reflected the effects of quantitative evaluation in the service sector, rather than results of monitoring per se. This desire to separate the effects of the evaluation system (a monitor) from those of the evaluation criteria (quantified performance) was instrumental in designing the exploratory work.

3.1 Case Study Design

We chose to use a case study design for the exploratory phase for several reasons. It enabled us to study the complex evaluation process and the work environment as it occurred. This revealed causal factors which were not clearly identified in the conceptual model. It provided a rich base of contextual data, which proved vital in interpreting unexpected findings from the survey and interviews. We could combine extensive qualitative data gathering with the opportunity to hold several variables (most notably job content and CPMCS design) constant.

Fifteen supervisors and 81 employees of a major insurance company participated in the case study. The non-supervisory employees comprised the staff of 12 units, nine of which processed claims. Each unit had its own supervisor. A CPMCS counted individual daily claims processed for staff in eight of the claims units, while a supervisor collected identical data for processors in the ninth unit.

Each participant answered open-ended questions during a 45-minute individual interview and completed a survey based on the conceptual model presented in Figure 1. The research was conducted one month after annual reviews, so that all participants were aware of their performance ratings and the evaluation process. We also collected actual productivity, customer service and overall evaluation data for the participants. We used this evaluation data to test how well perceived importance of productivity and customer service corresponded to their actual role in determining performance ratings.

3.2 Effects of Monitoring

The first objective of the case study was to determine whether or not computerized monitoring had a different impact on attitudes than manual performance measurement. The case study showed that differences did exist between monitored and manually measured groups.

All employees believed that production (measured as average claims processed per day) was an important part of their evaluation. However, monitored employees discussed productivity in personal terms. Every monitored employee emphasized the importance of meeting individual production quotas, while unmonitored employees did not believe they had an individual quota. (In fact, all employees were subject to the same minimum quota.) Unmonitored employees referred to the importance of *unit* performance and their contribution to it. Regressing overall performance ratings onto production, attendance, and customer complaints demonstrated that production was actually a hygiene factor in this company. While processors had to meet the 65 claims per day quota for a satisfactory rating, exceeding the quota did not contribute to a higher rating. Thus, we could not attribute the difference in perceptions in actual performance ratings.

Second, monitored employees tended to describe customer service in terms of fast and accurate turnaround of claims. These were the performance dimensions most directly measured by quantitative systems. Yet, employees whose counts were kept manually emphasized courtesy in handling telephone calls, willingness to spend extra time finding an answer to a customer's question, and helping a co-worker with a difficult problem.

3.3 Reflections on the Conceptual Model

The interviews also pinpointed areas where the conceptual model needed refinement. First, interviews revealed that employees differed greatly in terms of their perceptions of the measurement system. The interviews showed that employees' perceptions did not necessarily match the actual characteristics or use of the CPMCS and that these perceptions seemed to differ between units. This supported other evidence that the supervisor and various elements of the qualitative system played a major role in determining reactions to monitoring. Second, the case study suggested that the degree to which the job was perceived as "quantitative" was the only job characteristic which belonged in the final model. Third, the interviews and preliminary surveys did support a number of the proposed relationships. Foremost among these were the links between use of monitoring and attitudes toward the importance of productivity and service.

3.4 Survey Design

The case study contributed to a number of revisions in the final survey design and the design of the low level, testable research model. Factor analysis of responses to the survey demonstrated the existence of "Acceptance of Quantitative Measures" and "Perceived Reliance on Quantitative Measures" as components of perceptions of the monitoring system. Survey responses also gave evidence of relationships among these components, the employer performance message, and attitudes toward importance of production and customer service.

The case study indicated problems with the reliability and validity of some scales in the survey. The Cronbach's alphas of the computer credibility and appropriateness scales were unacceptably low (less than 0.70), as was the reliability of the "Personal Importance of Customer Service" scale. During interviews, it became clear that researchers and participants defined service differently, resulting in the low reliability of the personal importance scale. We revised all three scales extensively prior to hypothesis testing.

The case study survey also failed to ask employees whether they had a customer service job. We used management's description of the position and observed the amount of direct customer contact to confirm that participants held a service job. Interviews showed that this view was not necessarily shared by monitored participants. This difference in job definition seemed to be an important indicator of the perceived importance of production versus service. We revised the survey to include measures of this perception.

Finally, the section of the case study survey dealing with acceptance of and reliance on quantitative measures received attention. Scale reliability was high, but participants described the five pages of questions as tedious and repetitive. To avoid an instrumentation effect (Cook and Campbell 1979) and make the survey more appealing, we reformatted this section.

4. HYPOTHESIS TESTING

4.1 The Hypotheses

Qualitative data from the exploratory study led to a more explicit causal model and an improved research instrument. This data also helped explain apparent anomalies in case study survey data. The quantitative and qualitative contributions were next integrated into the testable causal model of CPMCS impact and the survey used to test it.

Figure 2 presents the causal model developed from the literature review and the findings of the case study. This model represents a bounded version of the conceptual model. The qualitative system, for example, can be described by the same dimensions as the quantitative system. However, we were most interested in the impact of monitoring. We set the boundary of our model to include the role that acceptance of and reliance on qualitative evaluation plays in affecting attitudes, but to exclude the paths representing how the design of the system leads to acceptance and reliance. We also included perceptions of one job characteristic ("Quantitative Nature of Work"), but excluded others from the survey. Finally, we excluded the evaluation messages -- that is, the precise wording of supervisor comments or content of CPMCS reports. In a field survey sent to hundreds of workers in different jobs and companies, it was virtually impossible to gather and encode actual evaluation comments.

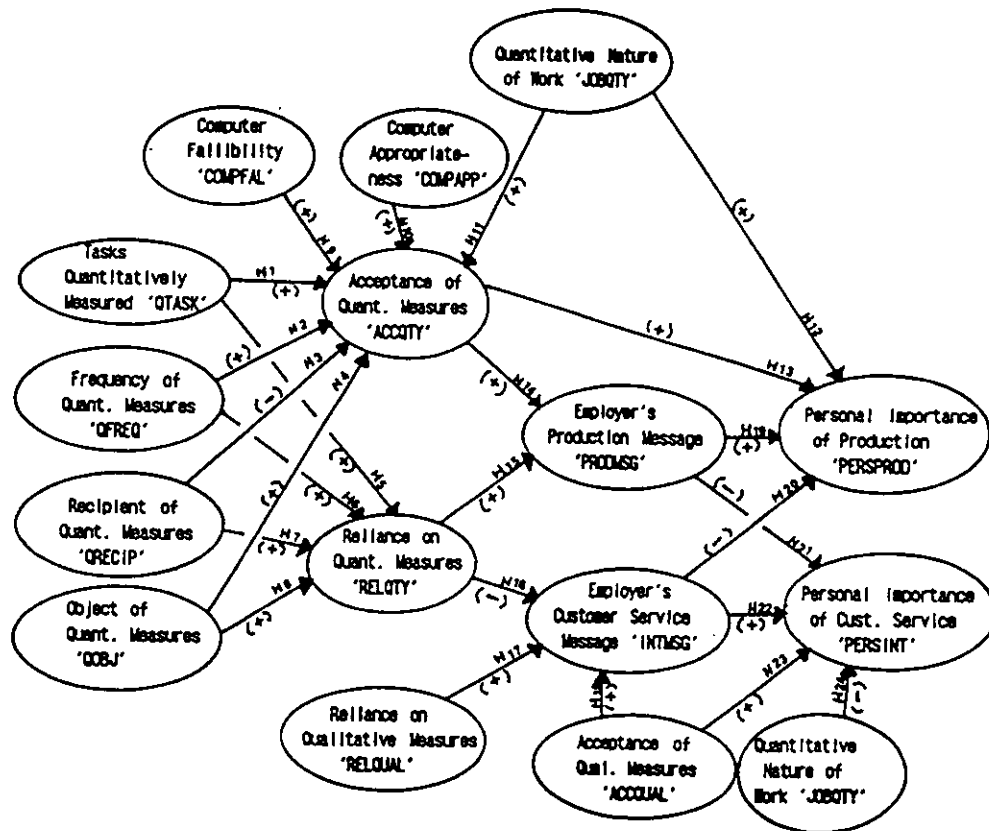


Figure 2. Original Causal Model

Tasks Qualitatively Measured (QTASK): The degree to which the employee believes the range of work activities is quantitatively measured, regardless of the measurement device.

Frequency of Quantitative Measures (QFREQ): The frequency with which quantitative performance data are available.

Recipient of Quantitative Measures (QRECIP): The degree to which performance data are made available to various parties in the workplace, ranging from employee only to posting for all staff to see.

Object of Quantitative Measures (QOBJ): The object of data collection – individual, group, department, etc.

Computer Fallibility (COMPFAL): The employee's belief that computers produce accurate information, independent of the application or environment.

Computer Appropriateness (COMPAPP): The employee's belief that a computer is an appropriate device for measuring his or her performance.

Quantitative Nature of Work (JOBQTY): The employee's perception of the degree to which the job is quantitative and routine in nature.

Acceptance of Qualitative Measures (ACCQUAL): The employee's perception of the accuracy, completeness and appropriateness of qualitative evaluation system data.

Reliance on Qualitative Measures (RELQUAL): The employee's perception of the degree to which the employer relies on qualitative system data in assessing performance.

Acceptance of Quantitative Measures (ACCQTY): The employee's perception of the accuracy, completeness and appropriateness of quantitative evaluation system data.

Reliance on Quantitative Measures (RELQTY): The employee's perception of the degree to which the employer relies on quantitative system data in assessing performance.

Employer's Production Message (PRODMSG): The employee's assessment of the absolute importance of production factors to the employer.

Employer's Customer Service Message (INTMSG): The employee's assessment of the absolute importance of customer service and non-production factors to the employer.

Personal Importance of Production (PERSPROD): The absolute importance the employee attaches to production factors as part of the job.

Personal Importance of Customer Service (PERSINT): The absolute importance the employee attaches to service and non-production factors as part of the job.

Figure 3. Definitions of Model Constructs

This boundary-setting meant that the causal model being tested was unlikely to explain *all* of the variance in the two dependent variables. However, we set the boundaries to include the key features of CPMCS design and thus to test their role in explaining the dependent variables. Figure 3 defines the constructs comprising the causal model.

Causal models, such as the one depicted in Figure 2, do not simply specify relationships among variables. They also indicate *direction* or causality embodied in the relationships (Asher 1983). Each path can thus be expressed as a causal hypothesis, where its sign indicates whether the relationship is positive or negative. For example, the path labeled H1 represents the hypothesis that:

More extensive quantitative measurement of tasks will lead to greater acceptance of quantitative measures.

As a result, testing the model tested the direction and strength of all paths in the network of relationships.

4.2 Survey and Data Analysis Methodology

The model testing stage of our research used data from a national survey of non-supervisory workers. Employers provided the names of 2,692 individuals, to whom we mailed personally addressed surveys. We received 1,498 usable responses (55.7 percent). The responses included surveys from employees in 50 companies and 14 industries (ranging from travel and tourism to financial services and government agencies). Respondents varied widely in terms of age, work experience, and evaluation systems used to measure their jobs.

We used a "holdout technique" with these surveys to test two versions of the proposed model. Each survey was assigned to one of two data sets, based on its identification number. The set of even-numbered surveys were used to test the research model of Figure 2. This test suggested areas in which the model could be improved, as discussed below. However, one should not use the same data to test subsequent models (Pedhazur 1982). The odd-numbered surveys, therefore, were used to test a revised model. There was no pattern to odd- and even-numbered survey distribution, so using these numbers to split responses was

virtually a random assignment. We did confirm, however, that there were no significant differences between the two sets in terms of independent variables and demographic characteristics.

The holdout technique simulates two complete hypothesis testing projects. It also provides the advantage of greater confidence when comparing results. Since all respondents received identical surveys during the same period of time, came from the same populations, and were randomly assigned to subsets, this method controlled for many of the validity threats inherent in comparing groups in quasi-experimental research.

4.3 Testing the Original Causal Model

We used Partial Least Squares (PLS) (Wold 1982) to analyze our research model. PLS is a second-generation, multivariate technique used to estimate the parameters of predictive-causal models. It focuses on the interrelationship of theory and data, while letting researchers specify which of the two will play a stronger role in estimating model parameters (Fornell 1984). This flexibility makes it particularly useful for early theory testing.

Second-generation techniques recognize two components of a causal model: the measurement model and the structural model. The measurement model consists of the relationships between the constructs and the items used to measure them. The characteristics of this model demonstrate the convergent and discriminant validity of the research instruments. The structural model consists of the unobservable constructs and the theoretical relationships among them (the paths). Together, the structural and measurement models form a network of constructs and measures. The estimated path coefficients indicate the strength and sign of the theoretical relationships, while the survey item weights and loadings indicate the strength of measures used to test the theory.

Fornell and Larcker (1981) recommended using three tests to assess convergent validity. "Item reliability" indicates the amount of variance in a measure due to the construct rather than to error. "Composite reliability" is analogous to Cronbach's alpha in measuring the overall reliability of a scale. "Average variance extracted" is the most conservative of the tests. It measures the amount of variance in the item explained by the construct relative to the amount due to measurement error. We used all three tests.

In general, the convergent validity of our survey measures was strong. Average variance extracted for all constructs exceeded 0.50, and the composite reliability of all scales exceeded the recommended 0.70 level (Fornell and Larcker 1981). All but two of the survey items loaded at, or above, 0.70 on their respective constructs. These results were not unexpected, since the case study had been used to extensively test and refine the survey.

Discriminant validity is the degree to which items differentiate between constructs, or measure distinct concepts. Poor discriminant validity reduces the ability to say whether two constructs represent a cause and effect or are actually the same construct. To assess it, one examines the correlations between constructs. The squared correlations (the shared variance) should be less than the average variance extracted by the items measuring the constructs. In other words, constructs should correlate more highly with their measures than with other constructs in the model (Fornell and Larcker 1981). As shown in Figure 4, the survey discriminated adequately between proposed cause and effect constructs.

Two correlations (underlined in Figure 4) caused some concern. As a result of a measurement artifact, the shared variance between "Recipient of Quantitative Measures" (QRECIP) and "Object of Quantitative Measures" (QOBJ) exceeded the average variance extracted for QRECIP. The high correlation between these two constructs suggested that one or the other could be dropped from the model without reducing its predictive power. There was also a high correlation between "Personal Importance of Production" (PERSPROD) and "Personal Importance of Customer Service" (PERSINT). They may have been highly correlated because of a methods effect, the employees' inability to distinguish between job factors in assessing importance, or the fact that attitudes toward production and interaction importance are closely related. Case study interview comments and observations suggested that the third explanation was the most likely. Since we did not propose a causal relationship in either of these cases, the multicollinearity did not threaten causal conclusions (Asher 1983).

Figure 5 depicts the path coefficients estimated by the PLS test of the research model originally proposed. To confirm or refute a hypothesis, the coefficient must be significantly different from zero. We used a nonparametric test of significance known as jackknifing (Fornell and Barclay 1988) to test the estimated coefficients. All but two paths (indicated by "***" in Figure 5) were statistically significant at $\alpha < 0.005$. This confirmed the *existence*, although not always the direction, of the hypothesized relationships.

The survey data supported much of the causal model originally proposed. Furthermore, the results demonstrated strong predictive and reasonable explanatory power. We did not stop with this test, however. Instead, we used the information it provided to revise and improve the model.

4.4 Revising the Model

One should not alter a causal model strictly on the basis of data (Asher 1983). A single test of any model could produce results which contradict the hypotheses. This need not mean that the theory underlying the model was flawed.

	QTASK	QFREQ	QRECIP	QOBJ	COMPFAL	COMPAPP						
QTASK	1.00											
QFREQ	0.24	0.72										
QRECIP	0.20	0.22	0.74									
QOBJ	0.18	0.22	0.79	1.00								
COMPFAL	0.42	0.13	0.07	0.07	1.00							
COMPAPP	0.41	0.22	0.07	0.06	0.59	1.00						
JOBQTY	0.00	0.13	0.00	0.01	0.00	0.05						
REQUAL	0.02	0.10	0.01	0.02	0.01	0.01						
ACCQUAL	0.00	0.00	0.00	0.00	0.01	0.00						
RELTQY	0.17	0.02	0.18	0.18	0.06	0.07						
ACCQTY	0.00	0.00	0.01	0.02	0.00	0.02						
PRODMSG	0.01	0.26	0.03	0.03	0.01	0.01						
INTMSG	0.00	0.00	0.00	0.00	0.00	0.00						
PERSPROD	0.00	0.00	0.00	0.00	0.00	0.00						
PERSINT	0.00	0.01	0.00	0.00	0.00	0.00						

	JOBQTY	REQUAL	ACCQUAL	RELTQY	ACCQTY	PRODMSG	INTMSG	PERSPROD	PERSINT
JOBQTY	0.06								
REQUAL	0.01	0.55							
ACCQUAL	0.08	0.03	0.76						
RELTQY	0.02	0.05	0.00	0.67					
ACCQTY	0.10	0.01	0.58	0.03	0.07				
PRODMSG	0.03	0.00	0.06	0.02	0.12	0.62			
INTMSG	0.07	0.02	0.23	0.00	0.19	0.52	0.61		
PERSPROD	0.10	0.01	0.14	0.00	0.11	0.30	0.44	0.58	
PERSINT	0.02	0.01	0.06	0.00	0.04	0.41	0.45	0.58	0.61

Figure 4. Discriminant Validity -- Original Causal Model

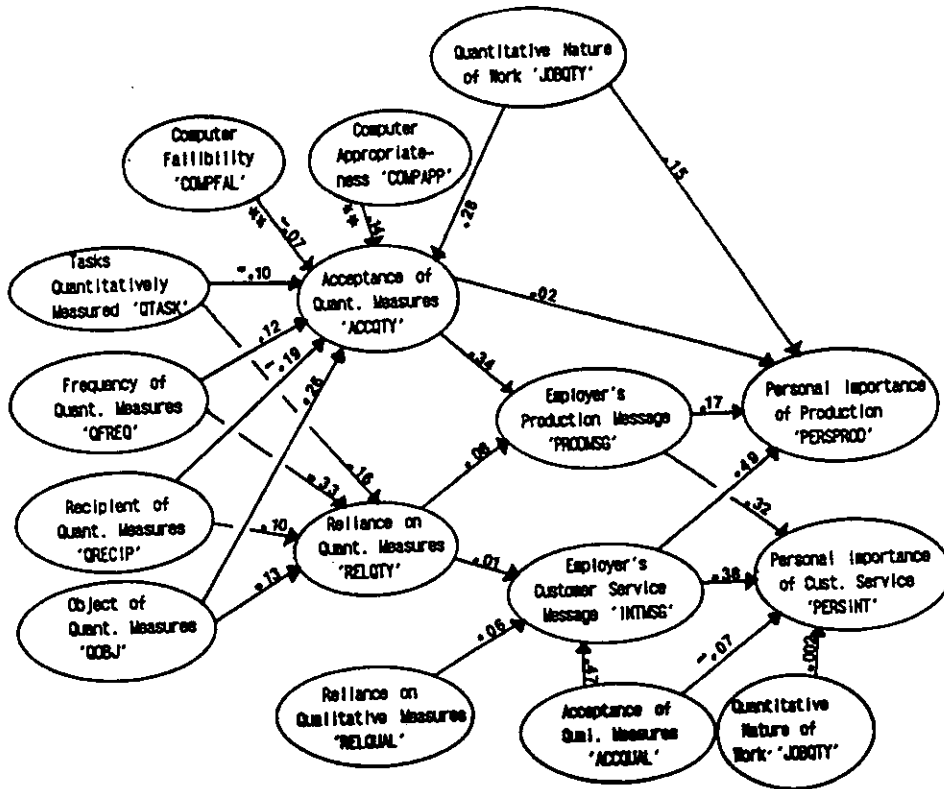


Figure 5. Path Coefficients -- Original Causal Model

There were noteworthy results which argued for revision of our model, however. They demonstrated areas where revisions would produce a more parsimonious model and where improved measurement would produce stronger results.

Survey data appeared to support the hypotheses related to "Recipient of Quantitative Measures" (QRECIP) and

"Object of Quantitative Measures" (QOBJ). However, the correlation ($r = .89$) between QOBJ and QRECIP showed that the survey did not discriminate between these constructs well. Thus, we could not interpret the results reliably. Dropping the two constructs entirely would have solved the problems resulting from their multicollinearity, but it also would have eliminated the information captured in the data collected for these two constructs. We chose

to merge them into a new construct, "Object and Recipient of Quantitative Measures" (QCOMB) (Asher 1983). In general, the more specific the target of measurement and the more public the results, the greater the value of QCOMB.

The literature (Lawler 1976; Lawler and Rhode 1979; Walton and Vittori 1983) indicated that "Recipient of Quantitative Measures" (QRECIP) and "Object of Quantitative Measures" (QOBJ) would have opposite effects on "Acceptance of Quantitative Measures" (ACCQTY). Therefore, the combination of the two would not have a predetermined effect. As a result, the revised model did not propose a path from QCOMB to ACCQTY. However, we had expected both QRECIP and QOBJ to increase "Reliance on Quantitative Measures" (RELQTY) (Lawler 1976). The revised model thus included a positive path from QCOMB to RELQTY.

"Computer Fallibility" also demonstrated measurement problems. It seemed most unlikely that a monitored employee who doubted computer accuracy would be more accepting of CPMCS measures. Thus, finding a negative path coefficient from "Computer Fallibility" (COMPFAL) to "Acceptance of Quantitative Measures" (ACCQTY) suggested a design error. The research had not captured the concept of fallibility well. Furthermore, the path from

"Computer Appropriateness" (COMPAPP) to "Acceptance of Quantitative Measures" (ACCQTY) was not statistically significant. Since COMPFAL and COMPAPP were correlated at $r=.77$, it was likely that the poorly measured COMPFAL had contaminated the model and confounded the effect of constructs on ACCQTY. As a result of these two problems, we decided to drop COMPFAL completely. The role of "Computer Fallibility" remains an area for future research.

The signs of three paths contradicted expectations: from "Tasks Quantitatively Measured" (QTASK) to "Acceptance of Quantitative Measures" (ACCQTY); from "Employer's Production Message" (PRODMSG) to "Personal Importance of Customer Service" (PERSINT); and from "Employer's Service Message" (INTMSG) to "Personal Importance of Production" (PERSPROD). The original hypotheses had been drawn largely from anecdotal evidence of CPMCS effects. Case study and contradictory survey results implied that the anecdotes could *not* be broadly generalized. Together, they made a strong case for hypotheses incorporating the signs indicated by the test of the first model.

Finally, mere statistical significance may not be grounds for leaving a path in the model. With such large samples, PLS

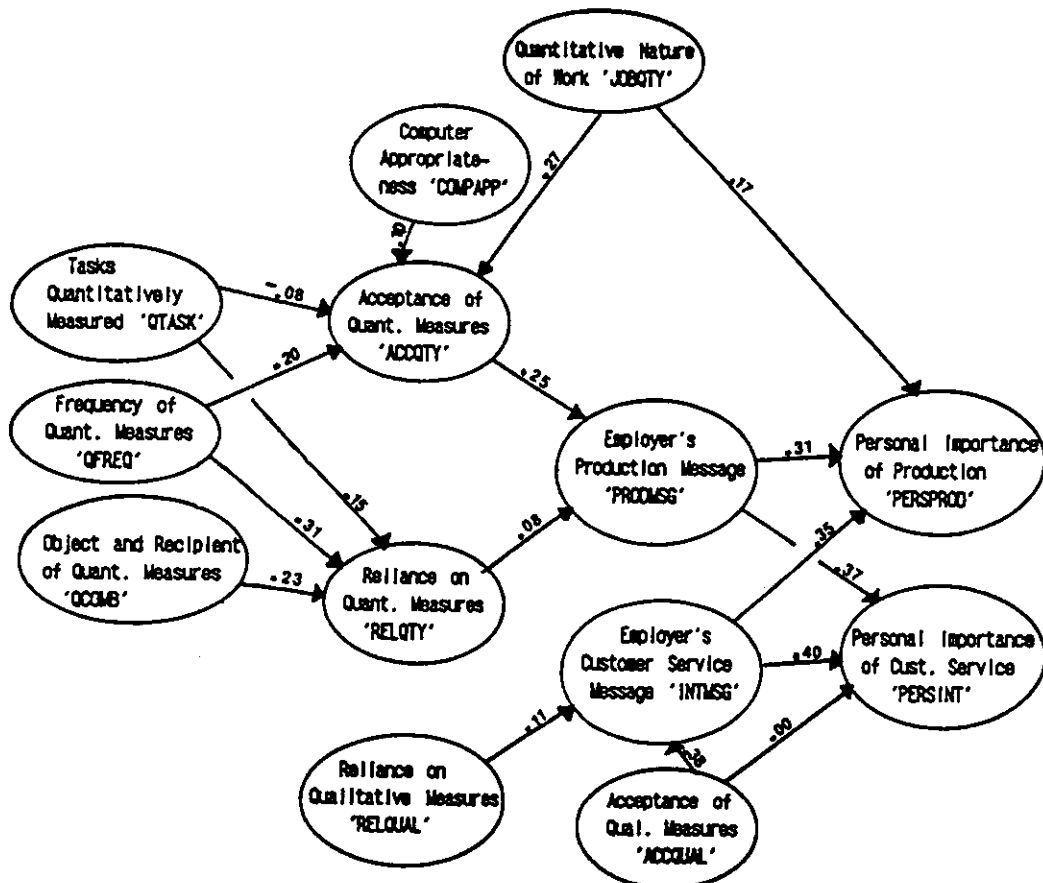


Figure 6. Path Coefficients: Revised Model

	QTASK	QFREQ	QCOMB	COMPAPP				
QTASK	1.00							
QFREQ	0.17	0.69						
QCOMB	0.21	0.15	0.79					
COMPAPP	0.45	0.08	0.08	1.00				
JOBQTY	0.12	0.00	0.00	0.06				
RELQUAL	0.12	0.01	0.00	0.00				
ACQUAL	0.01	0.00	0.00	0.00				
RELQTY	0.14	0.21	0.17	0.08				
ACCQTY	0.01	0.04	0.01	0.03				
PRODMMSG	0.00	0.01	0.01	0.01				
INTMSG	0.01	0.00	0.00	0.01				
PERSPROD	0.00	0.00	0.00	0.01				
PERSINT	0.00	0.01	0.00	0.01				

	JOBQTY	RELQUAL	ACQUAL	RELQTY	ACCQTY	PRODMMSG
JOBQTY	0.58					
RELQUAL	0.01	0.50				
ACQUAL	0.06	0.01	0.78			
RELQTY	0.02	0.00	0.01	0.70		
ACCQTY	0.09	0.00	0.52	0.06	0.70	
PRODMMSG	0.00	0.00	0.03	0.02	0.07	0.61
INTMSG	0.01	0.02	0.15	0.00	0.12	0.50
PERSPROD	0.05	0.02	0.07	0.00	0.07	0.31
PERSINT	0.01	0.02	0.05	0.00	0.03	0.42

	INTMSG	PERSPROD	PERSINT
INTMSG	0.58		
PERSPROD	0.34	0.56	
PERSINT	0.42	0.44	0.62

Figure 7. Discriminant Validity -- Revised Model

could detect effects which were so small as to be significant only in a statistical sense. Three coefficients fell into this category, and the associated paths were dropped: H13, "Acceptance of Quantitative Measures" (ACCQTY) to "Personal Importance of Production" (PERSPROD); H16, from "Reliance on Quantitative Measures" (RELQTY) to "Employer's Customer Service Message" (INTMSG); and H24, "Quantitative Nature of Work" (JOBQTY) to "Personal Importance of Customer Service"(PERSINT).

Although other paths had small coefficients, we chose to drop only those which were truly marginal. We had two reasons for this decision. First, the model represented a first test of the theory, subject to further improvement in measures of some of its constructs. Second, multicollinearity may have reduced some path coefficients. The relative size of the effect, rather than its absolute size, was meaningful in such cases and so the path remained.

These revisions produced the model shown in Figure 6. We then used the holdout sample of odd-numbered surveys to estimate path coefficients for the revised model and to test their significance. Convergent validity remained strong. Discriminant validity (Figure 7) improved slightly as a result of reducing multicollinearity among various constructs. All path coefficients in the revised model were significant at $\alpha < 0.005$.

The path coefficients cannot be interpreted as absolute measures of the degree to which modifying one construct would change another. First, some correlation among constructs will always remain, making the coefficient vary. Second, the constructs themselves represent perceptions captured by many measures: the notion of a "one-unit change" in such a construct has no meaning. However, the

path estimates indicated that the proposed relationships were highly significant and in the direction hypothesized.

4.4 Explanatory and Predictive Power

This research focused on *how* monitor design dimensions affected personal importance of two job dimensions. This is an explanatory objective. PLS estimation uses prediction as its primary objective (Wold 1985). However, the PLS process produces R^2 statistics indicating the portion of variance explained for each construct (Figure 8).

Original Model					
R^2 OF INTERVENING AND DEPENDENT CONSTRUCTS					
RELQTY	ACCQTY	PRODMMSG	INTMSG	PERSPROD	PERSINT
.33	.14	.13	.24	.47	.50

Revised Model					
R^2 OF INTERVENING AND DEPENDENT CONSTRUCTS					
RELQTY	ACCQTY	PRODMMSG	INTMSG	PERSPROD	PERSINT
.29	.13	.08	.17	.41	.50

Figure 8. Variance Explained

The causal model did a poor job of explaining impact, as originally hypothesized. It proposed negative relationships for H20 and H21, and positive for H1; the survey results showed opposite relationships. However, the proposed constructs *were* important in explaining impact. Each version of the model explained 50 percent of the variance in "Personal Importance of Customer Service" (PERSINT) and at least 40 percent of the variance in "Personal Importance of Production" (PERSPROD). Both models

suggested that monitor design affected acceptance of quantitative measures and perceived reliance on those measures. Acceptance and reliance in turn affected perceptions of employer messages.

The original model appeared to explain more of the variance in the dependent constructs as well as in the intervening constructs. Some of that explanation is misleading, however. It results from two, poorly defined constructs and areas of weak discriminant validity. The revised model explained slightly less variance, but its results could be interpreted more reliably.

Despite difficulties in using the first model to explain impact, it was a better predictor of personal importance attitudes. Predictive models tolerate misspecification and multicollinearity more easily, being concerned only with using correlations to forecast (Cook and Campbell 1979). The shared variance among constructs (Figures 4 and 7) represent the squared correlations. The higher values in the original model demonstrate that it would do a better job of predicting levels of intervening and dependent constructs than the revised version.

5. CONCLUSION

This research produced information for researchers interested in various monitoring issues. Other studies treated monitor design as unidimensional. We used four dimensions as independent variables. This work also developed and tested a causal theory of CPMCS impact. In so doing, it demonstrated areas where other researchers could build on this work.

The study showed that case and survey research can be combined to study and model a complex IS issue. Rapid changes in technology and system application will always exist. Case research is an effective way to study intricate, new issues in a natural environment (Benbasat, Goldstein and Mead 1987) and to develop grounded theory (Kaplan and Duchon 1988; Zaltman, LeMasters and Heffring 1982). However, broad-based tests of grounded theory require extensive data from multiple sites. Case research does not readily provide such data. This research demonstrated how a case study clarified issues and constructs, identified possible relationships, and pinpointed shortcomings in research instruments. The model development and field survey then built upon the case findings.

The research produced two tested causal models of CPMCS impact. Combining concepts drawn from the literature with the exploratory case study suggested four dimensions of automated control systems design. The hypothesis testing phase confirmed the usefulness of these dimensions in predicting and explaining monitor impact. These findings show that research schemes categorizing monitors as "low," "moderate," and "high" level should

specify the dimension on which the design is low, moderate or high. The schemes should also differentiate the systems studied along more than one dimension. Such practices would contribute to a *stream* of research into the effects of computer monitoring. Similar dimensions of design may apply when modeling the impact of other new information technologies.

The hypothesis testing also demonstrated that:

- The perceived importance of production to the employer has a positive effect on personal production *and* service importance.
- The perceived importance of service to the employer has a positive effect on personal service *and* production importance.
- Increasing the number of tasks monitored reduces the acceptability of CPMCS measures.
- The credibility of the computer as an appropriate measurement device has a positive effect on acceptance of CPMCS measures, among monitored employees.
- While design dimensions affect the perceived reliance on quantitative measures, they also affect the acceptance of those measures. This acceptance has the stronger effect on perceived production messages.

The hypothesis testing phase demonstrated the use of holdout samples to gain more knowledge in that stage of research. Many studies gather large quantities of data for hypothesis testing. This can represent analytic overkill when testing a causal model as an integrated unit. Even splitting our responses in half produced large enough samples to sense statistically significant, but essentially meaningless, coefficients. Leaving the responses in a single sample would have merely increased this effect. Large samples may be better used to expand the extent of hypothesis testing, rather than to enlarge the database in a single test.

Keen (1981) called for IS research that contributed to a cumulative tradition and that would "offer something that remains meaningful as technology changes." The models tested in this study work with constructs that can be generalized to the design of other technologies (e.g., tasks affected, acceptance of and reliance on the system, appropriateness of the technology to its intended role). It also draws from IS research on monitoring and systems design, tying these findings to other areas of interest in this field. By so doing, it contributes a set of immediate conclusions about monitoring in particular and "jumping off points" for ongoing research into this and other technological implementations.

6. ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance of Labour Canada, the Social Sciences and Humanities Research Council, the University of Western Ontario, and the University of Cincinnati in supporting this work. Chris Higgins, Don Barclay, Russ Knight and Peter Suttie of the University of Western Ontario, and Dan Robey of Florida International University also provided valuable comments and advice on the research.

7. REFERENCES

- Asher, H. B. "Causal Modeling (Second Edition)." *Sage University Paper Series on Quantitative Applications in the Social Sciences*, Number 07-003, Beverly Hills, California: Sage Publications, Inc., 1983.
- Benbasat, I.; Goldstein, D.; and Mead, M. "The Case Research Strategy in Studies of Information Systems." *MIS Quarterly*, Volume 11, Number 3, 1987, pp. 369-388.
- Cook, T. D., and Campbell, D. T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company, 1979.
- Eisenman, E. J. "Employee Perceptions and Supervisory Behaviors in Clerical VDT Work Performed on Systems that Allow Electronic Monitoring." In U.S. Congress, Office of Technology Assessment, *The Electronic Supervisor: New Technology, New Tensions - Contractor Documents*, Volume 1 (PB88-156351). Springfield, Virginia: U.S. Department of Commerce, NTIS, December, 1987, pp. 407-449.
- Fornell, C. "A Second Generation of Multivariate Analysis: Classification of Methods and Implications for Marketing Research." Working Paper, University of Michigan, April 1984.
- Fornell, C., and Barclay, D. "Jackknifing in PLS." Unpublished Notes, November, 1988.
- Fornell, C., and Larcker, D. "Evaluating Structural Equation Models With Unobservable Variables and Measurement Error." *Journal of Marketing Research*, Volume 18, Number 1, 1981, pp. 39-50.
- Garson, B. *The Electronic Sweatshop: How Computers are Transforming the Office of the Future Into the Factory of the Past*. New York: Simon & Schuster, 1988.
- Gregory, J., and Nussbaum, K. "Race Against Time: Automation in the Office." *Office: Technology & People*, Volume 1, 1982, pp. 197-236.
- Ilgen, D., Fisher, C., and Taylor, M. "Consequences of Individual Feedback In Organizations." *Journal of Applied Psychology*, Volume 64, Number 4, 1979, pp. 349-371.
- Irving, R.; Higgins, C.; and Safayeni, F. "Computerized Performance Monitoring Systems: Use and Abuse." *Communications of the ACM*, Volume 29, Number 8, 1986, pp. 794-801.
- Kaplan, B., and Duchon, D. "Combining Qualitative and Quantitative Methods in Information Systems Research." *MIS Quarterly*, Volume 12, Number 4, December 1988, pp. 571-586.
- Keen, P. G. W. "MIS Research: Reference Disciplines and a Cumulative Tradition." In E. R. McLean, Editor, *Proceedings of the First International Conference on Information Systems*. Philadelphia, Pennsylvania, 1980, pp. 9-18.
- Koepp, S. "The Boss That Never Blinks." *Time*, July 28, 1986, pp. 40-41.
- Lawler, E., III. "Control Systems in Organizations." In M. D. Dunnette, Editor, *Handbook of Industrial and Organizational Psychology*. New York: Rand McNally College Publishing Co., 1976, pp. 1247-1291.
- Lawler, E., III and Rhode, J. *Information and Controls in Organizations*. Pacific Palisades, California: Goodyear Publishing Co., 1976.
- Nussbaum, K. *The 9 to 5 National Survey on Women and Stress: Office Automation Addendum*. Cleveland, Ohio: 9 to 5 National Association of Working Women, 1984.
- Oreskovich, C. "Computer Monitoring Debate Rages." *Financial Post*, September 7, 1985, p. C14.
- Pedhazur, E. *Multiple Regression in Behavioral Research: Explanation and Prediction*, Second Edition. New York: Holt, Rinehart and Winston, 1982.
- Smith, M., Carayon, P., and Miezio, K. "Motivational, Behavioral, and Psychological Implications of Electronic Monitoring of Worker Performance." In U.S. Congress, Office of Technology Assessment, *The Electronic Supervisor: New Technology, New Tensions - Contractor Documents*, Volume 2, PB88-156369. Springfield, Virginia: U.S. Department of Commerce, NTIS, December, 1987, pp. 1-119.
- Taylor, M., Fisher, C., and Ilgen, D. "Individuals' Reactions to Performance Feedback in Organizations: A Control Theory Perspective." In K. Rowland and G. Ferris, Editors, *Research in Personnel and Human Resources Management*, Volume 2, London: JAI Press Inc., 1984, pp. 81-124.

Turkle, S. *The Second Self: Computers and the Human Spirit*. New York: Simon & Schuster, 1984.

U.S. Congress, Office of Technology Assessment, *The Electronic Supervisor: New Technology, New Tensions*. OTA-CIT-333. Washington, D.C.: U.S. Government Printing Office, September, 1987.

Walton, R., and Vittori, W. "New Information Technology: Organizational Problem or Opportunity?" *Office: Technology & People*, Volume 2, 1983, pp. 249-273.

Westin, A. "Privacy and Quality of Work-Life Issues in Employee Monitoring." In U.S. Congress, Office of Technology Assessment, *The Electronic Supervisor: New Technology, New Tensions - Contractor Documents*, Volume 1, PB88-156351. Springfield, Virginia: U.S. Department of Commerce, NTIS, December, 1987, pp. 1-260.

Whisler, R., and Shultz, G. "Information Technology and Management Organization." In G. Schultz and T. Whisler,

Editors, *Management Organization and the Computer*. Chicago: Graduate School of Business, The University of Chicago, 1960, pp. 3-36.

Wold, H. "Systems Analysis by Partial Least Squares." In P. Nijkamp, H. Leitner, N. Wrigley, Editors, *Measuring the Unmeasurable*. Dordrecht, The Netherlands: Martinus Nijhoff Publishers, 1985, pp. 221-252.

Wold, H. "Systems Under Indirect Observation Using PLS." In C. Fornell, Editor, *A Second Generation of Multivariate Analysis: Volume 1 - Methods*. New York: Praeger Publishers, 1982, Chapter 15.

Zaltman, G., LeMasters, K., and Heffring, M. *Theory Construction in Marketing*. New York: John Wiley & Sons, 1982, pp. 71-112.

Zuboff, S. "New Worlds of Computer-Mediated Work." *Harvard Business Review*, Volume 60, Number 5, 1982, pp. 142-152.