# DELETERIOUS EFFECTS OF INTERMITTENT INTERRUPTIONS ON THE TASK PERFORMANCE OF KNOWLEDGE WORKERS: A LABORATORY INVESTIGATION

by
Louis Coraggio



A Dissertation Submitted to the Faculty of the

COMMITTEE ON BUSINESS ADMINISTRATION

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 0

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Louis Coraggio

entitled DELETERIOUS EFFECTS OF INTERMITTENT INTERRUPTIONS ON THE TASK PERFORMANCE OF KNOWLEDGE WORKERS: A LABORATORY INVESTIGATION

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Nicholas J. Aquilano
Nicholas Aquilano — Date 11/14/90

George W. Summers
George W. Summers — Date 11/14/90

Christina E. Shalley
Christina M. Shalley — Date 11/14/90

Douglas R. Vogel
Douglas Vogel — Date 11/14/90

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Nicholas J. Aquilano
Dissertation Director Nicholas Aquilano — Date 11/14/90

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for an extended quotation from, or reproduction of, this manuscript, in whole or in part, may be granted by the copyright holder.

SIGNED: Louis Coraggio

## ACKNOWLEDGMENTS

I am grateful to all the members of my committee for their encouragement in pursuing a new line of research. Each deserves recognition for his or her contribution. My thanks to:

- Dr. Nicholas Aquilano for his steadfast dedication to the notion of applied research;
- Dr. George Summers for coming out of well deserved retirement to see that I did it right – despite my best efforts;
- Dr. Christina Shalley for taking a chance on a student from a strange discipline – and broadening his perspective;
- Dr. Joseph Valacich for sharing his enthusiasm, energy and experience in empirical research.
- Dr. Douglas Vogel for obtaining the resources and smoothing the way.

I would also like to thank my colleagues for their insights and my students for giving me their best effort. Finally, I want to thank my wife and partner, Beverly, whose moral, editorial, financial and spiritual support made all of this worth the effort.

## CONTENTS

## FIGURES

# TABLES

## ABSTRACT

Contemporary businesses compete in a highly reactive marketplace that demands a new breed of sophisticated knowledge workers. Managing these valuable people requires understanding the effects of their work environment on productivity. Frequent interruptions are an integral part of the knowledge worker's day.

In this laboratory study, two attributes of interruption, frequency and length, are examined in conjunction with two levels of task complexity. A 2x2x2 factorial design with a control group for each level of task complexity resulted in ten unique treatments.

Using personal computers, 122 student subjects from undergraduate courses in production management, took a practice, multiple-choice examination (the primary task) over course material. All subjects were allowed exactly 45 minutes to work on the primary task. Interruption episodes of trivia questions were generated at random intervals. Subjects were interrupted either two or six times with interruption lengths of either 30 seconds or 120 seconds. Performance on the primary task is measured using a point scoring system. A post experiment questionnaire was used to validate experiment manipulations.

Using a multiple regression approach to analysis of covariance, approximately 70% of variance in performance is explained. Final models include a covariate for prior classroom performance. Four significant conclusions emerged from the analysis: (1) In the high complexity version of the primary task, short interruptions result in an average performance reduction of 44% relative to control subjects. (2) For the low complexity task, long interruptions result in an 11% average performance improvement over control subjects. (3) Performance in long interruption treatments is significantly better than performance under short interruption treatments for both levels of task complexity. (4) No consistent effects for frequency of interruption occur at the levels used in the study.

# 1. INTRODUCTION

## 1.1 BACKGROUND

Contemporary business organizations face an accelerated pace of operations. The proliferation of computer / FAX networks, and "800" communications have forced businesses into a "real time" mode. To remain competitive, both manufacturing and service segments must become more responsive to customer demands. When coupled with an exponential growth in technological complexity, contemporary businesses must cope with shorter product life cycles, more complicated products, and more ongoing involvement with consumers.

In response to these pressures, the contemporary work force has an increased proportion of "knowledge" workers. Knowledge workers have primarily mental responsibilities and their task focus will be analytical and adaptive — as opposed to the repetitive and procedural focus of most manual/industrial workers ("manual" workers).

Beginning with Fredrick Taylor's *Principles of Scientific Management*, historical research has focused on structuring the time of manual workers. Job design and performance evaluation has focused on predictable allowances for learning, forgetting, fatigue, personal health and safety for manual workers [Barnes, 1980 and McCormick, 1976]. However, little work has been done on the knowledge worker.

A knowledge worker (e.g. manager, engineer, accountant, analyst, laboratory technician, customer service representative) typically represents a large investment in both salary and training. The knowledge worker is pivotal to growth of the service side of the economy. Service jobs represented about 70% of the work force in 1977 [Porat, 1977]. In 1986 the service workers comprised 75% of the total U.S. work force – with over half as "white collar, highly skilled occupations" [Chase and Aquilano, 1989, p96].

As contemporary organizations acquire more technical sophistication and consumer focus, the role of the knowledge worker expands. Training and knowledge acquisition for an individual worker becomes a significant asset to the organization. Increasing the productivity of these workers (and maintaining their mental and physical health) requires a better understanding of the environment in which they function. Interruptions are an integral part of that environment. Therefore, it is important that we understand the effects of interruptions.

## 1.2 RESEARCH FOCUS

Much popular management literature has focused on personal time management for those engaged in knowledge work. A theme that permeates almost all of these guides is minimizing interruptions [Douglass, 1987; Oseland & Kleiner, 1988; Moskal, 1986]. Plans for prioritizing and eliminating interruptions are presented by Davidson [1986]. Thus, the detrimental effects of interruption would seem to be taken as common knowledge.

Tansik and Smith [1990] suggest that many workers in reactive service businesses have a wide variety of "scripts" to handle customer interactions. A script is "a pattern of behavior or an operating routine that is triggered by an external stimulus." [Tansik and Smith, 1990]. These interactions can be viewed as predefined responses to "interruptions" created by customers. The notion of scripted behavior suggests that workers have developed formal tactics to better cope with interruptions.

Little formal research has been done on short term interruptions (less than one day). While there is certainly a time loss directly attributable to interruption, are there other effects on performance? If so, it would be advantageous to know what they are. This research will focus on identification of these hidden effects. If these effects can be identified and predicted, there are broad implications for management of both individuals and organizations. Some possible areas of application are:

- *Job Design* — If knowledge workers are to be subjected to interruptions, what "primary" tasks should be assigned to those persons who are subject to interruption? Should time be blocked into distinct time periods when interruptions will or won't be tolerated? Are some workers better than others at handling interruptions? Do some workers handle interruptions better than others? If so, can their attributes be identified, or their strategies isolated?

- *Project Management* — What allowances should be made for workers whose primary tasks are subject to interruptions? Should "person-hours" for a task be adjusted to reflect interruptions?

- *Office Layout* — What are the costs of shared offices and open layouts with respect to interruptions?

- *Budgeting and Job Costs* — What costs will be incurred when expedited or "emergency" tasks are imposed on planned activities?

## 1.3 DEFINITIONS

### 1.3.1 Interruption

The American Heritage Dictionary defines *interrupt* as:

> "1. To break the continuity or uniformity of. 2. To hinder or stop the action of (someone) by breaking in upon. 3. To break in upon an action or discourse. [1978]"

For purposes of this study, interruption will require a more specific definition. In the context of the knowledge worker:

- *an interruption is an externally-generated, randomly occurring, discrete event that breaks continuity of cognitive focus on a primary task.*

*Externally-generated* implies that the nature of the interruption is not controlled by the worker but by some other person or circumstance in the work environment.

*Randomly Occurring* means that the timing of the interruption is not specifically known to the worker before it happens. This does not imply that the worker expects to be free of interruption. The worker may know that an interruption is possible – without knowing when (or if) it will happen.

*Discrete event* indicates that the interruption has finite duration – a clear beginning and a clear end. This distinguishes an interruption from concurrent distracting conditions which may inhibit, but not break, worker concentration. Distracting conditions might be in the form of background noise or commotion that are part of the work environment.

A *primary task* is considered to be a well defined activity that constitutes the current "job" for the knowledge worker. The task should have a clear objective and some condition which constitutes its completion.

For example, the arrival of another plane would not be considered an interruption to an air traffic controller nor would a "911" call be an interruption to a police dispatcher. The primary task for these workers is to be in a response mode. These events are part of the cognitive focus for these workers and would cause no break in cognitive focus. However, a personal call from a spouse would be considered an interruption under this definition.

### 1.3.2 Knowledge Worker

There seems to be no universally accepted definition of knowledge worker. Kelley [1985, p. 8] suggests that knowledge workers "collect, process, analyze and disseminate information".

Dahms [1988] defines a knowledge worker as "any manager or related professional who deals with non repetitive tasks and who is involved with the review or processing of organized information".

Davis and Olson [1985, p409] contend that knowledge work "...involves thinking, processing information, and formulating analyses, recommendations and procedures." Davis and Olson also suggest a potentially useful distinction based on the source of knowledge:

> "Knowledge work tasks involve use of information. Some information derives from the knowledge and expertise of the knowledge worker (*internal knowledge base*); some derives from organizational and external data to which the knowledge worker has access (*external knowledge base*)." [ Davis and Olson, 1985, p409 (Italic definitions added) ]

These three authors come from divergent disciplines (Kelly from management, Davis and Olson from MIS and Dahms from industrial engineering). While they do not completely agree, a common set of attributes emerges. A knowledge worker:

- performs primarily mental and/or abstract tasks;
- focuses primarily on information processing;
- uses both internal and external knowledge bases;
- works on non-repetitive task; and
- possess a relatively high level of formal education.

## 1.4 PREVIEW OF STUDY

The remainder of this study details an exploratory empirical investigation into the effects of interruptions on the knowledge worker.

- Chapter 2 reviews salient literature on interruptions.
- Chapter 3 establishes a theoretical foundation for laboratory experiments and proposes hypotheses to be tested.
- Chapter 4 explains the design of the experiment and presents details of the methods used to conduct the study.
- Chapter 5 presents analyses of actual subject performance on the experiment.
- Chapter 6 examines self reported data provided by subjects on post experiment questionnaires.
- Chapter 7 proposes an explanation of the experiment results, formulates alternative working models and recommends formats for continuing research.

# 2. PRELIMINARY LITERATURE REVIEW

No literature has been found which directly addresses the measurement of task performance by knowledge workers under interruption. Related literature may be classified into two broad categories: Operations Management/Industrial Engineering and Psychology/Cognitive Sciences.

## 2.1 OPERATIONS MANAGEMENT/INDUSTRIAL ENGINEERING

### 2.1.1 Industrial Learning and Forgetting Curves

The use of the learning curve has been well documented in a variety of industries and tasks [See Yelle, 1979, for a comprehensive review of the literature]. Learning curves describe the cumulative effects of individuals and industrial groups whose costs or times (per unit of output) drop with increasing production volumes.

Recent literature has modified the original learning curve account for interruptions. Towill [1985] proposed a "Forgetting Factor" to be used for increasing cumulative production times for batch processing. Towill suggests that the forgetting factor functionally increases the true "setup" time for batch production. Duration of the interruption is presented as the primary reason for an increase in the "forgetting rate".

Globerson and Levin [1987] proposed a complex model for organizational forgetting based on mathematical models by Carlson and Rowe [1976]. The Globerson and Levin models predict an exponential decay in performance time per unit as a function of the duration of the interruption and of the degree of organizational learning.

Bailey [1989] used an experimental treatment to determine the effects of longer interruptions ( 7 - 114 days) on 35 manual workers. Workers were engaged in assembly and disassembly of a simple structure. The assembly task was considered a "procedural" task, and the disassembly a "continuous control" task. This essential distinction can be viewed as one of task complexity – procedural tasks essentially require more cognitive effort. (See the discussion of Schendel et.al [1978] in the "Task Classification" section below.)

Bailey used time per unit as a measure of proficiency. Subjects were first instructed in both tasks. During an eight hour period, each subject's time per unit was recorded to determine an individual learning curve. Forgetting was measured as the difference between the time predicted by the subject learning curve (for an incremental unit) and the actual time required for assembly of the incremental unit. Relearning occured when actual assembly times were equal to those predicted by the subject's learning curve.

Important results of the Bailey study are:

- There is no significant forgetting of continuous control tasks.

- Forgetting of procedural tasks is a function of the original learning level and the log of the duration of the interruption. (Subjects who were less proficient prior to interruption had less to forget.)

- There is no correlation between original learning rate and the relearning rate for an individual. (Subjects who acquired initial skills quickly did not necessarily reacquire those skills more quickly than subjects who had poorer initial learning rates.)

While Bailey makes no claims for tasks other than light assembly work, this study suggests that a distinction should be made between simple and complex mental activity.

Throughout this literature, the dependent variable measured was task completion time. The effects of forgetting or relearning are measured as an increase in the time to complete a discrete unit of production under repetitive task conditions. The underlying assumption here is that the quality of the work is a binary condition. The "product" is either built correctly or it is not. This assumption suggests that all effects of learning or forgetting are reflected in task completion time.

### 2.1.2 Measuring Performance of the Knowledge Worker

Measuring performance of knowledge workers in the work place is a fledgling science at best. Quantitative measurements are difficult because of the non-repetitive nature of knowledge work. Whitmore [1987] suggests two other aspects that complicate quantitative evaluation.

- Technical deficiencies – current methods are inadequate due to highly variable content and difficulty in measuring throughput.

- Psychological reaction – Knowledge workers may fear being measured, since they have not been traditionally subject to measurement. This real or imagined fear can result in resentment and atypical performance.

These difficulties are echoed by industrial practitioners who measure performance of knowledge workers by qualitative techniques. Below are some example methods used for qualitative performance evaluation:

- Nominal Group Techniques with Mutual Needs Inventory (Essentially a peer review and group decision method) [Gregerman, 1981]

- Identifying and formalizing a self reported "Best Work" accomplishment. [Helton, 1988]

- Creating time standards for "micro mental" activity, and establishing a standard time for bench mark comparisons. (A variant of predetermined motion time study.) [Boepple and Kelly, 1971]

- "The best – and only – method currently available is year to year comparison of a company's productivity figures." [Conn, 1984]

The diversity of these examples emphasizes the lack of standards for evaluating knowledge work. While not specifically stated by all of the authors above, they all acknowledge the existence of a completion time – performance quality tradeoff.

Typically, knowledge workers must do the best job possible in the time they have available. Often these two factors present a mutually constraining condition. Maximizing quality implies longer completion times. In practice, the knowledge worker often must do the best job possible in the time available. Combined with the non-repetitive nature of the work, evaluating performance must account for both the quality of work and the time available for completion.

Academic studies have recognized the need for evaluating both completion time and performance quality. In their experiments on supervisory control, Buck et.al. [1988] presented several varieties of cognitive tasks to subjects. Both latency times and accuracy scoring were used in evaluating results. (The latency time was defined as the time between presentation of the task and the answer by the subject.)

Okogbaa and Shell [1986] used an accuracy scoring system (percentage of correct responses) as a measure of performance in their experiments on knowledge worker fatigue. Two types of tasks, reading comprehension and arithmetic/logical, were presented. Times for the tasks were held constant.

Based on the paucity of studies in this area, and the alternative measuring methods, it seems clear that no universally accepted methods exist for evaluating knowledge worker performance. However, it seems clear that evaluating knowledge work must involve two dependent variables — task completion time and quality of output. Measurements of worker output must recognize the existence of this time-quality tradeoff.

## 2.2 PSYCHOLOGY/COGNITIVE SCIENCES

The psychological literature deemed most relevant, focuses primarily on learning and memory retention. This presupposes deleterious effects to be primarily a result of forgetting.

### 2.2.1 Individual Learning and Forgetting

There seems to be no universally accepted model for the forgetting mechanism. One detailed in Hulse et.al. [1975] suggest that there are three operations common to all memory systems: encoding, storage, and retrieval. Forgetting could be a result of dysfunction in any operation. Knowledge could be "lost" because it was:

- incorrect or incomplete when stored; or
- eroded while stored; or
- stored in insufficient space; or
- incorrectly accessed upon retrieval.

However, no explanatory mechanisms were provided.

Work by Christiaansen [1980] involved memorizing prose passages. After delays of 1 week, 1 month, and 2 months subjects were asked to recall the passages. Measures were provided by forced choice questions and four categories were evaluated: main character, theme, sentence gist and sentence format. This format would appear to somewhat parallel the tasks of knowledge workers in that it is primarily mental and involves some analytical skills. Christiaansen found the most important factor in forgetting is the original task proficiency. The difficulty in learning (original learning curve) seemed uncorrelated to the rate of forgetting. Slamecka and McElree [1983] confirmed the independence of forgetting rate and learning with nonsensical verbal lists. Brainerd et.al. [1985] concluded that learning and forgetting are "governed by quite different laws" and should be treated by different theoretical assumptions.

Thus, there seems to be agreement on only two characteristics of forgetting:

- the rate of forgetting decays exponentially over time — i.e. forgetting is a function of log(time); and
- the rate of forgetting is not correlated with the original learning rate.

### 2.2.2 Task Complexity and Skill Retention

The issue of task complexity is addressed by Schendel et.al. [1978] in a study for the U.S. Army. Although confining the classification to motor skills, the distinction of "procedural" vs. "continuous control" would seem to be made on the basis of mental involvement. Procedural tasks include: assembly of components; monitoring and adjustment; and setting dials in sequence. Continuous control tasks include target tracking and flight control.

> "Procedural tasks generally involve series of discrete motor responses... The responses themselves are easy to execute; it is deciding what responses to make, in what order, that pose the problem for the learner. ...
> Procedural skills appear to be highly susceptible to forgetting especially when contrasted with continuous control skills... " [Schendel and Hagman, 1982, page 605]

Shields [1979] also found that procedurally complex basic soldiering skills were forgotten at a faster rate than those with minimal procedural complexity. The results of these studies support contentions that increased mental complexity may aggravate forgetting.

### 2.2.3 Distractions and Performance

A branch of social facilitation research known as Distraction-Conflict theory was solidified by Zajonc [1965]. This theory contends that social conditions increase drive or arousal. This arousal facilitates performance of simple, well-learned tasks, but creates conflicts that impairs the performance of complex, counter-intuitive tasks.

In experiments by Sanders and Baron [1975], subjects performing tasks of varying complexity were distracted by noise. The measured performance of the subjects suggested that distractions actually improved performance of simple tasks and impaired performance of complex tasks.

Other studies have confirmed performance improvement in "simple" tasks as a result of distraction. Work by Hartley and Adams [1974], Houston and Jones [1967], and O'Malley and Poplawsky [1971] found reaction times to be improved under distracting conditions. This may be explained by an "arousal" effect for tasks that are monotonous [McBain, 1961] or tasks that require continual vigilance [Zuercher, 1965].

Other studies suggest that complex or poorly learned tasks will be impaired by distractions. For example, Woodhead [1965] used noise bursts to distract subjects performing mathematical tasks and found deterioration of quality but an increase in the rate of work. Eschenbrenner [1971] also found distractions of noise degraded performance of subjects simulating orbital tracking tasks.

Thus, some measure of task complexity should be included in any study of the effects of interruptions on the knowledge worker. Baron concludes:

> "In short, there seem to be at least 16 studies that demonstrate that distraction can either facilitate simple task performance, increase performance on tasks facilitated by other stressors, or impair complex task performance." [Baron, 1986, p13]

# 3. THEORETICAL BASIS

## 3.1 RESEARCH QUESTIONS

The literature review produced no specific evidence addressing the effects of interruptions on knowledge workers. Proceeding on the assumption that this research will define a narrow path of study, research questions to be answered are:

1. *Are there deleterious effects from interruptions on the performance of knowledge workers?*

2. *If the deleterious effects are present, can they be accurately measured and predicted?*

3. *If the deleterious effects are present, how will the length and frequency of interruption affect the size of these deleterious effects?*

Subsequent sections of this chapter will establish a theoretical framework designed to answer these questions. Topics to be covered are:

- selection of task/interruption characteristics;
- models for measuring deleterious effects;
- suggested mechanisms for the deleterious effects; and
- research hypotheses.

## 3.2 SELECTION OF STUDY VARIABLES

Based on related literature (reviewed below), three variables are of primary interest for this research:

- task difficulty;
- length of interruption; and
- frequency of interruption.

### 3.2.1 Task Complexity and Task Difficulty

Distraction-Conflict literature and work by Bailey [1989] suggest that some measure of task complexity be included. Campbell [1988] presents a comprehensive classification scheme for task complexity. In his work, he uses three basic categories for classifying complexity as :

- a primarily psychological experience;
- an interaction between task and person; or
- a function of objective task characteristics.

In comprising his scheme, Campbell suggests 16 unique types of complex tasks. These 16 types are defined by the presence (or absence) of the following attributes:

1. multiple paths to a desired end state;
2. multiple end states;
3. conflicting interdependence; and
4. uncertainty or probabilistic linkages.

The nature of knowledge work as defined here should always be "complex" under attribute 3. The objectives of meeting time deadlines and maximizing performance quality will frequently be conflicting. Depending on task specifics, the knowledge worker may also face introduced complexity under attribute 1. If work requiring completion exceeds time available, the worker must prioritize sub tasks — thus creating multiple paths.

The notion of task complexity thus becomes a relative comparison rather than an absolute scale of measure. However, Campbell suggests difficulty as a surrogate for complexity.

> "(T)he relation between task complexity and task difficulty is worth considering. Complex tasks are, by their nature, difficult. Thus, the two notions can be used interchangeably...but not always. ... The general point is that certain tasks can be difficult (i.e. require high effort) without being complex; in contrast, other tasks are difficult *because* they are complex." [Campbell, 1988] (Citations deleted)

For purposes of determining distinctions between tasks, the notion of difficulty may be easier to apply (on a relative basis) than complexity.

### 3.2.2 Duration of Interruption

Work by Bailey [1989] and Globerson and Levin [1987] suggest that industrial workers, and groups of workers, will experience an exponential decay of knowledge retention. This "decay" model has also been adopted for work in purely psychological experiments [Christiaansen, 1980; Schendel and

Hagman, 1982; Slemecka and McElree, 1983]. However, the interruption period of these studies were measured in days. For purposes of this study, interruptions will be measured in minutes. The inclusion of duration as an experimental treatment is intended to confirm the presence of a time related effect with these shorter interruption episodes.

With respect to learning and forgetting literature, interruption lengths will be an order of magnitude shorter than those used in memory/retention studies by Christiannsen [1980]; Schendel and Hagman [1982]; Slemecka and McElree [1983]; and Bailey [1989]. By contrast, noise distraction studies by Eschenbrenner [1971], Sanders and Barron [1975], and Woodhead [1965] used very short period of interruption (1-4 seconds). Interruptions used here are designed to simulate the lengths that might occur in an office setting.

### 3.2.3 Frequency of Interruption

Frequency was treated specifically by Baron and Sanders [1975]. Subjects were distracted 0,2,4,6 or 8 times. However, results were reported as "collapsed across trials" so no explicit relationship between frequency of interruption and performance degradation was reported.

The inclusion of frequency is logical for knowledge workers. Frequency is a countable variable. As pointed out by Voss [1983] jobs in many service organizations may be entirely based on frequent interruption. In this study, the actual cognitive mechanism is not as important as predicting performance. However, Kahneman [1973] suggests that a phenomenon known as "capacity interference" may be introduced by a high frequency of interruptions. This phenomenon suggests that individuals may simply "overload" their attention control mechanisms.

### 3.3 MEASURING TASK PERFORMANCE

The objective of this line of research is to identify and measure any deleterious effects of interruptions. Assuming that both quality of performance and completion time are components of knowledge worker performance in organizational settings, the effects of interruption could be measured as the:

- additional time required to achieve equivalent performance (Time Measured Model); or as the
- reduction in performance for a given time period (Quality Measured Model).

Statistically sophisticated techniques like MANOVA or Canonical Correlation [Tatsuoka,1971] could be used to evaluate multiple dependent variables. But, for initial exploration, one of the two factors (Quality or Completion

Time) will be held constant while the other is measured. This approach will simplify experimentation and application, and take advantage of a more established base of analytical techniques.

### 3.3.1 Time Measured Model (TMM)

By definition, the time directly attributable to interruptions is beyond the control of the worker. Of interest is the "extra" time attributable to the interruption. Consider the model below. (Figure 3.1)

**Figure 3.1:** Time Measured Model

$$TA = TU + \sum_{i=1}^{n} TI_i + Dt$$

Where:

$TA$ ≡ Actual time to complete task
$TU$ ≡ Time to complete uninterrupted task
$TI_i$ ≡ Time for interruption i (i = 1 to n)
$n$ ≡ Total number of interruptions
$Dt$ ≡ Deleterious time effects

The "hidden" effects will be captured in the size of the D term. If interruptions have an effect – other than the directly attributable time captured by $\Sigma I$ – then the size of the D term should be significantly larger than zero.

To use the TMM, TU would have to be established. The task would then be completed with interruptions, and Ta measured. With the TMM, the quality of performance should be captured in extended completion time. To effectively use this model, a task would not be complete until its quality is acceptable. The majority of previous research has used this form of model.

### 3.3.2 Quality Measured Model (QMM)

If the time for completing a task is held constant, and the change in quality examined, then the model will resemble Figure 3.2.

**Figure 3.2:** Quality Measured Model

$$SA = SU + Ds$$

Subject to:

$$TA = TU + \sum_{i=1}^{n} TI_i$$

Where:

- $SA$ ≡ Actual Score on task
- $SU$ ≡ Score of uninterrupted task
- $Ds$ ≡ Score attributable to deleterious effects
- $TA$ ≡ Actual time to complete task
- $TU$ ≡ Time to complete uninterrupted task
- $TI_i$ ≡ Time for interruption i (i = 1 to n)
- $n$ ≡ Total number of interruptions

The QMM assumes that the quality measure is a pseudo continuous variable. If there are deleterious effects from interruptions then Ds should be significantly less than zero. The primary limitations for this form of measurement model involve determining a measure of performance. The quality measure must be sensitive enough to pick up relatively subtle changes in performance. It must also be relevant to the final quality of the task.

## 3.4 INTERRUPTION MECHANISMS

Based on prior research, two mechanisms will be explored in the course of this research:

- Forgetting – Deleterious effects may be due to a deterioration of short term memory.
- "Changing Gears" – Deleterious effects may be function of the shift in cognitive focus.

### 3.4.1 Cognitive Model of the Knowledge Worker

Since a primary attribute of knowledge work is information processing (see Section 1.3.2), the Information Processing System (IPS) proposed by Newell and Simon [1972] is used as model for cognitive processing by the knowledge worker. The IPS model suggests that human information processing may be controlled by the same mechanisms that govern computer information processing.

For purposes of examining interruptions, a simplified IPS model would consist of three basic components:

- Short term memory – a limited area used for immediate processing of information. (Analogous to Random Access Memory in a computer)

- Working Memory – That portion of Short Term memory in actual use by the individual for performing the immediate task. (Basically, that portion of RAM being used by the current program)

- Long Term Memory – Much larger area of stored information, to be selectively searched and retrieved. Essentially the knowledge base. (Analogous to disk storage on a computer).

Newell and Simon suggest that humans engaged in problem solving and cognitive reasoning, essentially perform a task in working memory. Pending operations are stored in working memory until capacity limits are reached. Excess information, (or less frequently used information) is sent to (or retrieved from) long term memory. The mechanics of interruptions will be examined in the context of this IPS framework.

### 3.4.2 Forgetting Mechanism

In the context of the IPS model, forgetting may be attributable to imperfect storage/retrieval of information to/from long term memory, or due to an "overflow" of short term memory.

Work by Bailey [1989], Schendel and Hagman [1982], and Christiaansen [1980] agree that long term forgetting takes the form of an exponential decay function. Schendel et.al. [1978] found that the rate of forgetting was exacerbated by complexity of the task (a result confirmed by Shields [1979]). Yet, the rate of forgetting is generally expected to be uncorrelated to the original degree of learning. [Bailey, 1989; Slamecka and McElree, 1983; Brainerd et.al., 1985]

If the forgetting mechanism is the source of any deleterious effects ( the "D term" in the measurement model above), then the relatively complex tasks performed by knowledge workers should be subject to forgetting. Additionally, this mechanism would anticipate that longer interruptions would result in a larger "D term" in performance measurement.

### 3.4.3 Change Gear Mechanism

The lengths of interruption being examined in this study are an order of magnitude smaller than those used by Bailey [1989]. As a result, an alternative model may be needed to explain the "D term".

The IPS model suggests that externally generated interruptions will "load" an alternative processing program, execute it, and return to the original program. The existence of limited short term memory would require "swapping" programs from short term to long term memories and back to short term. If this model explains the interruption mechanism, then the time for "swapping" may be the "D term". This would suggest that the losses for an interruption are relatively constant.

Kahneman [1973] suggests that frequent interference may overload attention control mechanisms causing "capacity interference". Extending this to the Newell and Simon IPS model, interruptions could place the human IPS into a condition known as "thrashing" in computer operating systems. Thrashing occurs when the operating system has too many demands placed upon it and constantly jumps from one task to another — without actually performing much work on any of them.

### 3.4.4 Combined Mechanism

The forgetting mechanism suggests that the "D term" will increase as a function of the length of interruption. In a sense, the forgetting mechanism represents a "variable cost" for interruption. On the other hand, the change gears mechanism suggests that each interruption has a "fixed cost". It seems plausible that both mechanisms may be at work. If both are present, an effort should be made to distinguish and isolate the effects, and to examine their interactions.

## 3.5 RESEARCH HYPOTHESES

Six research hypotheses are presented below in relative order of confidence.

### 3.5.1 Effects of Interruption on Performance

Based on findings by Eschenbrenner [1971], Woodhead [1965] and Baron [1986], a decrease in performance should be expected for complex tasks under interrupted conditions. Using the classification scheme established by Campbell [1988], the very nature of a completion time /performance quality conflict (typical of knowledge work) should make tasks "complex" – if the quality/time conflict exists. By definition, tasks performed by knowledge workers are expected to be "complex" under the Campbell definition. The first, and strongest hypothesis is:

*H1: Interruptions will result in decreased performance levels for tasks performed by knowledge workers.*

Should subsequent research find that, the "D term" is not significantly different from 0, under any circumstance, then interruptions would seem to have no effect on performance – past the direct time loss attributable to the interruption.

If the "D term" is significant, but has the "wrong" sign (i.e. interruptions improve performance) then subsequent research should follow models proposed by Sanders and Baron [1975], Zuercher [1965] or McBain [1961]. This result may also support contentions by Weick [1985] who suggests that periodic breaks from computer based tasks will improve performance.

### 3.5.2 Interruption Effects and Task Complexity

Consider the difficulty-complexity equivocation proposed by Campbell [1988] in conjunction with the IPS model proposed by Newell and Simon [1972]. A relatively more complex (or difficult) task should be expected to occupy more cognitive capacity. If the IPS model holds true, then a complicated "program" should reasonably be expected to occupy more "short term memory". The occurrence of an interruption should have more impact on a "difficult" task than on a "simple" one – perhaps because there is more "program" to save and restore.

More complex tasks may also have more "pending" operations in working memory. Work by Miller [1956] first suggested the notion that humans have a fixed limit on the number of pending operations they can handle. If a complex task has too many pending operations, the occurrence of an interruption may require more extensive movement of information from short term memory to long term memory.

It seems reasonable to expect the occurrence of an interruption during a difficult task to be more disruptive to the "chain of thought". This research does not seek to model the cognitive behavior of the interrupted worker. However, the relative difficulty of a task is an external factor of interest to the knowledge worker, job designer, manager and project planner. Therefore, a test of this factor is warranted. On these bases then:

*H2: Reductions in performance from interruptions will be larger for difficult tasks than for simple tasks.*

If this hypothesis is rejected, then the mechanism for deleterious effects may be insensitive to difficulty level. More likely, the incremental difficulty of tasks may not be detectable or may be lost in other factors.

### 3.5.3 Frequency Effects

Work on external distractions by Woodhead [1965] and Eschenbrenner [1971] found that increased frequency of distraction decreased performance levels for abstract tasks.

Since other distraction work [Hartley and Adams 1974; Houston and Jones 1967; and O'Malley and Poplawsky, 1971] found reaction times to be improved under distracting conditions, inclusion of this parameter seems warranted.

Since the tasks included here are considered cognitively complex under Campbell's [1988] classification, performance degradation should be expected if the "change gears" mechanism applied.

*H3: Increased frequency of interruption will result in decreased performance levels.*

Outcomes where frequency effects are essentially zero may suggest that the change gears mechanism has minimal effect. A "D term" of the wrong sign would tend to support contentions by Weick [1985] or Sanders and Baron [1975].

### 3.5.4 Length of Interruption

As discussed in the "Forgetting Mechanism" section, the length of interruption is a logical inclusion for study. It certainly seems logical to assume that one of the possible effects will be due to workers forgetting where they were.

It should be noted that the length of time studied in forgetting experiments [Bailey, 1989; Christiaansen, 1980; and Schendel and Hagman, 1982] are fairly long (days to months.) For purposes of this study, interruption lengths will be considered in terms of minutes. Since the projected rate of forgetting varies as the log of the interruption length, the relatively short periods studied here may not result in detectable effects. However, if a "D term" is present, evidence points to it being of the correct sign.

*H4: Longer interruptions will decrease performance more than short interruptions.*

### 3.5.5 Combined Mechanism

As discussed in Section 3.4, both the forgetting and change gear mechanisms could be potentially applicable. These two mechanisms are not mutually exclusive. They could coexist and both be present in the "D term" (along with other yet unidentified effects). Since both mechanisms present plausible explanations of deleterious effects, their coexistence and interaction merits examination.

If either the forgetting or the change gears mechanism alone explains the presence of the "D term", then the other mechanism would be expected to contribute nothing to a combined treatment. Since the primary purpose of this experimental research is exploratory, the following hypothesis is advanced in support of a combined mechanism:

*H5: Simultaneously increasing both the length and the frequency of interruption will result in lower performance levels than equivalent increases for either factor alone.*

### 3.5.6 Presence of a Dominant Mechanism

The time dependency of the forgetting mechanism – when coupled with the time independence of the change gears mechanism – suggests that extremes in the relative length of interruption may cause one mechanism to dominate the other.

Consider the effects in the context of a single interruption. If the time for changing gears is fixed, then changing gears should dominate forgetting for very short interruption duration. For very long interruptions, forgetting should dominate change gears.

Now consider the situation where a knowledge worker has 30 minutes of interruptions within a four hour time span. If both mechanisms are present, then 15 two-minute interruptions should degrade performance more than two fifteen-minute interruptions. This would result from the additive effects of the change gear mechanism. This suggests:

*H6: If total interruption time is equivalent, many short interruptions will result in lower performance scores than a few long interruptions.*

If results do not confirm this hypothesis, then the forgetting mechanism may be dominating the change gears mechanism. If *H5* can be confirmed but *H6* cannot, then the proposed explanatory mechanisms may not be valid.

# 4. METHODOLOGY

A 2 x 2 x 2 factorial design with two control groups is used for the empirical study. A Quality Measured Model is employed with three variables of interest:

- task complexity (Hi and Lo values);
- frequency of interruption (Hi, Lo, Control); and
- length of interruption (Hi, Lo, Control).

To facilitate understanding the methodology employed, Section 4.1 will first discuss the primary task and its relationship to the subjects. Section 4.2 will explain the experimental design in detail – including discussions of several covariates included to validate results. Section 4.3 focuses on implementation details.

## 4.1 SELECTION OF TASK AND SUBJECT POOL

The population of junior/senior undergraduates in the business college provides a large pool of convenient subjects. In order to effectively simulate the knowledge work a task is required that falls within the knowledge domain of the subject group. The task should draw upon learned skills and assimilated expertise of the subjects. A simulated course examination was chosen as the primary task.

### 4.1.1 Primary Task Description

Subjects were offered an opportunity to take a "preview" final examination for classes in Production and Operations Management (taught by the researcher). As upper division students in a university characterized by large classes, subjects were expected to be proficient at taking multiple choice exams. This familiarity should minimize effects due to strategy selection. [Reder, 1987]

The use of a "sample final" exam as the primary task offered several advantages:

- Student subjects were familiar with the format.
- The sample exam should provide intrinsic incentives to participate.
- The format should stretch student study time.
- The exam could be readily graded for performance.
- The experimenter has extensive experience in designing and evaluating the effectiveness and difficulty of examinations.
- Material on the "sample final" has been covered in class and on two previous midterm examinations.

- Scores for two midterm exams taken in class were available on each subject as a measure of capability .

#### 4.1.2 Undergraduate Students as Knowledge Workers

With the basic task defined, this methodology presented some unique opportunities for study in that:

- The subject pool has equivalent exposure to the material.
- The subject matter is readily scored.
- The material requires computational expertise and an established knowledge base.
- The subjects are no more than 3 years away from potential employment in the knowledge worker setting.
- The "trial exam" represents a cognitively complex task.
- There is a completion time - performance quality tradeoff.

As with any laboratory study, additional research would be required to confirm that results actually extend to the workplace. However, these subjects, working this task, presented an reasonably homogeneous subject pool with respect to knowledge, ability, and familiarity with the task requirements. They presented an attractive opportunity to isolate the effects of interruptions – without many of the obscuring environmental factors present in an industrial setting.

#### 4.1.3 Introduction of Relative Task Difficulty

Based on feedback from a student focus group, and the researcher's past experience, difficulty in exams can be controlled by:

- introducing superfluous information;
- "burying" information in text rather than presenting it in tabular form; and
- introduction of complex scoring systems.

To allow for uniformity in scoring the "simple" version of the task presented 40 questions worth 10 points for each correct answer. Chain calculations are minimized and students are provided with cues to guide them in solving problems. Questions are grouped into sets which relate to a common set of problem information (See Figure 4.1).

For the difficult version, problem sets are aggregated into a single response. Necessary information is presented in textual rather than tabular format. Superfluous information is added to increase difficulty. (See Figure 4.1 for an example). When a problem set from the simple version is aggregated, the point totals are equivalent. For example, three 10-point questions will become one, 30-point question.

Thus, there is an opportunity to have an equivalent score for the same amount of computational effort.

In both the simple and complex versions, an unanswered question scores zero. An incorrect response subtracts 20% of the point value from the subject's cumulative score. The introduction of this penalty should discourage wild guesses, emphasize the "quality" component of the work in the minds of the subjects and provide for wider variability in scoring.

Several other researchers have used mathematically based tasks to present abstract, cognitively complex experimental presentations [Kintsch and Greeno,1985; Okogbaa and Shell,1986; and Woodhead,1965]. Bassock and Holyoak [1989] examined the notion of knowledge gained by analogy. This research examined transformations of isomorphic physics and algebra problems. Subjects readily recognized problems they knew how to solve – even when presented in an unfamiliar context. Thus, subjects here should recognize problems – even though the context differs from the usual exam format.

Figure 4.1: Typical Problem in Simple and Complex Formats

**Problem Set from Simple Version of Primary Task**

HELP (F1) FOR TABLES

Calculate the NPV @ 10% for the following purchases:(Q# 1-3)

| | MACHINE A | MACHINE B |
|---|---|---|
| Life : | 8 years | 4 years |
| Initial cost : | $30,000 | $14,000 |
| Annual Op cost: | $5,000/yr | $6,000/yr |
| Salvage Value : | $6,000 | $2,500 |

1. Approximate NPV of one Machine A is:
   A. $59,477 B. $62,675 C. $240,598
   D. $53,873 E. $50,675 F. None of these

2. Approximate NPV of one Machine B is :
   A. $33,020 B. $37,968 C. $88,373
   D. $32,035 E. $31,313 F. None of these

3. If you must choose between these two machines for an 8 year time span, what is the NPV of the better deal?
   A. $53,873 B. $66,040 C. $59,477
   D. $64,070 E. $50,675 F. None of these

**Equivalent Problem on Difficult Version of Primary Task**

You must choose between two electronic scanners for the warehouse. The ACME is a real Cadillac among scanners, with an 8 year life and an initial cost of $30,000, it has an estimated salvage value of $6,000 at the end of its life. The MR. SCAN unit is a lower end unit. MR. SCAN costs only $14,000 with an estimated salvage value of $2,500 at the end of its four year life. Annual cost of operation is similar, $6,000 for MR. SCAN and $5,000 for the ACME. Your cost of capital is 10%. ** HELP (F1) for Tables **

1. [30 points] Based only on a comparison of NPV, what is the approximate NPV of the better deal?
   A. $53,873 B. $59,477 C. $66,040
   D. $52,699 E. $50,675 F. None of these

### 4.1.4 Introduction of Interruptions

Three problems were considered critical to controlling the introduction of interruptions into the primary task:

1. Ensuring uniformity in length and content of interruptions among subjects and treatments.
2. Controlling task time across treatments and subjects.
3. Monitoring subjects to confirm a shift in cognitive focus – without relinquishing control of the time of the interruption.

To mitigate problems 1 and 2, personal computers were selected as the presentation medium. Custom software was written which presented the tasks in a format which closely approximated a "paper" exam.

In response to problem 3, interruptions are presented in the form of "trivia" questions. A screen with trivia questions obscures the text of the primary task – thus controlling direct access to the primary task. Subjects have 15 seconds to respond to each question before the question disappears from view. (A 30 second interruption is comprised of 2 questions.) The use of trivia questions for all interruptions maintains consistency among treatments and ensures that any differences among treatments are due to length or frequency of interruption.

The trivia questions are drawn from science, history, geography, art and literature. These areas should provide adequate shift in cognitive focus – provided subjects attempt to answer them. To provide an incentive for answering the trivia questions, subjects were told that correct trivia answers would add 1 point to their scores with no penalty for incorrect answers.

## 4.2 EXPERIMENTAL DESIGN

The experiment is a completely randomized, 2 x 2 x 2 factorial design with 2 control groups, and independent covariates. (See Table 4.3 for Variables.) The Linear Model is presented in Chapter 5.

**Table 4.1:** Experimental Treaments

| EXPERIMENTAL TREATMENTS | | | | |
|---|---|---|---|---|
| # | Treatment Name | Number of Interruptions | Seconds per Interruption | Task |
| 0 | SControl | 0 | 0 | Simple |
| 2 | SFLo DLo | 2 | 30 | Simple |
| 4 | SFLo DHi | 2 | 90 | Simple |
| 6 | SFHi DLo | 6 | 30 | Simple |
| 8 | SFHi DHi | 6 | 90 | Simple |
| 1 | CControl | 0 | 0 | Complex |
| 3 | CFLo DLo | 2 | 30 | Complex |
| 5 | CFLo DHi | 2 | 90 | Complex |
| 7 | CFHi DLo | 6 | 30 | Complex |
| 9 | CFHi DHi | 6 | 90 | Complex |

### 4.2.1 Tests of Research Hypotheses

This experimental design will allow testing of all research hypotheses and investigation of experiment specific effects. Assuming the "D term" will be significant in all cases, Table 4.2 depicts outcomes that support the hypotheses. The design also allows an appraisal of any dominance of frequency and duration effects. Treatments 4, 5, 6 and 7 have the same total interruption time — 3 minutes. A significant difference among these treatments suggests a dominance of one interruption mechanism.

Table 4.2: Experimental Outcomes Supporting Hypotheses

| Outcome | Supports |
|---|---|
| 3, 5, 7, 9 > 1<br>2, 4, 6, 8 > 0 | H1: Deleterious Effects from Interruption. |
| 3 > 2 ; 5 > 4<br>7 > 6 ; 9 > 8 | H2: Difficulty Amplifies Deleterious Effects |
| 6 > 2; 8 > 4<br>7 > 3; 9 > 5 | H3: Frequency Amplifies Deleterious Effects |
| 4 > 2; 8 > 6<br>5 > 3; 9 > 7 | H4: Duration Amplifies Deleterious Effects |
| 8 > 6 and 8 > 4<br>9 > 7 and 9 > 5 | H5: Interaction Effects |
| 6 > 4 and 7 > 5 | Frequency dominates duration |
| 4 > 6 and 5 > 7 | Duration dominates frequency |

### 4.2.2 Randomization and Subject Assignment

A schedule of tentative sign-up times was presented to subjects in advance of the experiments. Subjects wishing to participate signed up for individual sessions of their choice over the three day period of the experiments. Session sizes varied from 5 to 16 participants. Fourteen stacks of 10 index cards each were prepared in advance of the experiments. Each card contained a unique six digit number that represented one of the ten experimental treatments. Each stack of ten cards contained a number that represented each of the ten treatments. The first subject was offered the shuffled, face down cards in stack #1. The subject then chose one of the ten cards. The second subject was offered a choice from the remaining nine cards. Subject three then chose from the remaining eight cards and so on until stack #1 was distributed. Subject 11 then chose from the ten cards in stack #2 and the cycle was repeated until all subjects in all sessions had completed the experiment.

This method ensured that assignment to treatments was random, and that the number of subjects assigned to individual treatments also remained balanced throughout the experiment — a safeguard against no shows. Effects from differing numbers of session participants; subject interaction; and the laboratory envirionment; were also randomly distributed among treatments.

### 4.2.3 Mesurement Model

The experiment used the Quality Measured Model for evaluation. The QMM was selected because:

1. it controls the time subjects spend on the experiment;
2. the task selected lends itself to an objective, point scoring system;
3. it introduces the completion time - performance quality conflict; and
4. it minimizes any fatigue effects by controlling the subject participation time;
5. it helps to minimize social facilitation effects from differential completion times.

### 4.2.4 Control groups

Since a primary objective for the study is to determine the deleterious effects of interruptions, two control groups performed the primary task without interruptions — one group for the "simple" task and one for the "complex" task. This allows establishing a "benchmark" standard for task score. Initially, 45 minutes of access time is allowed under each complexity level. Longer access would allow for more sensitivity in scoring but may tend to frustrate the subjects. Task times vary between 45 minutes for Control treatments and 57 minutes for Hi Frequency - Hi Duration treatments.

### 4.2.5 Covariates

4.2.5.1 Class Score — Subjects had completed two midterm exams in the experimenter's course at the time of the experiments. A class average was computed for these two midterms. The class average was then subtracted from each student's combined cumulative point score for these exams. This score is included as a covariate for task performance expectations. The experiment will not specifically treat inherent capability. However, including this covariate may help differentiate effects of interruptions from the variations in subject "knowledge".

4.2.5.2 Extra Minutes — One possible strategy for coping with interruptions would be to ignore them completely. While there is no way to determine if the subject truly ignored the interruption, trivia questions left unanswered should be highly correlated. To provide some check on this possibility, an additional variable is included. Trivia questions are monitored to determine if the subject answered the question or left it blank. Each trivia question that was not answered adds 0.25 minutes to the "extra minutes" covariate.

4.2.5.3 Gender — While no performance differential is anticipated on the basis of gender, the inclusion of gender as covariate is included for completeness.

4.2.5.4 Previous Computer User — Discussions with a pre-experiment focus group suggested that some student subjects may have very limited computer experience — and that they may be intimidated by the machines used for the experiment. A covariate was included to check for prior computer experience.

### 4.2.6 Interruption Manipulations

Short (DLo) interruptions are 30 seconds in duration. They consist of two trivia questions. Long (DHi) interruptions are 1.5 minutes each and consist of six trivia questions. This should be long enough to provide a reasonable distinction, but should not unnecessarily lengthen the experiment.

Low Frequency (FLo) consists of two interruptions during the task. High Frequency (DHi) consists of six interruptions. With these Frequency and Duration settings, the total task completion times are:

| Treatment #s | Task Length (minutes) |
|---|---|
| 0,1 | 45 |
| 2,3, | 46 |
| 4,5,6,7 | 48 |
| 8,9 | 54 |

Allowing 15 minutes for debriefing and training, and 10 minutes for getting to/from the experiment site, this keeps total subject time to less than 90 minutes (the time between classes).

Table 4.3: Experimental Variables for Linear Model

| Variable Name | Scale/Levels |
|---|---|
| DEPENDENT VARIABLES | |
| Task Score | -80 to +400 |
| Correct Pts | 0 to +400 |
| Wrong Pts | 0 to +400 |
| INDEPENDENT CATEGORICAL VARIABLES | |
| Task Complexity | Hi,Lo |
| Interruption Length | Hi,Lo,Control |
| Interruption Frequency | Hi,Lo,Control |
| INDEPENDENT COVARIATES | |
| Class Score | -150 to +150 |
| Gender | Male, Female |
| Previous Computer User | Yes/No |
| Extra Minutes | 0 to 9 |

## 4.3 OPERATIONALIZING STUDY VARIABLES

### 4.3.1 Outline of Experiment Procedure

Experiments were conducted at the Park Center MIS Microcomputer Laboratories on April 24, 26 and 27,1990. (See Appendix A.1 for a layout of the Park Center Labs) As subjects arrived, they were checked against sign-up sheets. The subject then chose a computer workstation and arranged his or her reference materials and calculator for ready access.

With all subjects present, the proctor read the introductory script for the experiment (Appendix A.2) and subjects began a self-paced, software training period. During the training period, the proctor circulated through the lab to answer individual questions.

As individuals completed the training period, each was offered an index card, face down, with the six digit treatment code (see Section 4.2.2). The treatment code was then entered on the computer. Instructions (see Appendix A.3) were displayed and the subject began the primary task.

As subjects completed the task, the software reported the final Subject Score (Subject Score = Total Correct Points - 0.2*[Total Wrong Points] + 1 point for each Correct Trivia Question)[1]. The proctor then distributed the questionnaire (Appendix A.4) as each subject finished. After filling out the questionnaire and signing the index card, the subject returned both to the proctor. After attaching the card to the questionnaire, any questions were answered, and the subject was released.

---

1. Recall that there are 400 points on each exam and a maximum of 36 trivia questions.

### 4.3.2 Subject Motivation

Student subjects were asked to participate in the experiment on a voluntary basis for extra credit in the class. Extra credit (amounting to 2% of course weight) was awarded equally to all participants for completing the experiment. Course grades were based on a point system. Cuts for final grades were made <u>prior</u> to inclusion of extra credit points – thus participation could help the student, but nonparticipation did not hurt class standing.

Performance motivation, in the form of a cash reward, was offered to the top scorer in each of the ten treatments as a performance incentive. Additionally, the top three scores in each cell treatment were placed in a lottery. The winner of the lottery acquired a registered factory copy of a spreadsheet program. Further motivation was provided by keeping track of individual scores. Subjects were told that they would be individually evaluated to help minimize the effects of social loafing. [Harkins & Jackson, 1985]

### 4.3.3 Instruction and Training

To ensure uniformity across sessions, a formal script was used to brief subjects about the purpose of the experiments. Very little deceit was employed. Most knowledge workers who are subject to interruption are aware that the possibility exists. Thus it served no purpose to deceive subjects about the possibility – especially because the interruptions require a response and are part of the scoring system. To minimize the possibility of individual treatment clues, the proctor did not see the index card selected by the student until well after the experiment had begun. The six digit number was partially generated from a uniform random distribution to minimize the possibility of prior treatment knowledge.

The informal training in use of the software required subjects to work through all of the primary functions of the software. The proctor encouraged subjects to take their time during training. In order to move on to experiment task, subjects were required to demonstrate proficiency. Differences in training times also helped reinforce the notion that finishing times did not reflect performance.

### 4.3.4 Software Design

The software for the experiment consists of two modules – the training module and the experiment module.

The training module presents a small practice task. Subjects are required to move through a self-paced sequence that demonstrates the function of each of the keys used for scrolling text and answering questions. Questions can be answered by typing in a single letter. Help screens are provided to explain the functions of keys and to access necessary reference tables. The subject may then scroll through the practice task, answering questions and accessing help screens or tables. When the subject is satisfied that he or she has mastered the techniques (and demonstrates the capability by correctly exiting the program), the experiment module is accessed by typing in a six digit number provided by the experiment proctor.

The experiment module accesses a database of treatment parameters based on the six digit number. The subject is then presented with task specific instructions on scoring, the number of questions in the task, and reinforcement of conditions stated in the experiment script. (The text of the instructions can be found in Appendix A.3.) When the subject has finished reading the instructions, the Task Time begins.

The software ensures that all subjects have exactly the same Task Time (45 minutes) — thus score will be the only performance measure. For subjects with interruptions, the experiment module generates periods between interruptions from a uniform random distribution. No interruptions are scheduled for the first three minutes or the last three minutes of the Task Time.

The interruption mechanism constantly checks the current time. When an interruption is due, the program freezes the Task Time, covers the text of the primary task, and begins the sequence of trivia questions. At the end of the interruption, the primary task is reinstated and the remaining Task Time restored. Remaining Task Time is constantly displayed on the main screen. Subjects cannot stop the interruptions nor can they alter the time of the interruption.

After the experiment is completed, the software generates summary information files for later processing. Files contain information on:

- treatment/ subject
- date/time
- total points attempted
- total points for correct answers (Correct Points)
- total points for incorrect answers (Wrong Points)
- TOTSCORE (Correct Points - 0.2[Wrong Points] )
- trivia questions answered
- trivia questions correctly answered,
- total interruption time, and
- the correct answer and subject's answer for each question in the primary task.

Additionally the software generates a file containing the clock time for the following events:

- the beginning and end of the experiment
- each keystroke used for scrolling;

- each question answered;
- each trivia question answered; and
- the beginning and end of each interruption.

### 4.3.5 Post Experiment Questionnaire

After completing the experiment, subjects were asked to complete a questionnaire (Appendix A.4). The questionnaire provides demographic information, and self reporting information on experiment manipulations.

Other than demographics, questions are presented on an anchored scale of 1 to 7. Questions ask subjects to either: strongly agree or strongly disagree with statements; or estimate their performance on extreme scenarios.

Part I gathers information on age, sex, birthplace, and prior computer experience. Part II asks subjects to compare the difficulty of the experiment questions with those they have had in the experimenter's course. Subjects are also asked to how they would score on the equivalent exam if taken on paper. All subjects respond to Parts I and II.

Part III is directed only to those who experience interruptions. Three questions are directed to clarifying possible interruption mechanisms. Subjects are asked:

7. How do you think you would have scored if you were not interrupted? (Higher/Lower)

8. How did the interruptions affect your ability to concentrate on the practice exam? (Helped/Hurt)
9. When an interruption arrived, what was your initial reaction? (Pleased/Displeased)

Question 7 attempts to identify the subject's perception of performance degradation. Question 8 tries to confirm a shift in cognitive focus. Question 9 indirectly asks if the the interruption is a "pleasant break" from the exam.

Subjects are next asked how they believe they would score if they took an exam under typical control conditions (i.e. 100 minutes of uninterrupted time) versus 120 minutes with 20 minutes of interruption.

Finally, subjects are asked to express a preference for organizing the trivia questions. "Suppose you had to answer 12 trivial questions during the course of the practice exam. Which of the following interruption scenarios would you prefer?" Alternatives range from 1 interruption with 12 questions to 12 interruptions with 1 question each. The responses here are directed toward a formulating a coping strategy. All subjects are asked for comments.

# 5. ANALYSIS OF EXPERIMENT RESULTS

The experiment sessions resulted in 122 valid subjects. One subject left in the middle of the experiment, another experienced computer failure and elected not to reschedule. Results are statistically significant, but do not confirm theoretical hypotheses. Key findings:

- Complexity manipulations had the predicted effect
- Deleterious effects are present.
- Deleterious effects for Lo Frequency and Lo Duration treatments are significant for complex tasks.
- Performance levels under Hi Duration interruptions are superior to those under Lo Duration.

## 5.1 TASK SCORING RESULTS

Three performance measures are captured for evaluating performance under the QMM:

- Correct Points — total questions answered correctly
- Wrong Points — total questions answered incorrectly
- TOTSCORE — Correct Points - 0.2(Wrong Points)

Overall, the complexity manipulation proved quite successful. TOTSCORE averages are 141.1 for the simple and 42.2 for the complex tasks. The majority of difference stems from a much higher error rate (Wrong Points 150.7 for complex vs. 70.8 for simple). Total points attempted for both task types (223 vs 226) are equivalent.

For simple tasks, TOTSCORE followed a pattern consistent with the Class Score covariate. Low TOTSCORE appears to correspond to low Class Score in all cells except Hi Frequency-Hi Duration. For the complex task, Class Score appears to be less correlated to task performance.

Results of all performance measures (with Class Score, Gender, and Previous User covariates) are presented in Tables 5.1 and 5.2. These tables present numbers of participants by treatment with appropriate weighted averages. Overall 72 males and 50 females participated. Of the 72 males, 55 were previous computer users with 39 of 50 females having prior computer experience. Average Class Score for males is 1.1 and -0.8 for females. Males averaged 116.7, 95.9 and 103.9 on Correct Points, Total Score and Wrong Points. Equivalent scores for females are 109.6, 85.4, and 120.8.

Graphic presentations of scoring data are provided in Figures 5.1 and 5.2 along with Class Score + 100 for contrast. A listing of the raw data from the experiment can be found in Appendix B.1.

Table 5.1: Summary of Complex Task Results

| | NUM | Class Score | Correct Points | TOT-SCORE | Wrong Points | Prev User |
|---|---|---|---|---|---|---|
| Male | 7 | -2.6 | 94.3 | 70.9 | 117.1 | 6 |
| Female | 5 | -9.0 | 82.0 | 53.2 | 144.0 | 4 |
| C CONTROL | 12 | -5.3 | 89.2 | 63.5 | 128.3 | 10 |
| Male | 9 | -11.0 | 56.7 | 33.1 | 117.8 | 6 |
| Female | 3 | 40.0 | 80.0 | 50.0 | 150.0 | 2 |
| C FLo DLo | 12 | 1.8 | 62.5 | 37.3 | 125.8 | 8 |
| Male | 6 | 11.5 | 98.3 | 59.7 | 193.3 | 5 |
| Female | 6 | -4.5 | 30.0 | 0.3 | 148.3 | 5 |
| C FLo DHi | 12 | 3.5 | 64.2 | 30.0 | 170.8 | 10 |
| Male | 8 | -5.6 | 73.8 | 46.3 | 137.5 | 6 |
| Female | 5 | -11.2 | 46.0 | 13.6 | 162.0 | 5 |
| C FHi DLo | 13 | -7.8 | 63.1 | 33.7 | 146.9 | 11 |
| Male | 6 | 6.2 | 95.0 | 65.7 | 146.7 | 5 |
| Female | 6 | -32.5 | 71.7 | 28.3 | 216.7 | 5 |
| C FHi DHi | 12 | -13.2 | 83.3 | 47.0 | 181.7 | 10 |

TOTALS

| | NUM | Class Score | Correct Points | TOT-SCORE | Wrong Points | Prev User |
|---|---|---|---|---|---|---|
| Male | 36 | -1.6 | 81.1 | 53.2 | 139.4 | 28 |
| Female | 25 | -8.1 | 59.6 | 26.2 | 166.8 | 21 |
| COMPLEX | 61 | -4.2 | 72.3 | 42.2 | 150.7 | 49 |
| Male | 36 | 3.7 | 152.2 | 138.6 | 68.1 | 27 |
| Female | 25 | 6.5 | 159.6 | 144.6 | 74.8 | 18 |
| SIMPLE | 61 | 4.8 | 155.2 | 141.1 | 70.8 | 45 |
| Male | 72 | 1.1 | 116.7 | 95.9 | 103.8 | 55 |
| Female | 50 | -0.8 | 109.6 | 85.4 | 120.8 | 39 |
| TOTAL | 122 | 0.3 | 113.8 | 91.7 | 110.8 | 94 |

Table 5.2: Summary of Simple Task Results

| | NUM | Class Score | Correct Points | TOT-SCORE | Wrong Points | Prev User |
|---|---|---|---|---|---|---|
| Male | 9 | 18.8 | 157.8 | 142.4 | 76.7 | 7 |
| Female | 3 | 26.0 | 180.0 | 170.0 | 50.0 | 2 |
| S CONTROL | 12 | 20.6 | 163.3 | 149.3 | 70.0 | 9 |
| Male | 8 | -13.6 | 131.3 | 118.8 | 62.5 | 7 |
| Female | 4 | -12.3 | 112.5 | 95.0 | 87.5 | 3 |
| S FLo DLo | 12 | -13.2 | 125.0 | 110.8 | 70.8 | 10 |
| Male | 6 | -1.0 | 175.0 | 163.7 | 56.7 | 5 |
| Female | 7 | 14.4 | 168.6 | 156.3 | 61.4 | 5 |
| S FLo DHi | 13 | 7.3 | 171.5 | 159.7 | 59.2 | 10 |
| Male | 6 | 6.2 | 155.0 | 143.7 | 56.7 | 3 |
| Female | 6 | 15.5 | 153.3 | 135.7 | 88.3 | 5 |
| S FHi DLo | 12 | 10.8 | 154.2 | 139.7 | 72.5 | 8 |
| Male | 7 | 5.9 | 147.1 | 130.6 | 82.9 | 5 |
| Female | 5 | -12.2 | 180.0 | 163.6 | 82.0 | 3 |
| S FHi DHi | 12 | -1.7 | 160.8 | 144.3 | 82.5 | 8 |

TOTALS

| | | | | | | |
|---|---|---|---|---|---|---|
| Male | 36 | 3.7 | 152.2 | 138.6 | 68.1 | 27 |
| Female | 25 | 6.5 | 159.6 | 144.6 | 74.8 | 18 |
| SIMPLE | 61 | 4.8 | 155.2 | 141.1 | 70.8 | 45 |
| Male | 36 | -1.6 | 81.1 | 53.2 | 139.4 | 28 |
| Female | 25 | -8.1 | 59.6 | 26.2 | 166.8 | 21 |
| COMPLEX | 61 | -4.2 | 72.3 | 42.2 | 150.7 | 49 |
| Male | 72 | 1.1 | 116.7 | 95.9 | 103.8 | 55 |
| Female | 50 | -0.8 | 109.6 | 85.4 | 120.8 | 39 |
| TOTAL | 122 | 0.3 | 113.8 | 91.7 | 110.8 | 94 |

Figure 5.1
Scores For Complex Task by Treatment

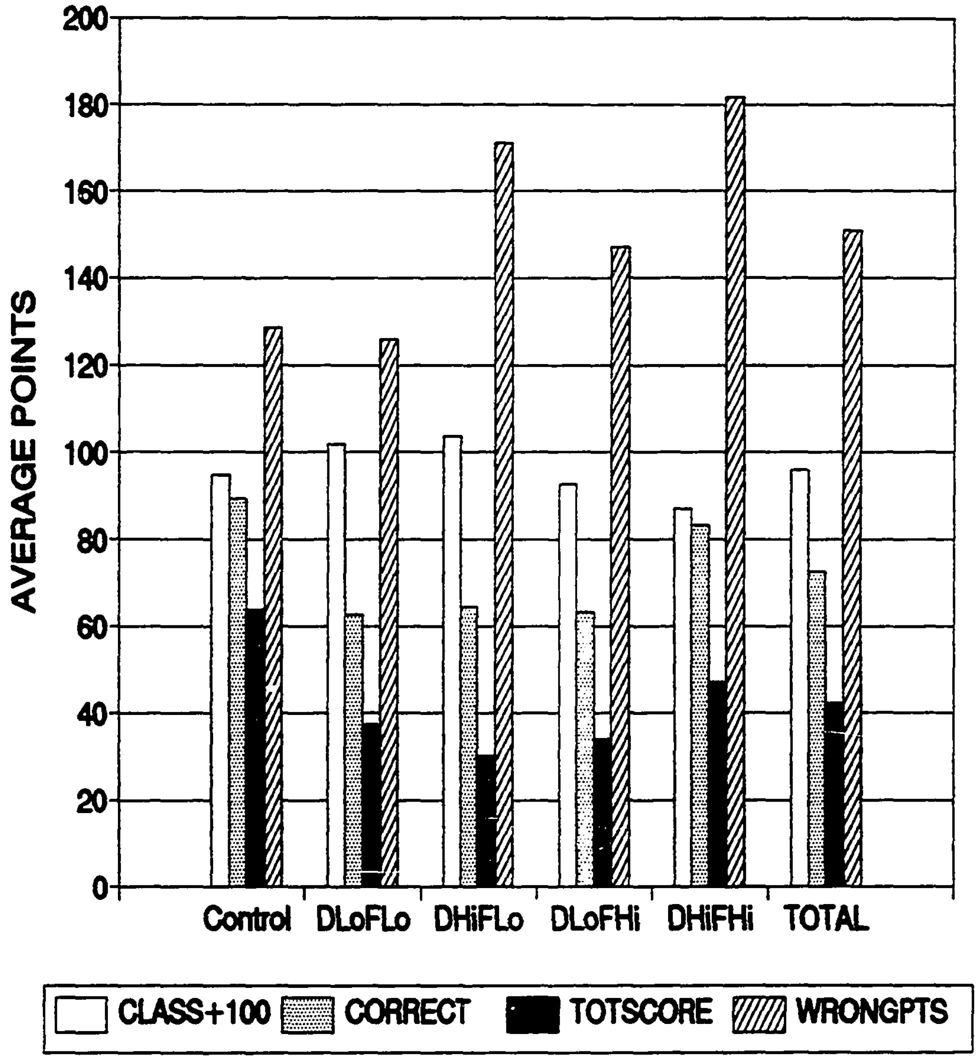
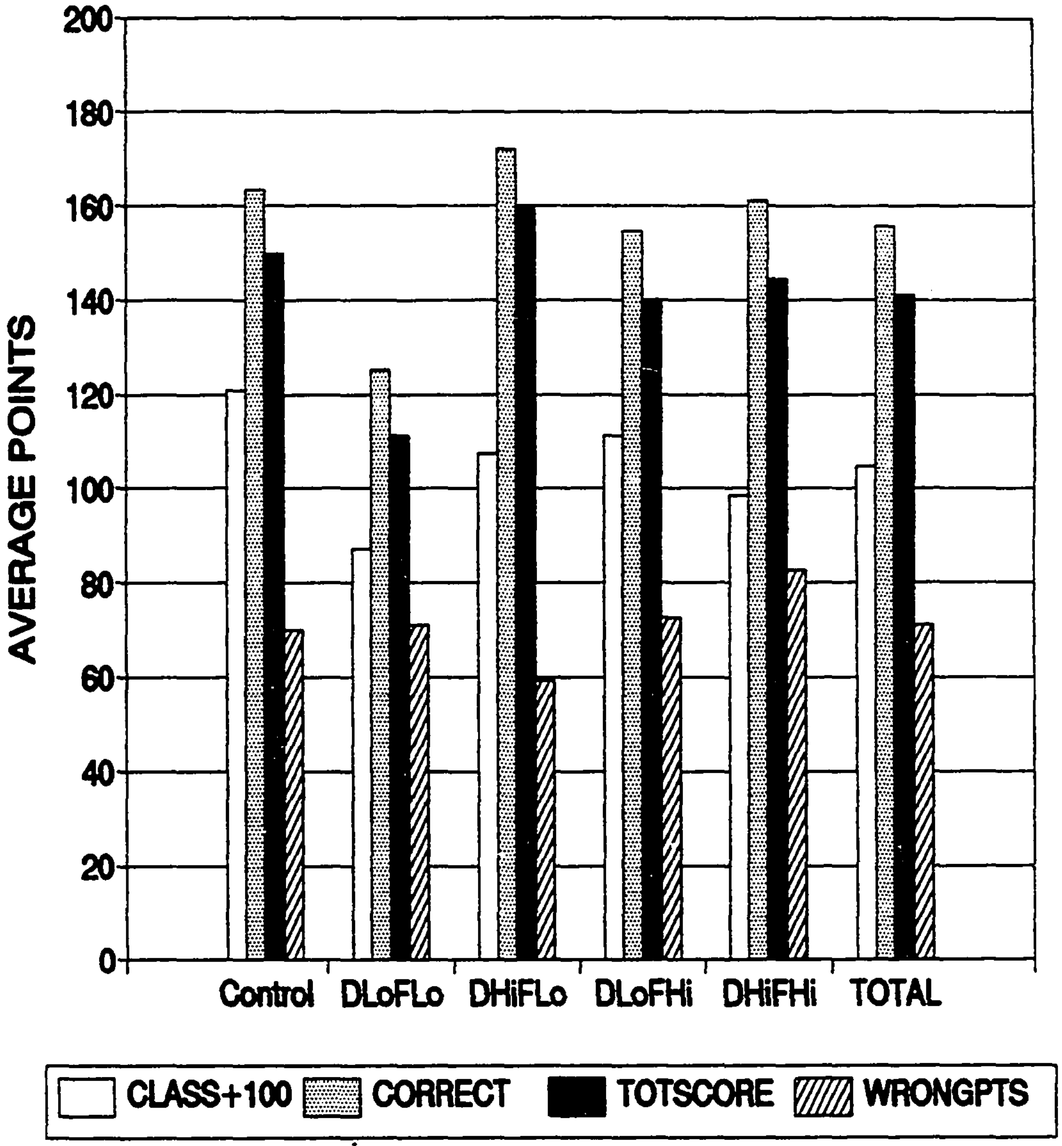


Figure 5.2
Scores for Simple Task by Treatment



## 5.2 LINEAR MODEL

### 5.2.1 Analytical Methodology

The general form of the linear model is described in Figure 5.3. The procedure used for analyses is as follows:

1. Perform a regression on the full data set with all covariates and interaction terms.

2. Remove three and four variable interaction terms. Check for significance of removed terms using Fischer's F ratio for reduced models ("Reduced F").

3. Test for influential observations on the reduced model. Successively eliminate 1 to 4 of the most influential subjects.

4. Perform a stepwise regression on remaining variables and subjects. Add variables until F to enter value indicates trivial significance for next candidate.

5. Construct a series of regression models by first introducing any covariates present in the stepwise regression to obtain reduced sums of squares for subsequent analyses. Remaining variables are introduced in the order of entry identified by the stepwise regression.

Figure 5.3: Linear Model

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 C_i + \beta_3 F_i + \beta_4 D_i + \beta_5 (X_i)C_i + \beta_6 (X_i)F_i + \beta_7 (X_i)D_i + \beta_8 (C_i)F_i + \beta_9 (C_i)F_i + \beta_{10} X_i (C_i)F_i + \beta_{11} X_i (C_i)D_i + \beta_{12} D_i (C_i)F_i + \beta_{13} X_i (C_i)F_i (D_i) + \beta_{14} S_i + \beta_{15} U_i + \beta_{16} M_i + \epsilon$$

WHERE:

$i \equiv$ Subject

$Y_i \equiv$ Score on primary task

$X_i \equiv$ Score on Midterms - Mean Score on Midterms

$M_i \equiv$ 0.25(Trivia Questions Not Answered)

$C_i \equiv$ Task Complexity $\{ C_i$ = +1 if complex; = -1 if simple $\}$

$F_i \equiv$ Interrupt Frequency $\{ F_i$ = -1 if control; = 0 if 2 (Lo); = +1 if 6 (Hi) $\}$

$D_i \equiv$ Interrupt Duration $\{ D_i$ = -1 if control; = 0 if 2 (Lo); = +1 if 6 (Hi) $\}$

$S_i \equiv$ Gender $\{ S_i$ = +1 if female; = -1 if male $\}$

$U_i \equiv$ Previous Computer User $\{ U_i$ = +1 if yes ; = -1 otherwise $\}$

### 5.2.2 Final TOTSCORE Model

TOTSCORE was chosen as the dependent performance measure because:

- TOTSCORE incorporates effects of both Correct Points and Wrong Points measures; and
- subjects were instructed that TOTSCORE would be the criterion for performance evaluations.

Three final regression models with TOTSCORE as the dependent variable are shown in Table 5.3.

Table 5.3: Regression Model Summary

| | MODEL I | | MODEL II | | MODEL III | |
|---|---|---|---|---|---|---|
| $R^2$ | 0.7025 | | 0.6995 | | 0.6517 | |
| Std Err | 37.89 | | 38.23 | | 42.12 | |
| TERM | Coeff | Tratio | Coeff | Tratio | Coeff | Tratio |
| CONST | 79.622 | 17.63 | 80.266 | 17.69 | 85.110 | 17.37 |
| C | -43.875 | 12.04 | -44.992 | 12.29 | -44.295 | 11.10 |
| X | 0.684 | 7.43 | 0.645 | 6.91 | 0.680 | 6.67 |
| CD | -20.509 | 3.36 | -11.750 | 2.51 | -10.170 | 1.98 |
| FD | 21.334 | 3.01 | 20.695 | 2.90 | 15.847 | 2.03 |
| CF | 12.441 | 2.03 | | | | |

Approximate T Critical Values for 113-115 degrees of freedom

| | | | | | |
|---|---|---|---|---|---|
| Percentile | 0.950 | 0.980 | 0.990 | 0.995 | 0.9995 |
| T Value | 1.660 | 2.080 | 2.361 | 2.619 | 3.376 |

Model I is developed using the 5 step procedure outlined in section 5.2.1. Three observations (subjects 511, 602 and 611) are eliminated at step 3.

Model II uses a stepwise elimination procedure based on lowest T Ratio for regression terms. The model is first run with all variables and observations. The variable with the least significant T value is then removed and the reduced model is rerun. This procedure is repeated until all remaining variables appear significant at the 95% level. After confirming elimination of all variables except those in Model I, a test for influential observations suggested eliminating subjects 511, 602 and 709. Elimination of these subjects reduces the significance of the CF term. The result is Model II.

Model III reintroduces subjects 511, 602, and 611 into Model I.

These three models are presented to demonstrate the robustness of the regression results. (These models are compared, in graphic form, for three values of Class Score in Figures 5.4, 5.5, and 5.6.) Model III confirms the presence of significant deleterious effects with all subjects present. Only Model I contains a Frequency effect which serves to partially counteract the effects of Duration because of opposing signs. Subsequent discussions will focus on Model I as the final form of the TOTSCORE model. (The Step Wise regression for MODELI can be found in Appendix B.2. Regression details for MODELI are found in Appendix B.3.)

Figure 5.4
Regression Models - Class Score = 0

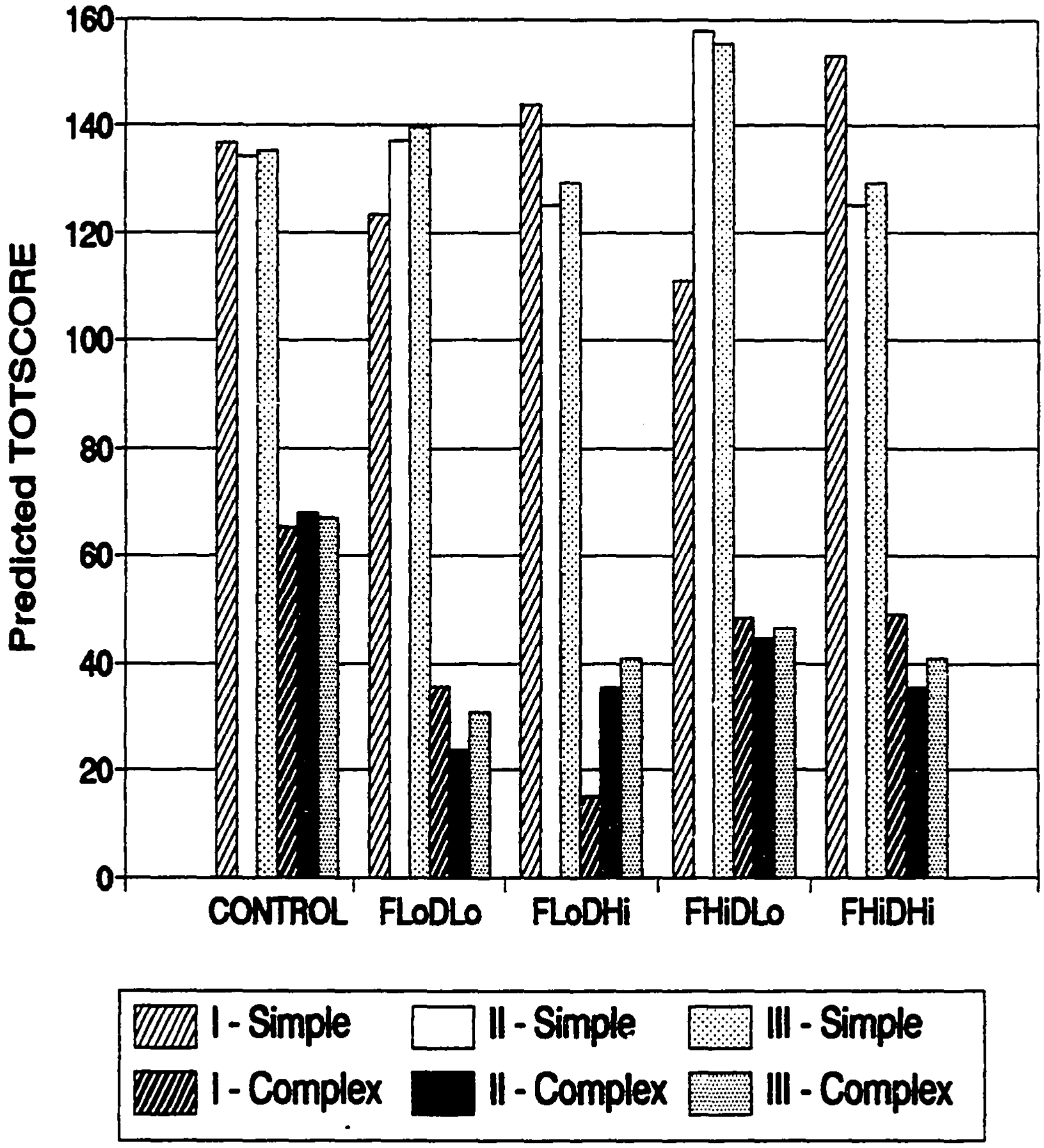


Figure 5.5
Regression Models - Class Score = -50

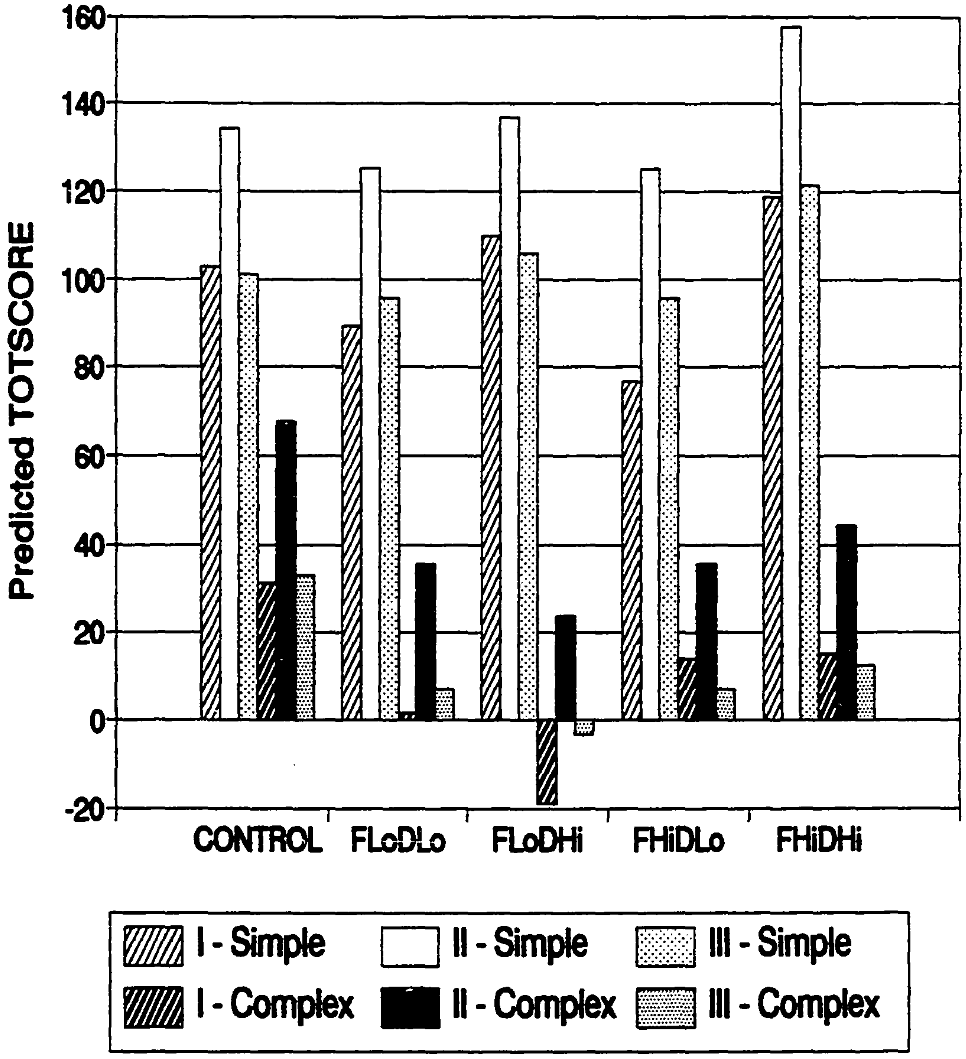
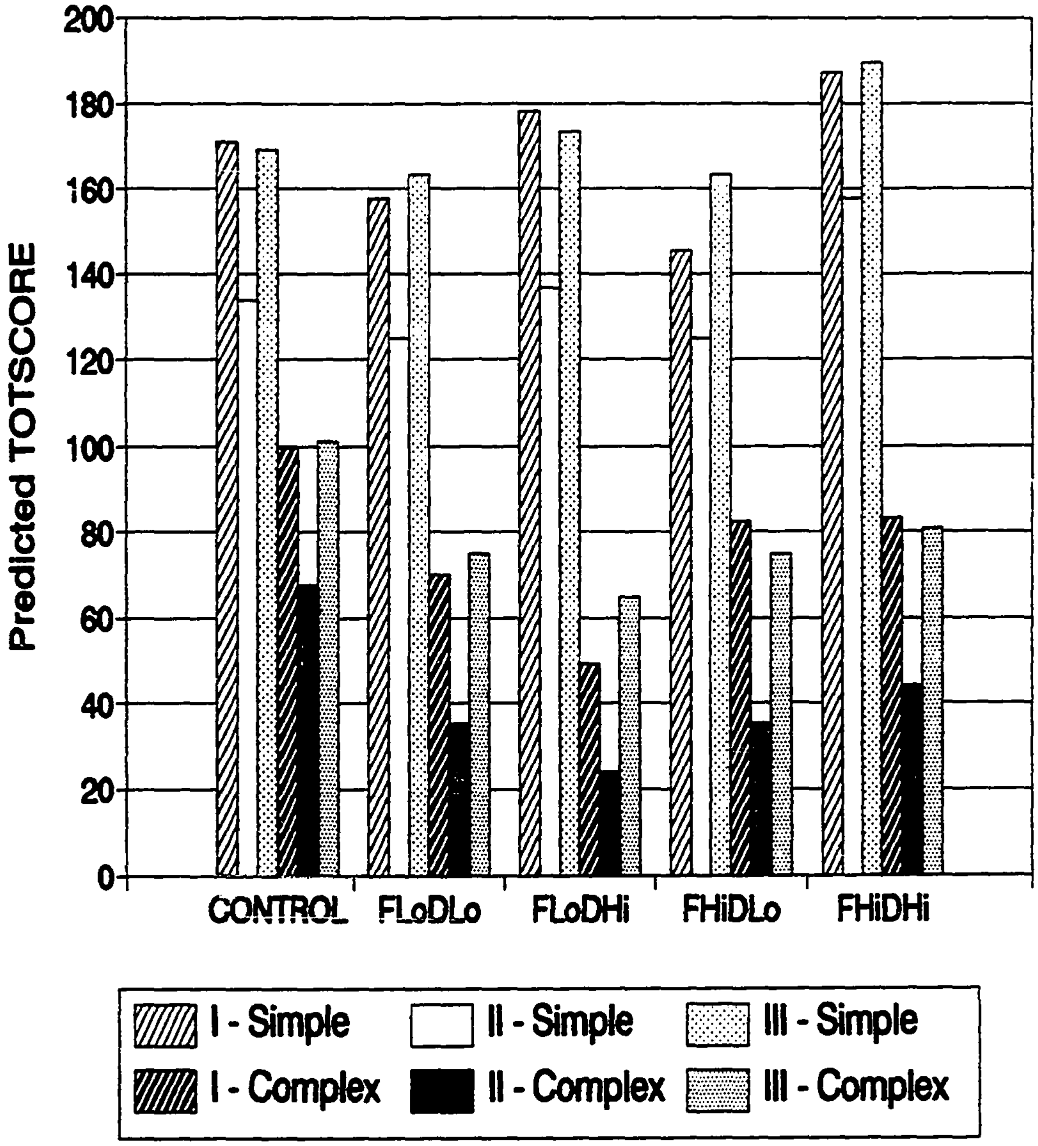


Figure 5.6
Regression Models - Class Score = +50



### 5.2.3 Influential Observations in Final Model

Three observations are eliminated from Model I - subjects 511, 602 and 611. In all three cases performance is significantly higher than predicted by Class Score and/or the treatment manipulations. In Table 5.4 these subjects are examined by their relative rankings of TOTSCORE and Class Score. (The scores for treatments involving these subjects are presented, in graphical form, in Figures 5.7 and 5.8.)

Table 5.4: Rank Comparison of Influential Observations

| Subject | 511 | 602 | 611 | (n) |
|---|---|---|---|---|
| TOTSCORE | 144 | 276 | 206 | |
| Rank in Task Type | 2 | 1 | 5 | (62) |
| Rank in Treatment | 1 | 1 | 2 | (12) |
| Class Score | -18 | 38 | -12 | |
| Overall Rank | 83 | 22 | 78 | (122) |
| Rank in Task Type | 39 | 13 | 41 | (62) |
| Rank in Treatment | 7 | 4 | 10 | (12) |

Subject 602 scored 276, 42 points higher than the 2nd best overall score. With an overall class rank of 22, this subject far exceeded predictions for the model.

Subjects 511 and 611 rank 83rd and 78th overall in class performance. Subject 511 had the 2nd best score on the complex task and subject 611 has the 5th best score on the simple task. Clearly, these subjects far exceeded expectations for both overall and treatment performance.

Figure 5.7
Outliers - Treatment 5 (CFLoDHi)

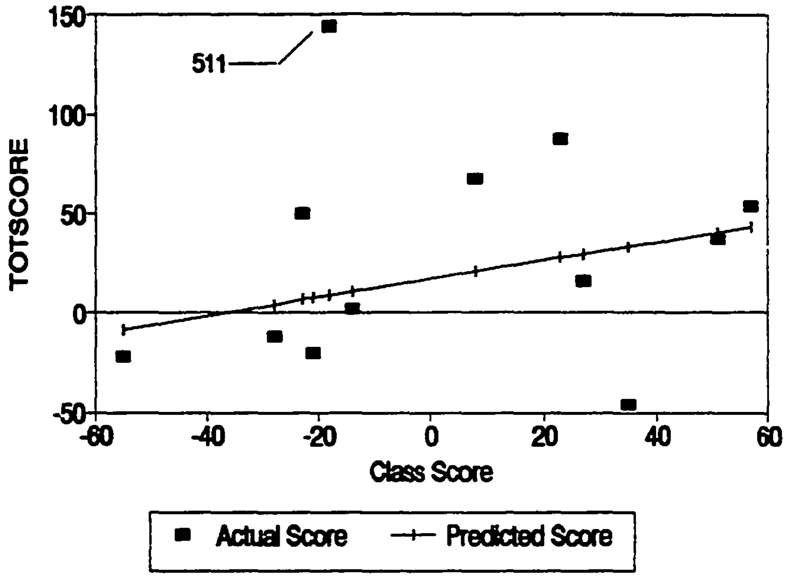
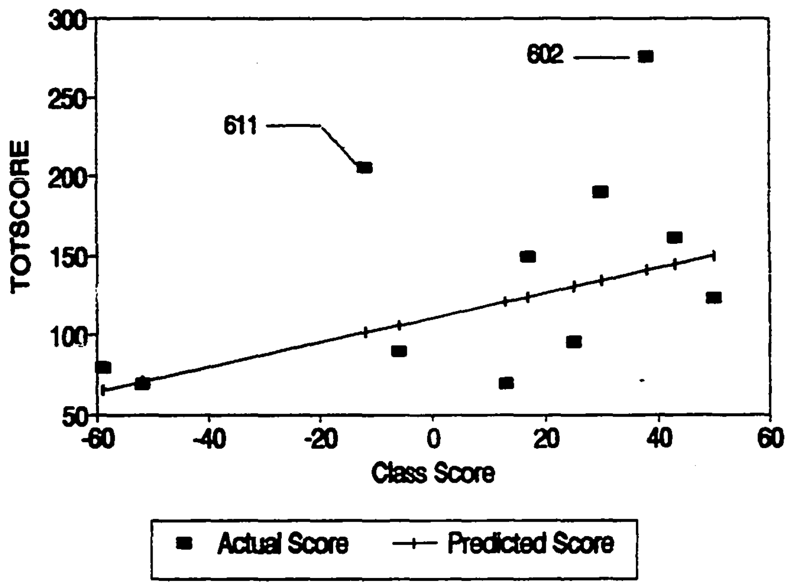



Figure 5.8
Outliers - Treatment 6 (SFHiDLo)



## 5.3 ALTERNATIVE ANALYSES

The need for alternative statistical analyses arises from two characteristics of this study:

- alternative performance measures; and
- a "hybrid" experimental design.

TOTSCORE is a combination of total points correct and total incorrect points. Each of these performance measures is addressed in Section 5.3.1.

The general form of MODELI does not fit any "pure" style of experimental design. The control groups are not truly part of the factorial design. When Frequency is at control level, so is Duration. To be thorough, two additional analyses are performed:

- single treatment covariance models to distinguish individual factor effects from control; and
- a covariance model of the 2 X 2 X 2 factorial subjects (i.e. all non-control subjects).

Single factor covariance models are addressed in Section 5.3.2 and the no control covariance model in Section 5.3.3.

### 5.3.1 Alternative Performance Measures

Two other measures of performance were gathered in the experiment — the total number of correct points (CORRECT) and the total number of points answered incorrectly (WRONG).

Both the CORRECT and WRONG models were analyzed using the procedure outlined in section 5.2.1.

<u>5.3.1.1 Correct Points Model</u> — The CORRECT model proved unremarkable. Significant terms are Class Score, Complexity, CD interaction, FD interaction. Subjects 511, 602 and 611 are also eliminated from this model. The CORRECT model provided lower values for $R^2$ and higher standard errors for the regression model and coefficients.

<u>5.3.1.2 Wrong Points Model</u> — The WRONG model contains only complexity (C), a third order interaction term (XFD) and the complexity - duration interaction (CD) term at significant levels. Explanatory power is relatively low ($R^2$=0.4). The CD term suggests that the simple task has a higher error rate under long duration interruptions while the complex task has a higher error rate under control conditions.

The XFD term suggests that, overall, error rates are not directly correlated to prior performance. The XFD term bears a negative sign. The interpretation suggests that better students have lower error rates under control and under the Hi Frequency Hi Duration treatments. Poorer students had higher error rates under the same treatments.

### 5.3.2 Single Treatment Covariance Models

Because the experiment used control groups, it is not a true factorial design. Neither frequency nor duration can exist at a control level while the other factor has a value greater than zero. As a result, the linear model used does not allow for complete differentiation of individual study variables. To be strictly correct, primary values for each variable should be compared to control conditions to completely distinguish the effects of manipulations. In other words, MODELI does not completely answer the question of each individual treatment score (Hi Frequency, Lo Frequency, Hi Duration, Lo Duration) being significantly different from control scores.

Four models were prepared to examine the individual effect of frequency and duration study variables. The Full data set was separated into Simple and Complex data sets based on the type of task performed. The Simple subset consists of subjects in treatments 0,2,4,6 and 8. The Complex subset includes those in treatments 1,3,5,7 and 9. Each of the two reduced data sets is then coded as a covariance model for Frequency effects and a covariance model for Duration effects. [Neter, Wasserman and Kuntner, 1985, 853-879]. ( General coding schemes for the four covariance models are shown in Figure 5.9 .) These models are designed to determine if individual study variables are different from control groups.

Models are first run with all interactions. If no terms appear significant, the model is checked for influential observations. After removing influential observations, terms are rechecked for significance. After employing a stepwise elimination procedure, the final model contains only those coefficients whose T values are significant at the 95% level. Neter, Wasserman and Kutner [1985] suggest checking for significance of factor effects (in covariance models) by using of a Bonferroni confidence interval. If the limits do not include zero (a conservative criterion), the term is considered significant. In this study, 95% Bonferroni confidence interval is reported for appropriate models.

For simple data neither of the Frequency terms is significantly different from control. Both Duration terms have T values significant at the 95% level after eliminating observations 404 and 611. However, only Hi Duration has a confidence limit that does not include zero. Class Score remains in all models.

Complex data models have less explanatory power than simple models. Lo Duration is significantly different from control. Lo Frequency is a significant term with all data present. After eliminating subjects 511 and 709, the Lo Frequency confidence limit does not include zero. The point change (from control conditions) for significant terms are summarized in Table 5.5.

Figure 5.9: Single Treatment Covariance Model

$$Y_i = \mu + \tau X_i + \alpha H_i + \beta L_i + \alpha\tau X_i(H_i) + \beta\tau X_i(L_i) + \epsilon$$

$X_i \equiv$ Score on Midterms - Mean Score on Midterms

$H_i \equiv$ High Treatment $\{ H_i = -1$ if control
$= 1$ if Hi(6)
$= 0$ otherwise $\}$

$L_i \equiv$ Low Treatment $\{ L_i = -1$ if control
$= 1$ if Lo(2)
$= 0$ otherwise $\}$

| Model | Data Subset | $H_i$ | $L_i$ |
|---|---|---|---|
| VSF | Simple Subjects Only | FHi | FLo |
| VSD | Simple Subjects Only | DHi | DLo |
| VCF | Complex Subjects Only | FHi | FLo |
| VCD | Complex Subjects Only | DHi | DLo |

Table 5.5: Single Treatment Covariance — Significant Terms

| MODEL | $R^2$ | REMOVED SUBJECTS | AVG SCORE | TERM | SCORE CHANGE |
|---|---|---|---|---|---|
| VSD | 0.47 | 404,611 | 133 | DHi | +37 |
| VSD | 0.47 | 404,611 | 133 | DLo* | -26 |
| VCF | 0.27 | 511,709 | 44 | FLo | -37 |
| VCD | 0.24 | 511,709 | 44 | DLo | -28 |
| VCD | 0.24 | 511,709 | 44 | DHi* | +36 |

* 95% Bonferroni confidence limit includes zero

### 5.3.3 Covariance Model Without Control Groups (COVNC)

The Single Factor Covariance Models do not allow complete differentiation of individual treatment cells from each other because Frequency-Duration interaction terms are not present. The Full model assumes "Lo" levels represent a mean response level for Frequency and Duration variables. To allow for complete differentiation of all treatments, a covariance model was prepared for the 2 X 2 X 2 factorial treatment of non-control subjects. (See Figure 5.10 for this model.)

The initial model was run with all subjects (201-912) using the procedure outlined in Section 5.2.1 and examined for influential observations. Subjects 511, 602 and 611 were deleted.

No interaction terms involving Class Score ($X_i$) are are significant in the revised model (a necessary condition to validate assumptions of the covariance model). The regression results for the COVNC model are depicted in Table 5.6. The Bonferroni 95% confidence interval for the CF term includes zero.

The COVNC model bears a strong similarity to MODELI. Complexity-Duration, Complexity-Frequency, Complexity and Class Score terms appear in both models. In the COVNC model, Duration replaces the FD term from MODELI. (Contrasts will be further explored in Section 5.4.)

**Figure 5.10: Covariance Model - No Control Subjects**

$$Y_i = \mu + \tau X_i + \alpha F_i + \beta D_i + \delta C_i + \alpha\delta F_i(C_i) + \beta\delta D_i(C_i)$$
$$\alpha\beta F_i(D_i) + \alpha\tau X_i(F_i) + \beta\tau X_i(D_i) + \delta\tau X_i(C_i) +$$
$$\alpha\beta\delta F_i(D_i)C_i + \beta\delta\tau X_i(D_i)C_i + \alpha\delta\tau X_i(F_i)C_i +$$
$$\tau\alpha\beta\delta F_i(D_i)C_i(X_i) + \epsilon$$

Where:

$X_i \equiv$ Score on Midterms - Mean Score on Midterms

$F_i \equiv$ Frequency $\{ F_i = -1$ if Lo(2); $= 1$ if Hi(6) $\}$

$D_i \equiv$ Duration $\{ D_i = -1$ if Lo(2); $= 1$ if Hi(6) $\}$

$C_i \equiv$ Complexity $\{ C_i = -1$ if Simple; $= 1$ if Complex $\}$

**Table 5.6: Significant Terms for COVNC Regression Model**

$R^2 = 0.694$ Std Error = 39.38 F = 40.3 w/ 5,89 dF

| Term | Value | Term | Value |
|---|---|---|---|
| Constant | 84.81 | Duration | 8.41 |
| Complexity | -48.07 | Complex-Freq* | 7.06 |
| Class Score | 0.632 | Complex-Dur | -8.85 |

* 95% Bonferroni confidence limit includes zero

## 5.4 INTERPRETATION OF RESULTS

### 5.4.1 Adjusted TOTSCORE Model

Using TOTSCORE Model I as a basis, the treatment means have been adjusted for differences in Class Score and eliminated subjects (see Table 5.7). The adjusted treatment means are used to construct the interaction diagrams in Figures 5.11, 5.12 and 5.13.

Table 5.7: Adjusted TOTSCORE Mean by Treatment

| # | TREATMENT | CLASS SCORE | AVERAGE TOTSCORE | ADJUSTED TOTSCORE |
|---|---|---|---|---|
| 0 | SControl | 20.58 | 149.33 | 135.46 |
| 1 | CControl | -5.30 | 63.50 | 67.33 |
| 2 | SFLoDLo | -13.17 | 110.83 | 120.04 |
| 3 | CFLoDLo | 1.75 | 37.33 | 36.34 |
| 4 | SFLoDHi | 7.31 | 159.69 | 154.90 |
| 5 | CFLoDHi | 5.45 | 19.64 | 16.11 |
| 6 | SFHiDLo | 10.43 | 134.00 | 127.07 |
| 7 | CFHiDLo | -7.77 | 33.69 | 39.21 |
| 8 | SFHiDHi | -1.70 | 144.33 | 145.70 |
| 9 | CFHiDHi | -13.20 | 47.00 | 56.23 |

### 5.4.2 Discussion Of Interaction Terms

The Duration-Complexity diagrams (Figure 5.11) show a consistent relationship between long and short interruptions

— and explain the presence of a positive D term in the COVNC model (Section 5.3.3).[2] However, long interruptions show improved scores over short interruptions. The control line crosses the long interruption line — indicating that long interruptions improve performance on simple tasks while degrading performance on complex tasks. This crossed line also helps in understanding the absence of Duration from MODELI.

The Frequency-Complexity diagrams (Figure 5.12) demonstrates why the CF term is present in Model I. Again the "inversion" effect is evident under long interruption conditions. On either of the diagrams, it can be seen that Frequency treatment means are quite close — so no distinctions are evident for high versus low Frequency. Thus no significant distinctions appear in the COVNC model. However, in both Duration treatments, the slopes of Frequency are steeper than control. These consistent divergent slopes, coupled with the inversion under long interruption, are enough to bring the CF term to marginal significance in Model I.

Interpretation of the Frequency-Duration interaction diagram (Figure 5.13) is some what more complicated.

---

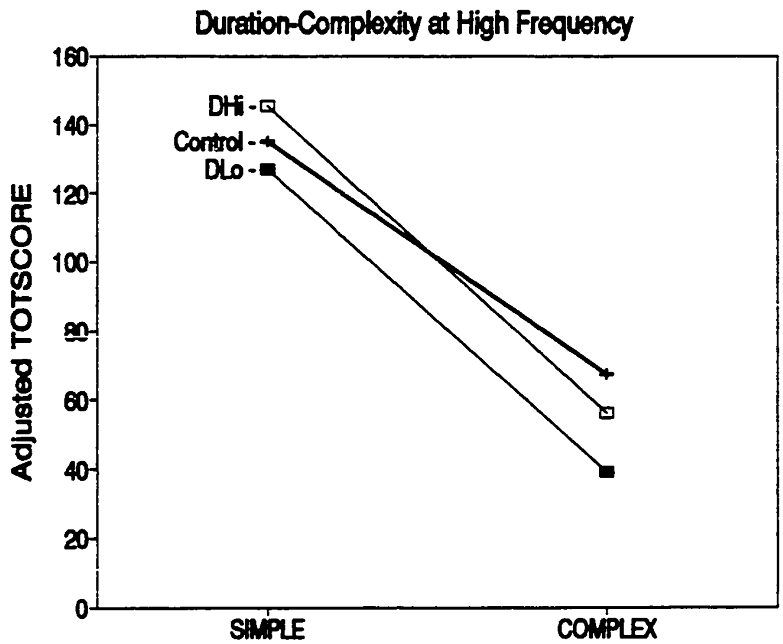
2. Parallel lines indicate a uniform effect for a study variable while divergent, or crossed, lines indicate interaction of terms.

Combined low values of Frequency and Duration result in lower scores for both task types. Essentially, the FD term differentiates Lo Frequency-Lo Duration treatments from remaining treatments. The absence of an FD term from the COVNC model reflects the tight treatment means and crossed slopes that remain when control conditions are ignored.

Since the short interruption condition is consistently below control conditions for complex tasks, it appears significant in single factor covariance models (Table 5.5). Long interruption conditions are consistently above control for simple tasks and also show up in the single factor models. Low Frequency conditions are consistently lower than control conditions for complex tasks as well.

In summary, Complex task performance is adversely affected by both Frequency and Duration conditions. Simple tasks appear to benefit from long interruptions. The consistent superior performance of subjects experiencing long interruptions (over those experiencing short interruptions) is an anomaly. Intuitively, subjects ignoring the interruptions might explain these results. However, the Extra Minutes covariate is not significant in any model. Nor does there seem to be any visible correlation between the total time and scoring. (Further discussion of results appear in Chapter 7.)

Figure 5.11 Duration-Complexity Interactions



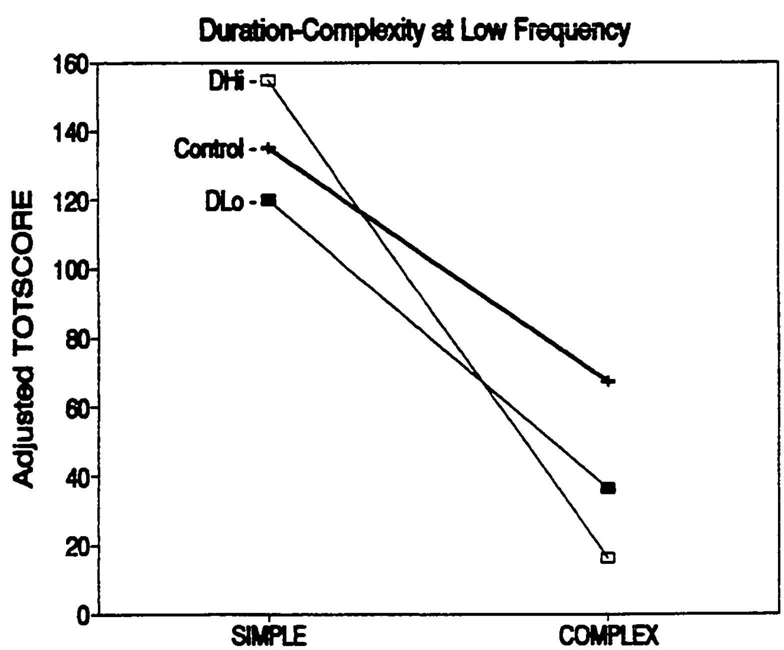



Figure 5.12 Frequency-Complexity Interactions

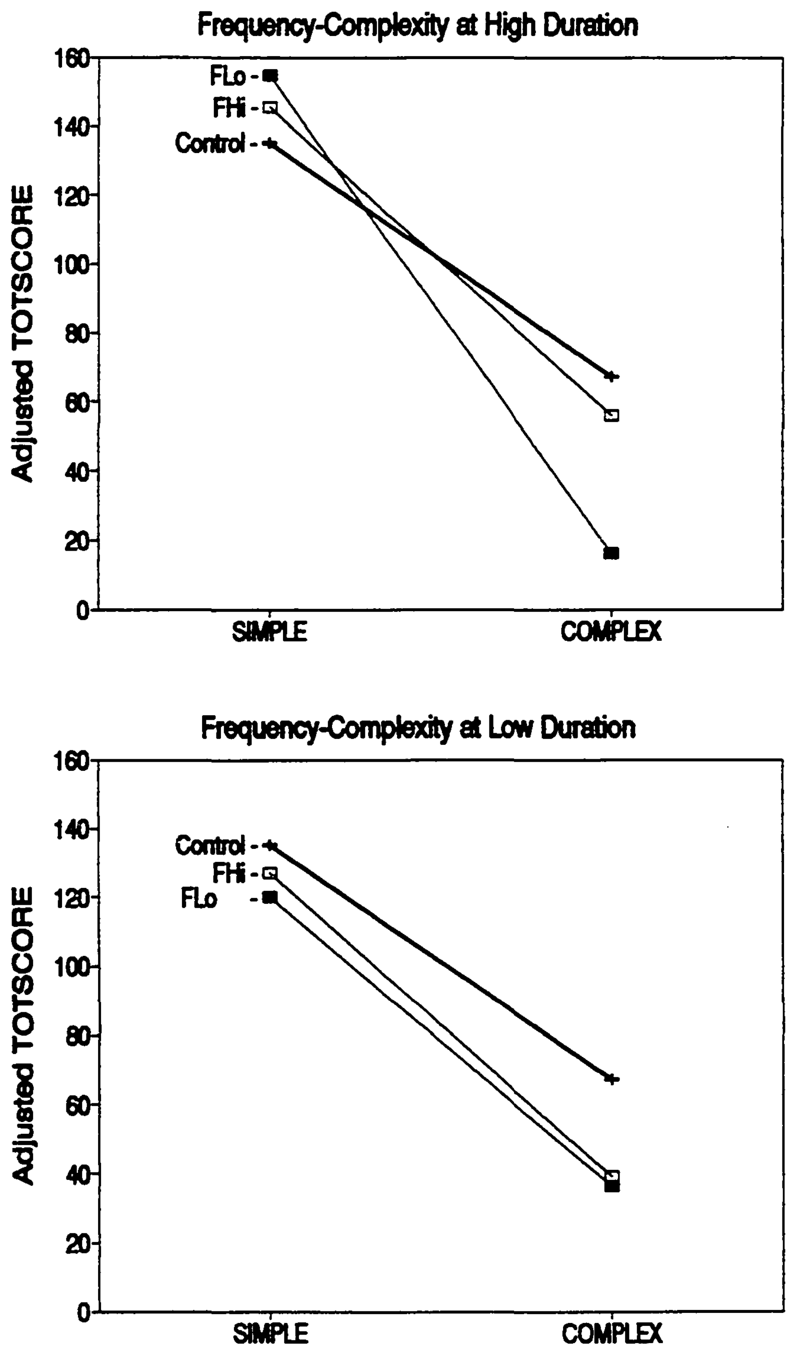
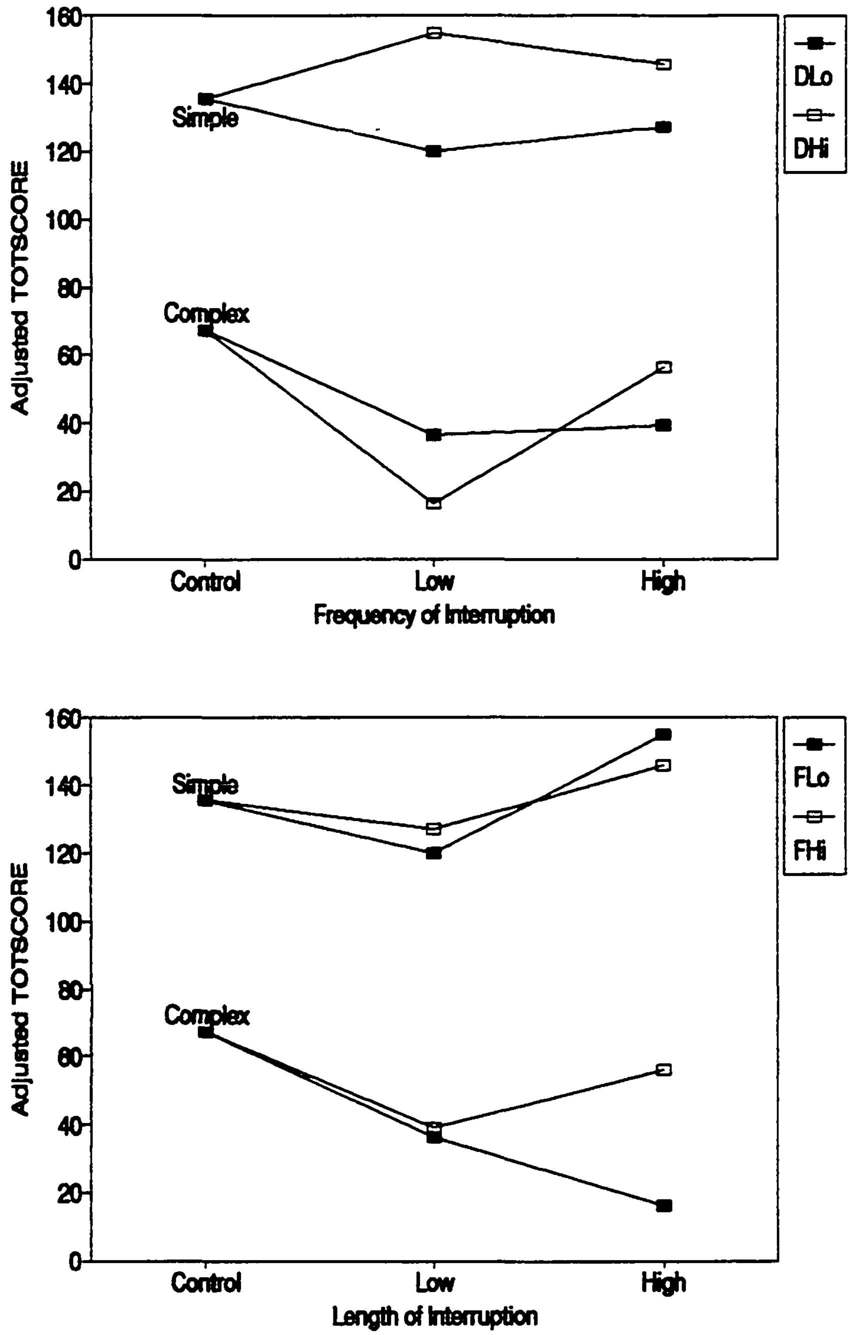


Figure 5.13 Frequency-Duration Interactions



## 5.5 TESTS OF RESEARCH HYPOTHESES

Because of the unexpected differences in results, formal tests of hypotheses are separately considered for simple and complex tasks where appropriate. Results are considered significant when the probability of a Type 1 error is held at 5% (i.e. $p > 0.95$).

These tests determine whether selected regression coefficients are different from zero. ( The general form of this test is described in Neter, Wasserman and Kuntner [1985] 290-293.) For purposes of subsequent analyses, tests used here use two forms of reduced models:

*For full data sets* : $Y_i = \beta_0 + \beta_1 X_i + \beta_2 C_i + \epsilon$

*For partial data sets*[3] : $Y_i = \beta_0 + \beta_1 X_i + \epsilon$

Formally stated the hypothesis test is of the form:

*Ho*: $\beta_3 = \beta_4 = \ldots = \beta_{p-1} = 0$

*Ha*: Not all $\beta$'s in Ho equal zero.

The $\beta$ terms refer to Frequency and Duration manipulations. Terminology used here will show $F^*$ as the actual value of the F Test with [$F_{.95}$, dfnumerator, dfdenominator] denoting the critical test value.

---

3. Partial data sets consist of simple task observations only or of complex task observations only.

### 5.5.1 Presence of Deleterious Effects

To determine if deleterious effects are present, formal tests of reduced regression models are used. Recall the statement of hypothesis 1:

*H1: Interruptions will result in decreased performance levels for tasks performed by knowledge workers.*

Tests of the general hypothesis for regression coefficients are summarized for the full data in models I and III. Each of the single factor covariance models described in section 5.3.2 are also depicted.

| MODEL | F Critical | F* | Outcome |
|---|---|---|---|
| MODELIII | 3.08 w/2 & 117df | 3.99 | Accept Ho |
| MODELI | 2.71 w/3 & 113df | 6.48 | Accept Ho |
| Simple Freq | 3.15 w/2 & 53df | 0.84 | Reject Ho |
| Simple Dur | 4.02 w/1 & 56df | 8.46 | Accept Ho |
| Complex Freq | 4.02 w/1 & 56df | 8.18 | Accept Ho |
| Complex Dur | 4.02 w/1 & 56df | 5.05 | Accept Ho |

Overall, deleterious effects are present — but only for complex tasks. Despite the significant effects in the simple-duration covariance model, the effects appear to improve performance. Therefore, we can conditionally accept *H1* for the complex task, and reject *H1* for simple tasks.

The conditional acceptance of H1 leads to accepting H2 as well:

*H2: Reductions in performance due to interruptions will be larger for difficult tasks than for simple tasks.*

Since the predicted deleterious effects are only evident in complex tasks, *H2* can be accepted as stated.

### 5.5.2 Effects of Manipulations

Frequency manipulations test for the presence of the "change gears" mechanism. *H3* suggests that performance will be worse under high Frequency conditions:

*H3: Increased frequency of interruption will result in decreased performance levels.*

For MODELI, the CF term is significant ( $[F_{.95,1,114df}]$ = 3.93, $F^*$ = 4.014). However, the net effect is performance deterioration only under Low Frequency conditions. The marginal effects of Frequency are confirmed by single factor covariance models and the absence of significance in the COVNC model. The instability of Frequency is confirmed by MODELII where the CF term failed to be significant at even a 90% level. As a result, unconditionally reject *H3*.

Duration effects are assuredly the stronger of the two manipulations. The theoretical "decay" mechanism suggests that deleterious effects will be increased by longer interruptions leading to the fourth hypothesis:

> *H4: Longer interruptions will decrease performance more than short interruptions.*

As with Frequency, Lo levels produced more dramatic results than high levels. However, Duration manipulations are much stronger and more consistent at Lo levels. The lack of significance for long interruptions in any of the single factor covariance models, or in the COVNC model, leads to rejection of *H4*.

In Hypothesis 5, an interaction effect is predicted for Frequency and Duration:

> *H5: Simultaneously increasing both the length and the frequency of interruption will result in lower performance levels than equivalent increases for either factor alone.*

The FD interaction term is present and is significant in all full models. For models I,II and III [ $F_{.95} \approx 3.94$ ]. $F^*$ values are (respectively) 9.05, 8.39, 4.12. However, the interaction term has the "wrong" sign. The net effect in full data models is to raise the predicted performance level of Hi Frequency-Hi Duration treatments. The FD term is not significant in the COVNC model — suggesting no true interaction. As a consequence, reject *H5*.

The final hypothesis is proposed to identify dominance of Frequency over Duration under equivalent total interruption time:

> *H6: If total interruption time is equivalent, many short interruptions will result in lower performance scores than a few long interruptions.*

Based on the discussions above, Duration appears to be the only manipulation resulting in deleterious effects so Frequency does not dominate Duration. Reject *H6*.

# 6. ANALYSIS OF QUESTIONNAIRE

The post experiment questionnaire was prepared and administered to:

- obtain a check on experiment manipulations;
- gather subject preferences and coping strategies; and
- to check the subject's perception of performance against measured performance.

Overall, reaction to the experiment was positive. Of the 122 subjects, 93 took the time to write comments. Of those writing comments, about 70 said they enjoyed the experiment, thought it was worthwhile, or expressed comment that suggested a positive reaction to the experience.

While individual questions will be addressed below, Table 6.1 summarizes the mean and standard deviation of responses for simple and complex task subjects. Results of questionnaires are presented in Appendix B.4. Histograms of responses by task to questions begin on page 116. A frequency distribution of responses by task is presented in Table 6.2, page 120. Frequency response distributions for individual questions (by treatment) begin on page 121.

Table 6.1: Mean Response to Questions by Task

| QUES | BOTH TASKS Mean | SdDev | SIMPLE TASK Mean | SdDev | COMPLEX TASK Mean | SdDev |
|---|---|---|---|---|---|---|
| 2 | 3.492 | 1.018 | 3.885 | 0.889 | 3.098 | 0.987 |
| 3 | 3.131 | 1.130 | 3.279 | 1.175 | 2.984 | 1.063 |
| 7 | 3.427 | 0.851 | 3.479 | 0.790 | 3.375 | 0.904 |
| 8 | 3.292 | 0.853 | 3.458 | 0.865 | 3.125 | 0.807 |
| 9 | 4.271 | 1.661 | 4.313 | 1.596 | 4.229 | 1.723 |
| 10 | 2.852 | 1.548 | 2.958 | 1.607 | 2.746 | 1.479 |
| 11 | 2.115 | 1.361 | 2.146 | 1.354 | 2.083 | 1.367 |

## 6.1 DEMOGRAPHICS

All subjects responded to questions on page one of the questionnaire. Part I gathers information on age, sex, place of birth, and previous computer use. Covariates for sex and previous computer use were prepared for the analysis in Chapter 5. Neither of these factors had a significant explanatory effect on performance.

To recap, 72 males and 50 females participated in the experiment. There were 94 subjects who had prior experience with a personal computer, 55 males and 39 females.

Average age of subjects is 23.25 years with a range of 19 to 47. An age covariate showed no significant relationship between age and task performance.

## 6.2 TASK AND MEDIUM APPRAISAL

To determine if subject answer patterns for questions 2 and 3 are correlated to performance, the responses are converted into "pseudo covariates"[4] by subtracting the subject response from 4. Individual response sets are then introduced into MODELI.

### 6.2.1 Relative Task Difficulty

Question 2 asks subjects to rate the difficulty of experiment questions against those they have experienced in course work. The inclusion of the Question 2 pseudo covariate (Q2) in MODELI, proved significant at the 99% level. On average, a one unit change in response corresponds to a 10 point change in TOTSCORE. When introduced into a model containing only Class Score, Q2 is significant at the same level as Class Score. Class Score - Q2 interactions are insignificant, indicating no correlation between prior

---

4. To be strictly correct, a covariate should be independent of any experiment manipulation. Since the questionnaire responses can be reasonably expected to be influenced by treatments, this is not a true covariate.

class performance and perception of relative task difficulty.

This result confirms that the subject's post-experiment perception of relative task difficulty is highly correlated to his or her performance on the task. While this statement seems unremarkable, it does indicate that subject response to the question is honest.

A simple regression with complexity as the independent variable and Q2 as dependent variable, confirms that subject's perceived the complex task as more difficult than the simple task. A frequency distribution for question 2 is presented in Table 6.2 (p 120) and depicted graphically in Figure 6.1 (p 116).

### 6.2.2 Computer Medium

Question 3 attempts to separate the effects of the presentation medium from the task. Subjects are asked to contrast their performance on the experiment with an equivalent task on paper. A pseudo covariate (Q3) for this response is also introduced into the MODELI regression model.

Q3 is significant at the 99% level. Interaction with class score is not significant. There is no significant correlation between complexity and Q3. The Q3 coefficient is about 9 points. Subjects who answered "3" scored about 9 points lower than those who answered "4". It appears that subjects who scored lower tend to perceive the computer format as a larger negative factor in their performance.

There is a correlation between Q2 and Q3 (covariance = 0.3), suggesting that some subjects who found the experiment questions to be more difficult also found the computer to be an impediment to performance. The absence of a significant relationship between complexity and Q3 tends support the independence of the medium and the task. (A frequency distribution for question 2 is presented in Table 6.2 and depicted graphically in Figure 6.1.)

## 6.3 INTERRUPTION MANIPULATIONS

To assess the impact of questions 7,8,9 and 10, pseudo covariates were prepared. None of these pseudo covariates proved to be significant when introduced into the COVNC model. Interactions of pseudo covariates with Class Score and Complexity are also insignificant in the COVNC model. Thus subject reactions to these questions would seem to have no correlation to actual performance. Remaining discussions in this section focus on insights provided by subject responses and comments.

### 6.3.1 Perceived Effects of Interruptions on Performance

Almost half the subjects indicated that interruptions had no effect on their performance. (Table 6.3 and Figure 6.2) The overall results described in Chapter 5 somewhat confirm this conclusion.

In order to check for any pattern in responses, a regression model was prepared. COVNC coding for Complexity, Frequency, Duration and first order interaction terms is used for the independent variables. The actual response to question 7 is used as the dependent variable. The resulting regression is marginally significant. Frequency remained as the only significant variable. High frequency levels are perceived to have a more detrimental effect on performance than other measures. There is no significant correlation between response and complexity.

The strong showing for the neutral response to this question may suggest that the interruptions used were not disruptive enough to have an effect. The correlation of frequency manipulations, although slight, tends to support the perception by subjects that Frequency dominates Duration. This perception is the inverse of actual performance where Duration dominates Frequency.

### 6.3.2 Effects on Concentration

Question 8 attempts to confirm a disruptive effect of the interruptions – the "shift in cognitive focus". Subjects confirmed the shift, 57 of 96 indicated that the interruptions hurt concentration, 4 felt the interruptions helped, with the remainder responding "no effect". (Table 6.4 and Figure 6.3.)

A regression model with the COVNC independent variables and responses to question 8 as the dependent variable proved marginally significant. Both Complexity and Frequency manipulation terms remained marginally significant suggesting that both factors amplified the negative effects of the interruptions on concentration.

### 6.3.3 Reaction To Interruptions

As explained in Chapter 5, the results of the experiment are significant, but deviate radically from those predicted in Chapter 3. No question provides more insight into possible explanations for the experiment outcome than does question 9.

The mean response here is very close to neutral but has an extremely wide range of responses – 29 subjects were pleased, 42 were displeased and the remaining 23 were indifferent. (See Table 6.5 and Figure 6.3.) No terms are significant in the COVNC model with question 9 response as dependent variable. The comment page of the questionnaire provided some additional insight.

After reading the comment page, it seems clear that reaction to the interruptions was quite strong – ranging from profane negative condemnation to sheer delight. Twenty subjects commented on the positive nature of the interruptions using terms like "relief", "fun", "eased the pressure" and "broke the tension" to describe their reactions. Some of the more insightful comments:

> "[Interruptions] made it easier for me by providing a break from what I was doing, time to clear my head [and by] forcing me to summarize what I was doing."– Subject 807

> "The interruptions were more interesting than the exam questions. I found myself hoping for another interruption... [they] gave my mind a short vacation from the stress of figuring out the questions." – Subject 502

> "...if you happen to be deadlocked on a problem, the trivia interruptions allow you clear your mind – even though, emotionally, you might be put off." – Subject 304

Other subject comments focused on the 15 second mandatory time for each trivia question. several found this period to be the most frustrating part of the experiment. Three subjects implied that they quickly answered the questions and used the remaining time to complete calculations. Several suggested that they "got used to it" or "figured out how to handle [the interruptions.]" These comments suggest some possible explanations for the results described in Chapter 5.

Recall that simple tasks had significantly better performance for Hi Duration conditions than for control. For complex tasks, both Lo and Hi Duration resulted in a significant reduction in performance relative to control groups. Discounting control subjects for the moment, in all four treatment permutations, average scores for Hi Duration subjects are higher than the average for Lo Duration subjects. While this difference is statistically significant only for simple tasks, it suggests the possibility of alternative explanations for results.

Subjects with long duration interruptions may have developed an adaptive strategy that used the interruption time for the primary task. Several subjects suggested that they took a quick guess and continued calculations on the sample exam. Others spoke of the need to write down intermediate results to minimize the impact of interruptions.

Another possible explanation might be a "mental vacation" effect. Perhaps those subjects with short interruptions didn't get a long enough "vacation" to gain any benefits. Weick [1985] suggests periodic, self induced interruptions will enhance performance and reduce fatigue in computer based tasks. (Discussions in Chapter 7 focus on alternative theories of explanation.)

### 6.3.4 Interruptions with Equivalent Task Time

Question 10 asks subjects if they believe the experiment conditions would affect their performance. The actual time manipulation is proposed as a hypothetical case — i.e. if you had equivalent performance time, would the interruptions affect your score.

The reaction is surprisingly strong. The modal response for both task types is "2" (Table 6.6 and Figure 6.2). Mean response for this question is the most extreme (2.85) of questions 2 through 10 (Table 6.1). Obviously, subjects feel that deleterious effects are present and have a strong effect. The inclusion of the Q10 pseudo covariate is not significant in the COVNC model. Nor is the response to this question significantly correlated to any of the treatment manipulations.

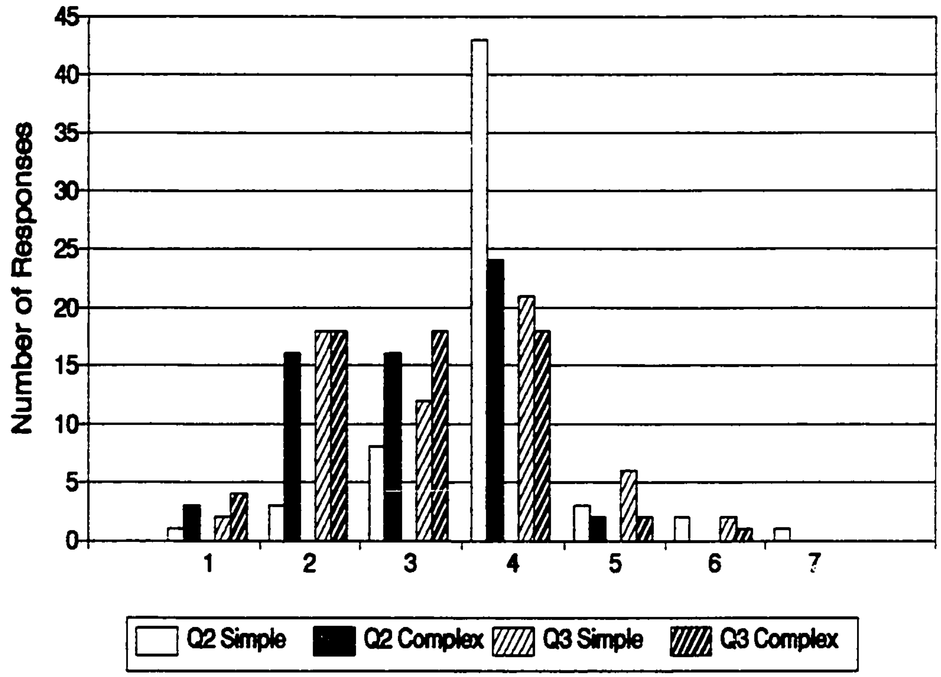
Contrast the responses to questions 9 and 10. For question 10, only 18 subjects said they would score higher under the interrupted conditions. Yet 29 subjects had a positive reaction to the interruptions in question 9. Despite reporting benefits from the actual experiment manipulations, subjects believe strongly in deleterious effects — perhaps as a consequence of cultural conditioning.

### 6.3.5 Preferred Interruption Scenario

Question 11 is included to ascertain strategies of the subjects for time management. Though not specifically stated, the response to question 11 implies subject preference for "grouping" interruptions. Almost half the subjects (46 of 96) responded that 1 long interruption would be best. (Table 6.2 and Figure 6.4 ). Only 15 subjects expressed a preference for four or more interruption events.

Subjects believe that grouping interruptions into one or two focused events would minimize the deleterious effects. This perception is partially supported by the results of the experiment. Long Duration interruptions did not degrade performance. But increased Frequency did not degrade performance. Because of the phrasing of the question, it is impossible to determine if subjects have a "high frequency" aversion or "long duration" preference.

Figure 6.1
Responses to Questions 2 and 3 by Task

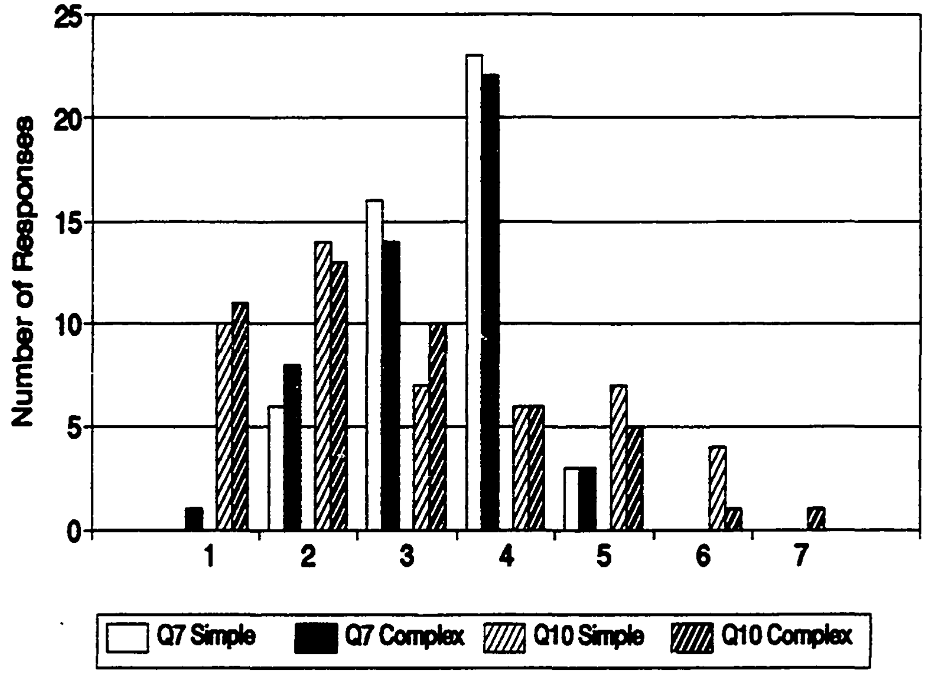



Q2 - Contrast difficulty of questions on practice exam with class exam
1 = Experiment Much Harder 7 = Class Much Harder

Q3 - Score on the same exam if taken on paper?
1 = Paper Score Much Higher 7 = Computer Score Much Higher

Figure 6.2

Responses To Questions 7 and 10 by Task

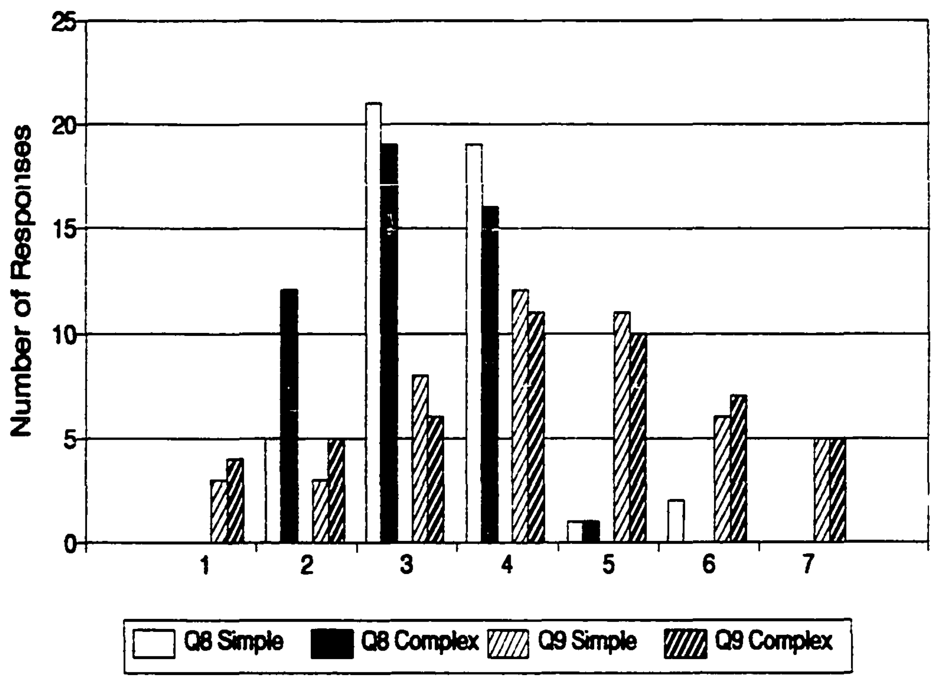



Q7 - How would you score without interruptions?
1 = Much Higher 7 = Much Lower

Q10 - 120 min w/20 min interruptions vs 100 min w/no interruption
1 = Uninterrupted Much Higher 7 = Interrupted Much Higher

Figure 6.3

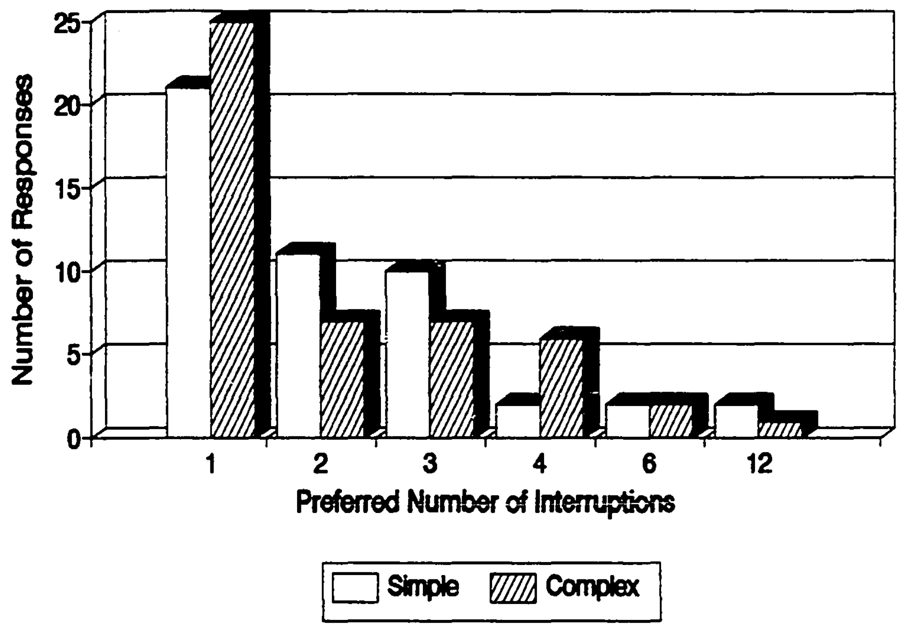
Responses to Questions 8 and 9 by Task




Q8 - How did interruptions affect your concentration?
1 = Really Hurt 7 = Really Hurt

Q9 - What was your reaction when an interruption occured?
1 = Really Pleased 7 = Really Displeased

Figure 6.4
Responses to Question 11 by Task




Q11 - If you had to answer 12 trivial questions during the course of the practice exam, how many times would you like to be interrupted?

Table 6.2: Frequency Distribution of Question Responses

| QUES | TASK | NUMBER OF RESPONSES 1 | 2 | 3 | 4 | 5 | 6 | 7 | TOT |
|---|---|---|---|---|---|---|---|---|---|
| 2 | SIMPLE | 1 | 3 | 8 | 43 | 3 | 2 | 1 | 61 |
| 2 | COMPLEX | 3 | 16 | 16 | 24 | 2 | 0 | 0 | 61 |
| 2 | TOTAL | 4 | 19 | 24 | 67 | 5 | 2 | 1 | 122 |
| 3 | SIMPLE | 2 | 18 | 12 | 21 | 6 | 2 | 0 | 61 |
| 3 | COMPLEX | 4 | 18 | 18 | 18 | 2 | 1 | 0 | 61 |
| 3 | TOTAL | 6 | 36 | 30 | 39 | 8 | 3 | 0 | 122 |
| 7 | SIMPLE | 0 | 6 | 16 | 23 | 3 | 0 | 0 | 48 |
| 7 | COMPLEX | 1 | 8 | 14 | 22 | 3 | 0 | 0 | 48 |
| 7 | TOTAL | 1 | 14 | 30 | 55 | 6 | 0 | 0 | 96 |
| 8 | SIMPLE | 0 | 5 | 21 | 19 | 1 | 2 | 0 | 48 |
| 8 | COMPLEX | 0 | 12 | 19 | 21 | 1 | 0 | 0 | 48 |
| 8 | TOTAL | 0 | 17 | 40 | 40 | 2 | 2 | 0 | 96 |
| 9 | SIMPLE | 3 | 3 | 8 | 12 | 10 | 5 | 5 | 48 |
| 9 | COMPLEX | 4 | 5 | 6 | 11 | 10 | 7 | 5 | 48 |
| 9 | TOTAL | 7 | 8 | 14 | 23 | 20 | 12 | 10 | 96 |
| 10 | SIMPLE | 10 | 14 | 7 | 6 | 7 | 4 | 0 | 48 |
| 10 | COMPLEX | 11 | 13 | 10 | 6 | 5 | 1 | 1 | 47 |
| 10 | TOTAL | 21 | 27 | 17 | 12 | 12 | 5 | 1 | 95 |
| 11 | SIMPLE | 21 | 11 | 10 | 2 | 2 | 2 | - | 48 |
| 11 | COMPLEX | 25 | 7 | 7 | 6 | 2 | 1 | - | 48 |
| 11 | TOTAL | 46 | 18 | 17 | 8 | 4 | 3 | - | 96 |

Table 6.3: Distribution by Treatment - Question 7

How do you think you would have scored on the practice exam if you had not been interrupted?

| Much Higher | | About the Same | | | Much Lower | |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| TREAT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | NUM |
|---|---|---|---|---|---|---|---|---|
| T2 | 0 | 2 | 2 | 7 | 1 | 0 | 0 | 12 |
| T4 | 0 | 1 | 3 | 7 | 1 | 0 | 0 | 12 |
| T6 | 0 | 1 | 7 | 4 | 0 | 0 | 0 | 12 |
| T8 | 0 | 2 | 4 | 5 | 1 | 0 | 0 | 12 |
| SIMP | 0 | 6 | 16 | 23 | 3 | 0 | 0 | 48 |
| T3 | 0 | 0 | 5 | 7 | 0 | 0 | 0 | 12 |
| T5 | 0 | 1 | 2 | 8 | 1 | 0 | 0 | 12 |
| T7 | 0 | 5 | 4 | 4 | 0 | 0 | 0 | 13 |
| T9 | 1 | 2 | 3 | 3 | 2 | 0 | 0 | 11 |
| COMP | 1 | 8 | 14 | 22 | 3 | 0 | 0 | 48 |

Table 6.4: Distribution by Treatment - Question 8

How did the interruptions affect your ability to concentrate on the practice exam?

| Really hurt | | Didn't Affect Me | | | Really Helped | |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| TREAT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | NUM |
|---|---|---|---|---|---|---|---|---|
| T2 | 0 | 2 | 4 | 6 | 0 | 0 | 0 | 12 |
| T4 | 0 | 1 | 4 | 7 | 0 | 0 | 0 | 12 |
| T6 | 0 | 1 | 8 | 2 | 1 | 0 | 0 | 12 |
| T8 | 0 | 1 | 5 | 4 | 0 | 2 | 0 | 12 |
| SIMP | 0 | 5 | 21 | 19 | 1 | 2 | 0 | 48 |
| T3 | 0 | 1 | 6 | 5 | 0 | 0 | 0 | 12 |
| T5 | 0 | 1 | 4 | 6 | 1 | 0 | 0 | 12 |
| T7 | 0 | 5 | 6 | 2 | 0 | 0 | 0 | 13 |
| T9 | 0 | 5 | 3 | 3 | 0 | 0 | 0 | 11 |
| COMP | 0 | 12 | 19 | 21 | 1 | 0 | 0 | 48 |

Table 6.5: Distribution by Treatment - Question 9

When an interruption arrived, what was your initial reaction?

| Really Pleased 1 | 2 | 3 | Indifferent 4 | 5 | 6 | Really Displeased 7 |
|---|---|---|---|---|---|---|

| TREAT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | NUM |
|---|---|---|---|---|---|---|---|---|
| T2 | 0 | 1 | 2 | 4 | 3 | 1 | 1 | 12 |
| T4 | 0 | 1 | 1 | 2 | 3 | 3 | 2 | 12 |
| T6 | 2 | 1 | 2 | 3 | 2 | 1 | 1 | 12 |
| T8 | 1 | 0 | 3 | 3 | 3 | 1 | 1 | 12 |
| SIMP | 3 | 3 | 8 | 12 | 10 | 5 | 5 | 48 |
| T3 | 0 | 1 | 1 | 5 | 4 | 1 | 0 | 12 |
| T5 | 2 | 1 | 3 | 1 | 2 | 1 | 2 | 12 |
| T7 | 0 | 1 | 2 | 2 | 4 | 4 | 0 | 13 |
| T9 | 2 | 2 | 0 | 3 | 0 | 1 | 3 | 11 |
| COMP | 4 | 5 | 6 | 11 | 10 | 7 | 5 | 48 |

**Table 6.6:** Distribution by Treatment - Question 10

Suppose that you had to take the practice exam under these two conditions: A.) 100 minutes of uninterrupted time or B.) 120 minutes with 20 minutes of interruption. Which of your two scores, A or B, would be higher?

| A much higher | | about the same | | | B much higher | |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| TREAT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | NUM |
|---|---|---|---|---|---|---|---|---|
| T2 | 5 | 2 | 0 | 1 | 3 | 1 | 0 | 12 |
| T4 | 3 | 4 | 2 | 2 | 1 | 0 | 0 | 12 |
| T6 | 2 | 5 | 3 | 1 | 0 | 1 | 0 | 12 |
| T8 | 0 | 3 | 2 | 2 | 3 | 2 | 0 | 12 |
| SIMP | 10 | 14 | 7 | 6 | 7 | 4 | 0 | 48 |
| T3 | 5 | 3 | 2 | 1 | 1 | 0 | 0 | 12 |
| T5 | 1 | 3 | 5 | 2 | 1 | 0 | 0 | 12 |
| T7 | 3 | 6 | 1 | 1 | 2 | 0 | 0 | 13 |
| T9 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 10 |
| COMP | 11 | 13 | 10 | 6 | 5 | 1 | 1 | 47 |

# 7. DISCUSSION OF RESULTS

## 7.1 OVERVIEW

There are four significant results to be explored in this chapter:

1. Long Duration interruptions result in an 11% increase in task performance for Simple tasks.

2. Short Duration interruptions result in a 44% reduction in task performance for Complex tasks.

3. Task scores appear to improve as the length of the interruption increases.

4. There appear to be no consistent effects attributable to Frequency at the levels used in the study.

An alternative model is proposed to incorporate this "Length Improvement" phenomenon in Section 7.2. Four possible explanations for the improvement are presented in Section 7.4.

As described in Section 7.3, the improved performance levels for simple tasks under interrupted conditions is somewhat supported by Distraction Conflict theory. However, no literature surveyed directly supports the superior performance of Hi Duration subjects over Lo Duration subjects.

In Section 7.5, the experiment is critiqued and future experiments to further define the effects of these variables are proposed. Alternative study variables for future research are explored in Section 7.6.

## 7.2 REVISED INTERRUPTION MODEL

It seems clear that the effects of interruptions are more complicated than originally proposed. For Complex tasks, there is clearly a negative component — interruptions degrade performance. The consistent superior performance of Hi Duration subjects over Lo Duration subjects suggests that there is also a positive component attributable to the length of the interruption.

If the Negative component is the "Change Gear" effect. then Hi Frequency — Lo Duration treatments should have the lowest scores. However, there appears to be no significant differences between Hi and Lo Frequency suggesting that the Negative component is either invariant across all Frequencies (a constant) or offset by the Positive component.

Now consider the Positive component. If the positive component is a function of Duration, then Hi Duration - Hi

Frequency treatments should have the highest performance level. However, Frequency does not affect Hi Duration performance — suggesting that the Negative term may be may be offsetting any beneficial effects of the positive component.

These effects are further compounded by the introduction of task complexity. Recall the Frequency-Duration Interaction diagram (Figure 5.13, page 99). Task Complexity not only introduces a large fixed change in score (approximately 86 points) but may also provide a "rotation" of the interruption curves. Complex tasks appear to amplify the effects of the Negative component. The overall complexity level may also affect the size of the Positive term.

Both Duration and Frequency can be measured as pseudo continuous variables — task complexity cannot. This suggests that a mathematical relationship might be established for a given type of task. (The general form of this relationship is described in Figure 7.1.)

Figure 7.1: Revised Interruption Model

$$\text{Score} = Dd(L) - Df(F)$$

Where:

- Score = Score for a Quality Measured Model
- Dd = Positive component
- Df = Negative component
- L = Length of Interruption
- F = Frequency of Interruption

To provide some confirmation of this revised structure, MODELI data set was recoded as MODELC. In MODELC, both Frequency and Duration variables use the following coding scheme: 0 if control, 2 if Lo, 6 if Hi (all other terms and covariates are as depicted in Figure 5.3, page 79.) This coding scheme implies that Frequency and Duration are pseudo continuous variables. A stepwise elimination procedure was used to produce the final results shown in Table 7.1.

Note that Frequency remains in the model when entered as a pseudo continuous variable and that it has the predicted negative value. Duration remains in the form of interaction terms with Task Complexity and with Frequency. The presence (and correct sign) of the Frequency term is consistent with the Revised Interruption Model (RIM) — suggesting that a Frequency effect may be present, and that it may be independent of Task Complexity.

The results of MODELC do not confirm the validity of the RIM. Further research and revised experimental designs will be required to validate the RIM (see Section 7.5). However, MODELC indicates that the results of this study are not inconsistent with the RIM.

Table 7.1: Results of MODELC Regression

$R^2 = 0.70$ Standard Error = 38.24 $F^* = 43.47$ w/6,112

| Primary Coeff | Value | Interaction Coeff | Value |
|---|---|---|---|
| Constant | 90.56 | Freq*Duration | 1.03 |
| Complexity | -31.82 | Complex*Duration | -4.14 |
| Frequency | -5.07 | Complex*Class Score | -0.208 |
| Class Score | 0.693 | | |

## 7.3 TASK COMPLEXITY EFFECTS

### 7.3.1 Social Facilitation

The notion of Social Facilitation was solidified by Zajonc [1965]. Simply stated, Social Facilitation theory predicts that performance of simple, well learned tasks will be improved by the social distraction of others while performance of complex, counter instinctive tasks will suffer from these distractions. Martens and Landers [1972] found that competitive situations amplified the Social Facilitation effect. Groff, Baron and Moore [1983] found that mechanical distractions had the same effect as distractions provided by other people. McBain [1961] found that monotonous task performance was significantly improved by mechanical distraction. Thus, taken at face value, Social Facilitation predicts that interruptions should improve scores on simple tasks and decrease scores on complex tasks.

Results somewhat follow this prediction. However, the social facilitation theories would also predict that Hi Frequency-Hi Duration treatments should show the most pronounced effects. Baron suggests that the inverse should be true:

> "[D]istracting individuals during an important (well learned) task should produce a curvilinear pattern, i.e. facilitation of the task at low to moderate levels of distraction, with impairment occurring at sufficiently intense levels of distraction." [Baron, 1986, p26]

Thus social facilitation may partially explain the score improvements for those who were interrupted while performing simple tasks. However, the relatively better performance of Hi Duration subjects, and the apparent insensitivity of subjects to Frequency effects, contradicts predictions from this theory.

### 7.3.2 Task Granularity

There seems to be no universally accepted taxonomy for classifying complexity. Campbell [1988] presents a classification scheme, but with sixteen different categories of task complexity, it may be too cumbersome for routine management use. A simple, more tangible measure, with more universal applicability, would be desirable.

One distinction between the Simple and Complex versions of primary task is structure. The sequence of questions and the number of subdivisions in the Simple version is more structured than the Complex version. For future research, a measure of structure may provide an attractive alternative to complexity. "Task Granularity[5]" may be that alternative measure.

For purposes of discussion, Task Granularity can be defined as the number of "natural" subtasks (or break points) in a given primary task. Recall that the Simple task had 40 questions while the Complex task had only 15. If answering one question is considered a natural subtask, then the Simple version has higher granularity than the Complex version of the practice exam.

To understand how this may affect the interruption action, consider the following scenario in the context of the experiment. Subjects commented that interruptions had less of an effect when they had just finished answering a question. The answering of a question would be a natural "break point". At that moment, the worker is mentally prepared to "change gears", and, there is a lower investment

---

5. The notion of granularity has its origins in systems analysis and design. Granularity, in the systems analysis context, is a measure used to describe the system design in terms of program module size. Overly large program modules are often difficult to maintain and revise. Too many small modules creates excessive system overhead.

in mental concentration than at any other time in the process of answering a question. An interruption, occurring at that time, would likely reduce the size of the negative component.

In tasks with high granularity, the probability of this scenario occurring would be higher than for those with lower granularity. It could be argued that low granularity is simply one characteristic of complex tasks. Fewer natural break points implies longer spans of concentration and possibly more factors to consider. Thus, granularity may serve as a more tangible descriptor for tasks than a complexity classification scheme — an advantage for application of research findings to management situations.

If granularity proves to be effective in predicting deleterious effects, then Miller [1956] may provide some insight into the shape of the interruption-granularity relationship. Miller suggests that humans can effectively deal with five to nine alternatives at one time. Too many alternatives causes a "cognitive overload" [Kahneman, 1973]. Too few underutilizes the human potential. The introduction of an interruption into a a task with low granularity may overload the worker. In tasks with very high granularity (approaching boredom), the interruption may actually heighten worker interest [ McBain,1961; Woodhead,1965; and

Zuerecher,1965]. If granularity acts as measure of "current mental load", then introducing interruptions may:

- result in improved task performance levels under high granularity;
- generate no discernible change under moderate granularity; and
- degrade performance of low granularity tasks.

## 7.4 POSITIVE EFFECTS OF INCREASED DURATION

Since short Duration interruptions result in lower expected performance levels than long Duration interruptions, the "forgetting" theory does not seem to apply. From an operating manager's point of view, the reason for this "Length Improvement" phenomenon is less important than its predictability. However, four possible explanations may account for this effect:

- the interruptions in the experiment were not adequately disruptive;
- subjects developed an adaptive strategy that favored longer interruptions;
- there is an incubation effect; and / or;

- there is a stress relieving effect.

### 7.4.1 Faulty Interruption Mechanism

Recall that each trivia question was presented to the subject for 15 seconds. The subject's only available action is to answer (or not answer) the question. If the subject "knows" the answer, (or chooses to take a wild guess), the additional time could be used to work on calculations or some other aspect of the primary task. The experiment software can only verify that that a subject either answered a trivia question correctly, answered incorrectly, or did not answer. There is no way to verify how the time was used or if the answer was a wild guess.

The Extra Minutes covariate checked for a relationship between unanswered trivia questions and changes in score. This term is not significant in any of the models. However, it is conceivable that subjects were able to divert interruption time toward the primary task after answering the trivia questions.

Subjects were allowed to bring any reference materials they wished. Subjects could be working with a calculator, class notes, or other reference materials when an interruption occurred. Since the only cue was visual, it was possible to completely miss an interruption. Subsequent experiments should include either an audible alarm, or require subject interaction before proceeding.

### 7.4.2 Adaptive Strategy

In 1988, Payne, Bettman and Johnson (PB&J) directly examined the notion of adaptive strategy – extending a framework established by Reder [1987] and Huber [1980]. The theoretical framework for the PB&J study contends that changes in time pressure would result in different strategies for making decisions.

In the PB&J study, a computer simulation was used to identify dominant heuristic strategies for a suite of 40 risk/payoff decisions. Subjects were then exposed to the 40 decisions and evaluated by comparing subject performance to the best results for the heuristics predicted by the simulation.

The PB&J experiments confirmed that time pressure resulted in a specific hierarchy of responses:

> " People may first accelerate their processing and simply try to do the same things faster. If the time pressure is too great for acceleration to suffice, individuals may next engage in filtration, focusing on a subset of information. Finally, people may change strategy when time pressure becomes extreme." [Payne, Bettman and Johnson, 1988 p551]

While the tasks and manipulations of the PB&J study are quite different from those presented here, the positive effects of longer interruptions may be a result of an adaptive strategy. PB&J subjects revised their strategies as the experiment progressed. Short interruptions may not provide adequate time to gain benefits from coping strategies devised by the subjects of this experiment. Alternatively, the short interruptions may not provide adequate opportunity to devise or revise a strategy to deal with the interruptions.

The results of the PB&J study also suggests that effects of each interruption may not be constant over the course of the session. Subjects may revise strategies as the experiment progresses and as the end of the experiment approaches.

### 7.4.3 Incubation

Another possible explanation for the positive effects of Duration is an "incubation" effect for problem solving first proposed by Wallas [1926]. Wallas suggests four phases to problem solving:

1. Preparation – gathering of information and preliminary attempts at solution;
2. Incubation – putting the problem aside to work on other activities or sleep;
3. Illumination – the key to the solution appears; and

4. Verification – checking out the solution to make sure it "works".

Incubation is usually thought of as "sleeping on it" – an indication that the incubation period is considered to be longer than two minutes [Ekstrand,1967].

The incubation theory does offer some explanatory power. Longer interruptions provide more incubation time. However, if incubation is part of the explanation, the effects do not seem to be cumulative, or are being offset by the negative component.

### 7.4.4 Stress Relief

Several subjects reported that the interruptions provided a stress relieving mechanism. This position is supported by Weick [1985] who suggests that periodic breaks will enhance performance in electronic contexts. Weick is blunt when he suggests "...pull the plug and go for a walk." The benefit of periodic break for meetings is advocated by Conn [1980]. Gregerman [1981] also emphasizes the beneficial effects of strategically utilized breaks.

The improvement shown with longer interruptions may provide some benefits to subjects in the form of relieving stress. It seems possible that the beneficial effects may be time dependent. Perhaps the two minute interruptions allowed the subjects adequate time to "wind down".

## 7.5 DEFINING FREQUENCY/DURATION EFFECTS

Frequency and Duration still remain as preferable interruption descriptors from a management point of view. They are "measurable" (i.e., they can be captured on some sort of numeric scale). To apply the results of this research to performance evaluation, project management, job design or personal time management, variables of this sort provide more universal applicability than the qualitative attributes described in Section 7.6. Because of this appeal, these two factors merit further experimental examination.

### 7.5.1 Alternative Performance Covariate

The specific use of course material somewhat limits the the ability to generalize results. The Class Score covariate also measured performance under different stress levels and a slightly different presentation formats than those used by the experiment. As a result, there may be less explanatory power in Class Score than could be obtained in an improved design. However, the high significance of the Class Score covariate suggests that any future work should have a covariate that reflects expected subject performance on the primary task.

These deficiencies may be addressed by using a more universally accepted problem solving task. Questions from GMAT, GRE, SAT, or other standardized aptitude tests could be used as the primary task [Okogbaa and Shell, 1986]. Subjects could then perform the primary task twice. Once to establish a "Task Capability" covariate, and again to provide scoring under experiment conditions. This approach would also help to familiarize subjects with the experimental software and allow pair-wise comparisons, by subject, of interrupted versus uninterrupted performance.

### 7.5.2 Revised Interruption Procedure

Many subjects commented on the trivia questions used for interruptions. The preponderance of comments fall into two areas:

- the inability to control the interruption time; and
- the "stress relief" effects of the trivia questions.

Additional comments suggested that some form of audible signal or visual alarm accompany the interruption to let the subject know it has arrived. Future modifications to the primary task should minimize use of external references and devices. The experimenter noticed occasional incidents where a short interruption was missed because a subject was absorbed in a calculation or in reference material.

For this initial exploration of interruptions, the researcher chose to tightly control the times for interruptions. This control facilitated use of the factorial design of the experiment and the distinctions between Frequency and Duration manipulations. If Frequency and Duration are explored separately, the need for this tight control can be somewhat relaxed.

The "fun" nature of the trivia questions may have some bearing on the results. Since no financial support was available, every effort was made to provide as much intrinsic motivation as possible. The lighthearted nature of the interruptions was introduced to enhance the experience of the subjects. It is possible that some of the positive component is attributable to the relatively "pleasant" nature of these interruptions.

Subsequent experiments could use a revised interruption mechanism. Assuming absolute control of interruption time is not necessary, a "look-up" type task could be used. This might take the form finding phone numbers, figures from an annual report, or references from a list. If the software required a correct response before proceeding, the interruption would reasonably be expected to command the attention of the subject. A rough control over the length of the interruption could still be provided by asking the subject to look-up a variety of references. However, by requiring subject response before proceeding, some control or motivation is necessary to inhibit using interruption time for working on the primary task.

### 7.5.3 Single Factor Experiments

7.5.3.1 Duration — Based on the results of this exploratory study, Duration should be the next variable investigated. A Duration only experiment would require holding the frequency level constant and varying the length of the interruption over a wider variety of lengths. Assume the interruption task is finding a phone number. An interruption episode might consist of finding 0,3,6,9... phone numbers. This form of experiment would help to define the shape of a regression relationship between length of interruption and change in performance. When used in conjunction with a subject paced interruption, this format would provide an opportunity to use both continuous variables (e.g. length of interruption) and indicator variables for the number of "look-ups".

7.5.3.2 Frequency — Experiments to define singular effects of Frequency are somewhat more challenging. Holding the length of the interruption constant suggests a format similar to the trivia questions used in this experiment. Allowing the subjects to determine the length of the interruption might introduce an interaction between Duration and Frequency that may be hard to separate. The format for these experiments should be similar to those described for Duration only experiments – hold the length of the interruption episodes constant while varying the number of interruptions. Again the objective is to identify a regression relationship with Frequency as the independent variable and performance as the dependent variable.

## 7.6 ALTERNATIVES FOR FUTURE RESEARCH

After considering the results of this exploratory investigation, three additional areas merit consideration for further research:

- varying the nature of the primary task;
- varying the nature of the interruption; and
- determining if prior knowledge alters interruption effects.

The remainder of the section offers speculative discussions on how these areas may be explored in future research.

### 7.6.1 Nature of Primary Task

From the standpoint of wide management application, the notion of granularity (Section 7.3.2) is the most attractive primary task descriptor. However, interruptions may have a different effect based on the type of primary task.

The distinction between abstract reasoning and reading comprehension merits consideration. Many standardized examinations are divided into these two areas – providing a wide variety of suitable experimental materials. Use of these types of examinations takes advantage of standardized performance measures, predefined granularity, and the design expertise of other experts in the field.

Considerable research has been devoted to the decision making process. Combining this research with empirical studies in decision theory offers an attractive opportunity to examine another broad area of primary task.

### 7.6.2 Nature of Interruption

While this writer claims no expertise in psychology, it seems that a brief interruption reporting a death in the family would be potentially more disturbing than recording a shopping list. While this example may seem drastic, it suggests that the nature of the interruption may be related to the size of its deleterious effect.

Manipulations of this variable will be more difficult than manipulating the nature of the primary task. The task distinctions for primary tasks (see previous section) may be equally applicable to interruptions. A classification scheme that could be used for defining the nature of both the primary task and of the interruptions would be beneficial in defining effects. For example, if the primary task is abstract reasoning, what type of interruption would produce the most significant effect? Another abstract reasoning task, or a reading comprehension task?

Interruptions need not be confined to strictly mental activities. Weick [1985] suggests "going for a walk" as a self initiated interruption. Perhaps there are different (or therapeutic) effects from interruptions requiring physical activity. Many Japanese manufacturing plants have scheduled calisthenics for workers. Perhaps the purported benefits of periodic physical exercise would also benefit the knowledge worker.

Whatever form the interruption takes, care must be exercised to isolate nature of interruption from length and frequency effects. In this study, the choice of trivia questions was motivated by a desire to change the length of the interruption without changing its essential nature. This will remain as a difficulty in future studies.

### 7.6.3 Prior Knowledge

Subjects were told that interruptions may be occurring during the course of the reported experiment. Many workers will also find themselves in situations where they expect to be interrupted.

Mintzberg [1973] found managerial work to be fragmented and characterized by a wide variation of brief activities. According to Mintzberg's findings, half of a chief executive's activities take less than nine minutes with only ten percent requiring more than an hour. Managerial work can be thought of as interruption driven. Managers, whose work is made up of a wide variety of disjointed tasks, should have a much different strategic approach to personal time management than those who have a more singular task focus (e.g. engineers, programmers, architects...). Interruptions may have less of an impact on managers than on other types of knowledge workers.

This observation suggests that there may be a different strategy for performing work under conditions with a higher probability of interruption. Perhaps acknowledging the likelihood of interruptions alters the strategy employed for the primary task. If so, "unexpected" interruptions should have more impact than those that are anticipated.

Testing this effect would require a straight forward, 2 by 2 factorial design and some deceit. Prior to beginning the experiment, the subjects would shown the format for potential interruptions, or told nothing about the possibility. Then the interruption event would either occur or not occur. The table below depicts the manipulations for each of the four treatments.

Table 7.2: Treatments for "Prior Knowledge" Experiment

| Treatment | Subject Told To Expect Interruption | Interruption Actually Occurs |
|---|---|---|
| I | No | No |
| II | No | Yes |
| III | Yes | Yes |
| IV | Yes | No |

If there is a different strategy being employed when the subjects expect to be interrupted, then there should be distinct differences in performance between treatments I and IV, and between treatments II and III.

### 7.6.4 Other Areas for Research

7.6.4.1 Individual Capability — It seems possible that coping strategies may be developed for coping with interruptions. Perhaps, through experience or some psychological predisposition, the ability to perform in an interrupted environment may vary by individual. If so, future research in this area may focus on screening potential workers for this capability. It may be possible to develop "aptitude" tests for identification of this capability.

7.6.4.2 Non-Electronic Contexts — The personal computer was selected for presenting the primary task and manipulating study variables. Computer based tasks offer compelling laboratory advantages for precisely controlling interruption mechanics and ensuring uniformity of manipulations among subjects. While knowledge workers will likely spend much time in front of a CRT, future work should validate results in non computer contexts.

7.6.4.3 Time Dependency and Fatigue — The effects of interruptions may be changed by the length of the task. As workers progress through the day, effects may be modified by fatigue or other time dependent effects.

## 7.7 CONCLUSION

The role of the knowledge worker in contemporary business will continue to expand. Competitive firms will prosper through quicker response to dynamically shifting customer demands. As products and manufacturing processes become more sophisticated, industry and government workers will require more intellectual sophistication as well. As the value of knowledge workers increases, managing their time effectively will require a better understanding of the reactive environment in which they function. Managing interruptions will be a significant component in making best use of valuable human resources.

This exploratory investigation has confirmed that interruptions have a deleterious effect on complex task performance. However, short interruptions have a stronger deleterious effect than do long interruptions. For simple tasks long interruptions actually improved performance.

The counter-intuitive results of this exploratory study suggest that the nature of these effects will prove to be of interest to researchers and practitioners alike.

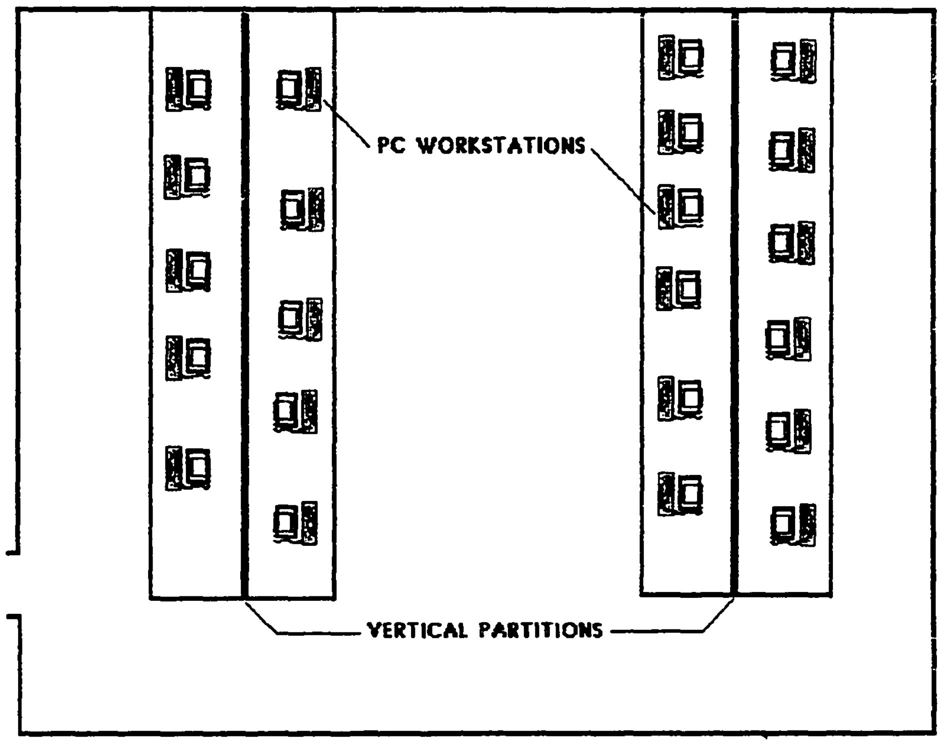
# Appendix A

# EXPERIMENT DESCRIPTION

## A.1 Laboratory Layout

Plan View of Park Center MIS MicroComputer Laboratory

(Not to Scale)



## A.2 Pre-Experiment Script

Welcome to the Park Center Labs. I see most of you found your way. Let me explain what you will be doing today.

These experiments are designed to find out how workers react to various conditions in the office environment. Your individual performance on a complex task will be evaluated, but your ideas how to cope with conditions are equally important.

Based on the random selection of these cards, [*show cards*] each of you will be asked to work on a sample examination. The exam is based on mathematical problems we covered on the first two midterms. There are ten unique versions of exams and times. The software will control the time you have to finish. Some of you will be allowed less time than others will. Your objective is to get the highest possible score in the time you have available. You may leave at any time but, credit will only be awarded for completed work.

A cash prize will be awarded to the best score in each of the ten task categories. In addition, the top three scores in each task category will be entered in a drawing for a copy of Borland's Quattro - a spreadsheet package with complete documentation. The assignment of tasks will be completely random — so you will be competing against 10 or 11 other people.

We will be simulating an office environment. Like an office, you may be interrupted at any time by trivial questions from colleagues or clients. You should answer these as best you can. Scoring systems are about the same for all tasks. Each question on the sample exam has a point value.

- A correct answer adds those points to your score.
- No answer scores zero.
- An incorrect answer subtracts 20% of the point value from your score.
- Trivial questions score 1 point each for correct answers and zero for a wrong or blank answer.

Thus, wild guesses are OK on trivial questions, but will hurt your score the practice exam. Details for your particular task will be presented on the computer.

Before we get on to the experiment, let's learn how the the software works. We will first have a practice session. When you feel you have mastered the software, raise your hand, and I will let you choose a card with your individual assignment.

Turn to your machine, and type "go" followed by the ENTER key. This will start our training session. Follow the instructions on the screen and feel free to ask any questions you wish. Take your time and ask any questions you may have.

[ TRAINING SESSION HERE ]

Print your name, section and 4 digit ID on the card to receive credit for participating. When you see a brown screen, type in the number exactly as shown on the card. This will initiate the experiment.

Leave the card where I can see it. I will be around to record the number so you will get credit for your work. When your task is complete, bring me the card, fill out a questionnaire, and you are done.

Thanks for helping me out and good luck.

## A.3 Task Instructions

---Instructions for Complex Task---

OBJECTIVE:

To get the highest posible score in the time you have available.

SCORING:

Practice Class Exam

- All questions are worth the points shown in braces [30] when answered correctly
- Questions left blank score 0
- Each incorrect answer costs -20%, so an incorrect answer to a 30 point question will subtract 6 from your score

Trivial Questions from Colleagues (if you have any)

- A correct answer scores 1, an incorrect or blank scores 0

YOUR TASK

There are 15 questions in your version of the practice exam. YOU ARE NOT EXPECTED TO FINISH! I suggest you scan the exam and answer the easiest questions first. Cash and/or prizes will be awarded to the highest combined score. Remember that you may be interrupted by trivial questions (or you may not.) You are only competing against people who have the identical task ( about 11 people). GOOD LUCK!

If you have questions, ask them now. Your time begins when you hit ENTER.

## Instructions for Simple Task

OBJECTIVE:

To get the highest posible score in the time you have available.

SCORING:

Practice Class Exam

- All questions are worth +10 points when answered correctly
- Questions left blank score 0
- Each incorrect answer costs you -2, so wild guesses will hurt your score

Trivial Questions from Colleagues (if you have any)

- A correct answer scores 1, an incorrect or blank scores 0

YOUR TASK

There are 40 questions in your version of the practice exam. YOU ARE NOT EXPECTED TO FINISH! I suggest you scan the exam and answer the easiest questions first. Cash and/or prizes will be awarded to the highest combined score. Remember that you may be interrupted by trivial questions (or you may not.) You are only competing against people who have the identical task ( about 11 people). GOOD LUCK!

If you have questions, ask them now. Your time begins when you hit ENTER.

## A.4 Post Experiment Questionnaire

The purpose of this questionnaire is to gather information on your skills and background, as well as information about your experience during the experiment. Your responses will be kept completely confidential.

PART I – Background

1. What is your age? __________ years old

2. What is your gender? (circle one) Female Male

3. Where were you born? ______________________________

4. Before today, have you used a computer system like this one for other experiments, school work, playing games etc?

   (circle one) Yes No

   If yes, about how often have you used a computer like this one in the last two months?

   About _______________ times.

Part II – The Practice Exam

This part of the questionnaire has several questions. Please indicate how each statement applies to your experience by circling the number that comes closest to your impression. There are no right or wrong answers. You may find some statements are similar to other statements. Don't be concerned about this. Work through quickly and circle your first impression.

1. How many questions were in your practice exam? _____

2. Compared to the exams you have taken in class, how would you the rate the questions on this practice exam?

| Much Harder | | About the Same | | | | Much Easier |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

3. If you had taken the practice exam on paper, instead of on the computer, (with all other conditions being the same), how do you think you would have scored?

| Much Higher | | About the Same | | | | Much Lower |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

4. Were you interrupted by trivial questions ? YES NO
   If YES, go on to page 2. If NO, skip to page 3.

Part III Interruptions ________________________________

5. About how many times were you interrupted by trivial questions?

About ________ times

6. On average, about how many questions were you asked each time you were interrupted with trivial questions?

About _________ questions were asked each time.

7. How do you think you would have scored on the practice exam if you had not been interrupted?

| Much Higher | | | About the Same | | | Much Lower |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

8. How did the interruptions affect your ability to concentrate on the practice exam?

| Really hurt | | | Didn't Affect Me | | | Really Helped |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

9. When an interruption arrived, what was your initial reaction?

| Really Pleased | | | Indifferent | | | Really Displeased |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

10. Suppose that you had to take the practice exam under these two conditions: A.) 100 minutes of uninterrupted time or B.) 120 minutes with 20 minutes of interruption. Which of your two scores, A or B, would be higher?

| A much higher | | | about the same | | | B much higher |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

11. Suppose you had to answer 12 trivial questions during the course of the practice exam. Which of the following interruption scenarios would you prefer?

A> 1 interruption with 12 questions
B> 2 interruptions with 6 questions in each
C> 3 interruptions with 4 questions in each
D> 4 interruptions with 3 questions in each
E> 6 interruptions with 2 questions in each
F> 12 interruptions with 1 question in each

Part IV Your Comments and Strategies

Please use this page for comments on the experiment and any observations you might have on how you deal with interruptions.

Thanks again for your help!

# Appendix B

# EXPERIMENT ANALYSIS AND RESULTS

## B.1 Raw Experiment Data

LEGEND:

| | | |
|---|---|---|
| F | = FREQUENCY | { 0 = Control, 2 = Lo, 6 = Hi } |
| D | = DURATION | { 0 = Control, 2 = Lo, 6 = Hi } |
| T | = TASK | { S = Simple, C = Complex } |
| CLSX | = CLASS SCORE | |
| PRE | = PREVIOUS USER | { Y = Yes , N = No } |
| SEX | | { M = Male, F = Female } |
| CPTS | = CORRECT POINTS | |
| TPTS | = TOTSCORE | |
| WPTS | = WRONG POINTS | |
| TAT | = Trivia Questions Attempted ( Max = C times F ) | |
| TC | = Trivia Questions Correct | |

| SUB | F | D | T | CLSX | PRE | SEX | CPTS | TPTS | WPTS | TAT | TC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | S | -1 | N | F | 180 | 174 | 30 | 0 | 0 |
| 2 | 0 | 0 | S | -5 | Y | M | 140 | 116 | 120 | 0 | 0 |
| 3 | 0 | 0 | S | 30 | Y | F | 160 | 150 | 50 | 0 | 0 |
| 4 | 0 | 0 | S | 25 | Y | M | 170 | 150 | 100 | 0 | 0 |
| 5 | 0 | 0 | S | 54 | Y | M | 210 | 186 | 120 | 0 | 0 |
| 6 | 0 | 0 | S | -25 | Y | M | 120 | 100 | 100 | 0 | 0 |
| 7 | 0 | 0 | S | 42 | N | M | 190 | 172 | 90 | 0 | 0 |
| 8 | 0 | 0 | S | 15 | Y | M | 90 | 80 | 50 | 0 | 0 |
| 9 | 0 | 0 | S | 47 | Y | F | 200 | 186 | 70 | 0 | 0 |
| 10 | 0 | 0 | S | -23 | Y | M | 90 | 80 | 50 | 0 | 0 |
| 11 | 0 | 0 | S | 45 | N | M | 240 | 234 | 30 | 0 | 0 |
| 12 | 0 | 0 | S | 41 | Y | M | 170 | 164 | 30 | 0 | 0 |
| 101 | 0 | 0 | C | 15 | Y | F | 90 | 58 | 160 | 0 | 0 |
| 102 | 0 | 0 | C | -47 | N | M | 80 | 66 | 70 | 0 | 0 |
| 103 | 0 | 0 | C | 26 | Y | M | 100 | 82 | 90 | 0 | 0 |
| 104 | 0 | 0 | C | -19 | Y | F | 110 | 94 | 80 | 0 | 0 |
| 105 | 0 | 0 | C | -30 | Y | M | 40 | 12 | 140 | 0 | 0 |
| 106 | 0 | 0 | C | 29 | Y | M | 150 | 136 | 70 | 0 | 0 |
| 107 | 0 | 0 | C | 15 | N | F | 70 | 32 | 190 | 0 | 0 |
| 108 | 0 | 0 | C | -22 | Y | F | 70 | 46 | 120 | 0 | 0 |
| 109 | 0 | 0 | C | 48 | Y | M | 100 | 72 | 140 | 0 | 0 |
| 110 | 0 | 0 | C | -34 | Y | F | 70 | 36 | 170 | 0 | 0 |
| 111 | 0 | 0 | C | 7 | Y | M | 140 | 114 | 130 | 0 | 0 |
| 112 | 0 | 0 | C | -51 | Y | M | 50 | 14 | 180 | 0 | 0 |

| SUB | F | D | T | CLSX | PRE | SEX | CPTS | TPTS | WPTS | TAT | TC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 201 | 2 | 2 | S | -91 | N | F | 30 | -12 | 210 | 4 | 0 |
| 202 | 2 | 2 | S | 27 | Y | F | 80 | 64 | 80 | 4 | 1 |
| 203 | 2 | 2 | S | -19 | Y | M | 120 | 110 | 50 | 3 | 0 |
| 204 | 2 | 2 | S | 29 | Y | M | 110 | 102 | 40 | 2 | 1 |
| 205 | 2 | 2 | S | -39 | Y | M | 90 | 82 | 40 | 3 | 1 |
| 206 | 2 | 2 | S | -24 | Y | M | 110 | 98 | 60 | 2 | 1 |
| 207 | 2 | 2 | S | -19 | Y | M | 120 | 104 | 80 | 4 | 3 |
| 208 | 2 | 2 | S | 28 | Y | F | 190 | 180 | 50 | 3 | 1 |
| 209 | 2 | 2 | S | -13 | Y | F | 150 | 148 | 10 | 4 | 2 |
| 210 | 2 | 2 | S | 53 | Y | M | 240 | 226 | 70 | 0 | 0 |
| 211 | 2 | 2 | S | -29 | Y | M | 170 | 162 | 40 | 4 | 3 |
| 212 | 2 | 2 | S | -61 | N | M | 90 | 66 | 120 | 4 | 1 |
| 301 | 2 | 2 | C | 31 | Y | F | 90 | 50 | 200 | 4 | 2 |
| 302 | 2 | 2 | C | 32 | N | F | 60 | 28 | 160 | 1 | 0 |
| 303 | 2 | 2 | C | 42 | Y | M | 70 | 38 | 160 | 4 | 1 |
| 304 | 2 | 2 | C | -10 | Y | M | 70 | 46 | 120 | 4 | 3 |
| 305 | 2 | 2 | C | 63 | Y | M | 100 | 68 | 160 | 2 | 1 |
| 306 | 2 | 2 | C | -43 | Y | M | 90 | 68 | 110 | 3 | 1 |
| 307 | 2 | 2 | C | 57 | Y | F | 90 | 72 | 90 | 2 | 0 |
| 308 | 2 | 2 | C | -29 | Y | M | 0 | -26 | 130 | 4 | 3 |
| 309 | 2 | 2 | C | -1 | N | M | 30 | 6 | 120 | 4 | 0 |
| 310 | 2 | 2 | C | -47 | N | M | 30 | 0 | 150 | 3 | 0 |
| 311 | 2 | 2 | C | -84 | N | M | 30 | 14 | 80 | 4 | 2 |
| 312 | 2 | 2 | C | 10 | Y | M | 90 | 84 | 30 | 4 | 0 |

| SUB | F | D | T | CLSX | PRE | SEX | CPTS | TPTS | WPTS | TAT | TC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 401 | 2 | 6 | S | 35 | Y | M | 210 | 198 | 60 | 11 | 8 |
| 402 | 2 | 6 | S | -31 | Y | M | 170 | 160 | 50 | 12 | 4 |
| 403 | 2 | 6 | S | 39 | Y | F | 230 | 218 | 60 | 3 | 1 |
| 404 | 2 | 6 | S | 35 | N | F | 90 | 78 | 60 | 12 | 4 |
| 405 | 2 | 6 | S | -74 | Y | M | 130 | 118 | 60 | 10 | 3 |
| 406 | 2 | 6 | S | -22 | Y | F | 140 | 122 | 90 | 11 | 4 |
| 407 | 2 | 6 | S | 30 | Y | M | 170 | 162 | 40 | 12 | 1 |
| 408 | 2 | 6 | S | 31 | Y | F | 180 | 174 | 30 | 10 | 2 |
| 409 | 2 | 6 | S | 4 | Y | F | 210 | 192 | 90 | 11 | 3 |
| 410 | 2 | 6 | S | -11 | Y | M | 150 | 138 | 60 | 12 | 8 |
| 411 | 2 | 6 | S | 6 | N | F | 170 | 162 | 40 | 11 | 6 |
| 412 | 2 | 6 | S | 45 | N | M | 220 | 206 | 70 | 12 | 4 |
| 414 | 2 | 6 | S | 8 | Y | F | 160 | 148 | 60 | 11 | 3 |
| 501 | 2 | 6 | C | 8 | Y | M | 100 | 68 | 160 | 11 | 6 |
| 502 | 2 | 6 | C | -55 | Y | F | 0 | -22 | 110 | 11 | 1 |
| 503 | 2 | 6 | C | -14 | Y | F | 40 | 2 | 190 | 12 | 1 |
| 504 | 2 | 6 | C | 51 | Y | F | 70 | 38 | 160 | 12 | 2 |
| 505 | 2 | 6 | C | 23 | Y | M | 110 | 88 | 110 | 12 | 8 |
| 506 | 2 | 6 | C | 57 | N | M | 90 | 54 | 180 | 10 | 8 |
| 507 | 2 | 6 | C | -21 | Y | F | 10 | -20 | 150 | 11 | 1 |
| 508 | 2 | 6 | C | 35 | N | F | 0 | -46 | 230 | 12 | 3 |
| 509 | 2 | 6 | C | 27 | Y | M | 80 | 16 | 320 | 9 | 4 |
| 510 | 2 | 6 | C | -23 | Y | F | 60 | 50 | 50 | 11 | 2 |
| 511 | 2 | 6 | C | -18 | Y | M | 160 | 144 | 80 | 12 | 5 |
| 512 | 2 | 6 | C | -28 | Y | M | 50 | -12 | 310 | 12 | 7 |

| SUB | F | D | T | CLSX | PRE | SEX | CPTS | TPTS | WPTS | TAT | TC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 602 | 6 | 2 | S | 38 | N | M | 290 | 276 | 70 | 10 | 7 |
| 603 | 6 | 2 | S | 43 | Y | M | 170 | 162 | 40 | 3 | 1 |
| 604 | 6 | 2 | S | -59 | N | M | 90 | 80 | 50 | 10 | 3 |
| 605 | 6 | 2 | S | 25 | Y | F | 110 | 96 | 70 | 10 | 3 |
| 606 | 6 | 2 | S | 13 | Y | F | 90 | 70 | 100 | 11 | 4 |
| 607 | 6 | 2 | S | 50 | Y | M | 130 | 124 | 30 | 5 | 1 |
| 608 | 6 | 2 | S | -52 | N | M | 80 | 70 | 50 | 3 | 1 |
| 609 | 6 | 2 | S | 17 | Y | M | 170 | 150 | 100 | 11 | 5 |
| 610 | 6 | 2 | S | -6 | N | F | 110 | 90 | 100 | 9 | 1 |
| 611 | 6 | 2 | S | -12 | Y | F | 220 | 206 | 70 | 12 | 9 |
| 612 | 6 | 2 | S | 30 | Y | F | 210 | 190 | 100 | 12 | 3 |
| 614 | 6 | 2 | S | 43 | Y | F | 180 | 162 | 90 | 11 | 0 |
| 701 | 6 | 2 | C | -8 | Y | M | 40 | -22 | 310 | 12 | 6 |
| 702 | 6 | 2 | C | 15 | Y | F | 90 | 56 | 170 | 10 | 3 |
| 703 | 6 | 2 | C | -104 | Y | M | 30 | 18 | 60 | 12 | 4 |
| 704 | 6 | 2 | C | -25 | Y | M | 20 | -6 | 130 | 9 | 2 |
| 705 | 6 | 2 | C | -24 | Y | F | 0 | -30 | 150 | 12 | 5 |
| 706 | 6 | 2 | C | -10 | Y | F | 20 | -30 | 250 | 12 | 5 |
| 707 | 6 | 2 | C | 29 | Y | M | 60 | 38 | 110 | 12 | 4 |
| 708 | 6 | 2 | C | 1 | Y | F | 100 | 80 | 100 | 12 | 7 |
| 709 | 6 | 2 | C | 36 | N | M | 170 | 158 | 60 | 9 | 8 |
| 710 | 6 | 2 | C | -38 | Y | F | 20 | -8 | 140 | 7 | 4 |
| 711 | 6 | 2 | C | -1 | Y | M | 70 | 56 | 70 | 11 | 6 |
| 712 | 6 | 2 | C | 12 | Y | M | 70 | 44 | 130 | 10 | 1 |
| 714 | 6 | 2 | C | 16 | N | M | 130 | 84 | 230 | 11 | 5 |

| SUB | F | D | T | CLSX | PRE | SEX | CPTS | TPTS | WPTS | TAT | TC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 801 | 6 | 6 | S | -24 | N | M | 140 | 110 | 150 | 36 | 11 |
| 802 | 6 | 6 | S | -16 | N | F | 180 | 162 | 90 | 36 | 9 |
| 803 | 6 | 6 | S | 36 | N | F | 220 | 202 | 90 | 33 | 7 |
| 804 | 6 | 6 | S | -19 | Y | F | 170 | 166 | 20 | 35 | 16 |
| 805 | 6 | 6 | S | 61 | Y | M | 200 | 180 | 100 | 17 | 13 |
| 806 | 6 | 6 | S | -82 | Y | F | 120 | 96 | 120 | 34 | 19 |
| 807 | 6 | 6 | S | -9 | Y | M | 100 | 94 | 30 | 35 | 11 |
| 808 | 6 | 6 | S | -4 | Y | M | 130 | 116 | 70 | 35 | 9 |
| 809 | 6 | 6 | S | 15 | N | M | 210 | 196 | 70 | 34 | 9 |
| 810 | 6 | 6 | S | -16 | Y | M | 130 | 104 | 130 | 36 | 13 |
| 811 | 6 | 6 | S | 18 | Y | M | 120 | 114 | 30 | 21 | 10 |
| 812 | 6 | 6 | S | 20 | Y | F | 210 | 192 | 90 | 35 | 16 |
| 901 | 6 | 6 | C | 46 | N | F | 120 | 72 | 240 | 36 | 21 |
| 902 | 6 | 6 | C | -55 | Y | F | 70 | 18 | 260 | 35 | 16 |
| 903 | 6 | 6 | C | 44 | Y | M | 50 | 10 | 200 | 36 | 16 |
| 904 | 6 | 6 | C | -54 | Y | M | 40 | -14 | 270 | 32 | 17 |
| 905 | 6 | 6 | C | -12 | Y | F | 60 | 30 | 150 | 33 | 13 |
| 906 | 6 | 6 | C | -1 | N | M | 140 | 122 | 90 | 32 | 6 |
| 907 | 6 | 6 | C | -4 | Y | M | 120 | 94 | 130 | 34 | 19 |
| 908 | 6 | 6 | C | -61 | Y | F | 20 | -14 | 170 | 22 | 8 |
| 909 | 6 | 6 | C | 6 | Y | M | 140 | 120 | 100 | 27 | 17 |
| 910 | 6 | 6 | C | 2 | Y | F | 90 | 54 | 180 | 35 | 6 |
| 911 | 6 | 6 | C | 46 | Y | M | 80 | 62 | 90 | 27 | 9 |
| 912 | 6 | 6 | C | -115 | Y | F | 70 | 10 | 300 | 8 | 2 |

## B.2 MODELI Stepwise Regression Results

RESULTS OF STEPWISE REGRESSION - MODELI
Step Wise Data Set - Omit subjects 511, 602, 611

Number of Variables = 14      Number of Observations = 119

F to Enter = 3.9000      F to Leave = 3.8000

Dependent Variable is TotScore

| VARIABLE(S) NOT IN MODEL | PARTIAL CORRELATION COEFFICIENT(S) |
|---|---|
| XSCORE | .1995567 |
| C | .5156741 |
| F | .0122012 |
| D | 2.419484E-04 |
| XC | 5.822057E-03 |
| XF | 1.291716E-02 |
| XD | 2.167205E-03 |
| CF | 3.479102E-02 |
| CD | 8.446172E-02 |
| FD | 2.257153E-02 |
| XtraMins | 1.574081E-04 |
| PREVUSER | 3.034231E-03 |
| SEX | 4.95008E-03 |

C HAS THE LARGEST PARTIAL CORRELATION COEFFICIENT
THE VARIABLE TO ENTER IS C

| VARIABLE | PARTIAL F | F-TO-ENTER |
|---|---|---|
| C | 124.5729 | 3.9 |

THE MODEL NOW CONTAINS THE FOLLOWING VARIABLE(S):
C

THE PERCENTAGE OF VARIATION EXPLAINED IS: 51.56741 %

THE % IMPROVEMENT OVER THE PREVIOUS MODEL IS: 51.56741 %

********** END STEP 1 ************

| VARIABLE(S) NOT IN MODEL | PARTIAL CORRELATION COEFFICIENT(S) |
|---|---|
| XSCORE | .1356464 |
| F | 7.718366E-03 |
| D | 1.105614E-03 |
| XC | 5.085505E-03 |
| XF | 1.152978E-02 |
| XD | 5.765457E-03 |
| CF | 5.12738E-05 |
| CD | 1.165327E-02 |
| FD | 2.110623E-02 |
| XtraMins | 1.266634E-04 |
| PREVUSER | 7.406942E-05 |
| SEX | 3.987561E-03 |

XSCORE HAS THE LARGEST PARTIAL CORRELATION COEFFICIENT
THE VARIABLE TO ENTER IS XSCORE

| VARIABLE | PARTIAL F | F-TO-ENTER |
|---|---|---|
| XSCORE | 45.12734 | 3.9 |

PARTIAL F VALUE FOR REMAINING VARIABLES

| VARIABLE | PARTIAL F | F-TO-LEAVE |
|---|---|---|
| C | 150.2944 | 3.8 |

– ALL VARIABLES REMAIN –

THE MODEL NOW CONTAINS THE FOLLOWING VARIABLE(S):
XSCORE
C

THE PERCENTAGE OF VARIATION EXPLAINED IS: 65.13204 %

THE % IMPROVEMENT OVER THE PREVIOUS MODEL IS: 13.56464 %

********** END STEP 2 ************

| VARIABLE(S) NOT IN MODEL | PARTIAL CORRELATION COEFFICIENT(S) |
|---|---|
| F | 2.456272E-03 |
| D | 8.264191E-05 |
| XC | 1.144272E-02 |
| XF | 1.114233E-03 |
| XD | 3.986347E-03 |
| CF | 2.028113E-04 |
| CD | 1.808265E-02 |
| FD | 2.142842E-02 |
| XtraMins | 1.967675E-03 |
| PREVUSER | 3.169644E-04 |
| SEX | 3.237828E-03 |

FD HAS THE LARGEST PARTIAL CORRELATION COEFFICIENT
THE VARIABLE TO ENTER IS FD

| VARIABLE | PARTIAL F | F-TO-ENTER |
|---|---|---|
| FD | 7.530204 | 3.9 |

PARTIAL F VALUE FOR REMAINING VARIABLES

| VARIABLE | PARTIAL F | F-TO-LEAVE |
|---|---|---|
| XSCORE | 47.78099 | 3.8 |
| C | 158.2493 | 3.8 |

— ALL VARIABLES REMAIN —

THE MODEL NOW CONTAINS THE FOLLOWING VARIABLE(S):
XSCORE
C
FD

THE PERCENTAGE OF VARIATION EXPLAINED IS: 67.27489 %

THE % IMPROVEMENT OVER THE PREVIOUS MODEL IS: 2.142845 %

********** END STEP 3 ************

| VARIABLE(S) NOT IN MODEL | PARTIAL CORRELATION COEFFICIENT(S) |
|---|---|
| F | 3.490955E-04 |
| D | 5.736959E-04 |
| XC | 8.355437E-03 |
| XF | 1.76254E-04 |
| XD | 1.418515E-03 |
| CF | 1.092899E-04 |
| CD | 1.890761E-02 |
| XtraMins | 5.918081E-04 |
| PREVUSER | 2.966863E-04 |
| SEX | 2.811793E-03 |

CD HAS THE LARGEST PARTIAL CORRELATION COEFFICIENT
THE VARIABLE TO ENTER IS CD

| VARIABLE | PARTIAL F | F-TO-ENTER |
|---|---|---|
| CD | 6.990474 | 3.9 |

PARTIAL F VALUE FOR REMAINING VARIABLES

| VARIABLE | PARTIAL F | F-TO-LEAVE |
|---|---|---|
| XSCORE | 52.71203 | 3.8 |
| C | 137.7133 | 3.8 |
| FD | 8.227461 | 3.8 |

— ALL VARIABLES REMAIN —

THE MODEL NOW CONTAINS THE FOLLOWING VARIABLE(S):
XSCORE
C
CD
FD

THE PERCENTAGE OF VARIATION EXPLAINED IS: 69.16565 %

THE % IMPROVEMENT OVER THE PREVIOUS MODEL IS: 1.890762 %

********** END STEP 4 ************

| VARIABLE(S) NOT IN MODEL | PARTIAL CORRELATION COEFFICIENT(S) |
|---|---|
| F | 3.02506E-04 |
| D | 5.208861E-04 |
| XC | 1.029167E-02 |
| XF | 1.942624E-03 |
| XD | 5.430767E-03 |
| CF | 1.085795E-02 |
| XtraMins | 1.212522E-03 |
| PREVUSER | 2.004221E-04 |
| SEX | 3.234098E-03 |

CF HAS THE LARGEST PARTIAL CORRELATION COEFFICIENT
THE VARIABLE TO ENTER IS CF

| VARIABLE | PARTIAL F | F-TO-ENTER |
|---|---|---|
| CF | 4.124397 | 3.9 |

PARTIAL F VALUE FOR REMAINING VARIABLES

| VARIABLE | PARTIAL F | F-TO-LEAVE |
|---|---|---|
| XSCORE | 55.16365 | 3.8 |
| C | 144.9202 | 3.8 |
| CD | 11.26495 | 3.8 |
| FD | 9.050663 | 3.8 |

— ALL VARIABLES REMAIN —

THE MODEL NOW CONTAINS THE FOLLOWING VARIABLE(S):
XSCORE
C
CF
CD
FD

THE PERCENTAGE OF VARIATION EXPLAINED IS: 70.25144 %

THE % IMPROVEMENT OVER THE PREVIOUS MODEL IS: 1.085793 %

********** END STEP 5 ************

| VARIABLE(S) NOT IN MODEL | PARTIAL CORRELATION COEFFICIENT(S) |
|---|---|
| F | 4.023862E-04 |
| D | 5.335649E-04 |
| XC | 8.636753E-03 |
| XF | 6.029455E-04 |
| XD | 4.376249E-03 |
| XtraMins | 1.461104E-03 |
| PREVUSER | 3.886848E-04 |
| SEX | 2.702491E-03 |

XC HAS THE LARGEST PARTIAL CORRELATION COEFFICIENT
THE VARIABLE TO ENTER IS XC

| VARIABLE | PARTIAL F | F-TO-ENTER |
|---|---|---|
| XC | 3.348867 | 3.9 |

SINCE THE PARTIAL F FOR XC IS <= THE F-TO-ENTER VALUE
THE STEPWISE REGRESSION IS COMPLETED

## B.3 MODELI Regression Results

REGRESSION REPORT: FINAL1 Dependent Variable: TOTSCORE
Class Score Covariate and TOTSCORE

| | | | |
|---|---|---|---|
| R Squared : | 0.199557 | SSTO: | 545338.21849 |
| Std Err of Est: | 61.080886 | SSR : | 108825.88760 |
| Determinant R : | 1.000000 | SSE : | 436512.33088 |
| F Statistic : | 29.169002 | MSE : | 3730.87462 |
| with | 1 and 117 dF | | |

| Variable | Coefficient | TScore | Variance Inf Fact | Std Err Coeff. | Beta Coeff |
|---|---|---|---|---|---|
| CONSTANT | 88.49197 | 15.8039 | | 5.5994 | |
| XSCORE | 0.79465 | 5.4008 | 1.0000 | 0.1471 | 0.4467 |

REMARKS

None

REGRESSION REPORT: FINAL2 Dependent Variable: TOTSCORE
Add C to FINAL1

| | | | |
|---|---|---|---|
| R Squared : | 0.651320 | SSTO: | 545338.21849 |
| Std Err of Est: | 40.487151 | SSR : | 355189.93043 |
| Determinant R : | 0.987355 | SSE : | 190148.28806 |
| F Statistic : | 108.341843 | MSE : | 1639.20938 |
| with | 2 and 116 dF | | |

| Variable | Coefficient | TScore | Variance Inf Fact | Std Err Coeff. | Beta Coeff |
|---|---|---|---|---|---|
| CONSTANT | 88.90748 | 23.9535 | | 3.7117 | |
| XSCORE | 0.65934 | 6.7177 | 1.0128 | 0.0982 | 0.3707 |
| C | -45.79245 | -12.2595 | 1.0128 | 3.7353 | 0.6764 |

REMARKS

F for Reduced Model = 66.034 with 1 and 117 dF

REGRESSION REPORT: FINAL3 Dependent Variable: TOTSCORE
Add FD to FINAL2

| | | | |
|---|---|---|---|
| R Squared : | 0.672749 | SSTO: | 545338.21849 |
| Std Err of Est: | 39.393506 | SSR : | 366875.66523 |
| Determinant R : | 0.987299 | SSE : | 178462.55326 |
| F Statistic : | 78.804023 | MSE : | 1551.84829 |
| with | 3 and 115 dF | | |

| Variable | Coefficient | TScore | Variance Inf Fact | Std Err Coeff. | Beta Coeff |
|---|---|---|---|---|---|
| CONSTANT | 80.75856 | 17.2725 | | 4.6755 | |
| XSCORE | 0.66013 | 6.9124 | 1.0128 | 0.0955 | 0.3711 |
| C | -45.72061 | -12.5797 | 1.0129 | 3.6345 | 0.6754 |
| FD | 20.20059 | 2.7441 | 1.0001 | 7.3614 | 0.1464 |

REMARKS

F for Reduced Model = 7.129 with 1 and 116 dF

REGRESSION REPORT: FINAL4 Dependent Variable: TOTSCORE
Add CD to FINAL3

| | | | |
|---|---|---|---|
| R Squared : | 0.691656 | SSTO: | 545338.21849 |
| Std Err of Est: | 38.405901 | SSR : | 377186.70699 |
| Determinant R : | 0.915758 | SSE : | 168151.51150 |
| F Statistic : | 63.929376 | MSE : | 1475.01326 |
| with | 4 and 114 dF | | |

| Variable | Coefficient | TScore | Variance Inf Fact | Std Err Coeff. | Beta Coeff |
|---|---|---|---|---|---|
| CONSTANT | 80.36686 | 17.6215 | | 4.5607 | |
| XSCORE | 0.67769 | 7.2603 | 1.0180 | 0.0933 | 0.3810 |
| C | -43.13561 | -11.7351 | 1.0900 | 3.6758 | 0.6372 |
| CD | -12.42855 | -2.6440 | 1.0781 | 4.7007 | 0.1428 |
| FD | 20.59010 | 2.8684 | 1.0005 | 7.1784 | 0.1492 |

REMARKS

F for Reduced Model = 6.644 with 1 and 115 dF

REGRESSION REPORT: FINAL5 Dependent Variable: TOTSCORE
Add CF to FINAL4 model

| | | | |
|---|---|---|---|
| R Squared : | 0.702514 | SSTO: | 545338.21849 |
| Std Err of Est: | 37.890183 | SSR : | 383107.96306 |
| Determinant R : | 0.493968 | SSE : | 162230.25543 |
| F Statistic : | 53.370069 | MSE : | 1435.66598 |
| with | 5 and 113 dF | | |

| Variable | Coefficient | TScore | Variance Inf Fact | Std Err Coeff. | Beta Coeff |
|---|---|---|---|---|---|
| CONSTANT | 79.62200 | 17.6373 | | 4.5144 | |
| XSCORE | 0.68440 | 7.4272 | 1.0193 | 0.0921 | 0.3847 |
| C | -43.87544 | -12.0383 | 1.1010 | 3.6447 | 0.6481 |
| CF | 12.44086 | 2.0309 | 1.8539 | 6.1259 | 0.1419 |
| CD | -20.50930 | -3.3563 | 1.8718 | 6.1106 | 0.2356 |
| FD | 21.33412 | 3.0084 | 1.0032 | 7.0914 | 0.1546 |

REMARKS

F reduced = 4.014 with 1 and 114 Df

## B.4 Questionnaire Results

LEGEND:

| | | |
|---|---|---|
| SUB | = | Subject Number |
| AGE | = | Age in Years |
| SEX | = | Male, Female |
| USE | = | Previous Computer User (Y/N) |
| XSC | = | Class Score |
| TOTSC | = | TOTSCORE |
| COM | = | Did subject provide additional comments on questionnaire (Y/N) |

QUESTION RESPONSES: ( 0 indicates subject did not respond)

Q2 - Compared to the exams you have taken in class, how would you the rate the questions on this practice exam? ( 1 = Much Harder)

Q3 - If you had taken the practice exam on paper, instead of on the computer, (with all other conditions being the same), how do you think you would have scored? ( 1 = Much Higher)

Q8 - How did the interruptions affect your ability to concentrate on the practice exam? ( 1 = Really Hurt)

Q9 - When an interruption arrived, what was your initial reaction? ( 1 = Really Pleased)

Q10- Suppose that you had to take the practice exam under these two conditions: A.) 100 minutes of uninterrupted time or B.) 120 minutes w/20 minutes of interruption. Which of your two scores, A or B, would be higher? ( 1 = A much higher)

Q11- Suppose you had to answer 12 trivial questions during the course of the practice exam. Which of these interruption scenarios would you prefer?
- 1 = 1 interruption with 12 questions
- 2 = 2 interruptions with 6 questions in each
- 3 = 3 interruptions with 4 questions in each
- 4 = 4 interruptions with 3 questions in each
- 5 = 6 interruptions with 2 questions in each
- 6 = 12 interruptions with 1 question in each

| SUB | AGE | SEX | USE | XSC | TOTSC | Q2 | Q3 | Q8 | Q9 | Q10 | Q11 | COM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 47 | F | N | -1 | 174 | 4 | 2 | 0 | 0 | 0 | 0 | Y |
| 2 | 20 | M | Y | -5 | 116 | 4 | 6 | 0 | 0 | 0 | 0 | Y |
| 3 | 21 | F | Y | 30 | 150 | 3 | 3 | 0 | 0 | 0 | 0 | Y |
| 4 | 0 | M | Y | 25 | 150 | 4 | 4 | 0 | 0 | 0 | 0 | Y |
| 5 | 20 | M | Y | 54 | 186 | 3 | 2 | 0 | 0 | 0 | 0 | N |
| 6 | 22 | M | Y | -25 | 100 | 4 | 4 | 0 | 0 | 0 | 0 | N |
| 7 | 0 | M | N | 42 | 172 | 4 | 4 | 0 | 0 | 0 | 0 | N |
| 8 | 22 | M | Y | 15 | 80 | 3 | 1 | 0 | 0 | 0 | 0 | Y |
| 9 | 0 | F | Y | 47 | 186 | 4 | 2 | 0 | 0 | 0 | 0 | Y |
| 10 | 24 | M | Y | -23 | 80 | 4 | 3 | 0 | 0 | 0 | 0 | Y |
| 11 | 21 | M | N | 45 | 234 | 4 | 4 | 0 | 0 | 0 | 0 | N |
| 12 | 21 | M | Y | 41 | 164 | 4 | 4 | 0 | 0 | 0 | 0 | Y |
| 101 | 19 | F | Y | 15 | 58 | 3 | 2 | 0 | 0 | 0 | 0 | Y |
| 102 | 26 | M | N | -47 | 66 | 4 | 2 | 0 | 0 | 0 | 0 | Y |
| 103 | 21 | M | Y | 26 | 82 | 2 | 3 | 0 | 0 | 0 | 0 | Y |
| 104 | 21 | F | Y | -19 | 94 | 5 | 6 | 0 | 0 | 0 | 0 | Y |
| 105 | 28 | M | Y | -30 | 12 | 3 | 2 | 0 | 0 | 0 | 0 | Y |
| 106 | 22 | M | Y | 29 | 136 | 4 | 3 | 0 | 0 | 0 | 0 | Y |
| 107 | 21 | F | N | 15 | 32 | 4 | 3 | 0 | 0 | 0 | 0 | N |
| 108 | 26 | F | Y | -22 | 46 | 3 | 4 | 0 | 0 | 0 | 0 | Y |
| 109 | 20 | M | Y | 48 | 72 | 3 | 2 | 0 | 0 | 0 | 0 | Y |
| 110 | 22 | F | Y | -34 | 36 | 4 | 4 | 0 | 0 | 0 | 0 | N |
| 111 | 22 | M | Y | 7 | 114 | 4 | 5 | 0 | 0 | 0 | 0 | Y |
| 112 | 27 | M | Y | -51 | 14 | 4 | 4 | 0 | 0 | 0 | 0 | N |

| SUB | AGE | SEX | USE | XSC | TOTSC | Q2 | Q3 | Q8 | Q9 | Q10 | Q11 | COM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 201 | 22 | F | N | -91 | -12 | 1 | 2 | 2 | 6 | 1 | 1 | Y |
| 202 | 20 | F | Y | 27 | 64 | 2 | 3 | 4 | 4 | 2 | 2 | Y |
| 203 | 22 | M | Y | -19 | 110 | 4 | 2 | 2 | 7 | 1 | 1 | Y |
| 204 | 22 | M | Y | 29 | 102 | 3 | 1 | 3 | 4 | 1 | 3 | N |
| 205 | 25 | M | Y | -39 | 82 | 2 | 2 | 4 | 4 | 4 | 5 | Y |
| 206 | 22 | M | Y | -24 | 98 | 4 | 4 | 4 | 4 | 5 | 5 | Y |
| 207 | 27 | M | Y | -19 | 104 | 4 | 2 | 3 | 5 | 2 | 1 | Y |
| 208 | 21 | F | Y | 28 | 180 | 6 | 4 | 4 | 3 | 5 | 3 | Y |
| 209 | 22 | F | Y | -13 | 148 | 4 | 6 | 3 | 5 | 1 | 2 | Y |
| 210 | 22 | M | Y | 53 | 226 | 4 | 5 | 3 | 5 | 1 | 1 | Y |
| 211 | 23 | M | Y | -29 | 162 | 4 | 5 | 4 | 3 | 5 | 1 | Y |
| 212 | 22 | M | N | -61 | 66 | 4 | 4 | 4 | 2 | 6 | 6 | Y |
| 301 | 21 | F | Y | 31 | 50 | 1 | 5 | 3 | 4 | 1 | 1 | N |
| 302 | 22 | F | N | 32 | 28 | 4 | 3 | 2 | 6 | 2 | 1 | Y |
| 303 | 23 | M | Y | 42 | 38 | 2 | 4 | 4 | 4 | 3 | 1 | Y |
| 304 | 24 | M | Y | -10 | 46 | 2 | 3 | 4 | 5 | 5 | 5 | Y |
| 305 | 22 | M | Y | 63 | 68 | 4 | 4 | 4 | 5 | 1 | 5 | N |
| 306 | 23 | M | Y | -43 | 68 | 4 | 3 | 4 | 4 | 1 | 1 | Y |
| 307 | 25 | F | Y | 57 | 72 | 4 | 4 | 3 | 4 | 1 | 1 | Y |
| 308 | 22 | M | Y | -29 | -26 | 3 | 4 | 4 | 3 | 4 | 3 | N |
| 309 | 25 | M | N | -1 | 6 | 4 | 3 | 3 | 5 | 3 | 4 | Y |
| 310 | 20 | M | N | -47 | 0 | 2 | 3 | 3 | 4 | 1 | 3 | Y |
| 311 | 23 | M | N | -84 | 14 | 3 | 3 | 3 | 5 | 2 | 1 | Y |
| 312 | 20 | M | Y | 10 | 84 | 4 | 2 | 3 | 2 | 2 | 2 | Y |

| SUB | AGE | SEX | USE | XSC | TOTSC | Q2 | Q3 | Q8 | Q9 | Q10 | Q11 | COM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 401 | 31 | M | Y | 35 | 198 | 4 | 2 | 3 | 6 | 2 | 1 | Y |
| 402 | 22 | M | Y | -31 | 160 | 4 | 4 | 3 | 7 | 1 | 1 | Y |
| 403 | 22 | F | Y | 39 | 218 | 4 | 5 | 4 | 5 | 2 | 1 | Y |
| 404 | 20 | F | N | 35 | 78 | 4 | 2 | 3 | 5 | 2 | 2 | Y |
| 405 | 21 | M | Y | -74 | 118 | 4 | 4 | 4 | 6 | 5 | 3 | N |
| 406 | 26 | F | Y | -22 | 122 | 4 | 4 | 3 | 5 | 4 | 2 | Y |
| 407 | 23 | M | Y | 30 | 162 | 4 | 2 | 4 | 4 | 2 | 2 | N |
| 408 | 21 | F | Y | 31 | 174 | 3 | 2 | 4 | 3 | 3 | 3 | Y |
| 409 | 21 | F | Y | 4 | 192 | 4 | 2 | 4 | 4 | 4 | 3 | Y |
| 410 | 25 | M | Y | -11 | 138 | 4 | 5 | 0 | 0 | 0 | 0 | N |
| 411 | 20 | F | N | 6 | 162 | 4 | 2 | 2 | 6 | 1 | 1 | Y |
| 412 | 20 | M | N | 45 | 206 | 4 | 4 | 4 | 2 | 3 | 2 | N |
| 414 | 21 | F | Y | 8 | 148 | 5 | 3 | 4 | 7 | 1 | 2 | N |
| 501 | 20 | M | Y | 8 | 68 | 4 | 4 | 3 | 2 | 2 | 1 | N |
| 502 | 20 | F | Y | -55 | -22 | 4 | 2 | 5 | 1 | 5 | 4 | Y |
| 503 | 20 | F | Y | -14 | 2 | 4 | 4 | 4 | 7 | 4 | 2 | Y |
| 504 | 23 | F | Y | 51 | 38 | 3 | 2 | 3 | 3 | 1 | 1 | Y |
| 505 | 21 | M | Y | 23 | 88 | 5 | 3 | 3 | 5 | 3 | 1 | Y |
| 506 | 21 | M | N | 57 | 54 | 4 | 1 | 4 | 1 | 3 | 1 | Y |
| 507 | 20 | F | Y | -21 | -20 | 2 | 2 | 3 | 6 | 3 | 1 | Y |
| 508 | 21 | F | N | 35 | -46 | 4 | 2 | 2 | 4 | 2 | 1 | Y |
| 509 | 23 | M | Y | 27 | 16 | 2 | 2 | 4 | 3 | 2 | 2 | N |
| 510 | 22 | F | Y | -23 | 50 | 4 | 4 | 4 | 5 | 3 | 1 | N |
| 511 | 23 | M | Y | -18 | 144 | 3 | 4 | 4 | 3 | 3 | 3 | Y |
| 512 | 37 | M | Y | -28 | -12 | 2 | 4 | 4 | 7 | 4 | 4 | Y |

| SUB | AGE | SEX | USE | XSC | TOTSC | Q2 | Q3 | Q8 | Q9 | Q10 | Q11 | COM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 602 | 20 | M | N | 38 | 276 | 3 | 3 | 3 | 7 | 1 | 1 | Y |
| 603 | 22 | M | Y | 43 | 162 | 4 | 2 | 3 | 1 | 2 | 2 | Y |
| 604 | 27 | M | N | -59 | 80 | 4 | 2 | 2 | 5 | 1 | 2 | Y |
| 605 | 21 | F | Y | 25 | 96 | 4 | 4 | 3 | 1 | 2 | 1 | N |
| 606 | 21 | F | Y | 13 | 70 | 4 | 3 | 3 | 3 | 3 | 2 | N |
| 607 | 23 | M | Y | 50 | 124 | 4 | 2 | 3 | 4 | 2 | 1 | N |
| 608 | 20 | M | N | -52 | 70 | 4 | 4 | 4 | 4 | 3 | 1 | N |
| 609 | 33 | M | Y | 17 | 150 | 3 | 4 | 4 | 4 | 4 | 6 | N |
| 610 | 26 | F | N | -6 | 90 | 4 | 3 | 3 | 5 | 2 | 2 | Y |
| 611 | 20 | F | Y | -12 | 206 | 7 | 4 | 5 | 2 | 6 | 3 | Y |
| 612 | 19 | F | Y | 30 | 190 | 4 | 4 | 3 | 6 | 3 | 1 | Y |
| 614 | 20 | F | Y | 43 | 162 | 4 | 4 | 3 | 3 | 2 | 4 | Y |
| 701 | 21 | M | Y | -8 | -22 | 2 | 3 | 4 | 4 | 5 | 4 | Y |
| 702 | 21 | F | Y | 15 | 56 | 2 | 2 | 2 | 6 | 1 | 2 | Y |
| 703 | 20 | M | Y | -104 | 18 | 3 | 4 | 3 | 3 | 4 | 3 | Y |
| 704 | 27 | M | Y | -25 | -6 | 4 | 1 | 2 | 5 | 2 | 1 | Y |
| 705 | 34 | F | Y | -24 | -30 | 2 | 1 | 2 | 5 | 1 | 1 | N |
| 706 | 26 | F | Y | -10 | -30 | 3 | 4 | 3 | 4 | 2 | 1 | Y |
| 707 | 22 | M | Y | 29 | 38 | 2 | 2 | 3 | 3 | 2 | 2 | N |
| 708 | 21 | F | Y | 1 | 80 | 4 | 2 | 3 | 6 | 3 | 1 | Y |
| 709 | 37 | M | N | 36 | 158 | 3 | 2 | 3 | 6 | 2 | 1 | N |
| 710 | 31 | F | Y | -38 | -8 | 4 | 3 | 4 | 2 | 2 | 1 | N |
| 711 | 20 | M | Y | -1 | 56 | 3 | 4 | 3 | 5 | 5 | 6 | Y |
| 712 | 20 | M | Y | 12 | 44 | 1 | 3 | 2 | 5 | 2 | 1 | Y |
| 714 | 28 | M | N | 16 | 84 | 1 | 3 | 2 | 6 | 1 | 1 | Y |

| SUB | AGE | SEX | USE | XSC | TOTSC | Q2 | Q3 | Q8 | Q9 | Q10 | Q11 | COM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 801 | 24 | M | N | -24 | 110 | 4 | 2 | 4 | 3 | 3 | 1 | Y |
| 802 | 19 | F | N | -16 | 162 | 4 | 3 | 3 | 4 | 2 | 1 | N |
| 803 | 24 | F | N | 36 | 202 | 2 | 2 | 2 | 7 | 2 | 1 | N |
| 804 | 20 | F | Y | -19 | 166 | 5 | 3 | 4 | 4 | 6 | 3 | Y |
| 805 | 26 | M | Y | 61 | 180 | 5 | 3 | 4 | 4 | 2 | 1 | N |
| 806 | 23 | F | Y | -82 | 96 | 4 | 5 | 3 | 5 | 5 | 1 | N |
| 807 | 28 | M | Y | -9 | 94 | 4 | 4 | 6 | 3 | 6 | 4 | Y |
| 808 | 25 | M | Y | -4 | 116 | 6 | 5 | 4 | 1 | 4 | 3 | Y |
| 809 | 21 | M | N | 15 | 196 | 4 | 3 | 3 | 5 | 5 | 3 | Y |
| 810 | 21 | M | Y | -16 | 104 | 3 | 3 | 3 | 6 | 3 | 1 | Y |
| 811 | 21 | M | Y | 18 | 114 | 4 | 4 | 3 | 5 | 4 | 1 | N |
| 812 | 20 | F | Y | 20 | 192 | 4 | 4 | 6 | 3 | 5 | 3 | Y |
| 901 | 45 | F | N | 46 | 72 | 2 | 2 | 4 | 4 | 4 | 4 | N |
| 902 | 22 | F | Y | -55 | 18 | 3 | 3 | 2 | 7 | 0 | 1 | Y |
| 903 | 22 | M | Y | 44 | 10 | 2 | 1 | 2 | 7 | 1 | 1 | Y |
| 904 | 21 | M | Y | -54 | -14 | 4 | 2 | 4 | 2 | 4 | 2 | Y |
| 905 | 27 | F | Y | -12 | 30 | 2 | 3 | 2 | 2 | 1 | 1 | Y |
| 906 | 24 | M | N | -1 | 122 | 3 | 2 | 2 | 7 | 5 | 3 | N |
| 907 | 30 | M | Y | -4 | 94 | 2 | 2 | 2 | 6 | 2 | 1 | Y |
| 908 | 23 | F | Y | -61 | -14 | 4 | 3 | 3 | 4 | 6 | 3 | Y |
| 909 | 21 | M | Y | 6 | 120 | 2 | 4 | 3 | 1 | 3 | 3 | Y |
| 910 | 20 | F | Y | 2 | 54 | 3 | 4 | 3 | 4 | 3 | 2 | Y |
| 911 | 22 | M | Y | 46 | 62 | 3 | 3 | 0 | 0 | 0 | 0 | Y |
| 912 | 21 | F | Y | -115 | 10 | 4 | 4 | 4 | 1 | 7 | 4 | Y |

# REFERENCES


1. *American Heritage Dictionary of the English Language*, W. Morris, Editor, Houghton Mifflin, Boston, 1978.

2. Barnes, R.M., *Motion and Time Study: Design and Measurement of Work*, 8th Edition, John Wiley & Sons, 1980.

3. Bassok, M. and Holyoak, K.J., "Interdomain Transfer Between Isomorphic Topics in Algebra and Physics", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 15, No. 1, 1989, 153-166.

4. Bahrick, H.P., Noble, M.E. and Fitts, P.M., "Extra-task Performance As a Measure of Learning a Primary Task", *Journal of Experimental Psychology*, Vol. 48, 1954, 298-302.

5. Bailey, C.D., "Forgetting and the Learning Curve: A Laboratory Study", *Management Science*, Vol. 35, No. 3, March 1989, 340-352.

6. Baron, R.S., "Distraction-Conflict Theory: Progress and Problems", *Advances in Experimental Social Psychology*, Vol. 19, 1986, 1-40.

7. Blyton, P., Hassard, J., Hill, S. and Starkey, K., *Time, Work and Organization*, Routledge, London, 1989.

8. Boepple, E.A. and Kelly, L.A., "How to Measure Thinking", *Industrial Engineering*, July 1971, 8-16.

9. Boggs, D.H. and Simon J.R., "Differential Effect of Noise on Tasks Of Varying Complexity", *Journal of Applied Psychology*, Vol. 52, 1968, 148-153.

10. Boker, J.R., "Immediate and Delayed Retention Effects of Interspersing Questions in Written Instructional Passages", *Journal of Educational Psychology*, Vol. 66, 1974, 96-98.

11. Brainerd, C.J, Kingma, J. and Howe, M.L., "On the Development of Forgetting", *Child Development*, Oct 1985, 1103-1119.

12. Buck, J.R, Badarinath, N.B. and Kachitvichyanukul, V., "Cognitive Task Performance Time and Accuracy in

Supervisory Process Control", *IIE Transactions*, Vol. 20, No.2, June 1988, 122-131.

13. Campbell, D.J., "Task Complexity: A Review and Analysis", *Academy of Management Review*, Vol.13, No.1, 1988, 40-52.

14. Carlisle, B.,"Job Design Implications for Operations Managers" *International Journal of Operations and Production Management*, Vol. 3, No.3, 1983, 40-48.

15. Carlson, J.G. and Rowe, A.J., "How Much Forgetting Costs?", *Industrial Engineering*, Vol. 8, No. 9, September 1976.

16. Chase, R.B. and Aquilano, N.A., *Production and Operations Management, A Life Cycle Approach*, 5th Edition, Irwin, Boston, 1989.

17. Chi, M.T.H., Feltovich, P.J. and Glaser, R., "Categorization and Representation of Physics Problems by Experts and Novices", *Cognitive Science*, Vol. 5, 1980, 121-152.

18. Chiles, W.D., "Complex Performance: The Development of Research Criteria Applicable in the Real World", *Measurement of Man at Work*, W.T. Singleton editor, Taylor & Francis Ltd., London, 1971, 159-164.

19. Christiaansen, R.E.,"Prose Memory: Forgetting Rates for Memory Codes", *Journal of Experimental Psychology: Human Learning and Memory*, Vol. 6, No. 3, 1980, 611-619.

20. Conn, H.P., "Improving Use of Discretionary Time Raises Productivity of Knowledge Workers in Offices", *Industrial Engineering*, July 1980, 70-77.

21. Dahms, A.R.,"Time Management and the Knowledge Worker", *Industrial Engineering*, March-April 1988, 26-28.

22. Davidson, J.P., "Improving the Quality of Interruptions", *Journal of Accountancy*, Oct. 1986, 176-179.

23. Davis, G.B. and Olson, M.H., *Management Information Systems, Conceptual Foundations, Structure and Development*, Second Edition, McGraw-Hill, New York, 1985.

24. Douglas, M.E.,"Do You Have to Suffer from All Those Interruptions?", *Management Solutions*, Vol. 32, No. 7, July 1987, 40-43.

25. Ekstrand, B.R., "Effects of Sleep on Memory", *Journal of Experimental Psychology*, Vol. 75, 1967, 64-72.

26. Eschenbrenner, A.J., "Effects of Intermittent Noise on the Performance of a Complex Psychomotor Task", *Human Factors*, Vol. 13, 1971, 59-63.

27. Gibbons, J.D., *Nonparametric Methods for Quantitative Analysis*, Holt Rinehart & Winston, 1976.

28. Gibbons, J.D., *Nonparametric Statistical Inference*, McGraw Hill, 1971.

29. Globerson, S. and Levin, N. "Incorporating Forgetting into Learning Curves", *International Journal of Operations and Production Management*, Vol. 7, No. 4, 1987, 80-92.

30. Gregerman, I.B., *Knowledge Worker Productivity*, AMACOM, New York, 1981.

31. Groff,B.D., Baron, R.S. and Moore, D.L., "Distraction, Attentional Conflict, and Drivelike Behavior", *Journal of Experimental and Social Psychology*, Vol. 19, 1983, 359-380.

32. Harkins, S.G. and Jackson, J.M., "The Role of Evaluation in Eliminating Social Loafing", *Personality and Social Psychology Bulletin*, Vol 11., No.4, December 1985, 457-465.

33. Harkins, S.G. and Petty, R.E., "Effects of Task Difficulty and Task Uniqueness on Social Loafing", *Journal of Personality and Social Psychology*, Vol. 43, No.6, 1214-1229.

34. Hartley, L.R., and Adams, R.G., "The Effect of Noise on the Stroop Test", *Journal of Experimental Psychology*, Vol. 102, 1974, 62-66.

35. Harris, M.F. and Vining, G.W., "The IE's Future Role in Improving Knowledge Worker Productivity", *Industrial Engineering*, July 1987, 28-32, 76.

36. Helton, B.R., "The 'Best Work' Method of Knowledge Worker Assessment", *Industrial Management*, Sept-Oct 1988, 19-21.

37. Houston, B.K. and Jones, T.M., "Distraction and Stroop Color — Word Performance", *Journal of Experimental Psychology*, Vol. 74, 1967, 54-56.

38. Huber, O., "The Influence of Some Task Variables on Cognitive Operations in an Information Processing Decision Model", *Acta Psycholgica*, Vol 45, 1980, 187-196.

39. Hulse, S.H., Deese, J., and Egeth, H., *The Psychology of Learning*, McGraw Hill, New York, 1975.

40. Kahneman, D., *Attention and Effort*, Prentice Hall, New Jersey, 1973.

41. Kelley, R.E., *The Gold Collar Worker*, Addison-Wesley, 1985.

42. Kintsch, W. and Greeno, J.G., "Understanding and Solving Word Arithmetic Problems", *Psychological Review*, Vol. 92, 1985, 109-129.

43. Martens, R. and Landers, D.M., "Evaluation Potential as a Determinant of Coaction Effects", *Journal of Experimental and Social Psychology*, Vol. 8, 1972, 347-359.

44. Mayer, R.E., *Thinking, Problem Solving, Cognition*, W.H. Freeman & Co., New York, 1983.

45. McBain, W.N., "Noise, the 'Arousal Hypothesis', and Monotonous Work", *Journal of Applied Psychology*, Vol. 45, 1961, 390-397.

46. McCormick, E.J., *Human Factors in Engineering and Design*, 4th Edition, McGraw Hill, New York, 1976.

47. Meister, D., *Behavioral Analysis and Measurement Methods*, Wiley, New York, 1985.

48. Miller, G.A., "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capability for Processing Information", *The Psychological Review*, Vol. 63, No. 2, March 1956, 81-97.

49. Mintzberg, H., *The Nature of Managerial Work*, Harper & Row, New York, 1973.

50. Morrison, D.F., *Applied Linear Statistical Methods*, Prentice Hall, New Jersey, 1983.

51. Moskal, B.S.,"No Interruptions, Please!", *Industry Week*, Vol. 231, No.3, Oct 1986, 61-64.

52. Neter, J., Wasserman, W. and Kutner, M.H., *Applied Linear Statistical Models*, Second Edition, Irwin, Illinois, 1985.

53. Newell, A. and Simon, H.A., *Human Problem Solving*, Prentice-Hall, New Jersey, 1972.

54. Okogbaa, O.G. and Shell, R.L., "The Measurement of Knowledge Worker Fatigue", *IIE Transactions*, Dec 1986, 335-342.

55. O'Malley, J.J. and Poplawsky, A., "Noise Induced Arousal and Breadth of Attention", *Perceptual and Motor Skills*, Vol. 33, 1971, 887-890.

56. Oseland, M. and Kleiner, B.H., "Slamming the Door on Interruptions", *Management World*, Vol. 17, No.1, 1988, 37-38.

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

58. Payne, J.W., Bettman, J.R. and Johnson, E.J., "Adaptive Strategy Selection in Decision Making", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 14, No.3, 1988, 534-552.

59. Porat, M., *Information Economy: Definition and Measurement*, OT Special Publication 73-11(1), U.S. Department of Commerce, Office of Telecommunications, Washington D.C., May 1977.

60. Pounds, W.F., "The Process of Problem Finding", *Industrial Management Review*, Vol. 11, No. 1, Fall 1969, 1-19.

61. Reder, L.M., "Strategy Selection in Question Answering", *Cognitive Psychology*, Vol 19, 1987, 90-138.

62. Rolfe, J.M., "The Secondary Task as a Measure of Mental Load", *Measurement of Man at Work*, W.T. Singleton editor, Taylor & Francis Ltd., London, 1971, 135-148.

63. Sanders, G.S. and Baron, R.S., "The Motivating Effects of Distraction on Task Performance", *Journal of Personality and Social Psychology*, Vol. 32, 1975, 956-963.

64. Sanford, A., *Cognition and Cognitive Psychology*, Bath Press, Avon, 1985.

65. Schendel, J.D. and Hagman, J.D., "On Sustaining Procedural Skills Over a Prolonged Retention Interval",

*Journal of Applied Psychology*, Vol. 67, No. 5, 1982, 605-610.

66. Schendel, J.D., Shields, J.L. and Katz, M.S.,"Retention of Motor Skills: Review", Technical Paper 313, U.S. Army Research Institute For Behavioral and Social Sciences, Alexandria, Virginia, 1978.

67. Schroeder, R.G., Anderson, J.C. and Scudder, G.D., "Measurement of White Collar Productivity", *International Journal of Operations and Production Management*, Vol. 5, No.2, 1985, 25-34.

68. Shields, J.L., Goldberg, S.L. and Dressel, J.D., "Retention of Basic Soldiering Skills", Research Report 1225, U.S. Army Research Institute For Behavioral and Social Sciences, Alexandria, Virginia, 1979.

69. Slamecka, N.J., "On Comparing Rates of Forgetting: Comment on Loftus", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 11, No. 4, 1985, 812-816.

70. Slamecka, N.J. and McElree, B., "Normal Forgetting of Verbal Lists as a Function of Their Degree of Learning", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 9, No. 3, 1983, 384-387.

71. Tansik, D.A. and Smith, W.L., "Dimensions of Job Scripting in Services Organizations", Working Paper, Karl Eller Graduate School of Management, University of Arizona, March 1990.

72. Tatsuoka, M.M., *Multivariate Analysis: Techniques for Educational and Psychological Research*, 1971, John Wiley and Sons, New York, 1971.

73. Towill, D.R., "The Use of Learning Curve Models for Prediction of Batch Performance Production", *International Journal of Operations and Production Management*, Vol. 5, No.2, 1985, 13-24.

74. Trumbo, D., Noble M. and Swink, J., "Secondary Task Interference in the Performance of Tracking Tests", *Journal of Experimental Psychology*, Vol. 73, 1967, 232-240.

75. Voss, C.A.,"The Service Dispatcher/Receptionist Role", *International Journal of Operations and Production Management*, Vol. 3, No.3, 1983, 35-39.

76. Wallas, G., *The Art of Thought*, Harcourt Brace Jovanovich, 1926.

77. Weick, K.E., "Cosmos vs Chaos: Sense and Nonsense in Electronic Contexts", *Organizational Dynamics*, 14, 1985, 50-64.

78. Whitmore, D.A., *Work Measurement (Second Edition)*, Institute of Management Services, London, 1987.

79. Wood, R.E., "Task Complexity: Definition of the Construct", *Organizational Behavior and Human Decision Processes*, Vol. 37, 1986, 60-82.

80. Woodhead, M.M., "The Effects of Bursts of Noise on an Arithmetic Task", *American Journal of Psychology*, Vol. 77, 1965, 627-633.

81. Woods, D.D. and Roth, E.M., "Cognitive Engineering: Human Problem Solving with Tools", *Human Factors*, Vol. 30, No.4, 1988, 415-430.

82. Yelle, L.E.,"The Learning Curve: Historical Review and Comprehensive Survey", *Decision Sciences*, April 1979, 302-328.

83. Zajonc, R.B., "Social Facilitation", *Science*, Vol. 149, 1965, 269-274.

84. Zuercher, J.D., "The Effects of Extraneous Stimulation on Vigilance", *Human Factors*, Vol. 7, 1965, 101-105.