PREDICTING SHOULDER FATIGUE FOR LONG DURATIONS USING PSYCHOPHYSICAL MEASURES OBTAINED FROM SHORT TRIALS

By

DEEPTI SOOD

Submitted to the Faculty of

Virginia Polytechnic Institute and State University in

Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE IN

INDUSTRIAL AND SYSTEMS ENGINEERING

Dr. Maury A Nussbaum

Dr. Thurmon Lockhart

Dr. Kari L. Babski-Reeves

May 11, 2004

Blacksburg, Virginia

Keywords: Fatigue, Perceived Discomfort, Fatigue Prediction, Endurance prediction, Personality type, Intermittent work, Overhead work

---

Predicting Shoulder Fatigue for Long Durations using Psychophysical Measures Obtained from Short Trials

Deepti Sood

ABSTRACT

Localized muscular loads have in many cases replaced whole body loads in the current mechanized industry. In highly automated automobile industries, the prevalence of upper extremity musculoskeletal disorders is a matter of continuing concern. Overhead work has especially been noted for its association with shoulder related musculoskeletal disorders. Research aimed at determining causal relationships between overhead work and risk of injury has increasingly used localized muscle fatigue as an indirect or surrogate measure. In this study, localized muscle fatigue was used as a primary measure for studying the effects of workload level while performing overhead work. Subjective (ratings of perceived discomfort) measures of fatigue were collected and their predictive potential was investigated. Effect of personality type was also examined to account for any inter-individual differences in fatigue perception.

While researchers have studied specific task conditions in controlled environments, the specific relationship between various risk factors and underlying injury mechanisms is largely unknown. Two main problems faced by researchers are limited resources and the large scope of potential ergonomic analyses. This study attempted to circumvent some of these limitations by examining the time-course of fatigue and the predictive potential of subjective measures. The feasibility of using shorter experimental durations to make deductions for a 2-hour work period was explored. Reductions in experimental duration means decreased experimental time, expenses and resources. Thus, in turn, the researcher can utilize available resources to study more factors and a more general scenario. Specifically, subjective measures of shoulder fatigue were used to determine the possibility of reducing experimental duration for an intermittent overhead task.

A laboratory-simulated intermittent overhead task was designed based on observations made at an automotive assembly unit. For this study, two treatment conditions were tested consisting of

different combinations of two tool masses and two duty cycles. The choice of the treatment conditions was made to simulate different task difficulty levels of occupational tasks and their effects on shoulder fatigue. Each experiment was conducted for 2 hours (a common duration in industries with job rotation) for these selected treatment conditions. Subjective measures of fatigue were collected to assess shoulder fatigue and relative acceptability of the overhead work.

Any observed trends in the subjective fatigue measure were determined and tested using statistical and mathematical models to determine how best to represent their salient characteristics. Derived qualitative and quantitative measures were also used to estimate the maximal acceptable task durations using certain formalized assessment techniques. Results of this research suggest possible reductions in the experimental duration. Short (8 to 26 minute) trials were found to be sufficient to predict performance measures for 2 hours. Results also indicated a strong influence of task difficulty level on the predictive performance of subjective measures though personality type did not show very consistent trends. Various unique analysis techniques used to look at the psychophysical data may prove useful for further investigation into predictive verification. A generalized mathematical model, a type of approach, was also developed to represent changes in the psychophysical measures over time. This research can find both industrial and research applications where resources are constrained and using psychophysical measures is feasible. In the following report, details on this work are presented, including a description of the factors that inspired this study, an outline of the relevant literature, methodology, results and their implications.

ACKNOWLEDGEMENTS

Many people helped me through this academic process and on the way enriched it in their own special way. Among these, my sincerest regards go to my advisor Dr. Maury A Nussbaum. He provided me with opportunities to grow and learn both intellectually and personally. I am extremely lucky to have him as my mentor and guide. Thanks also go to Dr. Kari Babski-Reeves and Dr. Lockhart for their insightful suggestions and encouragement.

Thanks to the entire VT-HFES group for some really fun times and sweet memories. Special mention is needed for Dadi, Grace, Pak, Melinda, Ravi, Kris, Yassierli, and Miguel for their valuable friendship and help with my research. I will also like to thank all the administrative staff particularly Randy Waldron, Will Vest, and Lovedia Cole for all their help.

My earnest thanks to some really wonderful friends I made in Blacksburg and to my family. I owe thanks to my dear awe-inspiring Nani and late Nana for their love. My gratitude goes to a very strong and idealistic lady, my mother Asha for her unwavering support and belief in me. So many thanks go to my brother Gaurav who has been my confidant and a cerebral stimulant! My deepest love and respects are for my husband Rohit, who gave me unconditional love. Most of all, he gave me a sense of belonging and completeness I lacked.

In the end, I will like to mention the dear beautiful Blacksburg. It has all the trees, birds, cows, horses, streams, hills, spirally roads, dilapidated barn sheds, quaint mansions (sans VT buildings!), pastures, nice people, and vistas that I had always dreamt to be a part of!!

I feel truly grateful today to have so many things and people to thank for.

Table of Contents

TABLES AND FIGURES

List of Tables

List of Figures

1.0 INTRODUCTION

Musculoskeletal injuries of the shoulder are very prevalent in industry. Shoulder injuries have been associated with various work related factors. Overhead work or working with hands above head has been repeatedly associated with increased incidence of shoulder injuries. Existing research, though, has been unable to quantify these relationships or recommend guidelines for situations where overhead work cannot be avoided. The multitude of task factors that may be responsible for shoulder related musculoskeletal disorders makes it difficult to find cause-effect relationships between them. This predicament has affected almost all avenues of human factors research, as it is extremely difficult to encompass all the related factors. Also, every addition of a factor in effect translates into added experimental costs and time. Thus, there is an acute need to enhance experimental procedures and protocols so that studies can be more comprehensive and expansive in their approach. Previous research for an automotive manufacturer involved studying of shoulder fatigue while doing overhead work was ridden with similar difficulties (HAM Report, 2002). A unique approach undertaken to transcend some of those difficulties formed an inspiration for this study.

Motivation

The central idea for this research had been inspired by previous research for an automotive manufacturer. This research consisted of a pilot experiment and a main experiment, where the pilot experiment was used to determine the experimental protocols for the main research. The main research involved determination of recommended limits for the combinations of work height, duty cycle and tool mass for upper-extremity overhead work related to automotive assembly tasks (HAM Report, 2002; Sood et al., 2002). An intermittent overhead task was simulated in the laboratory to represent assembly-type work, and to capture features common to such work. Experimental design involved studying the effect of two task factors, tool mass and duty cycle, on shoulder fatigue. Three levels each of the two task factors were chosen, generating a total of 3^2 or 9 treatment combinations. The large number of treatments required a very large sample pool for counterbalancing the treatment presentation and for increasing the experimental power. Also, the prevalence of two-hour work shifts (worker worked at a subassembly for a maximum of 2 hours per day before being rotated to another job) in the automobile industry required the results to signify changes for that time period. Due to time constraints it would have been impossible to test a large participant pool for the entire two-hour duration on all the treatment conditions. A pilot experiment was conducted before the main research to circumvent this problem.

The pilot experiment was designed to explore the possibility of reducing the experimental duration for the main experiment. Also, an attempt was made to find a feasible method to extrapolate data obtained from experiments of short duration (i.e. < 2hours) to predict results for 2-hour work periods. Based on the results of the pilot experiment, the experiment treatment length was reduced to 30 minutes. The data collected in 30 minutes was later extrapolated to determine performance measures for two hours using linear regression method. Use of this approach raises questions about the validity of using short experimental durations in fatigue research, the soundness of extrapolation methods, and the reliability of the conclusions drawn. These questions formed the basis of the verification process addressed in this study.

2.0 BACKGROUND

Scope of the Chapter

Literature review helps broaden perspective by exposing researcher to the answered, continuing and unasked questions in the interested research area. This chapter will provide the reader with an overview of past research in the area of overhead work tasks and accentuate the relevance of the proposed study. This chapter begins with a general review of musculoskeletal risk factors and later branches into more specific areas relevant to this study. It is hoped that the need and importance for this study will be underscored with the help of cited research and examples.

2.1 Work Related Musculoskeletal Disorders

Musculoskeletal disorders (MSDs) are characterized by an abnormal condition of the muscle, tendon, tendon sheath, nerve, bursa, blood vessel, bone, joint, or ligament often accompanied by discomfort or pain, which can potentially result in impaired motor or sensory function. Work related musculoskeletal disorders (WMSDs) are caused when an aspect or attribute of work or the work environment leads to the development of MSDs (Kuorinka and Forcier, 1995).

As early as the 18th century, musculoskeletal disorders were related to occupational factors. In 1713, Bernardini Ramazzini in his treatise 'De Morbis Artificum Diatriba' or 'Diseases of Workers' (as translated by Wright, 1940) documented that musculoskeletal disorders were associated with workplace factors. But it was not until 1970s that epidemiological methods were used for examining occupational factors. Since then however, more than 6,000 scientific articles have been published addressing WMSDs (NIOSH, 1997).

WMSDs are prevalent in many industries and result in both monetary and human costs (Putz Anderson, 1988). The Occupational Safety and Health Administration (OSHA) noted that the WMSDs are 'the most prevalent, most expensive and most preventable' work-related injuries and illnesses in America (OSHA, 1999). WMSDs have increased from 20% of all occupational illnesses in 1981 to 65% in 2001. High costs are also associated with WMSDs, amounting to about $32 billion a year in medical costs, over $20 billion a year in workers compensation and more than $100 billion in total costs, including lost productivity (NIOSH, 1995). Attention is thus warranted towards development of preventive measures. The upper extremity is responsible for much of these costs, as it has been shown to account for more than 60% of all occupational illnesses (BLS, 2001).

2.2 Upper Extremity – Shoulder WMSDs

Many researchers have noted the high prevalence of upper extremity WMSDs (e.g. Chiang et al., 1993; McCormick et al., 1990). WMSDs of the upper extremity include tendonitis of the shoulder, elbow, wrist and hand, rotator cuff injuries, and carpal tunnel syndrome among others. Pathological evidence does indicate that mechanical strain may play an important role in the development of upper extremity MSDs (NIOSH, 1997), but uncertainty persists regarding factor/s leading to stress on the biomechanical structures.

In upper extremity research, much attention has been devoted to the study of shoulder. Work-related injuries to the shoulder are very common in the construction and manufacturing trades. The BLS reported 96,117 instances of shoulder related injuries and illnesses that resulted in days away from work (BLS, 2000). Olson (1987) found the prevalence of shoulder tendonitis to be as high as 30 to 40% in the construction and manufacturing trades.

The complex structure of the shoulder makes it a common site for many soft tissue injuries. The shoulder joint is a multipart biomechanical structure composed of three bones: the clavicle (collarbone), the scapula (shoulder blade), and the humerus (upper arm bone), which are held in place by the soft tissues: muscles, tendons, and ligaments. These soft tissues can get injured as a result of overuse related to job demands, which over time can gradually result in the development of WMSDs. Epidemiological research has associated specific soft tissue disorders of the shoulder region such as tendon-related disorders, prolonged shoulder muscle pain, nerve related disorders, neurovascular disorders, and syndrome disorders with various occupational risk factors (Sommerich et al., 1993).

2.3 Shoulder WMSDs - Risk factors

2.3.1 Extrinsic task factors

Factors including repetition, forceful exertions, deviated/extreme postures and duration have been associated with work related musculoskeletal disorders (Armstrong et al., 1986; Putz Anderson, 1988; Moore et al, 1991; Silverstein et al, 1986, 1987). Many studies have associated overhead work with an increased incidence of shoulder injuries, particularly work requiring arm elevations above shoulder height.

Björkstén et al. (2001), studying female blue-collar workers in metal and food industries, determined that working with hands and arms lifted and unsupported increased the risk of shoulder WMSDs. Bjelle et al. (1979), in a study on occupational health clinic patients, reported that 68.8% of patients with shoulder pain worked with their hands primarily at or above shoulder level. Rosecrance et al. (1996) found a 41% incidence rate of work-related shoulder pain in construction workers in the pipe trade, who often performed overhead work. In a study on orchard workers, the workers who repetitively reached overhead had higher incidence of shoulder stiffness and muscular tenderness than workers who did similar tasks at lower heights (Sakakibara et al., 1995).

Relationships between elevated arm postures and associated shoulder injuries have some physiological basis. Increased stresses on tendons, ligaments and capsular tissues have been observed with elevation of the arm even without a hand load (Engin, 1980). Flatlow et al. (1994) showed that working overhead, especially with shoulder elevation in the 60 –120^{o} range or at the end of range of motion, can lead to mechanical impingement of the bicipital and supraspinatus tendons in the subacromial region and be harmful to the shoulder girdle. Tasks requiring frequent reaches above shoulder height may lead to shoulder injuries and can create degenerative tendonitis (Bjelle et al., 1981). It is also hypothesized that if overhead work is done repeatedly, it can lead to tendonitis or tendon rupture (Lohr and Uthoff, 1990; Swinontkowski et al., 1990).

Real work situations, industrial or occupational, rarely have single risk factor acting in isolation. Instead, it is the compounding effect of work related risk factors that is responsible for increased risk and subsequent injuries. Overhead work, when done at high repetition rates and excessive loads, can substantially increase the risk of shoulder injury. Hand held weight, as little as 0.95kg, has been reported to induce rapid fatigue with repetitive arm movements (Wiker et al., 1989). Anderson (1984) noted that workers performing manual work with excessively and/or repeatedly loaded upper limbs often develop painful symptoms due to excessive fatigue in the muscles, which may lead to reduction in work capacity and increase in days away from work. A study on young males indicated rapid increases in shoulder fatigue with sustained elevated arm work, especially if supporting a load (Chaffin, 1973).

The etiological and pathological mechanisms for the shoulder WMSDs are unclear. Difficulties in studying the shoulder arise not only due to the complexity of the structure but also because most of the muscles are activated to some extent during any movement of shoulder or arm. Since there are often situations wherein overhead work cannot be avoided, it is of interest to determine if there are task factors that can be modified to minimize shoulder fatigue and associated musculoskeletal risks.

2.3.2 Intrinsic risk factors – Personality type

Personality type is a psychosocial risk factor, which may affect the job stress induced on and the stress management techniques employed by an individual. Type A individuals have been linked to extreme competitiveness, ambition, high performance standards, extreme aggressiveness, time urgency, restlessness, hurried motor and speech patterns. Sparacino (1979) showed that Type A people exhibit tense hyperactive movements. Also, Jenkins et al (1971) noted that Type A people are more alert, restless and prone to hurried motor movements. Personality type warrants attention due to its influence on an individual's technique of performing a task, which may lead to inter-individual differences in muscle recruitment patterns and activity level. However, it still remains an ignored topic among industrial ergonomics researchers.

In pioneering work by Glasscock et al. (1999), the effect of personality type on the neuromuscular control strategies and biomechanical loading of the upper extremity were investigated. The muscle coactivation patterns of seven major muscles activated during elbow flexion were studied for four elbow flexion tasks. Their measure of relative antagonism, the computed sum of scaled extensor torques (antagonists) to the sum of scaled flexor torques (agonists), was found to be significantly higher for the Type A participants (18%) as compared to Type B (10%) participants. Averaging across all conditions, Type A individuals showed significantly higher antagonist muscle activity than Type B individuals. This indicates an association between the personality type and the biomechanical stress imposed.

Salminen et al. (1991), in their study on metal industry workers, found that Type A people more often reported tenderness in shoulder and neck. Wickstrom et al. (1989) reported similar results: Type A people had more incidences of back pain than Type B people. In 1993, Bonger et al. also proposed that personality type and coping styles may influence worker response to external psychosocial risk factors and work-related stress and may further affect the subsequent development of musculoskeletal symptoms into chronic and disabling disorders. Malchaire et al. (2001) found associations between neck pain and Type A behavior: neuroticism, urgency of time, and extraversion. These results may indicate higher levels of muscle activity in Type A personality people as compared to their Type B counterparts, which may make them more susceptible to an early onset of fatigue and musculoskeletal symptoms. Higher levels of muscle activity may in turn indicate higher loads being placed on the biomechanical structures, which may increase an individual's risk to musculoskeletal injuries, stress symptoms, and mental stress. The observed difference in the muscular and motor behavior of Type A and Type B individuals highlights the importance of considering personality type in any psychophysical study.

2.4 Fatigue: a surrogate measure of biomechanical demands and injury risks

Various methods have been developed over the years to quantify and measure the exposure to risk factors. Among them, a variety of fatigue measures are quite extensively used. Fatigue is a well studied though relatively less understood phenomenon, its perplexity only increased by its multi-causal nature. Fatigue has been defined as a physiological and biomechanical process that results in discomfort and impaired performance. It is a time dependent process occurring when muscle contraction is maintained for a long time and is identifiable through externally observable changes in mechanical performance (DeLuca, 1984) and physiological processes (Basmajian, 1978). Fatigue is associated with reduced work output, decreased mental concentration and alertness, localized pain and tremors, increased load on circulatory, respiratory and neuromuscular function, increased lactate accumulation and increased body core temperature

(Astrand and Rodahl, 1986; Basmajian and De Luca, 1985). Researchers have used fatigue as surrogate measure of WMSDs, though its definitive role is still not understood.

Localized muscle fatigue (LMF) as a specific indicator of WMSD risks is increasingly gaining importance in contemporary studies. With the advent of automation, many modern industrial tasks involve specialized, low to moderate force, high-repetition work. This kind of work requires repetitive straining of a small muscle group in a localized area, leading to localized muscle fatigue. LMF is localized to the synergistic muscle groups involved in performing the contraction (De Luca, 1994). LMF is argued as a comprehensive measure incorporating the influences of biomechanical (moment arms, forces, external loads) and physiological (energy depletion, metabolic accumulation) exposures.

LMF has been shown to indicate the presence of risk factors such as repetition, physical load and prolonged postures (Zollinger, 1927; Baidya et al., 1988; Hagberg et al., 1995). In two separate studies, Carpenter et al (1998) and Skinner et al (1986) independently found decreases in proprioceptive sense with localized muscle fatigue. In addition, they found decreases in the threshold to detection of movement and decreased joint position sense with fatigue, both of which can lead to more mistakes and decreases in efficiency and accuracy.

Researchers have explored the LMF mechanism at both micro cellular levels and at macro physiological, biomechanical and psychosocial levels. Approaches for measuring LMF are based on these underlying mechanisms. LMF can be measured using both objective measures (e.g. changes in the muscle activity and recruitment pattern, decreased endurance, decreased maximal force output), observed through changes in myoelectric activity and subjective measures (e.g. increased discomfort and pain, decreased endurance), observed through subjective ratings of perceived exertion and discomfort. Objective measures help to quantify the exposures but may lack the flexibility and subjectivity of psychophysical measures, which enable the incorporation of inter/intra individual differences. Thus, where possible, it might be best to use both the methods. Complex interactions between the objective and subjective measures may be difficult to classify, but when applied together may in turn hold the answer to better understanding of LMF.

In the following section, relevant measurement tools are reviewed. These measurement tools have been widely used to study static isometric exertions. Relatively less support is available to justify the appropriateness of these tools to study complex dynamic intermittent sub-maximal tasks, which constitute most of modern industrial work. Researchers have used several methods, but limited evidence is available in the existing literature to provide logical guidance and justification for using one method over another.

2.5 Objective and Subjective measures of LMF

2.5.1 Electromyography – an objective measure

EMG has been used extensively in the past to assess LMF. Cobb and Forbes (1923) initially reported fatigue-induced changes in the EMG amplitude and frequency. Since then, EMG has been used widely used to measure LMF under a range of task conditions. Researchers have employed a variety of measures (e.g. changes in EMG signal amplitude; shifts in EMG spectral analysis observed through changes in the mean and median power frequency) to quantify changes in the EMG signal with fatigue. Among them, EMG spectral changes have been extensively used in the ergonomics field.

Chaffin (1973) observed shifts in the EMG power spectrum towards lower frequencies while studying the effects of elevated arm postures on fatigue in young males. Spectral changes with fatigue may indicate changes in the conduction velocity of the muscle fibers, increases in the average duration for which the motor units are active, and changes in the muscle recruitment patterns. Rapid increases in trapezius fatigue have also been observed, using EMG analysis, when the arm was flexed and held at 90^{o} (Hagberg, 1981). In a study of visual display units, the effect of shoulder posture on performance was investigated. Poorer performance was observed in postures requiring 30^{o} shoulder flexion compared to 0^{o} shoulder flexion. Also, EMG analysis indicated significantly greater fatigue at 30^{o} shoulder flexion compared to 0^{o} shoulder flexion (Straker et al., 1997). Changes in the EMG activity, magnitude, and spectral shifts have been associated with LMF. In overhead work (among welders), the supraspinatus muscle showed significant spectral changes in EMG, indicating that the muscle is under sustained strain (Herberts et al., 1976). In another study with welders working at a shipyard, high EMG activity in the deltoid, trapezius, and supraspinatus muscles was observed (Kadefors et al., 1976).

Most of the studies using EMG, however, have been conducted under isometric conditions. Under these conditions, changes in the power spectrum have indeed been shown to indicate muscle fatigue (Hagberg, 1981). Comparatively less evidence is available on the effectiveness of EMG to indicate fatigue in low effort, repetitive, dynamic tasks. These tasks by definition are comprised of non-isometric and non-isotonic contractions. Under these conditions, EMG has been shown to give inconsistent and even contradictory results (Hagberg, 1981). Gerdle et al. (1988) showed confounding of EMG signal due to changes in muscle length. Changing muscle force may also potentially confound the EMG measures (Shankar et al., 1989). In addition, surface EMG can best be used to study few specific prominent and easily accessible muscles. This can be a potential hindrance in studying a multi-muscle group like shoulder. Studying only a few shoulder muscles might not provide complete information on shoulder activity (Anderson and Galinsky, 1993). Thus, caution should be exercised while using EMG for studying shoulder fatigue during dynamic intermittent tasks.

2.5.2 Rating Scales – a subjective measure

Humans react to their perception of the surrounding environment (Borg, 1970). Both internal factors (e.g. stress on the muscles and joints, on cardio respiratory system and/or on central nervous systems) and external factors (e.g. high work pace, organizational policies, decision latitude, possibilities of making a career, personality, social support) may influence an individual's perception of fatigue. Subjective ratings of fatigue may help to combine these factors together into an overall exposure. Borg (1982) noted that perceived exertion is the best indicator of degree of physical strain. Anderson and Galinsky (1993) argued that presently psychophysical techniques that aid in quantifying worker's perceptual experiences are the only way to obtain a global assessment of local fatigue sensations.

Common measures of subjective exertion or discomfort include rating scales and self-report (Borg, 1980; Snook, 1978). Borg's 10-point scale is a commonly used rating scale in the human factors research. In a simulated assembly line task, Ulin (1990) found that the Borg scale and two other visual analogue scales produced similar trends, but participants preferred the Borg scale. The advantage of using rating scales is that they are easy to administer. However, the ratings are likely to be biased by inter- and intra-individual variability. The ratings are also influenced by factors such as the motivation level (Kilbom, 1990) and decision latitudes (Levi et al, 1986).

2.5.3 Psychophysical approach and relevant research

Researchers have extensively used psychophysical techniques to find the maximal acceptable weights or forces for various task conditions like lifting, pushing, pulling, lowering and carrying task. In psychophysical studies, it is common to use short observation periods of 0.5 to 4 hour durations to determine maximum acceptable weight of lift (MAWL) for an 8 -12 hour workday (Ciriello et al., 1990; Leg and Pateman, 1984). There are some difficulties associated with interpreting results obtained from such studies. Only a few studies have attempted to validate the reliability of results drawn from shorter duration experiments to longer time periods. Among these studies there is no general agreement or consensus concerning the validity of these estimates. This limits the applicability of results obtained from such experiments.

Ljungberg et al. (1982) found that participants were able to estimate the maximum acceptable weight they could lift for 8 hours, in a short duration of 5-10 minutes. Also, the case weight selected in 5-10 minutes was about the same as the one selected after 1 hour of lifting work, indicating that participants were able to estimate their limits in a short duration of time. In another study on psychophysical determination of acceptable weights and forces for 8 hours, no significant difference was found between weights and forces selected after 40 minutes from those selected after 4 hours (Ciriello et al., 1990). In this study, 18 varieties of lifting, lowering, pushing, pulling and carrying tasks for different task frequencies were investigated. Twelve females and ten males performed these tasks or a combination of these tasks for a 4-hour period. Results suggested that for frequencies up to 4.3 per minute, the maximum acceptable weights selected during the first 40 minutes could be maintained for more than 4 hours without physiological indications of fatigue.

Morrissey and Liou (1988), in a study on acceptable loads for load carriage, found no appreciable difference in the weight chosen in 20 minute adjustment period and the ones actually carried for 1 hour. Mital and Manivasagan (1983) used psychophysical methods to ascertain the amount of weight that a participant could carry comfortably for a specified carrying distance.

Participants then carried the selected acceptable weight for the same distance to confirm their prediction. Males could carry 0.5% more and females 0.5% less than the estimated weight. These results demonstrated that participants were able to accurately estimate acceptable carrying loads, within ±0.5 % error range, without actually carrying them. In a recent study, force limits for single-digit tasks were determined for a 2-hour period using a 25-minute adjustment period (Nussbaum and Johnson, 2002). Force data collected in the first 5 minutes of the adjustment period differed only minimally from that collected in the 25-minute period. This might indicate that a shorter 5-minute adjustment period might have given similar results but a verification study needs to be conducted to ascertain that. These studies support the use of short duration experiments, but caution should be exercised while making that deduction, as results from some other studies seem to indicate otherwise.

A few authors have suggested that estimations made in smaller experimental duration may not be representative of the performance over longer periods of time. Mital (1983) was the first to question the validity of using 20 - 25 minute adjustment periods to determine MAWL for 8-hour durations using psychophysical methods. In his study, the participants initially estimated MAWL for 8 and 12-hour shifts and later verified their estimations by actually performing the task for 12 hours. The results obtained indicated overestimation of MAWL values. In the 8-hour period, male participants could lift only 65% while female participants lifted 84% of their estimated weight. For 12-hour duration, males and females could lift 70% and 77% of the estimated weights, respectively.

Similar results were reported in another analogous experiment using a 4-minute adjustment period, albeit maximum acceptable lifting frequency for 2 hours was selected. When the experiment was conducted for 2 hours, it was found that the maximum acceptable frequency for 2 hours was less than 51% of the maximum acceptable frequency selected in 4 minutes (Mital and Asfour, 1983). Thus, though attractive to use short duration experiments, it is important to verify the results obtained from short-term experiments. The importance of verification is even more pronounced for the proposed study, as an intermittent task will be examined. With limited research available on intermittent tasks, and even less that verifies the use of short duration experiment, such investigation is necessitated.

2.6 Relevance to industry

In 2001, the BLS reported that incidence rate of WMSDs among motor vehicle manufacturers was as high as 21 per 100 full time workers. Based on data from 1989 OSHA 200 logs, motor vehicle manufacturers were among the top three industries with the highest reported rates of repetitive trauma disorders in the manufacturing sector (NIOSH, 1995). A healthy workforce is an important factor in attaining high morale, work satisfaction, low workforce turnover and better quality of production.

Numerous industries like construction, automotive assemblies, and several manufacturing industries have a number of tasks that require people to work overhead. The task situations there often require working overhead using tools and may involve lifting, rotating or exerting force. *(For example, in some automobile assembly tasks, the exhaust pipe installation requires overhead work with drills and pliers for extended periods of time).* Limited guidelines are available for such task situations involving greater fatigue and MSD risk during overhead work.

The results from this research will help to verify the procedures used in previous research for developing guidelines on overhead work. This investigation will also assist the research in this area by verifying the use of shorter duration experiments. If the use of shorter assessment durations are verified, then this would indicate the possibility of studying additional task factors within the same time and with the same limited resources. There is also limited evidence available in the existing literature to provide verification for using shorter duration experiments. This study will help to fill these gaps, first by exploring the possibility of reducing experimental durations, and second by verifying the use of results obtained from such experiments.

3.0 RATIONALE & OBJECTIVES

3.1 Rationale

Different aspects of the verification process using psychophysical methods made this study necessary, especially as an intermittent task was studied. A unique challenge was presented since, for dynamic tasks, psychophysical methods have not been extensively used. Existing literature documents psychophysical studies that have used shorter duration experiments for predicting load lifting capabilities for 8-hour workdays using 20 - 25 minute experiments (Snook and Irvine, 1967; Snook, 1978; Karwowski and Yates, 1984). Nussbaum and Johnson (2002) even suggest that 5-minute experimental durations are sufficient to determine maximum acceptable limits for hand intensive tasks. Among these, very few studies have been conducted for predictive verification of the results drawn from shorter duration experiments (Mital, 1983; Mital and Manivasagan, 1983).

In the studies on predictive verification, shorter trials have been used to determine acceptable loads for the longer task durations. Later, the estimated acceptable loads in shorter trials are tested for the longer duration to check their correctness. None of these studies verified the use of the shorter experiment trial duration itself when loads were given or fixed. In many human factors research experiments, the task conditions are predefined and cannot be selected from a range of possible values. In such cases, it is usually the effect of some defined task conditions on certain dependent measures that are of primary concern. For such situations the reduction of experimental duration can facilitate faster results and increase capability to study more task conditions. This study investigated the possibility of reducing this experimental duration.
In any experiment requiring human participants, the constraint on time and resources constrains the number of independent variables that can be studied. More independent factors to study may imply the necessity of using a large participant pool, to increase power and to obtain a balanced experimental design. Larger sample pools are also required to account for the inter-individual differences arising due to differences in participant's personality. Reduction in experimental duration can enable the experimenter to test more experimental conditions and test more

participants within the same time framework. Here resided the justification for this study, as shorter experimental durations can imply savings in terms of time, effort, money and resources.

To find the shortest trial durations that can be used to predict fatigue and performance over longer task durations, it is important to understand the inter-correspondences between the task factors and the induced biomechanical and perceptual loads. Understanding these could help to derive certain mathematical relationships, which might describe certain useful characteristics or trends. These relationships could be subsequently used to characterize the temporal progression of fatigue, which consequently might help reduce experimental duration in fatigue studies.

There are many unanswered questions associated with using shorter experimental durations or treatment lengths and also with using them as a basis for making prediction for a longer duration.

<u>Three pivotal questions to this issue being:</u>

- ☑ *What is the shortest trial duration for which the fatigue measures need to be collected to predict fatigue measures for 2-hour duration?*
- ☑ *What is the effect of task characteristics and participant's personality type on the selected shortest trial duration?*
- ☑ *What errors result from making deductions for 2 hours using predicted fatigue measures? Do the benefits justify the cost of error accrued from using shorter trial durations?*

These questions formed the underlying basis for this research.

3.2 Objectives

The focus of this study was on determining plausible reductions in the experimental duration of a fatiguing intermittent overhead task. Specifically, the research objectives of this study were:

i) Collect subjective measures of shoulder fatigue for an intermittent overhead task, and represent the relationships between fatigue measures and endurance time using a mathematical model. Using this mathematical model, determine the shortest trial duration for which an objective and subjective fatigue need to be measured, in order to characterize fatigue progression and to subsequently extrapolate that progression to longer trial durations.

ii) Compare the required shortest trial duration for different task conditions and for different personality types.

iii) Compare the predicted measures obtained from shorter duration experiments against the actual measures obtained in the 2-hour trial duration, and using this data and existing evidence make recommendation for future work.

The objectives are illustrated in figure 3.1 for better comprehensibility of the process. The fatigue measures were obtained for 2 hours, and later using portions of the available 2 hours of data temporal relationships were derived. The predicted fatigue measure for two hours (Y_2) will then be compared to those actually obtained after 2-hour duration (Y_1).

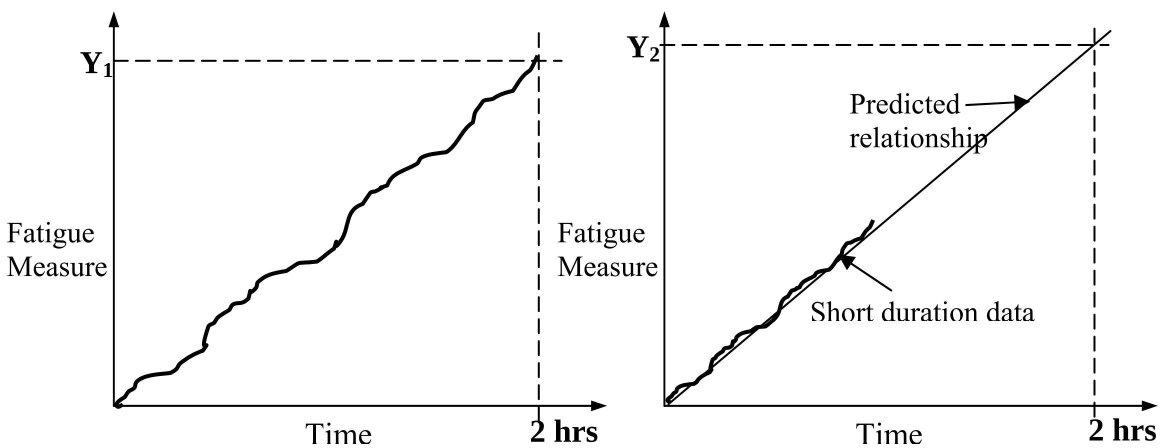


Figure 3.1 Fatigue measures obtained in shorter duration used to predict results for 2 hours.

3.3 Research Hypotheses

i) The consistent relationships between subjective fatigue measures and time can be described using a mathematical model. This mathematical model can be used to predict fatigue measures and endurance times for a 2-hour period using data collected in shorter trial durations of $\leq$ 30 minutes.

ii) The required shortest trial duration will differ for different task conditions and for participants with different personality types.

iii) Predicted results will correspond to those actually obtained over a 2 hour trial and the recommendations based on this and existing studies will favor the use of shorter duration trials.

4.0 METHOD

Research Plan

A laboratory study was planned to answer the research questions proposed earlier. For this, a laboratory task was designed to simulate common overhead assembly tasks in the automotive industry. Participants were required to complete an overhead-tapping task using a modified hand tool. Tasks involved tapping of fixed targets overhead at a fixed pace. The overhead task was performed intermittently, i.e. the entire work duration was interspersed with rest breaks. Two task conditions were imitated by using different combinations of duty cycle (work-rest ratio) and drill tool mass. Duty cycle denotes the percentage of cycle time in which the overhead task was being performed. For example, if cycle time is 60 seconds, then a 50% duty cycle indicates that the work is being performed for 50% of cycle time (30 seconds) and rest is given in the remaining 50% of cycle time (30 seconds). Ratings of perceived discomfort (RPD) using the Borg-10 point scale were obtained during the course of the experiment. Electromyography data was not used for this study as from the pilot study data no significant trends were observed to add to the proposed research objectives.

Sixteen participants were selected for completing the experimental procedures (see section 4.2 for sample size estimation). Every participant performed the experimental task for 2 hours on each of the treatment conditions. RPD data collected was later analyzed using formal statistical procedures. Details on the independent variables (experimenter defined treatment conditions and personality type), dependent variable (subjective measure of shoulder fatigue), experimental design, procedures and data analysis are provided.

4.1 Experimental Design

A mixed factor experimental design was used for this study to determine the effect of treatment conditions (combination of tool mass and task duty cycle) and personality type (Type A and Type B) on the predictability of subjective fatigue measure. Treatment condition was studied as a within subject factor and personality type was studied as a between subject factor.

The linear model expressing the main and interaction effects of these factors is defined as:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_{k(i)} + \alpha\beta_{ij} + \beta\gamma_{jk(i)} + \varepsilon_{l(ijk)}$$

Where Y_{ijkl} = one of the observations in any one of the treatment groups

μ = grand mean of the treatment populations

α_i = main effect of the personality type

β_j = main effect of the task treatment condition

$\gamma_{k(i)}$ = participant effect with nested personality type

$\alpha\beta_{ij}$ = interaction effect associated with the personality type and treatment condition

$\beta\gamma_{jk(i)}$ = interaction effect associated with the treatment condition and participant

$\varepsilon_{l(ijk)}$ = experimental error

4.1.1 Independent Variables

Treatment condition and personality type were the independent variables. A personality type measure (PM) was used to categorize participants into two groups, Type A and Type B, based on their responses on a checklist (Appendix A). This checklist was developed using descriptions of Type A people by Friedman and Ulmer (1984), Mathews et al. (1982) and Musante et al. (1983). Each question on the personality checklist and all the 5 anchors of every question had equal value (Figure 4.1). A score of '1' point was assigned to every anchor and consequently each question had a maximum of '5'point value. The participant's total score on the checklist was used to determine their personality group. On the basis of total scores obtained on the checklist, participants were said to exhibit more Type A behavior if their scores were above 62.5 and more Type B behavior if the scores were below 62.5. This criterion was selected as opposed to the one-third criterion used in many other psychology studies to reduce the experimental difficulties associated with finding participants that exhibit strong Type A and Type B behaviors.

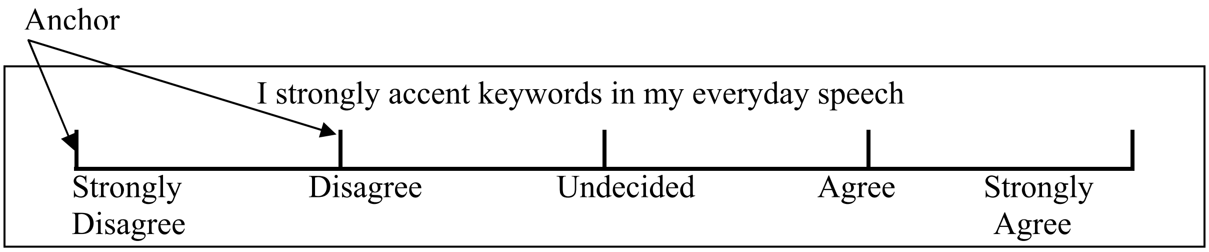


Figure 4.1 Sample personality question from personality checklist in Appendix A. Each question had five anchors: Strongly Disagree, Disagree, Undecided, Agree and Strongly Agree.

Two levels of treatment condition, T1 (1.25 kg tool mass and 50% duty cycle) and T2 (2.0 kg tool mass and 67% duty cycle), were selected based on task analyses of several automobile industrial tasks requiring overhead work (HAM Report, 2002). Also, previous research indicated that these treatments might have substantial effects on shoulder fatigue. Two commonly available drill tools were used, differing only in their masses (1.25 kg and 2.0 kg), into which a wooden bit had been inserted. The cycle time was chosen as 54 seconds, based on task analyses.

The selected treatments differed in difficulty levels with T1 being moderately difficult and T2 a very difficult task condition. The difficulty levels of the treatment condition were predicted based on the results from previous research discussed earlier. Previous results provided estimates of the percentage of participants who could complete the treatment condition for 2 hours without getting excessively fatigued or without reaching ratings of perceived discomfort (RPD) of more than 7 on Borg's scale (Table 4.1).

Table 4.1 The estimated percentage of participants who could perform the overhead task for two hours without reaching high levels of fatigue (RPD≥7)

Treatment Condition	**Pass Percentage**
T1 (50% of 54 seconds / 27 sec)	80.6%
T2 (67% of 54 seconds / 36 sec)	*

* Results indicated that the estimated percentage completed should be close to zero for this treatment condition.

These percentages were derived using the RPD data from 36 participants, collected every other minute, for a maximum of 30 minutes. The RPD data was extrapolated using linear regression to predict performance over a two-hour work period. The participants were assigned a 'Pass' or a 'Fail' based on their predicted performance over the two-hour period. A Pass was assigned if results indicated that the participant could perform the treatment condition for 2 hours without reaching high levels of fatigue (RPD ≥ 7) and a Fail otherwise. The Pass percentage was then derived for each treatment condition. Lower percentages indicate that fewer participants can perform that treatment condition for 2 hours without reaching high levels of fatigue and vice versa.

Treatment condition T1 was selected as an indicator of the condition that most participants could most likely perform for 2 hours without substantial difficulty or shoulder fatigue. T2 was chosen as an indicator of an extreme condition or what might be a very difficult treatment condition to perform. These treatments helped characterize changes in the dependent fatigue measures as a function of treatment difficulty levels. This was especially important for the present study, where it was of interest to determine the effect of task difficulty level on the predictability of fatigue measures.

Both treatment conditions were performed at the same participant-specific overhead task height (Figure 4.2). To determine the task height, H, two anthropometric measures were taken for the dominant hand/arm. These measures were: 1) hand height when the upper arm was held horizontal and the elbow flexed at 90^o (H1), and 2) hand height with arm in full extension (maximum overhead reach) and with shoulders parallel to ground (H2). Hand height was measured to the center of the grip, with participants holding the drill tool used for the task (Figure 4.2). The task height, H, was taken as the sum of 'H1' and 40% of the difference between H1 and H2, (H1 + 0.40[H2-H1]).

This task height was selected based on results from previous research which indicated that working on elevations higher than this height should be avoided as it could lead to very rapid development of fatigue and rapid increases in error rates (Sood et al., 2002). These results were also equivocal regarding the demands placed at chosen task height H versus a lower height, H1, which forms the boundary of when work begins to be categorized as overhead, and did not indicate any substantial decrease in fatigue with choice of height lower than H. Thus, there was no motivation to choose a lower overhead task height. These relative effects were also consistent with the study of overhead work and fatigue in which no consistent increases in fatigue were noticed with different task heights (Nussbaum et al., 2001).

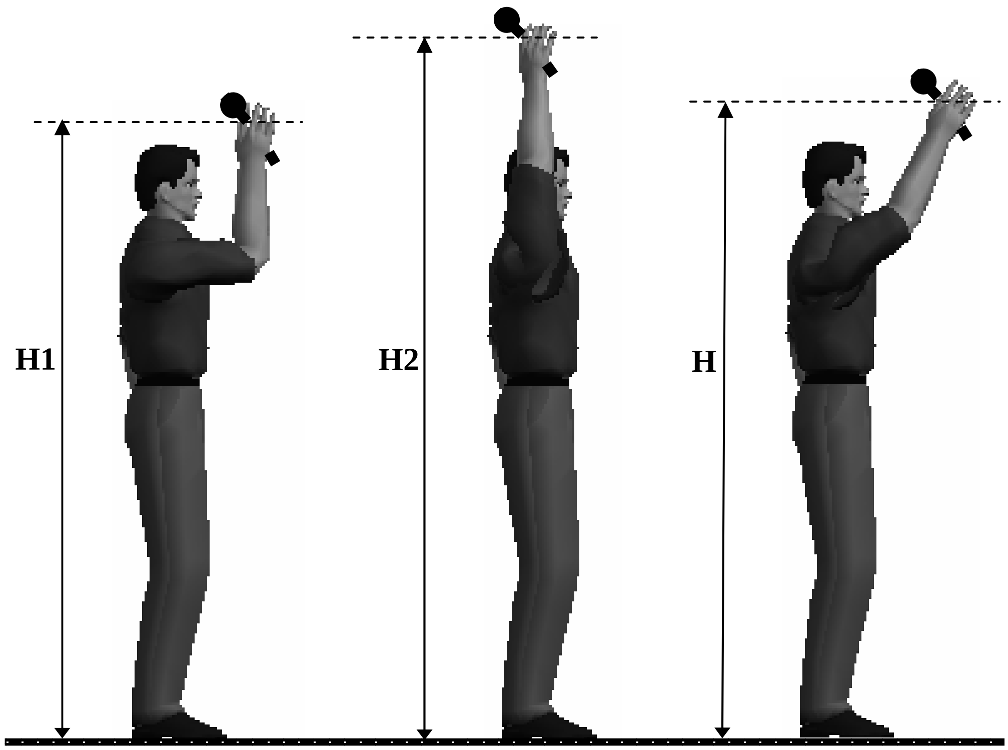


Figure 4.2 Participant's posture while measuring the three heights. All heights corresponded to the vertical location of the hand, measured to the center of the grip, wile holding the drill tool.

4.1.2 Dependent Variables

Ratings of perceived discomfort (RPD) for the dominant shoulder were examined. Ratings of perceived discomfort were collected using Borg's CR-10 scale (Borg, 1980; 1982). On this scale, ratings range from 0 to 10, with 0 being no noticeable level of pain or discomfort and 10 being extremely high levels of pain or discomfort.

Participants were given standardized instructions on rating shoulder discomfort using the scale (Appendix B). Practice on using the Borg's scale was also given at the beginning of the first experimental session. For the rating scale practice, the participant held a mass in their non-dominant hand, with their arm abducted to the side and parallel to the floor. The mass was selected so that total shoulder loads (moments) were approximately 40-50% of individual capacity (Appendix C). This mass was held in a static posture until the limit of endurance. During this exercise, participants provided their RPDs using the Borg 10-point scale at roughly 5-second intervals. Participants were not allowed to change their posture or release the load until they reported near exhaustion (RPD ≥ 9). Thus, during the practice, participants progressed through the entire Borg Scale, which required about 1-2 minutes. By following this procedure, it was believed that the participants could better conceptualize the RPD scale, and thereby provide more reliable values. During the experiment, RPD was recorded every other work cycle (every 1.8 minutes) on a pre-prepared data collection sheet.

4.2 Participants

Based on a power analysis (power = 0.80, $\omega^2 = 0.15$, $\alpha = 0.05$), a sample size of sixteen participants was used. Treatment presentation order was balanced by gender and personality type. Eight males and eight females were selected in order to obtain a complete repetition of the balanced design for each gender (Table 4.2). To balance the design by personality type, four of the eight participants in each gender were of personality type A and another four of type B.

Table 4.2 Treatment presentation order for 8 male participants. Same presentation order was used for 8 female participants.

P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**
T1	T2	T1	T2	T1	T2	T1	T2
T2	T1	T2	T1	T2	T1	T2	T1

* P denotes participants

All potential participants were screened using the personality checklist and only those exhibiting the required personality type were considered. A demographic questionnaire was used to limit the participation to individuals who either had recent manual work experience (within the past six months) or who did upper extremity exercises on a regular basis. Potential participants were also screened for any current or recent injuries or musculoskeletal disorders of the upper extremity that could have affected their performance in the experiment. This selection process ensured a participant pool representative of industrial workers in terms of familiarity with physical work and level of conditioning. This selection process was extremely important, as the psychophysical data obtained from participants constitutes the central part of this study. All participants were recruited from the local Blacksburg area. Participants were compensated at $10/hour and given a $10 bonus for completing the experiment.

4.3 Procedures

A laboratory simulated intermittent overhead task was used for this study. The task involved performing overhead work at two different treatment conditions on two different days. Both treatment conditions were performed at height H, determined as described earlier (Figure 4.1). Each treatment condition was performed for 2 hours or until the endurance limit was reached.

4.3.1 Experimental Task

A height-adjustable (up to 263cm) overhead platform was used for this study. Height adjustability was required, as the overhead work height was set for every participant specific to his or her anthropometry. A keyboard was attached to the bottom of this overhead platform (Figure 4.3). Participants stood underneath the platform and used the drill tool to tap four designated buttons on the keyboard. Two thin wires were strung over the keyboard, 3 cm away from the middle of the keyboard and 10 cm apart (Figure 4.3). These wires forced the participant to move the drill vertically in order to move between keys on the keyboard. Fixed pacing for key tapping was achieved by having participants follow a digital metronome set at 80-beats/ min. This pace was representative of the real task demands and pacing at the automotive assembly unit.

The simulated overhead task, including precise movements to targets and obstacle avoidance, was designed based on observations of several overhead tasks in the automotive industry and was considered representative of typical task demands in a variety of industries. For example, while constructing the roof of a house, trusses are fixed to the inner walls using drill-in bolts. Due to the confined areas around places where the bolts have to be fixed, the construction worker has to move his hand up and around barriers.

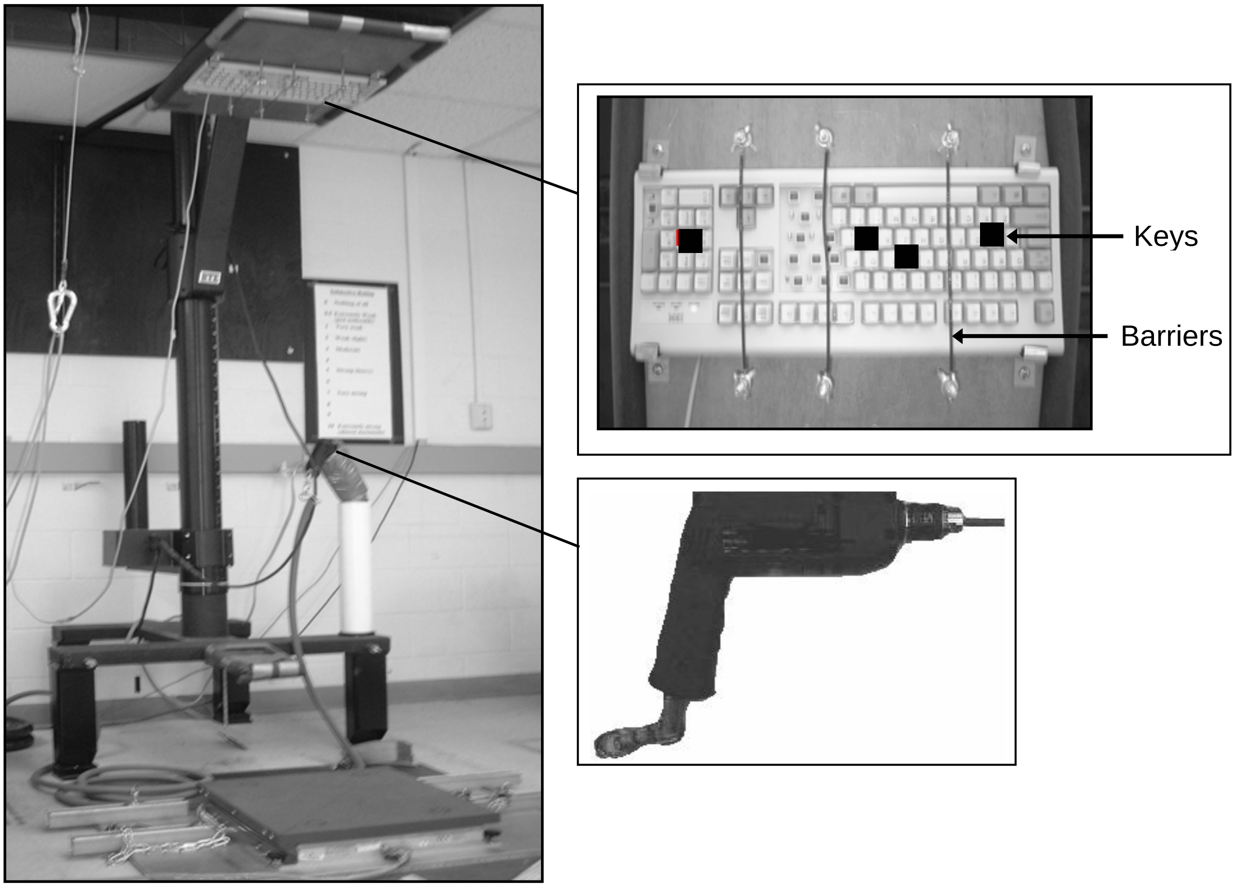


Figure 4.3 Experimental setup, keyboard with tapping keys marked in black and the drill tool

The experimental task was comprised of repeated cycles of work and rest periods with the duration of each determined by the duty cycle. The metronome paced key-tapping task was performed during the work period (Figure 4.4). A self-paced light manual task involving screwing and unscrewing of bolts and nuts was done during the rest period. This task was representative of the work performed by workers during rest period at an automotive assembly line. Prior to starting each work period, a warning beep warned the participant to prepare (take drill tool in hand) for the task. A second warning beep (start beep) indicated the start of the work period. A third warning beep (stop beep) indicated the end of the work period and start of the rest period. At this time the drill tool was placed in a tool holder and the participant moved to the right side of the workstation to do the light manual task. A warning beep was then provided to indicate the end of rest period. Each of these elements was repeated for the entire experimental duration.

Figure 4.4 Participant shown performing the overhead task using drill tool.

The weight of the drill tool used and the duration for which the participant worked at each of the task portions was governed by the duty cycle of the treatment condition. Instructions were given to emphasize that keeping pace with the metronome was more important than the accuracy of hitting the correct keys, but at the same time participants were asked to try their best to hit the correct keys in the correct order.

4.3.2 Experimental procedures

On arrival, participants were given a brief introduction to the experiment and a written informed consent was obtained, using procedures approved by the Virginia Tech Institutional Review Board (Appendix D). Participants then filled out a demographic questionnaire. Following this, a set of anthropometric measurements were taken: weight, stature, shoulder height (measured from floor to acromion), upper arm length (measured from the acromion to lateral epicondyle with arm held horizontally and in the frontal plane), and lower arm length (measured from the lateral epicondyle to the wrist crease with the arm held horizontally and in the frontal plane). These measures were taken to facilitate a description of the participant pool, and to allow for possible future normalization of the results. Except for weight and stature, all other anthropometric data were collected from the dominant arm. All the measurements were taken with participants wearing shoes, as the participant performed the simulated task with shoes. To guard against any intra-subject variability arising from differences in shoe heights, participants were asked to wear the same shoes for both experimental sessions. A videotape recording of various overhead assembly tasks was also shown to emphasize the relevance and importance of the study. A short task practice of 5 minutes was used to familiarize the participant with the experimental task and the treatment condition for that day. From the previous automotive study, this time was considered sufficient for getting the participant familiar with the procedures without adding any substantial shoulder fatigue. The rating scale practice followed to provide practice in using the Borg's scale. After the practice, participants began the experimental task described earlier. RPD data was collected intermittently throughout the experiment.

Experimental tasks were performed for either two hours or until the participant felt near exhaustion i.e. reports RPD ≥ 9. To ensure that RPD ≥ 9 had actually been reached, the trial was continued until such values were consecutively reported twice, or until an RPD of 10 was reported. This procedure was followed to account for the normal variability in RPD ratings. The data from participants who stopped before they reported substantial fatigue or before reaching RPD ≥ 9 was discarded as that data might be difficult to use for predictive verification.

Each participant was exposed to only one treatment condition per day. Thus, all participants were required to come for two experimental sessions (one for each treatment condition). The two sessions were separated by at least 48 hours, to minimize the chance of residual fatigue from the previous session, which might otherwise confound data obtained in a subsequent session. The described procedures remained same for both experimental sessions. To ensure against time of day bias, both the experimental sessions were conducted at approximately the same time of day. The laboratory conditions were controlled to standardize environment conditions for every participant. Also, to guard against variability arising due to experimenter bias, the instructions for the experiment were presented in the same manner in each session.

4.4 Data processing and analysis

A variety of derived subjective dependent measures were analyzed. Temporal changes in the magnitude of dependent measures were analyzed using analysis tools in Excel (Microsoft, Redmond, WA) and JMP (SAS, Cary, NC). RPD data was analyzed qualitatively and quantitatively to determine the minimum trial duration (MTD) sufficient to allow extrapolation and make predictions for a longer trial.

4.4.1 Qualitative Analysis

Categorization matrices were used for this analysis. To assist with the categorization process, RPD results were compiled in a tabular form, as in Table 4.3. For analysis purposes, a 'cut-off' value of RPD = 7 (R_7) was selected. On the 10-point Borg scale, a value of 7 corresponds to 'Very Strong' perceptions of discomfort or pain, and it was argued that this is the maximal level at which work should be performed in order to substantially reduce risks of injury or performance decrements.

Table 4.3 Sample RPD observations to be used for forming the Categorization matrix. The 1.8-minute time gradation is due to the data being collected only every 1.8 minutes.

Reported RPD Level	Participant Number	Ratings after							
		0.9 min	2.7 min	4.5 min	6.3 min	8.1 min	9.9 min	....	119.7 min
	1	0	0.5	1	1.2	1.5	2	....	6
	2	0.1	0.5	0.5	1	2	2.5	....	5
	3	0.5	1	1.5	2.1	2.5	2.7	....	6.6
	4	0.5	0.7	0.8	1.2	2	2	....	6
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	16	1	1.5	2.2	2	2.2	2.5	....	5.5

Table 4.4 Categorization matrix for determining critical trial duration for a given RPD level. The cells with 'X' denote correct categorization.

		Completed 2 hours without reaching R_C?	
		Yes	No
Reported 'R' by 'T' minutes?	Yes		X
	No	X	

Two examples of the categorization matrices are shown in Table 4.5. These examples have been constructed for two different combinations of R and T using Table 4.3. For complete analysis, 737 such matrices were constructed to study all possible combinations of 11 R-values (0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10) and 67 T-values (0.9, 2.7, 4.5,.…. 117.9, and 119.7).

<u>Table 4.5</u> Sample matrices for performing qualitative analysis of temporal changes in RPD. Matrix A is constructed for RPD = 1 and Time = 5 minutes, and the matrix B for RPD = 3 and Time = 10 minutes. In matrix A, only 2 of the 5 trials are correctly categorized. All the 5 trials are correctly categorized in matrix B.

Matrix A

		Completed 2 hours without reaching R_7?	
		Yes	No
Reported RPD = 1 by 5 minutes?	Yes	3	
	No	2	

Matrix B

		Completed 2 hours without reaching R_7?	
		Yes	No
Reported RPD = 3 by 10 minutes?	Yes		
	No	5	

4.4.2 Quantitative Analysis

Quantitative changes in the RPD data were studied using regression analysis. From the data, the relationships between RPD (R) and Time (T) appeared linear, for all participants for both T1 and T2 treatment conditions. Thus, only linear regression was used for analysis. Functional regression relationships were derived between R and T. Extrapolation using regression analysis was used to determine the predictability of the RPD data. For this, a small portion of the RPD data chosen from the beginning of the 2 hours session was used. Data from the chosen duration (e.g. first 5 minutes) was used to derive a functional relationship between R and T. To systematize the process a stepwise iteration method was employed by using sub-samples of increasing size (Table 4.6). Stepwise iteration method involved determination of a succession of elements (numbers or functions) by performing operations on one or more preceding elements according to a formula or rule. The functions/values thus derived were called cumulative functions/values and denoted by F_n (Table 4.6). In the present case, application of the iteration method involved conducting the analysis procedure or the mathematical operation 66 times to go from T= 0.9 to T= 119.7 in steps on 1.8 minutes (based on the time at which RPD was collected).

Table 4.6 A sample table is shown to demonstrate the stepwise iteration technique used in deriving the cumulative slopes (β_T) among others

Time (min)	**RPD**	**Stepwise Iterations**						
0.9	0.5							
2.7	1.5	$F_{2.7}$						
4.5	1.6		$F_{4.5}$					
6.3	1.7			$F_{6.3}$				
8.1	1.9				$F_{8.1}$			
:	:					....		
:	:							
T	R_t						F_T	
119.7	3.7							$F_{119.7}$

where T = 2.7, 4.5,…, 119.7

Using this methodology, regression equations were derived and then extrapolation was used to predict the RPD after 2 hours. Similar procedures were used to predict the time taken to reach R_7 and then compared with the actual duration. Also, it was maintained that RPD should be less than R_7 at the end of 2 hours to ensure that the participant can perform the task for 2 hours without excessive fatigue or without substantially increasing their risk of injury. Therefore, in this case the experimental trial was called a 'Pass' and 'Fail' otherwise. *Example:* RPD data obtained in the first 20 minutes of a 2 hour experimental session is used to obtain a linear relationship, R = 0.042 T – 0.129 (Figure 4.5). Using extrapolation, R at the end of 2 hours is found to be 4.9. Many such extrapolations are obtained for various chosen shorter durations (Table 4.7).

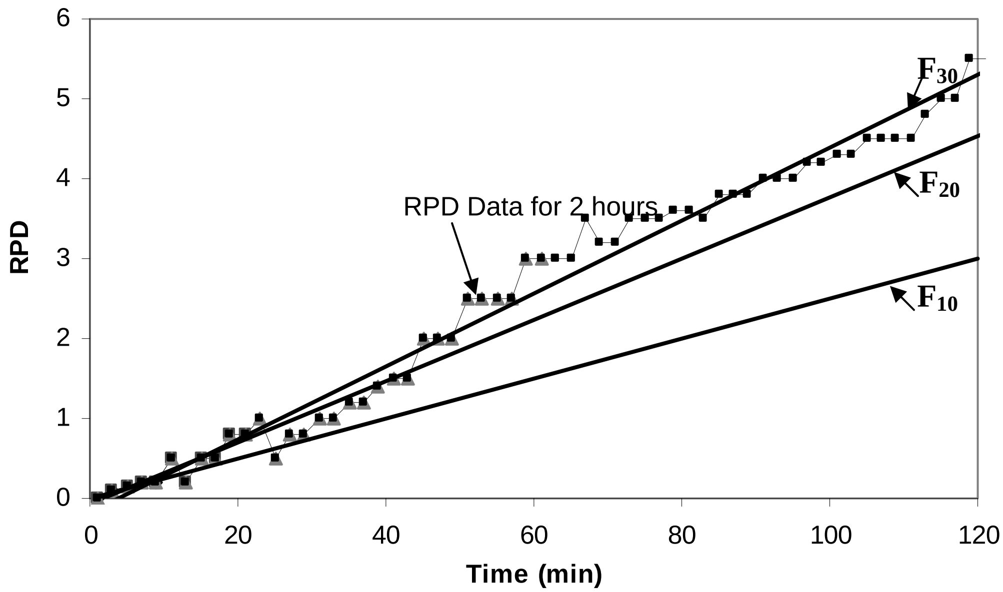


Figure 4.5 Sample RPD data obtained in 2 hours. Linear regression has been used to extrapolate RPD data obtained in first 10 min (F_{10}), 20 min (F_{20}) and 30 min (F_{30}) of the experiment to 2 hours or 120 minutes.

Table 4.7 Sample actual and predicted RPD data for certain durations

		RPD at t = 120min				RPD at t = 120m	
R_T	F_T	Actual	Predicted	R_T	F_T	Actual	Predicted
R_5	0	5	0	R_{35}	0.0338x - 0.0537	5	4.0
R_{10}	0.03t - 0.07	5	3.53	R_{40}	0.0357x - 0.0786	5	4.2
R_{15}	0.0357x - 0.0857	5	4.2	R_{45}	0.0385x - 0.12	5	4.5
R_{20}	0.0423t - 0.1286	5	4.9	R_{50}	0.0433x - 0.1997	5	5.0
R_{25}	0.0374x - 0.0857	5	4.4	R_{55}	0.0459x - 0.2454	5	5.3
R_{30}	0.0337x - 0.0513	5	4.0	R_{60}	0.0491x - 0.309	5	5.6

* R_T is the RPD after time 'T' and F (T) is the function of time.

Errors in the cumulative RPD slope and error in predicted RPD were also determined. Using the stepwise iteration method cumulative slopes (β_T) were found for every participant and treatment condition (Table 4.6). Errors in slope (e_T) were measured as the difference in the cumulative slopes (β_T) from the final 2-hour slope ($\beta_{2\text{-hour}}$) for each participant and for each treatment condition respectively.

4.5 Statistical Analysis

The selection of statistical analysis method depended on the applicability and appropriateness of the statistical model on which the test was based to the experimental data, and the type of measurement (nominal, ordinal, interval or ratio). Parametric analysis was used when the observations were independent, could be assumed to be derived from normally distributed populations with same variance and their effects were additive. Geary's test of normality, used to determine if the normal distribution could be a good approximation to the true unknown distribution, was used to determine the applicability of the parametric analysis. A nonparametric approach was undertaken when the parameters of the population from which the sample was drawn were unknown. For all statistical tests, the 0.05 level of probability was used as a criterion of statistical significance.

4.5.1 Qualitative Analysis of RPD

The RPD categorization matrix was used to determine the shortest trial length (MTD) for which maximum participants in a treatment were correctly categorized (Table 4.4). For this, 737 categorization matrices (11 R-values X 67 T-values) were constructed for T1 and T2 respectively. Later, using the results from both T1 and T2 matrices, matrices representing results of both T1 and T2 were derived. For these matrices, predictive validity measures were derived to help evaluate the ability of a categorization matrix to correctly categorize the performance in a trial. Predictive validity measures studied here included sensitivity, specificity, positive predictive value (PV), negative PV, and % correct and they are explained below.

Sensitivity equals the fraction of trials in which participants did not report R by T min from those in which they completed the 2-hour session without reaching R_7. If an R and T combination has sensitivity = 1, then for that combination all the participants who did not report R by T min could complete the 2-hour session without reaching R_7. Sensitivity of less than 1 implies that participants who completed the 2-hour session may/may not have reported R by T minutes.

Specificity equals the fraction of trials in which participants that reported R by T min from those in which they could not complete the 2-hour session without reaching R_7. If an R and T combination has specificity = 1, then for that combination all the participants who did report R by T min could not complete the 2-hour session without reaching R_7. Specificity of less than 1 implies that participants who could not complete the 2-hour session may/may not have reported R by T minutes.

Positive PV equals the fraction of trials in which participants that completed the 2-hour session from those trials in which they did not report R by T minutes. If an R and T combination has positive PV = 1, then for that combination all the participants who did not report R by T minutes completed the 2-hour session without reaching R_7. Positive PV of less than 1 implies that participants who did not report R by T min may/may not have completed the 2-hour session without reach R_7.

Negative PV equals the fraction of trials in which participants that could not complete the 2-hour session from those in which they did report R by T minutes. If an R and T combination has negative PV = 1, then for that combination all the participants who did report R by T minutes could not complete the 2-hour session without reaching R_7. Positive PV of less than 1 implies that participants who did report R by T min may/may not have completed the 2-hour session without reach R_7.

Percentage Correct equals the percentage of trials in which participants when did not report R by T minutes could complete the 2-hour session without reaching R_7 and who when reported R by T min could not complete the 2-hour session without reaching R_7 from the total trials. If an R and T combination has incorrect = 0, then for that combination all the participants who did not report R by T min could complete the 2-hour session without reaching R_7 and who did report R by T minutes could not complete the 2-hour session without reaching R_7. Incorrect of greater than 0% implies that participants who did/did not report R by T min may/may not have completed the 2-hour session without reach R_7.

For all the 737 matrices with combined results of T1 and T2, predictive validity measures were determined using methods specified in Table 4.8 (Knox and Moore, 2001; Lee et al., 1991).

Results of this analysis were used to find the combinations of R and T for which the maximum number of treatment trials were correctly categorized.

Table 4.8 Formulae for calculating sensitivity, specificity, positive PV, negative PV, and % correct

Parameter	Formula	Min Value	Max Value
Sensitivity	C / (A + C)	0	1
Specificity	B / (B + D)	0	1
Positive Predictive Value	C / (C + D)	0	1
Negative Predictive Value	B / (A + B)	0	1
Percentage Correct	[(B + C) / (A + B + C + D)]*100	0	100

Where A, B, C, and D are:

		Completed 2 hours without reaching R_C?	
		Yes	No
Reported ‘R’ by ‘T’ minutes?	Yes	A	B
	No	C	D

Example: For the categorization Matrix A in Table 4.5, sensitivity is = 0.4, which indicates that 40% of the participants who could complete the 2-hour task without reaching R_7 did not report RPD 1 by 5 min. A higher Sensitivity (~ 1) indicates good predictive capability or greater effectiveness of a combination of R and T and vice versa. Similar deductions can be made about the other predictive validity measures.

For the categorization matrices with high predictive validity, further analysis was conducted. To find MTD, first it needs to be proved that the frequency with which the participants were assigned to one of the 4 cells of the categorization matrix was not randomly distributed. Second, it needs to be shown that for the selected MTD the correctly categorized cells contained significantly larger number of trials than the incorrectly categorized cells. Kolmogorov-Smirnov goodness of fit test was used to ascertain the nonrandom assignment to the cells and Phi coefficient test was used to determine the presence of significantly large number of trials in the correctly categorized cells.

4.5.2 *Quantitative Analysis of RPD*

For the chosen MTD, no significant differences should be observed between the actual[1] and predicted[2] observations. Kruskal-Wallis test was used to determine if the actual and predicted RPD values at the end of 2 hours were equivalent. This test was also used to determine any differences between the actual and predicted times to reach R_7, RPD linear regression line slopes, and the number of Pass or Fail trials for both the treatments, T1 and T2. The reliability of the predicted results when compared to actual data was also determined. For this purpose, Intraclass correlation (ICC (2, 1)), Pearson product moment correlation (r), and standard error of measurement (SEM) were calculated.

Linear regression analysis was used to find the correspondence between the actual and predicted RPD derived measures. Studentized t-test was used to determine the level of significance (P-value), $\{P(\beta_T); t = 2.7, 4.5, \ldots, 119.7\}$, of the cumulative slopes (β_T) of the linear regression fit. Using the stepwise recursive methodology explained in Table 4.7, this test was conducted 66 times for each of the 16 participants and for T1 and T2 respectively. Sigma (slope), $s(\beta_T)$; $T = 2.7, 4.5, \ldots, 119.7\}$, and normalized sigma (slope), $ns(\beta_T)$: $s(\beta_T) / \beta_T$; $T = 2.7, 4.5, \ldots, 119.7\}$, were also obtained (Equation 1 and 2). The P-value, sigma (slope) and norm-sigma (slope) were tabulated as shown in Table 4.8 for every participant and for treatment T1 and T2 respectively. These were used for further analysis and development of the general model.

If linear regression fit is: $Y_T = \gamma_T + \beta_T X_T$ $\quad T = 0.9, 2.7, \ldots, 119.7$

Where, Y_T = value of the response variable RPD in the t^{th} iteration

X_T = value of the predictor variable Time in the t^{th} iteration

γ_T = cumulative intercept in the t^{th} iteration, β_T = cumulative slope in the t^{th} iteration

Then, Residual (e_i) or the error in the predicted value of $Y_i = Y_i - Y'_i = e_i$

Residual sum of squares or error sum of squares = SSE = $(Yi - Y'i)^2 = (e_i)^2$

Since, SSE has n-2 degrees of freedom associated with it,

[1] Actual values refer to the data collected in the 2-hour trial duration.
[2] Predicted values refer to the derived measures from shorter trials.

Hence, Mean square error (MSE) = SSE / (n - 2) = $(e_i)^2$ / (n-2)

If X' = $\sum X_i$ and β_T = cumulative slope

Then, sigma (slope) = $s(\beta_T)$ = MSE / $\sum(X_i - X')^2$.................................... Equation 1

norm-sigma (slope) = $ns(\beta_T)$ = $s(\beta_T)$ / β_T.................................... Equation 2

Table 4.9 Sample table showing P-value, sigma (slope) and norm-sigma (slope) for treatment T1. A similar table was constructed for treatment T2.

	Participant 1				Participant 16		
Time (min)	P-value	$s(\beta_T)$	$ns(\beta_T)$	....	P-value	$s(\beta_T)$	$ns(\beta_T)$
2.7	0.0995	0.96	0.173	....	0.0995	0.96	0.119
4.5	0.0182	0.76	0.007	....	0.0182	0.76	0.012
6.3	0.0034	0.66	0.001	....	0.0034	0.66	0.001
8.1	0.0034	0.61	0.002	....	0.0034	0.61	0.002
9.9	0.0007	0.57	0.0005	....	0.0007	0.57	0.0003
⋮	⋮	⋮	⋮	....	⋮	⋮	⋮
114.3	0.000003	0.19	0.00007	....	0.000003	0.19	0.00001
116.1	0.000001	0.18	0.00003	....	0.000001	0.18	0.00002
117.9	0.000000	0.17	0.00002	....	0.000000	0.17	0.00002

4.5.3 Effect of personality type (PM)

Absolute differences between the actual and predicted RPD regression line slope obtained from Type A participants were compared to those obtained from the Type B participants for both T1 and T2 treatment conditions. It was expected that the difference would be larger for Type A as compared to the Type B personality participants as Type A participants would tend to give lower RPD levels initially. Spearman's Rho test was used to determine the correlation between the PM score and the difference between the actual and predicted RPD regression line slopes, RPD onset times, actual and predicted RPD at 2 hours, and the time to reach RPD 7.

4.5.4 Effect of treatment condition

Wilcoxon signed ranks test was used to ascertain that the slope of RPD regression line, indicating the rate of RPD progression with time, was different for T1 and T2.

5.0 RESULTS

5.1 Participants

Sixteen (8 males and 8 females) participants gave their informed consent to participate in this study. Eight of these 16 participants were of Type A personality and the remaining 8 were of Type B personality. Their mean (SD) age, weight and height were 26.1 (9.9) years, 73.6 (16.9) kg and 175.6 (9.4) cm, respectively. All of the participants were right hand dominant and their detailed anthropometric information is provided in Table 5.1. All participants had average or above average levels of general fitness, based on their reported fitness and daily level of physical exertion. None of the participants reported any musculoskeletal problems that might have impeded their performance on the experimental task. All participants had previous manual work experience involving either lifting heavy equipment, general shop tasks, construction work, or work as a mechanic. The mean length of employment as reported by participants was 5.3 (8.9) years, though the distribution was skewed towards work experience of more than 2.9 years.

Table 5.1 Age and anthropometric data from 16 participants (8 males and 8 females)

				Percentiles	
Anthropometric Parameter	**Mean**	**Median**	**SD**	**5th**	**95th**
Age (years)	26.1	22.0	9.9	20.0	45.3
Weight (kg)	73.6	70.7	16.9	55.2	98.6
Stature (cm)	175.6	177.9	9.4	162.3	187.7
Shoulder Height (cm)	146.4	147.6	8.6	134.6	158.2
Upper Arm Length (cm)	30.0	29.9	2.5	25.7	33.2
Lower Arm Length (cm)	26.5	26.5	2.5	22.8	29.7
Arm in full extension (cm)	200.9	201.4	15.7	175.4	222.2
Arm at 90degrees (cm)	168.0	170.3	13.3	143.7	183.1

Each participant completed two treatment conditions: T1 (1.25 kg tool mass and 50% duty cycle) and T2 (2.0 kg tool mass and 67% duty cycle). The distribution of working heights was roughly normal. From a fitted normal distribution (mean = 181.2, SD = 14.2 cm), quartile values (25^{th}, 50^{th}, and 75^{th} percentiles) were determined to be 172.5, 182.7, and 192.6 cm, respectively.

5.2 Ratings of Perceived Discomfort (RPDs)

Fourteen of the 16 participants were able to complete the treatment condition T1 for two hours. None of the participants were able to work on the treatment condition T2 for the entire two hours. The average times, at which selected values of shoulder discomfort (RPD) were reported, were different for the two treatment conditions (Figure 5.1). A given level of discomfort was reported earlier when the T2 treatment condition was performed. Twelve of 16 participants reported highest RPD levels of less than 5 while performing the T1 treatment condition. On T2, 14 of 16 participants reported RPD levels of greater than 7 before 20 minutes elapsed, thus indicating that T2 was a much harder treatment condition than T1.

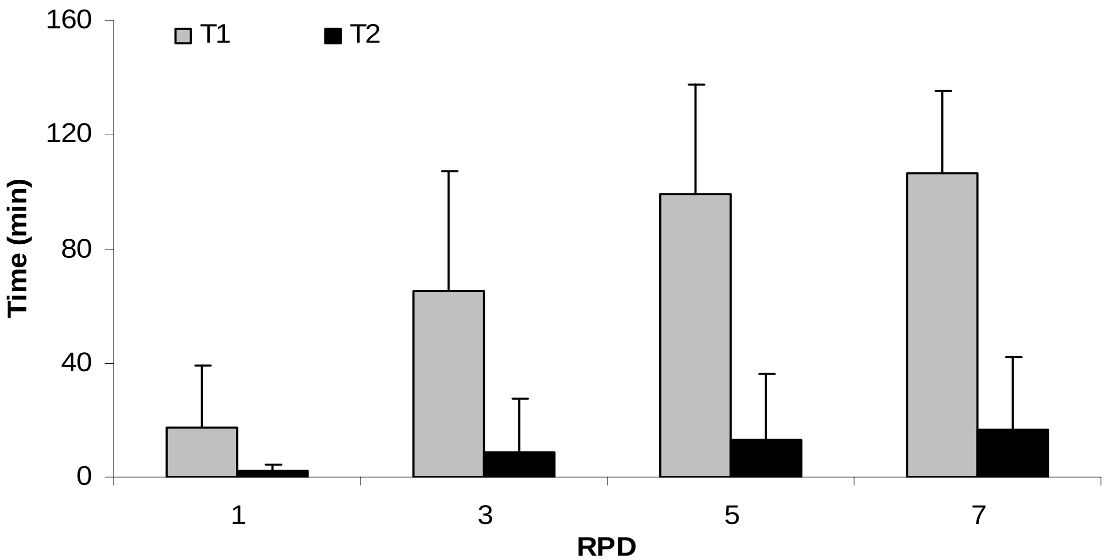
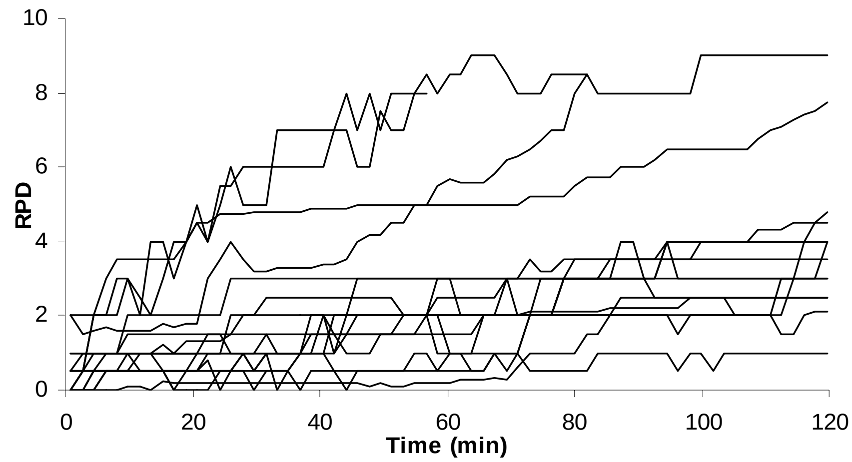


Figure 5.1 Relationship between ratings of perceived discomfort (RPD) and treatment condition. Mean times (SD) at which RPD levels were reached are shown.

RPD trends as a function of time were roughly monotonic and a significant linear trend was observed using bivariate analysis for both treatment conditions, T1 and T2 (Figure 5.2). In one case for T2, the participant consistently reported a low level of discomfort (RPD = ~2 and ~3) for a prolonged period but later exhibited a sudden increase in the RPD levels. It seems that this participant had some difficulty in conceptualizing the RPD scale and applying it to his/her perceptions (e.g. over-reliance on previous RPD ratings). For this case, the RPD trend was still monotonic and a linear fit still appears to capture the general data trends. Thus, this trial for T2 has been included in the qualitative and quantitative analysis.



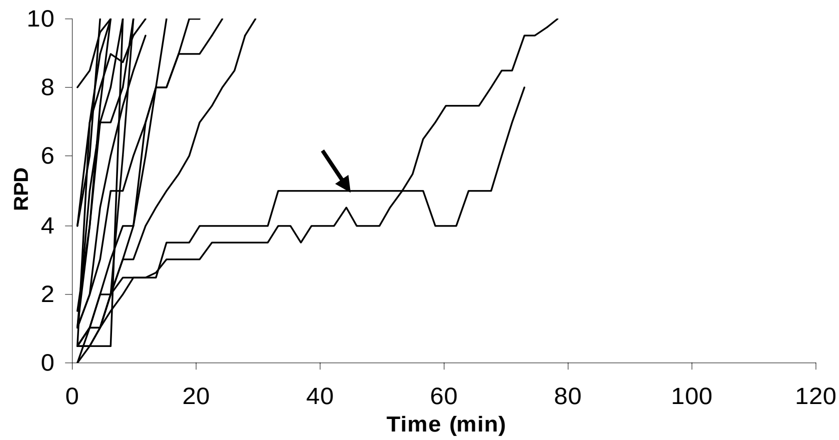


Figure 5.2 RPD levels as a function of time for sixteen participants. The top and bottom graphs indicate results from T1 (easy) and T2 (difficult) treatment conditions, respectively. One exception for T2 has been indicated and discussed in text.

5.3 Qualitative Analysis of RPDs

Qualitative analysis was undertaken to determine if RPD values were of sufficient reliability to adequately characterize the progression of fatigue. Using a 'cut-off' value of R_7, Categorization Matrices were developed for different values of time (between 0 and 120 minutes) and reported RPD levels. From these procedures, it was found that for treatment T1, twelve of 16 participants did not report high level of discomfort (RPD ≥ 7) within the 2 hours experimental duration. For treatment T2, ten of 16 participants reached high levels of discomfort by 10 minutes. For all the 737 (combination of 11-R values and 67 T-values) categorization matrices, sensitivity, specificity, positive PV, negative PV, and % correct were calculated and tabulated as shown in

Table 5.2. When all trials are correctly categorized, the values of sensitivity, specificity, positive PV, and negative PV for this perfect categorization are '1', and for % correct '100', respectively.

Table 5.2 Partial table showing values of sensitivity, specificity, positive PV, negative PV, and % correct for R=0.5 and several values of T obtained by combining results of both T1 and T2 treatment conditions. Similar tables were drawn for other RPD values.

RPD	Time (min)	Categorization Matrix Cells				Percent Correct	Sensitivity	Specificity	Positive P V	Negative P V
		A	B	C	D					
0.5	0.9	5	17	7	3	0.75	0.58	0.85	0.70	0.77
0.5	2.7	7	20	5	0	0.78	0.42	1	1	0.74
0.5	4.5	9	20	3	0	0.72	0.25	1	1	0.69
0.5	6.3	11	20	1	0	0.66	0.08	1	1	0.65
0.5	8.1	11	20	1	0	0.66	0.08	1	1	0.65
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.5	116.1	12	20	0	0	0.63	0.00	1	1	0.63
0.5	117.9	12	20	0	0	0.63	0.00	1	1	0.63
0.5	119.7	12	20	0	0	0.63	0.00	1	1	0.63

Where A, B, C, and D are:

		Completed 2 hours without reaching R_C? Yes	No
Reported 'R' by 'T' minutes?	Yes	A	B
	No	C	D

These tables were used to develop contour graphs that helped in visually assessing the changes in the predictive validity measures for different combinations of R and T. These were later used to determine the R and T combinations for which all predictive validity measures were at their maximum optimum values.

The s*ensitivity* contour graph shows the variation in the sensitivity values, shown as different levels of shading, as a function of time and RPD level (Figure 5.3). For lower RPD levels ($0 \leq R \leq 1$) sensitivity is always less than unity. For higher RPD levels ($5 \leq R \leq 10$), sensitivity is unity irrespective of the time. In mid ranges of RPD ($1 < R < 5$) sensitivity of the categorization matrix is dependent on duration. For example, for R = 2 and T = 20min, the sensitivity is 0.92, implying that from the trials in which participants could complete the 2-hour session without reaching R_7, 92% of participants did not and 8% of the participants did report RPD 2 by 20 minutes.

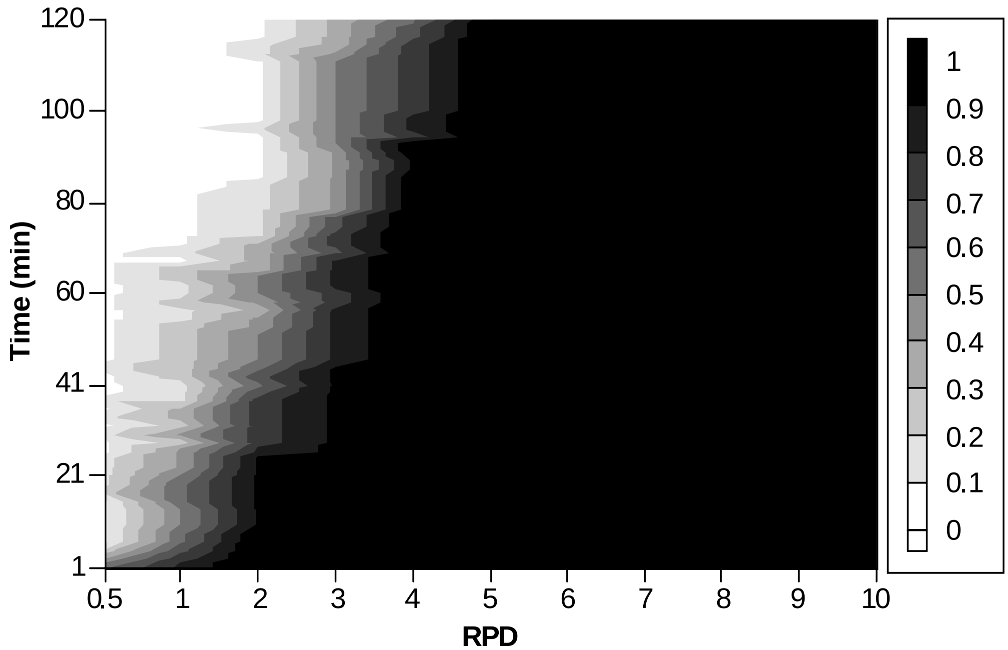


Figure 5.3 Sensitivity of categorization matrices for various combinations of R and T for both treatments and all 16 participants

The *specificity* contour graph shows a decrease in the specificity value with an increase in the RPD levels (Figure 5.4). Specificity increases with duration for all the RPD values. With increase in RPD level, more time is required for specificity to attain its maximum. For example, for R = 3 and T = 8 min, the specificity is 0.75, implying that of the trials in which participants reached R_7 before 2 hours, 75% of participants did and 25% of the participants did not report RPD 3 by 8 minutes.

The *positive PV* contour graph shows that for higher RPD higher duration yield a better positive PV (Figure 5.5). For example, for R = 5 and T = 45 min, the positive PV is 0.92, implying that of trials in which participants did not report RPD 5 by 45 minutes, 92% of participants could and 8% could not complete the 2-hour session without reaching R_7.

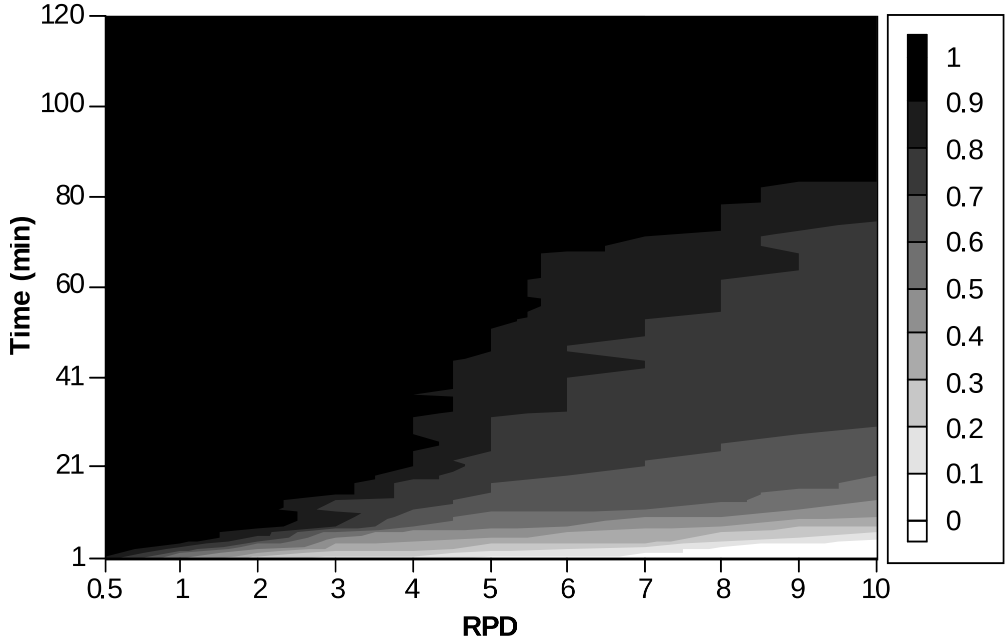


Figure 5.4 Specificity of categorization matrices for various combinations of R and T for both treatments and all 16 participants

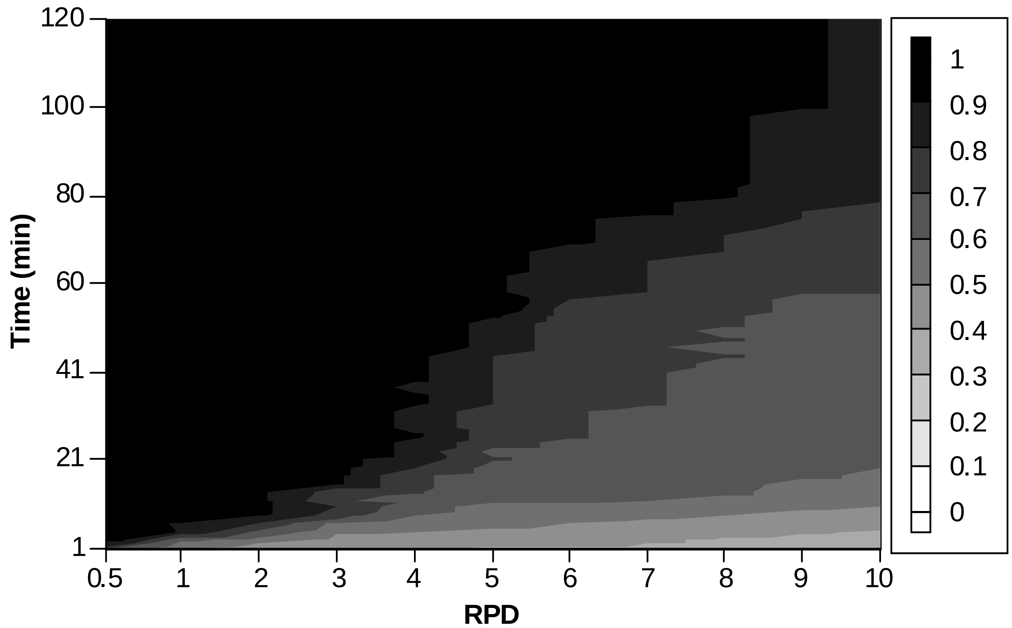


Figure 5.5 Positive PV of categorization matrices for various combinations of R and T for both treatments and all 16 participants

The *negative PV* contour graph shows that for higher RPD ranges ($5 \leq R \leq 10$) positive PV is unity irrespective of the duration (Figure 5.6). For lower RPD ranges ($0 \leq R < 5$), positive PV increases with the RPD level and decreases with duration. For example, for R = 1 and T = 30 min, the negative PV is 0.36, implying that of trials in which participants reported RPD 1 by 30 minutes, 36% of participants could not and 64% could complete the 2-hour session without reaching R_7.

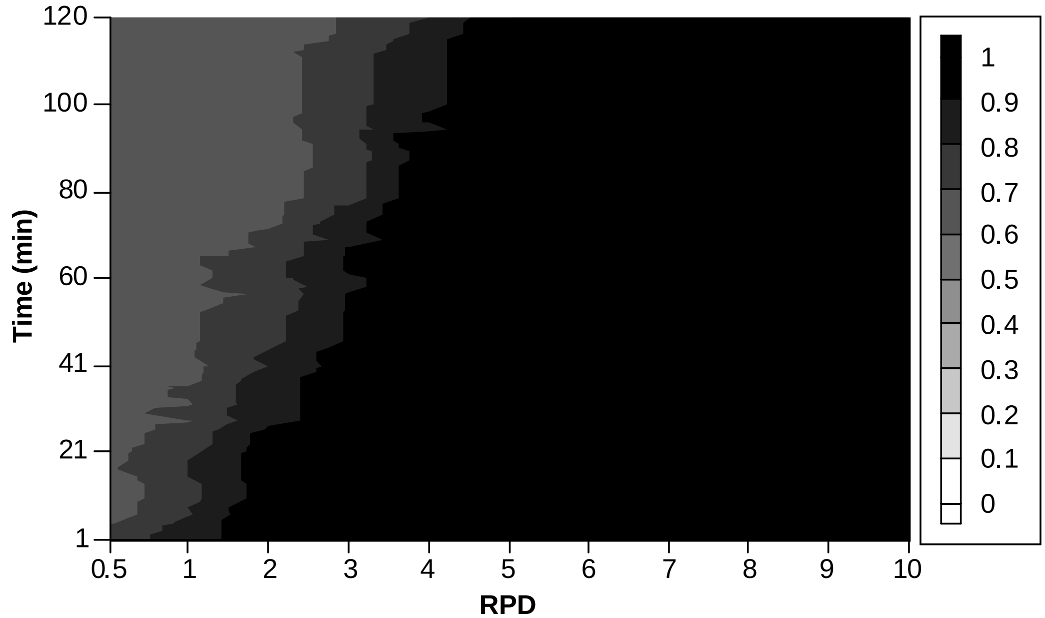


Figure 5.6 Negative PV of categorization matrices for various combinations of R and T for both treatments and all 16 participants

The *percentage correct* contour graph indicates the combinations of R and T for which all the trials were correctly categorized (Figure 5.7). For example, for R = 3 and T = 10 min, the % negative is 94%, implying that 94% of the trials were correctly categorized and 6% were incorrectly categorized.

Based on the predictive validity measures, ranges of R and T for which maximum number of trials were correctly categorized were determined. The combinations of RPD (R) and Time (T) for which all the participants were correctly categorized (indicated by 'X' in Table 4.4) are given in Table 5.3. Thus, using a combination of RPD and the duration ranges given in Table 5.3, it is possible to obtain categorization matrices that correctly categorize all the participants according to their performance (Pass/Fail based on whether participant can/cannot complete the task for 2 hours without reaching R_7) in 2 hours.

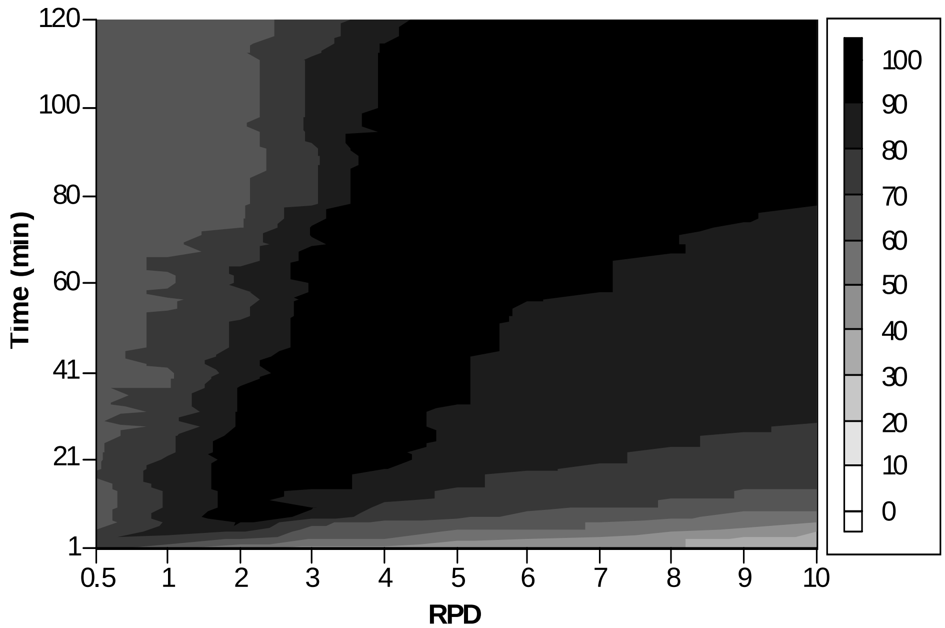


Figure 5.7 % Correct of categorization matrices for various combinations of R and T for both treatments and all 16 participants

Table 5.3 Time ranges for which the all the participants are correctly categorized in a categorization matrix for a certain RPD level. The predictive validity for these combinations of R and T was excellent (described by sensitivity = 1, specificity = 1, positive PV = 1, negative PV =1, and % correct = 100%)

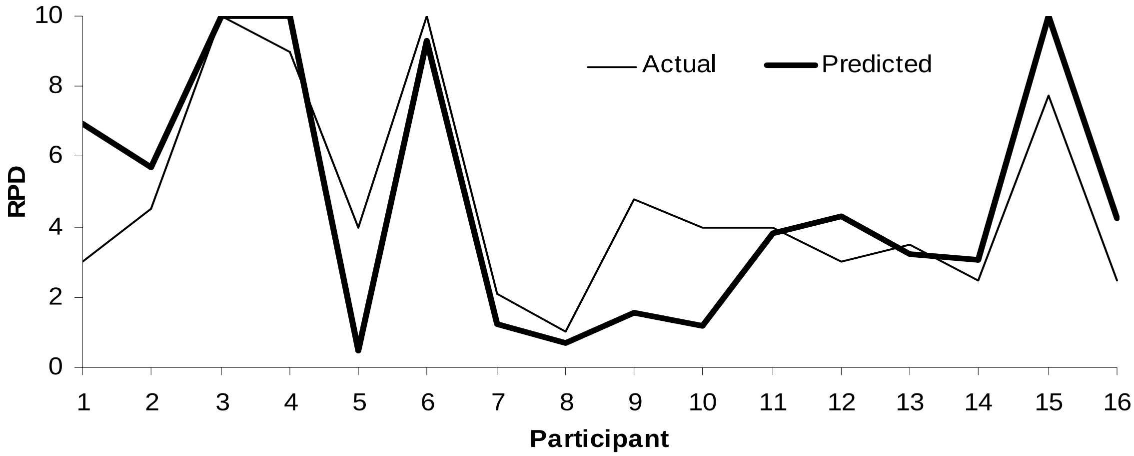
	Time Intervals (min)		
RPD	T1	T2	T1 and T2
0.5	NE	2.7 - 119.7	NE
1	NE	8.1 - 119.7	NE
2	NE	8.1 - 119.7	NE
3	22.5 – 27.91	15.3 - 119.7	22.5 – 27.91
4	45.9 - 92.7	33.3 - 119.7	45.9 – 92.7
5	54.9 - 119.7	63.9 - 119.7	63.9 - 119.7
6	87.3 - 119.7	69.3 - 119.7	87.3 - 119.7
7	110.7 - 119.7	71.1 - 119.7	110.7 - 119.7
8	NE	72.9 - 119.7	NE
9	NE	74.7 - 119.7	NE
10	NE	78.3 - 119.7	NE

*NE or Not Excellent indicates that for that RPD value, no value of T gave excellent predictive validity.

5.4 Quantitative Analysis of RPDs

Quantitative analysis involved extrapolation using regression analysis. Reported levels of RPD after 2 hours were compared to the predicted RPD levels obtained by extrapolating the data obtained in shorter time durations. Stepwise iteration method used increasing sequences of data in regression to obtain regression function for all the possible 67 sequences. From this analysis, best results were obtained by using data from the first 26.1 minutes for T1 and the first 8.1 minutes for T2. At 26.1 min for T1 and 8.1 min for T2, the actual and predicted RPD slopes, the RPD values at 2 hours, and the time taken to reach R_7 were highly correlated ($\rho > 0.8$) for both T1 and T2.

Using the Kruskal-Wallis test, the actual and predicted RPD at the end of 2 hours, and time taken to reach R_7 were compared. The predicted values were obtained using the first 26.1 minutes data for T1 and the first 8.1 minutes data for T2. The numbers of Pass/Fail participants were found. As the highest possible value of RPD was 10 (corresponding to maximum pain or discomfort), if the regression equation predicted a value higher than 10 at 2 hours the value was reset at 10. Similarly, as here only 2-hour performance was studied, if the regression equation predicted a time to reach R_7 as higher than 2 hours then it was reset as 2 hours. The values at the end of 2 hours, time taken to reach R_7, RPD linear regression line slopes, and the numbers of Pass/Fail participants were found to be equivalent for both the treatment conditions. For T1, the predicted RPD values at 2 hours were overestimated for 8 participants and underestimated for 8 participants, and for T2 they corresponded perfectly for all the participants (Figure 5.8). The Pass/Fail status derived using predicted values corresponded to the actual data in all trials for T1 and T2.



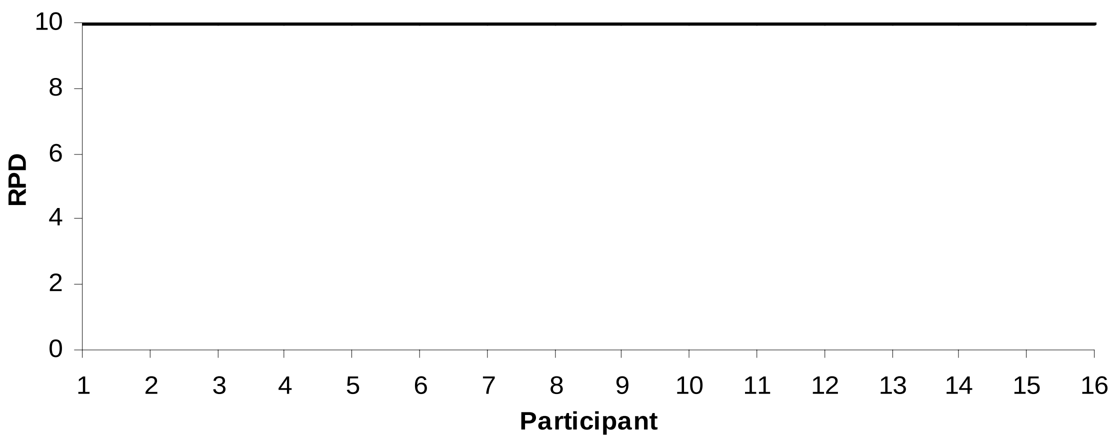


Figure 5.8 Actual and Predicted RPD levels for sixteen participants. Predicted RPD values were derived using first 26.1 minutes data for T1 (easy), shown on top and 8.1 minutes data for T2 (difficult), shown at bottom treatment conditions.

Using the 26.1 minutes data for T1 and 8.1 minutes data for T2, actual time taken to reach R_7 was compared to the predicted time (Figure 5.9). When the participant did not reach R_7 within the 2-hour experimental session, it was assumed that the participant would take at least 2 hours to reach R_7. Also, when the predicted time was more than 2 hours, it was assumed that the participant could at least do the task for 2 hours without reaching R_7. For T1, the predicted time to reach R_7 was overestimated for 13 participants and underestimated for 3 participants and correctly predicted for 2. For T2, the time was underestimated for 8 participants and overestimated for 8 participants (Figure 5.9). The Pass/Fail status derived using predicted time values corresponded to the actual data in all trials for T1 and T2.

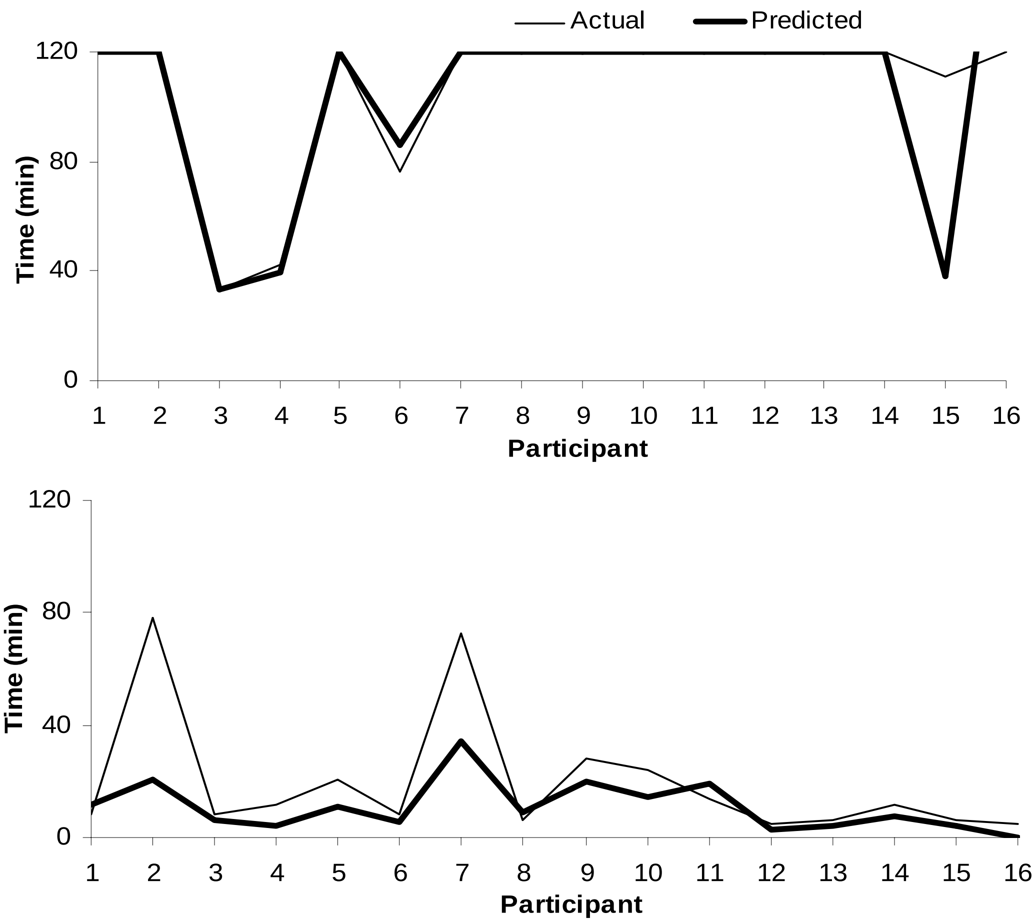


Figure 5.9 Actual and Predicted time to reach R_7 for sixteen participants. Predicted RPD values were derived using first 26.1 minutes data for T1 (easy) shown on top and 8.1 minutes data for T2 (difficult) shown at bottom, treatment conditions.

Reliability of the predicted RPD values was quantified by using the ICC, SEM and Pearson's r. ICC value ranges between 0 and 1 and a value between 0- 0.4 indicates fair, 0.4-0.75 good and 0.75-1 excellent reliability (Fleiss, 1986). Good reliability was observed for the predicted RPD values at 2 hours and excellent reliability was observed for RPD slopes for both treatment conditions (Table 5.4). Low values of SEM and high Pearson's r values also support strong correspondence between the predicted and actual RPD measures.

Table 5.4 ICC, SEM and Pearson's r-values for comparing the actual and predicted measures for treatments T1 and T2

	RPD Measure	ICC	SEM	Pearson's r
T1	RPD at 2 hours	0.87	1.20	0.80
	RPD slope	0.71	0.06	0.75
T2	RPD at 2 hours	0.89	0.48	0.99
	RPD slope	0.71	0.25	0.80

5.5 Minimum Trial Duration (MTD)

Based on qualitative and quantitative analysis, MTD for T1 was determined to be 26.1 minutes and critical RPD level as 3 (Table 5.3). For T2, the MTD was selected as 8.1 minutes and critical RPD level as 2 (Table 5.3). RPD level of 2 was selected instead of 1, as here selecting a higher RPD level cutoff will help ensure the accuracy of deductions without any loss of time. The categorization matrices for T1 with T = 26.1 minutes and R = 3, and for T2 with T = 8.1 minutes and R = 2, yield the results in Table 5.5 and Table 5.6. Kolmogorov goodness of fit test determined that the cell values were not uniformly distributed for both T1 and T2. A significant negative correlation was found between categorizing parameters (reported R = 3 by 26.1 minutes and completed 2 hours without reaching R_7) using the Phi coefficient test. This implies that the completion of the 2 hours trial is associated with not reporting of R = 3 by 26.1 minutes and not competing the 2 hours trial is associated with the reporting R = 3 within the first 26.1 minutes. Similar results were found for treatment T2 using the Phi coefficient test for the combination of R = 2 and T = 8.1 minutes.

Table 5.5 Categorization Matrix for determining critical trial durations and RPD levels for treatment condition T1

		Completed 2 hours without reaching RPD=7?	
		Yes	No
Reported RPD=3 by 26.1 minutes?	Yes	0	4/16 (T1)
	No	12/16 (T1)	0

Table 5.6 Categorization Matrix for determining critical trial durations and RPD levels for treatment condition T2

		Completed 2 hours without reaching RPD=7? Yes	No
Reported RPD=2 by 8.1 minutes?	Yes	0	16/16 (T2)
	No	0	0

5.6 Personality Measure and Treatment Type

The personality measure value for all the 16 participants is given in Figure 5.10. Type A personality people on average had lower RPD onset times than Type B people for both T1 and T2. However, no significant effects of personality were observed on the RPD onset times for RPD 1 (p = 0.59) and RPD 7 (p = 0.6) but effect was significant on the onset times of RPD 3 and 5. No differences were found between the actual and predicted RPD values at 2 hours (p = 0.77) and the time to reach R_7 (p = 0.59) using the sign test. Personality type was not considered for further analysis as no significant effects were found on RPD prediction.

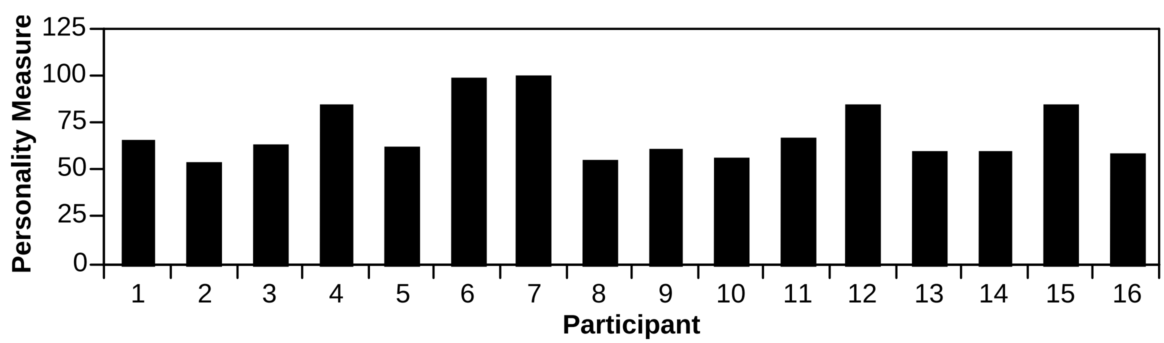


Figure 5.10 Personality measure calculated by adding the scores on the personality checklist for each of the 16 participants.

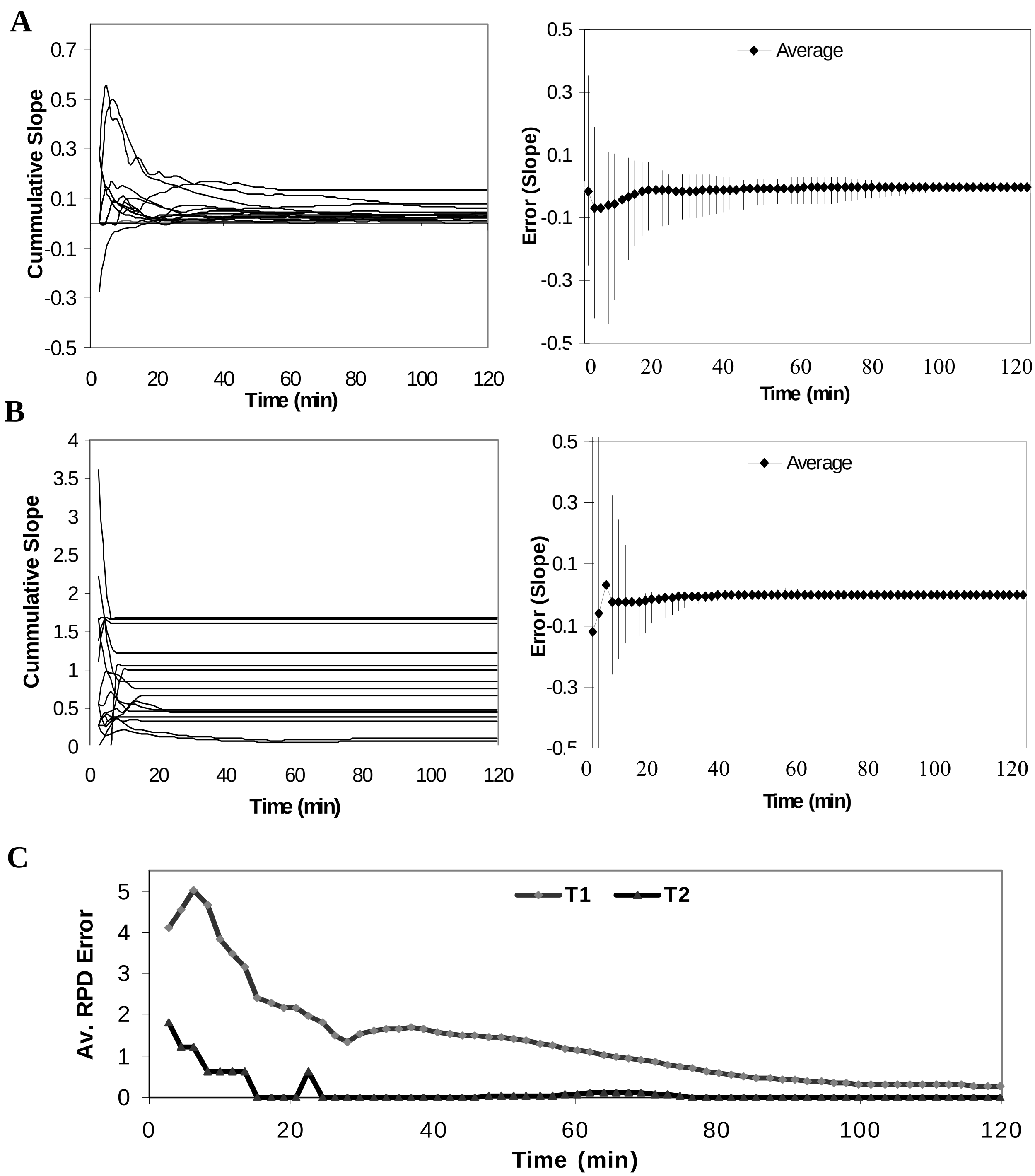
The RPD regression line slope for treatment T2 was found to be significantly greater than T1 using the Wilcoxon sign test indicative of a faster increase in RPD with time for T2 than T1. The RPD onset times were also found to be significantly higher for T2 than T1 using the sign test. Both the results indicated that RPD increased at a much faster rate for T2 than T1.

5.7 General Trends

Trends were observed in certain derived measures when averaged across participants for a treatment condition. These trends were difficult to generalize and are specific to the treatment conditions studied. For this analysis, errors in the cumulative RPD slope and error in predicted RPD were studied. Errors in slope (e_T), measured as the difference in the cumulative slopes (β_T) from the final 2-hour slope ($\beta_{2\text{-hour}}$), were determined for each participant and for each treatment condition respectively. Maximum, minimum and average errors in slope across the sixteen participants were calculated for each treatment condition $\{E_T = (\Sigma\, e_T)/16;\ T = 2.7, 4.5, \ldots, 119.7\}$. A positive error in slope ($e_T > 0$) indicated that the 2-hours slope was greater than the cumulative slope, and a negative error in slope ($e_T < 0$) that the cumulative slope was greater than the 2-hour slope at that 'T'. The cumulative slopes were positive for all but one participant for treatment condition T1.

For T1, average errors in slope indicate that initially the cumulative slope was higher than the 2-hour slopes, and then the difference between them decreases before leveling off to zero (Figure 5.11 A). Also, the errors were skewed towards negative values for all 'T'. For T2, average errors in slope show that initially the 2-hour slope was greater than the cumulative slope before it approaches zero (Figure 5.11 B). The errors were skewed towards positive values until T = 12min, shifting to negative values until T = 45min before they taper off to zero. For both treatment conditions, the error ranges (maximum and minimum values) decreased with time, at a faster rate for T2 than T1 (Figure 5.11 C).

Cumulative predicted RPD was calculated from using cumulative slopes and cumulative intercepts. Errors in RPD were calculated by subtracting the cumulative predicted RPD from the 2-hour RPD (RPD at 2hours - Cumulative predicted RPD). Errors in RPD corroborate the errors in slope (Figure 5.11). Average errors in slope and RPD indicate the time required for the slope to become sufficiently characterized so that no more data collection would effectively add to our understanding of the 2-hour results.



<u>Figure 5.11</u> Cumulative slopes and the minimum, maximum and average error ($e_T = \beta_{2\ hour} - \beta_T$) values for moderate treatment T1 (**A**) and difficult treatment T2 (**B**). Also errors in RPD (Final RPD – Predicted RPD) have been shown for T1 and T2 (**C**).

5.8 General Prediction Model

Various approaches were tried to classify trends across participants and treatment conditions. Developing a general modality required capturing changes (within a parameter or in one parameter with respect to another) where the absolute magnitudes may be task specific. Regression analysis emerged as the best tool allowing for measurement of changes/rate of changes in parameters over time. Three approaches were used, one looked at the changes in the significance level, second at the standard deviation, and a third at the normalized standard deviation of the cumulative RPD slopes with time. The level of significance of the cumulative slope was used to quantify the strength of the relationship between RPD and time and the standard deviation to measure the variability in the cumulative slope. The normalized standard deviation was determined to facilitate comparison of variability in cumulative slope across participants and treatment conditions.

Probability value (P-value), the observed level of significance, of RPD slope $\{P(\beta_T); T = 2.7, 4.5,\ldots, 119.7\}$, was obtained using the F-test. P-value thus obtained using inferential statistics is the probability of obtaining a statistics as or more different from zero. Standard deviation or sigma (slope), $\{s(\beta_T); T = 2.7, 4.5,\ldots, 119.7\}$, and normalized standard deviation or norm-sigma (slope), $\{ns(\beta_T): s(\beta_T) / \beta_T ; T = 2.7, 4.5,\ldots, 119.7\}$, were calculated for every participant and for T1 and T2 respectively (Equation 1 and Equation 2). The 'Pass/Fail' based on final predicted RPD, 'Pass/Fail' based on time to reach R_7 and errors in final predicted RPD corresponding to the P-value, sigma (slope) and norm-sigma (slope) were determined and tabulated as shown in Table 5.7.

Tables as the one shown in Table 5.7 were then used to determine the percentage of correct 'Pass/Fail' predictions based on final predicted RPD (% CPFR) and time to reach R_7 (% CPTR7) for a treatment condition and for various P-value, sigma (slope), and norm-sigma (slope). *Example:* From Table 5.7, for P-value = 0.0003, % CPFR is 87.5 %. This implies that for P-values 0.0003, 87.5 % of the total 16 participants could complete the T1 task for 2 hours without reaching R_7. Similarly, % CPFR and % CPTR7 were determined for various P-values, sigma (slope), and norm-sigma (slope) calculated for T1 (Figure 5.12) and T2 (Figure 5.13). Maximum possible errors in predicted final RPD (Errors PFR) were also determined as the maximum error in predicted RPD at 2 hours across participants for a treatment condition (Figure 5.12 and Figure 5.13).

Table 5.7 Partial sample table shows the Pass/Fail deductions based on the predicted final RPD for 16 participants for treatment T1. Predictions were called correct (1) if they concurred with those obtained from the 2-hour session and incorrect (0) otherwise. Similar tables were drawn for P-value, sigma, norm-sigma and corresponding errors in RPD and Pass/Fail deductions for T1 and T2 respectively.

	Pass/Fail Deduction based on the Predicted Final RPD															
P-Value	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
0.00001	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.00002	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.0002	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.0003	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.006	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1
0.007	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.1	0	0	1	1	1	0	1	0	0	0	0	0	1	1	1	0
0.2	0	0	1	1	1	0	1	0	0	1	0	0	0	1	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.9	0	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0
1	0	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0

For T1, results did not indicate a trend in the P-values with changes in % CPFR, %CPTR7, and Error PFR. For T2, P-values decreased with increases in % CPFR and %CPTR7, and decreases in Error PFR. Results indicated that sigma (slope) values decreased with increases in % CPFR and %CPTR7, and decreases in Error PFR. Similar trends were seen in norm-sigma (slope) values, which decreased with increases in % CPFR and %CPTR7, and decreases in Error PFR. Also for T2, errors in final RPD were very small ranging from 0.75 - 1.81 for P-values, 0 - 0.05 for sigma (slope), and 0.01 - 0.59 for norm-sigma (slope).

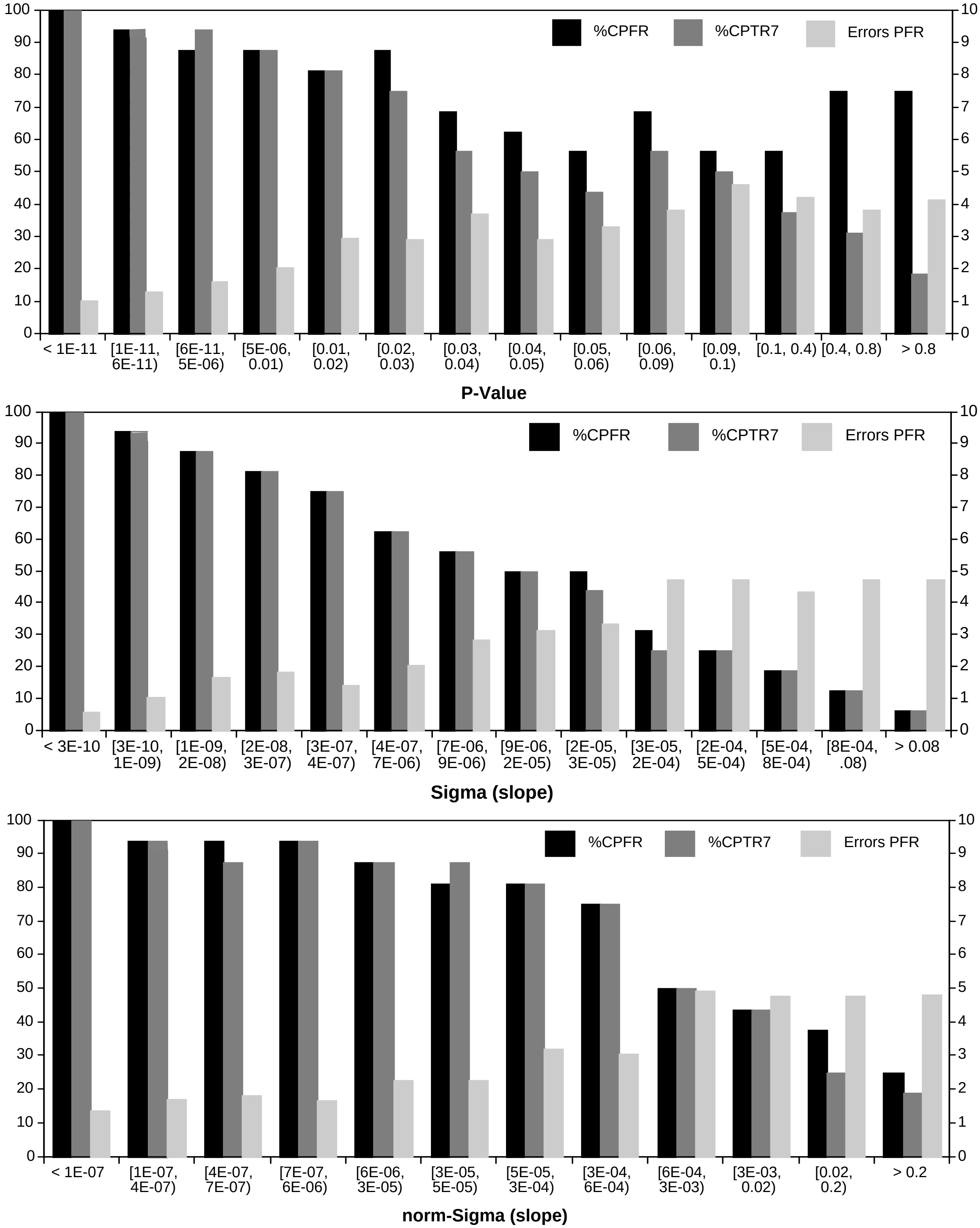


Figure 5.12 The % correct predictions based on Time to reach R_7 (% CPTR7), % correct predictions based on final RPD (% CPFR), and Errors in predicted Final RPD (Errors PFR) have been graphed against the ranges of P-value, sigma (slope), and norm-sigma (slope) in which they occur for treatment condition T1.

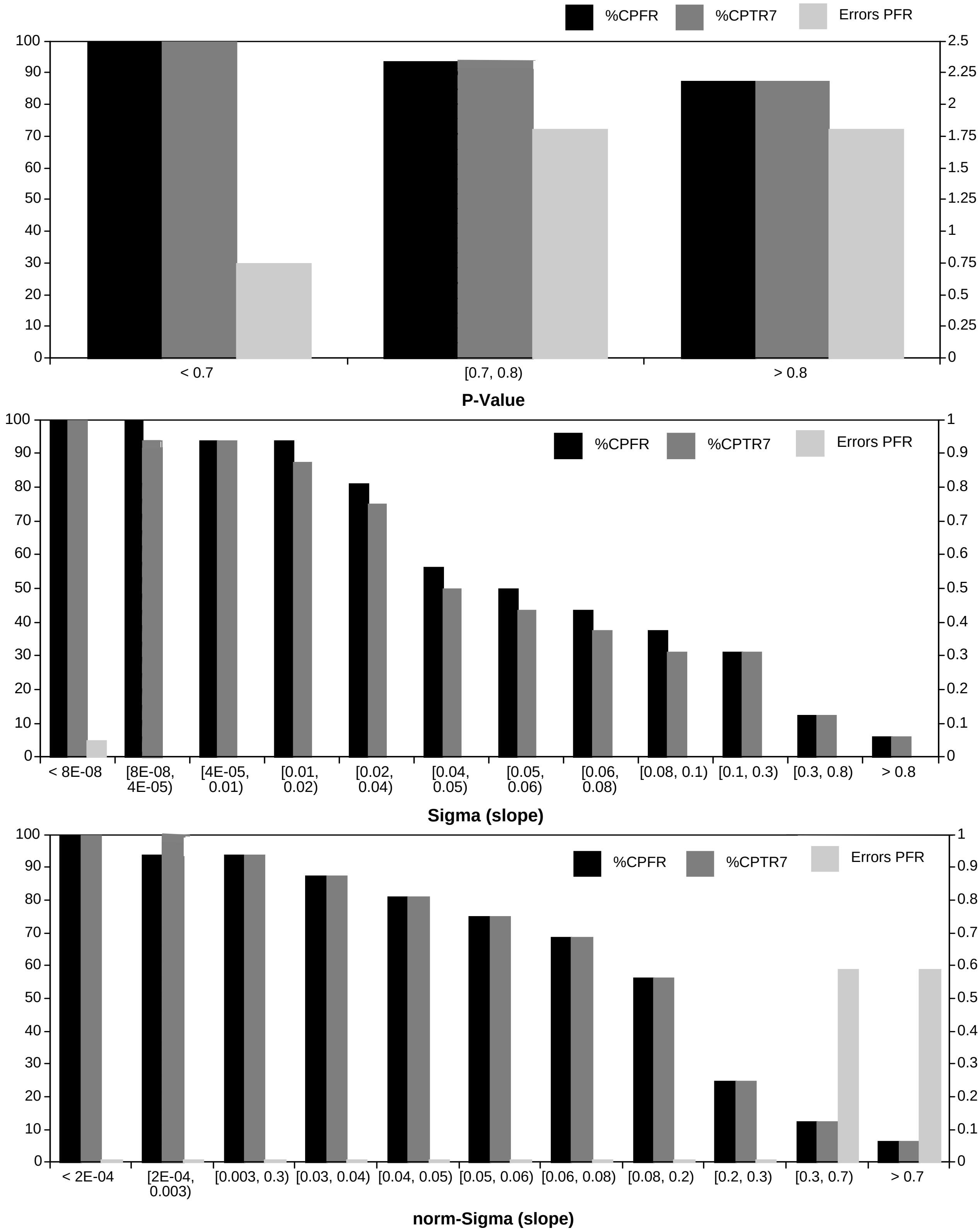


Figure 5.13 The % correct predictions based on Time to reach R_7 (% CPTR7), % correct predictions based on final RPD (% CPFR), and Errors in predicted Final RPD (Errors PFR) have been graphed against the ranges of P-value, sigma (slope), and norm-sigma (slope) in which they occur for treatment condition T2.

Using the reverse approach, the results shown in Figure 5.12 and Figure 5.13 were used to determine the ranges of P-value, sigma (slope), and norm-sigma (slope) for which a certain % CPFR and % CPTR7 could be made. Essentially, this process tried to answer questions such as ‘what is the P-value for T1’ when % CPFR is 80 %. Also, the P-values, sigma (slope), and norm-sigma (slope) corresponding to a certain % correct predictions were also determined for both T1 and T2. The % correct predictions for a treatment condition were based on both % CPFR and % CPTR7 results. These results were then used to determine the % combined-correct predictions. The % combined-correct predictions were determined to coalesce the results across the treatment conditions. The % combined-correct prediction refers to the combined results of % CPFR and %CPTR7 for both T1 and T2. For this, ranges of P-value, sigma (slope), and norm-sigma (slope) were obtained for the % combined-correct predictions by using results of both % CPFR and % CPTR7. There were three assumptions made while combining the results of % CPFR, %CPTR7 and errors PFR. The three assumptions have been demonstrated using an example.

Example: In the table for T1, Row 1 is used to explain assumption 1, Row 2 for assumption 2, and Row 3 for assumption 3.

	% Correct Predictions based on		Errors in	% Correct Predictions based on		Errors in
	Final RPD	Sigma (slope)	Final RPD	Time to reach R_7	Sigma (slope)	Final RPD
Row 1	100	< 0.01	1	100	<0.001	2
Row 2	100	[0.01, 0.02)	2	95	[0.01, 0.02)	3
Row 3	95	[0.02, 0.03)	3	95	[0.02, 0.03)	4

Row 1: For 100% combined predictions sigma (slope) chosen was < 0.001. While combining across criterion, here sigma (slope) for a % correct prediction, the lower value of sigma (slope) was chosen as a more stringent condition, thus providing a more conservative estimate. This assumption is based on the monotonicity observed sigma (slope) and norm-sigma (slope). The categories were not combined for P-value due to the absence of this monotonicity.

Row 2: When Sigma (slope) $\in$ [0.01, 0.02), atleast 95% combined correct predictions are made. While determining the % of combined conditions relating a sigma (slope) range, lower percentages of the two, % CPFR and % CPTR7, were selected as a conservative measure and it holds true for both the prediction methods.

Row 3: For 95% combined predictions, sigma (slope) $\in$ [0.02. 0.03), maximum error in final RPD is 4. While determining the errors PFR for a certain % combined correct prediction and sigma range, the higher error value was selected again as a conservative estimate.

Table 5.8 and Table 5.9 contain the results obtained from the % combined correct analysis. Together these tables form the general model. General model is a type of approach that looks at the significance level and the variability in slope to determine the costs associated with the desired predictions accuracy. It indicates the sigma (slope), norm-sigma (slope), and P-value ranges for which a certain % correct predictions can be made regarding the final RPD and time to reach R_7. The P-value, sigma (slope), and norm-sigma (slope) ranges corresponding to percentage correct predictions are independent of each other indicating that any of the criterions can be used separately.

Table 5.8 Percent correct predictions corresponding to P-value or sigma (slope) or norm-sigma (slope) ranges, for treatment T1. Also, given are the corresponding maximum errors in the final RPD values

% Combined	**T1**			
Correct	Sigma (slope)	Norm-sigma (slope)	P-Value	Error in Final RPD
100	[0 - 3E-10)	[0, 9E-07)	[0, 1E-13)	1.38
[90, 100)	[3E-10, 1E-09)	[9E-07, 0.03)	[1E-13, 5E-04)	1.71
[80, 90)	[1E-09, 3E-07)	[0.03, 0.05)	[5E-04, 0.02)	2.25
[70, 80)	[3E-07, 4E-07)	[0.05, 0.06)	[0.02, 0.03)	3.06
[60, 70)	[4E-07, 6E-07)	[0.06, 0.08)	[0.02, 0.03)	3.06
[50, 60)	[6E-07, 2E-05)	[0.08, 0.2)	[0.03, 0.05)	4.91
[40, 50)	[2E-05, 3E-05)	[0.2, 0.3)	[0.05, 0.2)	4.78
[30, 40)	[3E-05, 2E-04)	[0.2, 0.3)	[0.2, 0.8)	4.78
[20, 30)	[2E-04, 5E-04)	[0.2, 0.3)	[0.2, 0.8)	4.78
[10, 20)	[5E-04, .08)	[0.3, 0.7)	> 0.8	4.81
[0, 10)	[0.08, 0.09)	> 0.7	NA	4.81
0	> 0.09	NA	NA	4.81

* NA implies that the % correct was never in that range

Table 5.9 Percent correct predictions corresponding to P-value or sigma (slope) or norm-sigma (slope) ranges, for treatment T2. Also, given are the corresponding maximum errors in the final RPD values

% Combined	**T2**			
Correct	Sigma (slope)	Norm-sigma (slope)	P-Value	Error in Final RPD
100	[0, 8E-08)	[0, 1E-07)	< 0, 0.6	0.75
[90, 100)	[8E-08, 0.01)	[1E-07, 4E-07)	[0.6, 0.7)	1.22
[80, 90)	[0.01, 0.02)	[4E-07, 0.0003)	> 0.7	1.81
[70, 80)	[0.02, 0.04)	[3E-04, 6E-04)	NA	1.81
[60, 70)	[0.02, 0.04)	[3E-04, 6E-04)	NA	1.81
[50, 60)	[0.04, 0.05)	[6E-06, 3E-03)	NA	1.81
[40, 50)	[0.05, 0.06)	[3E-03, 0.01)	NA	1.81
[30, 40)	[0.06, 0.3)	[0.01, 0.2)	NA	1.81
[20, 30)	[0.06, 0.4)	> 0.2	NA	1.81
[10, 20)	[0.3, 0.8)	NA	NA	1.81
[0, 10)	> 0.8	NA	NA	1.81
0	NA	NA	NA	1.81

* NA implies that the % correct was never in that range

6.0 DISCUSSION

Constraints on resources is a concern in any research but especially so in industrial ergonomics studies requiring human participation. Usually, these studies involve consideration of several task factors to simulate real task conditions and need the results to be applicable for 2-8 hour work periods. In addition, these studies entail large sample sizes to increase the experimental power and account for the large variability in human response. Thus, some important considerations while designing an ergonomics research study tend to be the accessibility of participants and their time availability. In this scenario, any reductions in the duration of experiments would allow for the study of a larger sample pool using the same amount of resources, but this had to be done without compromising the applicability of results to 2-8 hour work periods. Thus, the results from this study are important in their practical significance for developing more efficient experimental protocols.

6.1 Qualitative Analysis

This approach mainly used categorization matrices to determine the best combination of RPD and duration. RPD ratings at an instant in time and the final 2-hour RPD ratings were used to derive the categorization matrix elements. Various predictive validity measures were derived including sensitivity, specificity, positive PV, negative PV, and % correct.

Figures 5.3 - 5.7 indicate that best results were found with a combination of mid-ranges of RPD (R) and duration (T). Combinations of very small duration with either low RPD (*example:* R = 0.5, T = 1 min) or high RPD (*example:* R = 8, T = 1min) did not provide information about the performance at 2-hours. The most plausible reason for this was the insufficient time to capture the temporal trends in RPD. On average, data also indicated that low RPDs could be maintained for a long time, but if a high RPD is reached fast then RPDs keep on increasing. This relationship between RPD and the times at which it is reported could be used to determine the minimum trial duration (MTD) to predict the 2-hour performance.

Predictive validity measures provide us with an objective measure to judge the performance prediction effectiveness of a certain RPD and duration combination. Results from this experiment indicate that for the T1 treatment combination R = 3 and T = 26.1 minutes yielded the best 2-hour predictions. For T2, R = 2 and T = 8.1 minutes yielded the best 2-hour predictions. These results can be applied by conducting the experiment for only the determined 'T' duration instead of 2 hours. The RPD at the end of T min could then be used to determine the final outcome after 2-hours. However, although it seems extremely attractive to use shorter experimental durations of less than 30 minutes or even less than 10 minutes, as results from this study reveal, there may be certain costs attached with this method.

There are two types of categorizing errors (CE) associated with the incorrect categorization of trials by the categorization matrix. *CE 1:* A trial where the participant does not report R by T minutes but is unable to complete the task; *CE 2:* A trial where participant reports R by T minutes but completes the 2 hours without reaching R_7. CE 1 leads to overestimation and CE 2 underestimation of work capacity. These errors manifest themselves as costs: CE 1 can result in injury (C_I) and CE 2 results in efficiency loss (C_E). In an industry both of these costs will manifest themselves as monetary losses.

There can be various sources for these errors including presence of an outlier (participant whose RPD trend significantly affected the average values) and insufficient data that is unable to average out any non-conforming behavior, and most importantly non-monotonic trends in RPD. Errors can result in misinterpretation of certain predictive validity measures. This can lead to choice of an R and T combination where one or more of those measures were not at their best level (1 for sensitivity, specificity, positive PV, negative PV, and % correct). In these cases, it is important to identify the associated cost and segregate those that may to be preferred over others while making decisions on the optimal combination of R and T. Errors occurring due to over or underestimation of sensitivity, specificity, positive PV, and negative PV are given in Table 6.1.

Table 6.1 Type of Errors and costs associated with over or underestimation of sensitivity, specificity, positive PV, and negative PV. Please refer to Table 4.7.

Over Estimation	No. of trials that result in a Pass*	No. of trials that result in a Fail*
Sensitivity	Not affected	Cannot say
Specificity	Cannot say	Not affected
Positive PV	Estimating more (C_I)	Estimating less (C_I)
Negative PV	Estimating less (C_E)	Estimating more (C_E)

Under Estimation	No. of trials that result in a Pass*	No. of trials that result in a Fail*
Sensitivity	Not affected	Cannot say
Specificity	Cannot say	Not affected
Positive PV	Estimating less (C_E)	Estimating more (C_E)
Negative PV	Estimating more (C_I)	Estimating less (C_I)

* Pass = # of trials in which participant could complete the task without reaching R_7

* Fail = # of trials in which participant could not complete the task without reaching R_7

It is important to note that the implications of over/underestimation of % correct are equivalent to combined implications of sensitivity and negative PV, and/or of combined implications of specificity and positive PV. As no costs are incurred due to errors in sensitivity and specificity, therefore an over/underestimation of % correct may indicate an over/underestimation of negative PV, and/or over/underestimation of positive PV. Thus positive PV and negative PV are more important measures than sensitivity and specificity in terms of cost analysis.

% Correct can provide a complete measure to check the robustness of the entire categorization matrix. But positive PV and negative PV together give more specific information into the categorization matrix elements. % Correct can only indicate the number of elements that are correctly or incorrectly categorized but together with positive PV and negative PV details can be derived on the % elements in each cell of the categorization matrix. For the present categorization matrix application, positive PV and negative PV together are the most important measures. As evident from the cost analysis, error in these measures is also the costliest.

Based on this analysis, combinations of RPD and duration, which gave excellent predictive validity, were obtained. These results indicated the duration for which RPD data need to be collected and the deductions that can be obtained from the last reported RPD in that duration. This approach relies principally on the monotonicity of the relationship between RPD and time. In the case where trends are not monotonic, reliance on this approach can lead to errors. Also, it requires that the results at 2-hours should be known. This in part defeats its application to derive MTD to find performance at 2-hours. Though, this method cannot be applied directly to determine the MTD in absence of any 2-hour data. But it can be used to analyze data from a 2-hour pilot study, results of which can be used to determine MTD for a succeeding larger study. This method was applied in a similar manner in the previous automotive research.

6.2 Quantitative Analysis

Regression analysis allowed for quantification of RPD trends over time. These trends were then used to extrapolate from shorter duration trials to predict for 2-hours. In this study, based on the observed trends, linear regression was used to represent the temporal changes in RPD. Linear regression also helped to decrease the complexity of the data analysis process.

Based on participant's average RPD ratings and RPD onset times, it was found that the treatment conditions T1 and T2 were very different in their difficulty levels. As evident from the cumulative slopes of RPD versus time, the rate of change in RPD was slower for T1 than T2. Subsequently, changes in the RPD slope over time took less time to characterize in the case of T2 than T1 (Figure 5.11). Minimum trial duration (MTD) of 26.1 minutes for treatment T1 and relatively short 8.1 minutes for the harder treatment T2 were found to be sufficient for making 2-hour predictions. MTD required for T1 was much longer than for T2, indicating the dependence of MTD on the treatment's difficulty level. The dependence of MTD on the task difficulty level restricts the direct applicability of these results to other task conditions but the methodology used for determining MTD may provide equally reliable results under different task conditions. Considering that attributes of a task can affect its difficulty level, it follows that they should also be taken into account while determining the MTD.

The selection of MTD was based on the accuracy of predictions at 2-hours. There was some over/under estimation of Predicted RPD and Predicted time to reach R_7, using 26.1 minutes data for T1 and 8.1 minutes data for T2. This indicated changes in the rate at which RPD changed over time. Overestimation indicated that rate of change in RPD decreased with time and underestimation indicated otherwise. Though, there was some over/under estimation of actual final RPD and time to reach R_7 but the appropriateness of this approach is justified by the accuracy of predicted performance (Pass/Fail results) measures using the MTD data.

6.3 General Prediction Model

Quantifying the changes in objective and subjective measures with fatigue progression has been an ongoing quest. It presents an interesting problem from the research as well as application point of view. Various researchers have tried to quantify fatigue progression using various different approaches. Many of these studies have also tried to put forward fatigue indices that use percentage changes in a certain fatigue measures with time. These have included observing changes in the EMG power spectrum and amplitude (Hagberg, 1981, Öberg et al., 1990), decrements in muscle strength (Fitts, 1996), decline in oxygen availability (Murthy et al., 2001), decline in conduction velocity (Krogh-Lund and Jørgensen, 1992), and occlusion of blood flow (De Luca, 1997) among others. These studies allow for comparisons between tasks by providing quantifiable measures to follow fatigue progression but these approaches do not provide any direct indication of fatigue onset nor give any indication about the duration for which the task can be performed (Nussbaum, 2001).

Some other studies have looked at the changes in variability of the fatigue measures. Among others, this approach has included looking at t-tests on the linear fit interactively to increasing sequences of data (Lindström et al, 1977) and looking at the changes in variability with respect to a specified baseline (Gamet et al., 1993). Nussbaum (2001) used a combination of these approaches to come up with a new fatigue index ‘time to fatigue’ (TTF). The general model developed here is a type of approach based on this approach. The variability in the RPD data seemed to be linked to the ability to make correct predictions. It decreased for both T1 and T2 as the ability to make correct predictions increased. Similar trends were observed for the normalized variability. Normalized variability was studied mainly to see if the differences in variability ranges for T1 and T2 could be chaffed out to obtain a more generalizable result. But the results showed that differences existed in the normalized sigma ranges indicating a strong influence of the treatment condition. Also P-value did not show consistent changes, which makes it difficult for it to be used. The reasons for the inconsistency in P-value might have arisen from the presence of some non-monotonic behavior in the data. Other factors that could have affected the general could be the number of the samples present in each category and the time at which the P-value, sigma (slope) and norm-sigma (slope) values were reached. Also, variability inherent to the RPD data might also need to be investigated before more general application of this model could be used.

6.4 Overall Results

Results from this study indicated that shorter experimental durations of less than 8 or 26 minutes, depending on the task, might be sufficient to determine performance for 2 hours. These results are consistent with psychophysical studies that have suggested short 20-45 minutes for predicting load lifting/carrying capabilities for the 2-hour durations (Snook and Irvine, 1967; Legg and Myles, 1981; Willis, 1994). Nussbaum and Johnson (2002) even suggested that for hand intensive tasks 5-minutes is sufficient to determine maximum acceptable limits for a 2-hour task. However, there are some difficulties associated with interpreting results obtained from shorter experiments due to a lack of consensus among studies that have validated these estimates (Mital, 1983; Mital and Manivasagan, 1983). Also, studies have indicated that psychophysical measures obtained from such experiments might be influenced by length of adjustment period (Wu and Chen, 2002) and task characteristics (Mital, 1984). This demands investigation of various factors in search of an optimal protocol for obtaining psychophysical estimates. Thus, further investigation on determining the necessary adjustment periods was warranted, and was undertaken in this research.

Nussbaum and Johnson (2002) suggested that a minimum adjustment period could be estimated for a specific task using psychophysical estimates from a longer period. Following a similar approach, 2-hour data was analyzed using both qualitative and quantitative analysis till the MTD was determined. Results corroborated this approach as performance measures (RPD at 2 hours, time taken to reach R_7, Pass/Fail status on a task) predicted using the MTD were similar to those obtained from the 2-hour trial. Specifically, results for T1, MTD of 26.1 minutes and for T2 with critical RPD level of 3 and for T2, MTD of 8.1 minutes and critical RPD level of 2 was found to be sufficient to make 2-hour predictions. Also, overall results indicated that if a participant reaches RPD = 3 quickly (within 10 minutes), fatigue will continue to increase rapidly.

Additionally, RPD parameters, RPD at 2 hours and the RPD slope, showed good reliability. The ICC and SEM levels concurred with low SEM observed with high values of ICC. Results suggested that both RPD slope and final rating of Borg CR-10 scale for discomfort are reliable parameters. Effect of personality type was not evident but some trends were observed. The average time at which an RPD level was reached was higher for Type A participants. This could be due to Type A people being more competitive and aggressive than their Type B counterparts.

Overall, the results of this study indicate that shorter experimental duration can be used to determine 2-hour performance. The qualitative and quantitative results indicate that the results from shorter duration estimates can provide researcher with good 2-hour performance estimates. The general model developed here provides guidelines to select a balance between desired prediction accuracy and costs. Also, the analysis applied here to study psychophysical estimates has shown promise and might prove useful to researchers for evaluating dynamic task conditions.

6.5 Limitations

The scope of this study was restrained by certain boundaries. These constraints made the study more feasible but at the cost of some loss of accuracy and completeness. Small sample size might have been a limitation for this study. This is also one of the important limitations with most Human Factors studies that this research is trying to prevent. Larger sample size helps lower the inter subject variability and thus help provide more generalizable results. Also, sample size increase might have enhanced the trends showing effect of personality type.

Another important limitation is that only two variation of an intermittent overhead-tapping task were studied. This limits the applicability of quantitative results from this study to other task conditions. Additionally, the two treatments studied varied vastly in their difficulty level, which influenced the RPD trends. Varying levels of treatment difficulty levels need to be studied to understand their influence on the RPD trends.

Using a fixed experimental duration of 2 hours rather than studying the task until the endurance limit might have also affected the comparisons drawn between treatments T1 and T2. Most of the participants did not reach their endurance limit for T1 in 2-hours but they reached their endurance limit fairly quickly ($t < 30$ minutes) for T2. Thus, RPD data for T2 reflected the changes over time spanning the entire range of the RPD scale, but this wasn't the case for T1. Studying T1 until endurance limit might have provided a more comprehensive picture of RPD trends against which to evaluate the general prediction model.

Quantitative results were based on the linear model fit of the RPD data. The assumptions (e.g. residuals, predicted minus observed values, should be normally distributed) required for the application of the linear regression analysis were not checked. If unsatisfied, these could have influenced the results. In the case where linear regression is not applicable, it might be difficult to use some of the results, as some of the deductions might be dependent on the monotonicity of the linear regression relationship. Another limitation hindering the applicability of the results is the lack of validation. This can be an important consideration for this study where trial duration reduction is proposed, especially as literature documents limited and even then conflicting results regarding the use of shorter trials to derive long durations estimates. Additionally, results of this study supporting the use of shorter trials were derived from the 2-hour data. If the experiment is only conducted for short MTD duration the results may vary. Thus, validation is necessary to prove the correctness of the measures derived from short MTD trials.

Another limitation pertains to the protocols of the experiment. In this study an intermittent task was studied where the posture was unconstrained. The participants were allowed to adjust their posture during the task. This could lead to the reduction in shoulder discomfort as measured through RPD. Though this makes the task more close to real life it might have influenced the RPD data.

6.6 Applications

Research applications: Results of this study can be used to study other similar tasks in shorter trials. The MTD derived from this study could be directly applied to study a similar task after validation. Specifically, a study could be undertaken to validate the MTD and then the results applied to larger study. The methodology used can be applied to design a pilot study for another task and then the results used to design the parameter for studying other larger study. This can mean that a larger task conditions might be studied in a relatively shorter time. Shorter trials can imply that participants are performing different tasks in the same duration leading to savings in money and time while increasing the scope of a study.

Industrial applications: Shorter trials can mean faster evaluation of task parameters of a job. Also, this study indicates that RPD might be a very economical and no-intrusive tool, which can very effectively be used for intermittent tasks evaluation. Also, many industries have begun employing 2-hour job rotation, 2-hour work periods are becoming very common. This increases the applicability of the results from this study.

6.7 Future Research

This study envisages shortening trial duration, increasing experimental variables, testing larger sample sizes faster, and showing RPD as a reliable measure for studying intermittent tasks. Results of this study do provide support for that vision but its many limitations restrict its scope and applicability. Future research needs to be conducted to address these.

Many different tasks spanning a range of difficulty levels should be studied using similar procedures. Also, it might help to conduct the experiments until endurance limit is reached to help crystallize the observed patterns in the data. Observing the tasks until endurance limit would also make comparison between treatments more appropriate and robust. Also, increasing the sample size might help to typify, if any, effects of gender and personality. Before the results could be applied, they need to be validated by examining the prediction made from a short trial against the ones obtained from a 2-hour trial.

7.0 CONCLUSIONS

This research provides support for possible reductions in the experimental durations. Gain due to increased efficiency of the data collection process may offset the relatively minor sacrifice of accuracy of the results. Good correspondence between the predicted and actual measures shows that careful selection of the trial duration can allow for excellent estimations of the predicted measures. The general model indicates some unifying trends, which might be further explored for other task situations. Use of subjective ratings as the measure of performance might be extremely useful in situations where it is difficult to obtain objective measures or where the objective measures do not yield good results. RPD provides a more comprehensive measure of the fatigue and is much easier and efficient to collect than objective measures like EMG and heart rate. Also, this study provided indications of the minimum trial duration. Although only a specific type of overhead tasks was studied, these results do indicate that brief trial duration might aid in rapidly determining dependable psychophysical estimates.

REFERENCES

Anderson HA (1984) Arthrosis and its relation to work. Scandanivian Journal of Work, Environment, and Health 10:429-433

Armstrong TJ, Radwin RG, Hansen DJ, Kennedy KW (1986) Repetitive trauma disorders: job evaluation and design. Human Factors 28:325-336

Asfour SS, Ayoub MM, Mital A, Bethea NJ (1983) Perceived exertion of physical effort for various manual handling tasks. American Industrial Hygiene Association 44(3):223-228

Astrand PO, Rodahl K (1986) Textbook of Work Physiology, 3rd ed. McGraw-Hill, New York

Baidya KN, Stevenson MG (1988) Local muscle fatigue in repetitive work. Ergonomics 31(2):227-239

Basmajian JV (1978) Cyclobenzaprine hydrochloride effect on skeletal-muscle spasm in lumbar region and neck – 2 double-blind controlled clinical and laboratory studies. Arch Phys Med Rehab 59(2):58-63

Basmajian JV, DeLuca CJ (1985) Muscles Alive: Their Functions Revealed by electromyography. Williams & Wilkens 1-20.

Bureau of Labor Statistics (2001) US Department of Labor.

Bjelle A, Hagberg M, Michaelsson G (1979) Clinical and ergonomic factors in prolonged shoulder pain among industrial workers. Scandanivian Journal of Work, Environment, and Health 5(3):205-210

Bjelle A, Hagberg M, Michaelsson G (1981) Occupational and individual factors in acute shoulder-neck disorders among industrial workers. British journal of industrial medicine 38(4):356-363

Björkstén MG, Boquist B, M. Talbäck, Edling C (2001) Reported neck and shoulder problems in female industrial workers: the importance of factors at work and at home. International Journal of Industrial Ergonomics 27(3):159–170

Bongers PM, Dewinter CR, Kompier M, Hildebrandt VH (1993) Psychosocial factors at work and musculoskeletal disease. Scandinavian Journal of Work, Environment, and Health 19(5):297-312

Borg G (1962) Psychophysical basis of perceived exertion. Medical Science Sports Exercise 14:371-381

Borg G (1990) Psychophysical scaling with applications in physical work and the perception of exertion. Scandanivian Journal of Work, Environment, and Health 16(1):55-58

Carpenter JE, Blasier RB, Pellizzon GG (1998) The effects of muscle fatigue on shoulder joint position sense. American Journal of Sports Medicine 26(2):262-265

Chaffin DB (1973) Localized muscle fatigue: definition and measurement. Journal of Occupation Medicine 15:346-354

Chiang HC; Ko YC, Chen SS, Yu HS, Wu TN, Chang PY (1993) Prevalence of shoulder and upper-limb disorders among workers in the fish-processing industry. Scandinavian Journal of Work, Environment, and Health 19(2):126-131

Cobb S and Forbes A (1923) EMG studies of muscular fatigue in man. American Journal of Physiology 64(2):234-251

DeLuca CJ (1984) Myoelectrical manifestation of localized muscular fatigue in humans. CRC Critical Reviews in Biomedical Engineering 11(4):251-279

DeLuca CJ, Erim Z (1994) Common drive of motor units in regulation of muscle force. Trends Neurosci 17(7):299-305

De Luca CJ (1997) The use of surface electromyography in biomechanics. Journal of Applied Biomechanics 13:135-163.

Engin AE (1980) On the biomechanics of the shoulder complex. Journal of biomechanics 13(7):575-590

Friedman M, Rosenman RF (1974) The Type A behavior and your heart. New York: Knopf.

Fleiss JL (1986) The design and analysis of clinical experiments. John Wiley & Sons, New York:2-12.

Fitts RH, Widrick JJ (1996) Muscle mechanics: adaptations with exercise training. Exercise and Sports Science Review 24: 427-473.

Gamet D, Duchene J, Bar-Garapon C, Goubel F (1993) Surface electromyogram power spectrum in human quadriceps muscle during intermental exercise. Journal of Applied Physiology 74:2704-2710.

Gerdle B, Eriksson NE, Brundin L, Edstrom M (1988) Surface EMG recordings during maximum static shoulder forward flexion in different positions. European Journal of Applied Physiology and Occupational Physiology 57(4):415-419

Glasscock NF, Turville KL, Joines SB, Mirka GA (1999) Biomechanics-The Effect of Personality Type on Muscle Coactivation during Elbow Flexion. Human factors 41(1):51-60

Hagberg M (1981) Muscular endurance and surface electromyogram in isometric and dynamic exercise. Journal of Applied Physiology 51(1):1-7

Herberts P, Kadefors R (1976) Study of painful shoulder in welders. Acta Orthopaedica scandinavica 47(4):381-387

Jenkins CD, Zyzanski SJ, Rosenman RH (1971) Progress towards validation of a computer-scored test for type-A coronary-prone behavior pattern. Pychosomatic Medicine 33(3):193-202

Jenkins, CD, Zyzanski SJ, Rosenman RH (1979) Jenkins Activity Survey. JAS Manual C 1-31.

Kadefors R, Peteren I, Herberts P (1976) Muscular reaction to welding work electromyographic investigation. Ergonomics 19(5):543-558

Krogh-Lund C, Jørgensen K (1992) Modification of myo-electric power spectrum in fatigue from 15% maximal voluntary contraction of human elbow flexor muscles to limit of endurance: reflection of conduction velocity variation and/or centrally mediated mechanism. European Journal of Applied Physiology 64:359-370.

Kuorinka I, Forcier L (1995) Work-related Musculoskeletal Disorders (WMSDs): A Reference Book for Prevention, London, England: Taylor & Francis

Lindstrom L, Kadefors R, Peterson I (1977) An electromyographic index for localized muscle fatigue. Journal of Applied Physiology 43:750-754.

Ljungberg AS, Gamberale F, Kilbom A (1982) Horizontal lifting-physiological and psychological responses. Ergonomics 25:741-757

Lohr JF, Uhthoff HK (1990) The Microvascular Pattern of the Supraspinatus Tendon. Clinical orthopaedics and related research 254:35-38

Malchaire B, Roquelaure Y, Cock N, Piette A, Vergracht S and Chiron H (2001) Musculoskeletal complaints, functional capacity, personality and psychosocial factors. International archives of occupational and environmental health 74(8):549-557

McCormack RR, Inman RD, Wells A, Berntsen C, Imbus HR (1990) Prevalence of tendonitis and related disorders of the upper extremity in a manufacturing workforce. Journal of. Rheumatology 17:958–964

Mital A, Manivasagan I (1983) Subjective estimates of one-handed carrying tasks. Applied Ergonomics 14(4):265-269

Mital A (1983) The psychophysical approach in manual lifting-A verification study. Human Factors 25:485-491

Mital (1984) Comprehensive maximum acceptable weight of lift database for regular 8-hour work shifts. Ergonomics 27(11):1127-1138

Moore A, Wells R, Ranney D (1991) Quantifying exposure in occupational manual tasks with cumulative trauma disorder potential. Ergonomics 34(12):1433-1453

Morrissey SJ, Liou YH (1988) Maximum acceptable weights in load carriage. Ergonomics 31(2)217-226

Murthy G, Kahan NJ, Hargens AR, Rempel DM (1997) Forearm muscle oxygenation decreases with low levels of voluntary contraction. Journal of Orthopaedic Research 15:507-511.

NIOSH (1997) Musculoskeletal disorders and workplace factors: A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back. National Institute for Occupational Safety and Health (NIOSH) Pub. No. 97–141.

Nussbaum MA (2001) Static and dynamic myoelectric measures of shoulder muscle fatigue during intermittent dynamic exertions of low to moderate intensity. European Journal of Applied Physiology 85(3-4):299-309

Nussbaum MA, Clark LL, Lanza MA, et al. (2001) Fatigue and endurance limits during intermittent overhead work. American Industrial Hygiene Association Journal 62(4):446-456

Nussbaum MA, Johnson H (2002) Determination and Evaluation of Acceptable Force Limits in Single-Digit Tasks. Human factors 44(4):545-556

Nussbaum M, Sood D, Hager K (2002) Honda Technical Report: Effects of differing overhead heights on shoulder fatigue during a repetitive intermittent task. Honda Automotive Manufacturing Inc., Maysville, Ohio.

Nussbaum MA, Johnson H (2002) Determination and Evaluation of Acceptable Force Limits in Single-Digit Tasks. Human Factors 44(4):545-556

Öberg T, Sandsjo L, Kadefors R (1990) Electromyogram mean power frequency in non-fatigued trapezius muscle. European Journal of Applied Physiology 61: 362-369

Occupational Safety and Health Administration (1999). Report of the Small Business Advocacy Review Panel on the Draft Proposed Ergonomics Program Rule.

Putz-Anderson V (1988) Cumulative trauma disorders: manual for the musculoskeletal diseases of the upper limbs, Taylor and Francis, Bristol, PA

Putz-Anderson V, Galinsky TL (1993) Psychophysically determined work durations for limiting shoulder girdle fatigue from elevated manual work. International Journal of Industrial ergonomics 11:19-28

Rosecrance JC, Cook TM (1996) Ergonomic intervention in newspaper production in Orthopaedic and physical therapy clinics of North America. 5(2):273-286

Rosecrance JC, Cook TM (1996) Ergonomic intervention in newspaper production. Orthopaedic and Physical Therapy Clinics of North America 5(2):273-286.

Sakakibara H, Miyao M, Kondo T, Yamada S (1995) Overhead work and shoulder-neck pain in orchard farmers harvesting pears and apples. Ergonomics 38(4):700-706

Salminen JJ, Pentti J, Wickstrom G (1991) Tenderness and pain in neck and shoulders in relation to type A behavior. Scandinavian journal of rheumatology 20(5):344-350

Shankar S, Gander RE, Brandell BR (1989) Changes in myoelectric signal (MES) power during dynamic contractions. Electroen Clin Neuro 73(2):142-150

Sparacino J (1979) Type A behavior pattern - critical assessment. Journal of human stress 5(4):37

Silverstein BA, Fine LJ, Armstrong TJ (1986) Hand wrist cumulative trauma disorders in industry. British Journal of Industrial Medicine 43:779-784

Silverstein BA, Fine LJ, Armstrong TJ (1987) Occupational factors and carpal tunnel syndrome. British Journal of Industrial Medicine 11:343-358

Skinner HB, Wyatt MP, Hodgdon JA, Conard DW, Barrack RL (1986) effect of fatigue on joint sense of the knee. Journal of Orthopaedic Research 4(1):112-118

Snook S (1978) The design of manual handling tasks. Ergonomics 21:963-985

Sommerich CM, McGlothlin JD, Marras WS (1993) Occupational risk factors associated with soft tissue disorders of the shoulder: a review of recent investigations in the literature. Ergonomics 36(6):697-717

Sood D, Hager K, Nussbaum M (2002) The effects of differing overhead heights on shoulder fatigue during a repetitive intermittent task. Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting: 1081-1085.

Straker LM, Pollock CM, Mangharam JE (1997) The effect of shoulder posture on performance, discomfort and muscle fatigue whilst working on a visual display unit. International Journal of Industrial Ergonomics 20(1):1-10

Ulin SS, Ways CM, Armstrong TJ, Snook SH (1990) Perceived exertion and discomfort versus work height with a pistol-shaped screwdriver. American Industrial Hygiene Association Journal 51(11):588-594

Wickstrom G, Pentti J, Hyytianen K (1989) Type A behavior and back pain. Work and Stress 3(2):203-207

Wiker SF, Langolf GD, Chaffin DB (1989) Arm postures and human movement capability. Human Factors 31(4):421-441

Wright WC (1940) Diseases of Workers (The Latin text De Morbis Artificum Diatriba of Ramazzini, 1713, translated by Wright). University of Chicago Press.

Wu SP, Chen JP (2003) Effects of the adjustment period on psychophysically determined maximum acceptable weight of lift and the physiological cost. International Journal of Industrial Ergonomics 31:287-294

APPENDICES

Appendix A Personality Type Determination Form

This checklist was developed using descriptions of Type A people by Friedman and Ulmer (1984), Mathews et al. (1982) and Musante et al. (1983).

1. I strongly accent keywords in my everyday speech

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

2. I eat and walk quickly

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

3. I believe that children should be taught to be competitive

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

4. I feel restless when watching a slow worker

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

5. I hurry other people to get on with what they are trying to say

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

6. I find it highly aggravating to be stuck in traffic or waiting for a seat at a restaurant

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

7. I continue to think about my own problems and business even when listening to someone else

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

8. I try to eat and shave, or drive and jot down notes at the same time

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

9. I catch up on my work on vacations

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

10. I bring conversation around to topics of concern to me

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

11. I feel guilty when I spend time just relaxing

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

12. I find that I am so wrapped up in my work that I no longer notice office decorations or the scenery when I commute

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

13. I find myself concerned with getting more things done rather than developing my creativity and social concerns

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

14. I try to schedule more and more activities into less time

Strongly Disagree | Disagree | Undecided | Agree | Strongly Agree

15. I always appear for appointments on time

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

16. I clench or pound my fists, or use other gestures, to emphasize my views

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

17. I credit my accomplishments to my ability to work rapidly

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

18. I feel that things must be done now and quickly

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

19. I constantly try to find more efficient ways to get things done

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

20. I insist on winning at games rather than just having fun

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

21. I interrupt others often

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

22. I feel irritated when others are late

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

23. I leave table immediately after eating

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

24. I feel rushed

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

25. I feel dissatisfied with my current level of performance

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree

Appendix B Instructions for Using RPD scale

You should imagine that you are working on an assembly line. You work for 8 hours 5 days a week, with a 30 minutes lunch break and two short 15-minute tea breaks. You want to go back home every day without feeling excessively tired, sore or weekend in your shoulder.

Instruction for ratings of perceived discomfort (RPD)

- You have to rate the pain or discomfort that you feel in your dominant shoulder at the end of the working portion of the duty cycle.
- Rating scale should be used **only** to rate the pain in the dominant shoulder. Pain or discomfort in any other body part (e.g. forearm, neck, back) should be reported separately to the experimenter.
- ‘0’ on the scale indicates no pain or a completely relaxed shoulder.
- ‘10’ on scale means excruciating pain or extreme discomfort. You are close to the point where you cannot do the task any more.
- ‘3’ means that shoulder is a bit painful but it is not affecting the performance.
- ‘8’ means that your shoulder is very painful and uncomfortable but you can still do the task.
- You are not in any competition. Your ratings are not compared to that of others.
- You are encouraged to rate honestly without being influenced.
- Your ratings are very important to us and will be used to determine guidelines for acceptable workloads.

Appendix C Mass Determination for RPD Practice

Mass used for RPD practice was determined using the maximum voluntary exertion of medial deltoid muscle. Maximum voluntary exertion (MVE) data was collected using surface electromyography (EMG) data was collected but was not used towards the present research

EMG data collection preparation

EMG was collected from medial deltoid, an easily accessible shoulder muscles. EMG data was collected using pre-gelled bipolar Ag/AgCl disposable surface electrodes. Two electrodes were placed parallel to each other on the belly of the muscle, with 1-2 cm inter-electrode distance. The location for electrode placement was determined using guidelines as described by SENIAM (2001). A ground electrode was also placed on the bony process of the clavicle.

During electrode placement participants maintained 90° shoulder flexion with arm weight supported by a platform. The electrodes were placed midway between the acromion and the deltoid insertion point. Skin preparation was done before electrode placement, by shaving, abrading and cleansing the skin with alcohol. The electrodes were placed on the muscle belly perpendicular to the length of the muscle fibers. Inter-electrode resistance was measured after 5 minutes using an ohmmeter. If the resistance was high (> 10 ohm), the skin was prepared again. Once the resistance had been determined to be low, the lead wires were attached to the electrodes. EMG data was routed through a preamplifier to amplifying (x 100) the data closer to the source. An amplifier was then be used to further amplify the collected data by a set gain, determined by the experimenter for each experimental session. Gain was set for each muscle separately such that the signal is within ±10V, a limitation imposed by the analog to digital signal board. A time constant of 110 ms was used to get the RMS data. Raw EMG signal was be sampled at 2048 Hz and recorded in a computer using LabView software.

MVE data collection

Maximum Voluntary Exertion (MVE) was performed using the dominant hand. MVE allowed for estimation of strength and maximum electromyographic (EMG) activity levels for the medial deltoid, which was taken as the estimate of the arm strength. MVE was performed in the posture that is consistent with those used by Nussbaum et al. (2001) in a similar overhead study. MVE was obtained using the ramp up–hold–ramp down procedure in 5-second exertion (Figure A). Participant was given standardized instructions before starting the MVE data collection. Participants were given approximately 5 minutes practice on the MVE procedures. This time was considered sufficient based on the observations made during the Pilot study. After this, the MVE data collection began. The data was recorded using a customized LabView program with a user interface that displays exerted force and the EMG activity on the main screen.

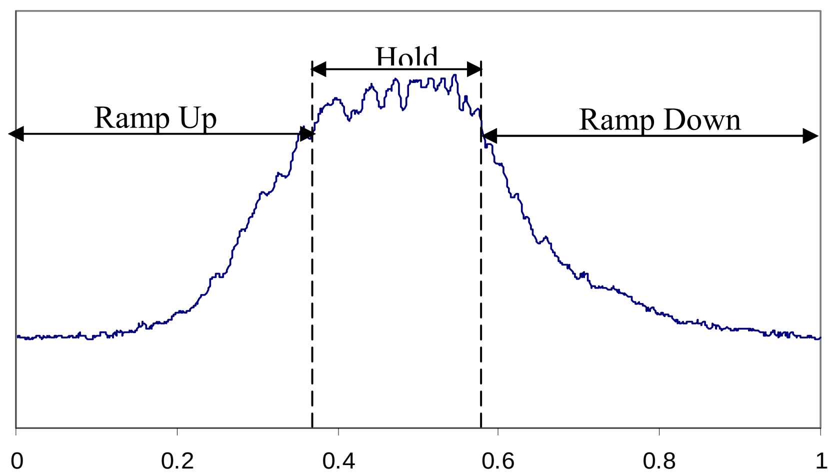


Figure A EMG data from a sample MVE trial obtained using the ramp up – hold – ramp down procedure.

Participant used muscle specific fixtures (harnesses and handgrips), to standardize the posture during MVE data collection. Participant held their dominant arm in 90^0- shoulder abduction with elbow flexed at 90^0 and palm facing downwards. Strap attached to an adjustable chain, connected to the load cell, were placed just behind the elbow or at distal portion of the humerus. Length of the chain was adjusted till it had minimum slack. Participants pushed against the strap during the exertion. Each participant completed at least 3 repetitions of MVEs for each of the three muscles and the task MVE. Additional repetitions were done if the exerted force showed an increase in the third trial, otherwise 3 trials were considered sufficient. MVE trials were stopped only when the exerted force values had stabilized or had begun to decrease. Peak EMG values obtained during these MVE were used to determine the hand weight used for RPD practice.

MVE Instruction given to participant

- Each MVE exertion has to be completed in 5 seconds.
- Multiple readings will be taken for each muscle.
- You can take as much time as required for practice for each exertion before data collection begins.
- Maintain the designated posture throughout the (Ramp up - Hold - Ramp down) process.
- After hearing the start beep, wait for a second. Then take about a second to slowly build up your maximum. Hold your maximum force for a second. Then slowly start decreasing your force till you feel completely relaxed.
- You are encouraged to try exerting your maximum force.

Appendix D IRB Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING (ISE)

Informed Consent for Participants of Investigative Projects

Title of Project: Predicting Shoulder Fatigue using long durations using psychophysical measures obtained her short trials

Principal Investigators: Dr. M. A. Nussbaum, Assistant Professor, ISE Department
Deepti Sood, Graduate Research Assistant, ISE Department

I THE PURPOSE OF THIS RESEARCH

You are invited to participate in a study to determine limits for upper-extremity overhead work related to automotive assembly tasks. To obtain this information, two experiments are to be conducted. The first is designed to determine the influence of task height. The second experiment will involve simulation of overhead work in a variety of conditions. It is anticipated that there will be approximately 6 participants for the first experiment, and 46 participants for the second experiment (6 of which are pilot studies).

II PROCEDURES

The procedures used in this study are as follows:

1) You will have electrodes placed on three muscles, which move the shoulder. These electrodes are used to collect information from the muscles, which can indicate fatigue levels. The procedure for each electrode involves cleansing a small patch of skin (approximately the size of two quarters) over the muscle area. The electrodes are then placed on the skin and remain in place with an adhesive.
2) The investigator will demonstrate the data collection procedures, which involve performing overhead work tasks at various heights, or performing overhead work tasks at various work-cycle durations and exertion levels.
3) You will conduct simulated overhead work cycles as demonstrated by the investigator with rest periods.
4) For this experiment, each participant will perform simulated overhead work for a maximum of one hour at three different heights.

The total estimated time of participation is 3 to 4 hours (including rest periods) for this experiment.

III Risks and Benefits of this Research

Your participation in this study will provide information that will be used to develop design guidelines for overhead work. It is the objective of this study to contribute design information for improving worker safety, comfort, and productivity.

The primary focus of this study is to measure muscle fatigue. Therefore, you may experience some discomfort related to extended use of some muscles. The muscle fatigue will occur due to use over a long period of time with regular breaks, and not due to generation of large forces. In addition, an investigator will continuously monitor your condition to minimize any opportunity of strain.

There is minimal risk involved in this study.

IV Extent of Anonymity and Confidentiality

It is the intent of the investigators of this project to report the findings of this study. The information you provide will have your name removed and only a subject number will identify you during analysis and any written reports of the evaluation.

V Compensation

If you decide to participate in this study, you will be paid $10.00 per hour for the time you participate and $10 as bonus at the end of all the experimental sessions. The evaluation is expected to last 3-4 hours depending on the experiment. You will be paid at the conclusion of the testing session.

VI Freedom to Withdraw

You are free to withdraw from this study at any time for any reason without penalty. If you choose to withdraw during the study, you will be compensated for the portion of the testing which has been completed.

VII Approval for This Research

This research project has been approved, as required, by the Institutional Review Board for projects involving human participants at Virginia Polytechnic Institute and State University, and by the Grado Department of Industrial Engineering.

VIII PARTICIPANT RESPONSIBILITIES

I know of no reason why I cannot participate in this study. I have the following responsibilities:

- To notify the investigator at any time about a desire to discontinue participation.
- To notify the investigator of any medical conditions that may be negatively influenced by extended muscular exertion. This may include heart disease, conditions influenced by blood sugar levels, or any other medical problems that may interfere with results or increase the risk of injury or illness.

____

Signature of Participant

IX PARTICIPANT'S PERMISSION

Before you sign the signature page of this form, please make sure that you understand, to your complete satisfaction, the nature of the study and your rights as a participant. If you have any questions, please ask the investigator at this time. Then, if you decide to participate, please sign your name above and on the following page (please repeat for your copy).

SIGNATURE PAGE

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I choose to do so.

Signature ___________________________

Printed Name ___________________________

Date ___________________________

The research team for this experiment includes Dr. M. A. Nussbaum, Assistant Professor and Deepti Sood, Graduate Research Assistant. Research team members may be contacted at the following address and phone number:

Grado Department of Industrial and Systems Engineering Department
250 New Engineering Building
Virginia Tech
Blacksburg, VA 24061
(540) 231-6053

In addition, if you have detailed questions regarding your rights as a participant in University research, you may contact the following individual:

Dr. David Moore
Chair, Institutional Review Board
CVM Phase II (Pathobiology)
Virginia Tech
Blacksburg, VA 24061
(540) 231-4991

RESUME

DEEPTI SOOD
(Email: dsood@vt.edu, Phone: 848-459-3464)

EDUCATION

Virginia Polytechnic Institute and State University, VA
M.S. in Industrial and Systems Engineering (GPA: 3.89/4.0) Aug'01-present
(Biomedical Engineering Option)

National Institute of Fashion Technology, New Delhi, India
Post Graduate Diploma in Garment Manufacturing Technology May'98-Jun'00

University of Delhi, Delhi, India
Bachelor of Arts in Mathematics Apr'94-May'97

RESEARCH EXPERIENCE

- **Graduate Research Assistant**

Industrial Ergonomics Laboratory, Dept. of Industrial and Systems Engineering, Virginia Tech
Advisor: Dr. Maury A Nussbaum

Involved with three research projects for Honda of America Manufacturing Inc. (HAM). Worked in team with other graduate students/s. Job included help with designing and conducting experiments, supervising graduate/undergraduate students, and writing technical reports.

Recommendations for manual torquing tasks for HAM assembly line May'03-present
Developing guidelines to reduce shoulder fatigue in industrial torquing tasks, based on laboratory simulations, using heart rate monitor and psychophysical measures

Recommended limits for overhead tasks for HAM assembly line May'01-Jul'02
Developed recommendations for overhead tasks from laboratory simulated experiments using electromyography, psychophysical measures, and force place data

Video task analysis of moving industrial carts for HAM May'01-Nov'01
Determined instantaneous position, which was used to determine push/pull forces for moving industrial carts in an assembly line, by digitizing HAM assembly videos

- **Masters Thesis** Jan'03-May'04

Industrial Ergonomics Laboratory, Dept. of Industrial and Systems Engineering, Virginia Tech

Predicting shoulder fatigue for long durations using psychophysical measures obtained from short trials
Determine plausible reductions in the experimental duration of a fatiguing intermittent overhead task and then verify the results obtained from shorter duration experiments

- **Graduate Diploma Project** Jan'00-May'00

National Institute of Fashion Technology, New Delhi, India

Ergonomics, an application to the garment industry
Developed and implemented ergonomic changes to enhance productivity in a garment assembly line at Jaipur Polo Company, India. Up to 15% reduction in cycle time and qualitative increase in worker satisfaction was achieved

TEACHING EXPERIENCE

Graduate Teaching Assistant (Industrial Ergonomics ISE 3624) Fall'02 and Spring'03
Responsibilities included holding office hours, help setting and evaluating homework and exam questions

PROJECTS

- Estimated **Endurance Times** and **Onset of Fatigue** for Painting Task
- Developed **Smart Route Guide System,** an application of GPS
- Suggested **improvements in Blacksburg Transit Information System** for VT Students
- Designed **Work Aids** to incorporate visually challenged into Garment Industry
- Prepared **Feasibility Study** for a garment manufacturing unit (Ministry of Textiles, India)
- Studied **Work Cultures** under the Interactive Industry Learning Module

INDUSTRIAL EXPERIENCE

- **Management Internee**, SHIVAM (Garment Export House) Jun'99-Aug'99
  Worked as a Production Supervisor and Assistant Merchandiser
- **Management Trainee**, FICCI Ladies Organization (F.L.O.) Aug'98-Oct'98
  Facilitated business conferences

PUBLICATIONS

- Deepti Sood, Kris Hager, and Maury A. Nussbaum (2002) The effects of differing overhead heights on shoulder fatigue during a repetitive intermittent task. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting.* Baltimore, MD. pp 1081-1085
- Deepti Sood, Maury A. Nussbaum, and Kari L. Babski-Reeves (2004) Effects of work conditioning and adjustment period on psychophysical estimates in manual torquing tasks (*Submitted to Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting)*

ACHIEVEMENTS AND AWARDS

- **President**, Human Factors and Ergonomics Society (HFES) Virginia Tech. Chapter (2003)
  - *Chapter was awarded* '**Outstanding Student Chapter of the year**' *that year*
- **Vice President**, American Society of Safety Engineers (ASSE) Virginia Tech. Section(2003-04)
- Awarded **Outstanding Teaching Assistant** of the year (2002-03)
- Awarded the **UDEL Fellowship** for the year by University of Delaware (2000-01)
- Awarded **Best Diploma Project** at national level, NIFT, India (May 2000)
- **Treasurer**, HFES VT Chapter (2002)
- **Secretary**, ASSE VT Section (Aug'02-May'03)
- **President** and **Founding member**, Social Service Society, NIFT (Oct'98-Apr'00)
- **Anchored** shows for All India Radio, the National Broadcasting Channel (Spring'99)

SKILLS

- Proficient in ergonomics tools: Electromyography, Force Plate, Heart Rate Monitor
- Statistical Packages: JMP, SAS
- Software: LabVIEW, MSOffice, Arena, Adobe PhotoShop, CAD/CAM

MEMBERSHIP

- Student member, Human factors and Ergonomics Society
- Student member, American Society of Safety Engineers
- Student member, Association for Women in Mathematics
- Member of Student Government Association, Virginia Tech

INTERESTS

- Reading and Writing (contributed reviews and poetry to school and college magazines)
- Dancing (learnt Indian classical dance 'Bharatnatyam' for 5 years)