

**Development and Evaluation of an Observational Method for
Quantifying Exposure to Hand Activity and Other Physical Stressors
in Manual Work**

by

Wendi Ann Latko

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in The University of Michigan
1997

Doctoral Committee:

Professor Thomas J. Armstrong, Chair
Associate Professor Alfred Franzblau
Professor Gary D. Herrin
Associate Professor Bernard J. Martin

© Wendi Ann Latko 1997
All Rights Reserved

ACKNOWLEDGMENTS

Primary financial support for this research was provided by NIOSH/CDC grant #R01-OH02941 and gifts from private industry. A University of Michigan Regents Fellowship and a GE Predoctoral Fellowship provided support during the first and last years of this research. Other funds and laboratory facilities were provided by the Center for Ergonomics.

I wish to thank my advisor, Professor Thomas Armstrong, for his guidance and support throughout this research project. I would also like to thank the other members of my dissertation committee, Professors Al Franzblau, Bernard Martin, and Gary Herrin, for their advice and input during this research.

I also wish to acknowledge and thank the core team which has worked together to develop and refine the job analysis method presented in this work: Tom Armstrong, Jim Foulke, Randy Rabourn, and Sheryl Ulin. Jim Foulke provided unique insight and perspective in the job analysis process, as well as invaluable advice concerning equipment and instrumentation. Randy Rabourn contributed job analysis expertise as well as encouragement during the process. Sheryl Ulin spent many days on the road with me traveling to plants to select jobs, performed the data analysis for Chapter 5, and generally offered advice, support and encouragement. As a whole, the team's insight, cooperation, and patience during the sometimes seemingly endless rating sessions allowed the method to evolve into its current form. Their professional expertise as well as their friendship have been invaluable to me.

I must also thank the entire Center for Ergonomics staff for their support in this research, especially Chuck Woolley, Rod Capps, and Kelly Cormier. In addition, the other graduate students in the labs have done much to make this work possible. Alisa Bono, Mary Kelly Aylward, and Julia Greenwald all assisted with various aspects of the studies. Michelle Homan, Kim Monroe, and Xudong Zhang provided encouragement, support, and helped me put the research into perspective.

My parents, Stan and Sondra, have always expected the best from me and supported me in all my endeavors. Finally, I wish to thank my husband, Michael Gerard, for his support during this research. His advice, encouragement, and ability to make me laugh have enabled me to get through this process. I am grateful for his unending patience and support.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF FIGURES.....	vi
LIST OF TABLES	viii
LIST OF APPENDICES	x
CHAPTER	
1. INTRODUCTION.....	1
Problem Statement.....	1
Background.....	1
Research Objectives	12
Dissertation Organization.....	12
References.....	14
2. DEVELOPMENT AND EVALUATION OF AN OBSERVATIONAL METHOD FOR ASSESSING REPETITION IN HAND TASKS.....	22
Abstract.....	22
Background.....	23
System development	27
Evaluation (Methods/results).....	31
Discussion.....	36
Conclusions/future work:	40
References.....	43
3. EVALUATION OF INTERRATER RELIABILITY AND TEST-RETEST RELIABILITY OF OBSERVATIONAL RATINGS OF PHYSICAL STRESSORS IN MANUAL WORK.....	47
Abstract.....	47
Background.....	48
Methods.....	50
Results	55
Discussion.....	64
References.....	73

4. COMPARISON OF OBSERVATIONAL AND INSTRUMENTAL BASED MEASUREMENTS OF REPETITION, FORCE, AND WRIST POSTURE IN MANUAL WORK	80
Abstract.....	80
Introduction	81
Methods.....	82
Procedure.....	84
Results	87
Discussion.....	93
References.....	99
 5. A CROSS-SECTIONAL STUDY OF THE RELATIONSHIP BETWEEN REPETITIVE WORK AND THE PREVALENCE OF UPPER LIMB MUSCULOSKELETAL DISORDERS.....	103
Abstract.....	103
Background.....	104
Methods.....	107
Results	116
Discussion.....	122
References.....	128
 6. DISCUSSION.....	135
Summary of findings	135
1) Ratings of hand activity/repetition exhibit varying levels of agreement with previous metrics for assessing repetitive work.	135
2) Inter- and intra-rater reliability are strong for hand activity ratings, low to moderate for other stressors. The use of the group consensus process generally increases the consistency of ratings.....	136
3) Observational ratings of repetition, force, and wrist posture are consistent with direct measurements of the same parameters.....	137
4) Increased levels of hand activity is associated with increased prevalence of upper limb musculoskeletal disorders in workers.....	137
Discussion of Findings.....	138
Recommendations for future work.....	147
 APPENDICES	153

LIST OF FIGURES

Figure 1.1:	Proposed exposure-response curve for work-related upper extremity musculoskeletal disorders.	4
Figure 2.1:	Proposed exposure-response curve for work-related upper extremity musculoskeletal disorders. Studies have shown higher risk at higher exposure levels (Andersson, 1995, Stock, 1991). The exact shape of the curve is not known.	25
Figure 2.2:	Flow chart illustrating the steps involved in job selection and system development and application.	28
Figure 2.3:	Visual-analog scale for rating repetition/hand activity, with verbal anchors.	31
Figure 2.4:	Relationship between team ratings of repetition/hand activity and number of exertions per second in 33 jobs.	32
Figure 2.5:	Typical work/recovery profiles for jobs at various levels of the rating scale.	33
Figure 2.6:	Relationship between team ratings of repetition/hand activity and % recovery time within the task cycle for 33 jobs.	34
Figure 2.7:	Plot of cycle time and team rating of repetition/hand activity for 33 jobs. There was no significant relationship.	35
Figure 2.8:	Average team ratings of repetition on two different occasions for 33 jobs ($r^2 = 0.88$)	36
Figure 3.1:	Average initial and final repetition ratings (± 1 SD), for 17 raters and 12 jobs. Job labels correspond to the jobs listed in Table 3.4, and are shown in order of increasing mean magnitude.	56
Figure 3.2:	Average initial and final force ratings (± 1 SD), for 17 raters and 12 jobs.	56
Figure 3.3:	Average initial and final finger contact stress ratings (± 1 SD), for 17 raters and 12 jobs.	57

Figure 3.4:	Average initial and final wrist posture ratings (± 1 SD), for 17 raters and 12 jobs.	57
Figure 3.5:	Agreement in average group ratings between Groups 2 and 3 for 45 jobs.	61
Figure 3.6:	Agreement in Group 2 average ratings at times 1 and 2 for 12 jobs.	64
Figure 4.1:	Ten cm visual-analog scales for rating repetition/hand activity (a), forceful exertions (b), wrist posture (c), and wrist movement speed (d).	85
Figure 4.2:	Repetition ratings as a function of average angular velocity (a), and repetition ratings as a function of % time muscles were active below 2.5% MVC (b).....	89
Figure 4.3:	Observational rating of peak and average force as a function of (a) peak and average finger flexor EMG and (b) peak and average grip force.....	90
Figure 4.4:	Observational ratings of peak and average wrist posture as a function of peak and average wrist angle.....	92
Figure 4.5:	Observational ratings of average wrist movement speed as a function of average wrist angular velocity.....	92
Figure 5.1:	Proposed exposure-response curve for work-related upper extremity musculoskeletal disorders.....	104
Figure 5.2:	Visual-analog scale for rating repetition/hand activity, with verbal anchors.....	108
Figure 6.1:	Illustration of high reliability but low accuracy (a), and high reliability/high accuracy (b) (adapted from Montgomery, 1996).....	142
Figure 6.2:	Prevalence of five WRMSDs by repetition/hand activity category.....	145
Figure 6.3:	Proposed process for enhancing rating guidelines for force, posture, etc.....	149

LIST OF TABLES

Table 1.1:	Work-related factors which have been identified as risk factors for musculoskeletal disorders and selected references.	3
Table 1.2:	Instrumental techniques commonly used to assess exposure to physical stressors in manual work and examples of applications.....	5
Table 1.3:	Several observational job analysis methods and the upper limb stressors they consider.....	8
Table 2.1:	Selected epidemiological studies which showed a relationship between repetitive work and upper extremity musculoskeletal disorders.	23
Table 2.2:	Thirty-three jobs included in this study. Ratings are on a 10-cm scale, and are expressed as the average of five raters, rounded to the nearest 0.5 cm. In all cases, the range in ratings between the five raters was less than or equal to 1 cm.	30
Table 3.1:	Work-related factors which have been identified as risk factors for musculoskeletal disorders and selected references.	49
Table 3.2:	Physical stress parameters included in complete rating system.....	52
Table 3.3:	Composition of the four rating teams, including work organization, primary educational background, and years of experience in ergonomic job analysis.....	53
Table 3.4:	Industry and job descriptions of the 12 jobs which were rated by all four teams.....	54
Table 3.5:	Variance components and intra-class correlations for initial and final ratings of seven stressors (12 jobs, 17 raters in 4 groups).....	59

Table 3.6:	Variance components and intra-class correlations for initial and final ratings of seven stressors (45 jobs, 7 raters).....	60
Table 4.1:	Pooled main effects of each experimental parameter. Different superscripts indicate significant effects at $p < 0.05$	87
Table 4.2:	Model coefficients and p values for repetition rating as a function of movement speed and exertion time.	88
Table 5.1:	Work-related factors which have been identified as risk factors for musculoskeletal disorders and selected references.	105
Table 5.2:	Definitions of "repetition" that have been used in previous epidemiological studies.....	106
Table 5.3:	Physical stress parameters included in complete rating system (Latko, 1997b).....	109
Table 5.4:	Jobs included in study, stratified on repetition level. Average repetition ratings are in parentheses.....	110
Table 5.5:	Average exposures to selected stressors stratified by repetition (mean $\pm$ std. dev.).....	112
Table 5.6:	Hand diagram scoring definitions.....	113
Table 5.7:	Case definitions of health outcome measures presented in this paper.	115
Table 5.8:	Age, gender, and BMI distribution by plant and repetition category.....	117
Table 5.9:	Linear trend of symptoms, tendinitis, and CTS in the dominant extremity.	117
Table 5.10:	Predictors of discomfort in the wrist, hand, or fingers.....	118
Table 5.11:	Predictors of tendinitis in the dominant elbow, forearm, wrist, hand, or fingers.....	119
Table 5.12:	Predictors of CTS based on hand diagram score of 2 or 3.....	120
Table 5.13:	Predictors of CTS based on median mononeuropathy (0.5 ms threshold).....	120
Table 5.14:	Predictors of CTS based on median mononeuropathy (0.5 ms threshold) combined with hand diagram scores of at least 2.....	121

Table 5.15: Odds ratios based on changes in repetition level for the
five health outcome variables presented.122

LIST OF APPENDICES

Appendix A:	Job analysis rating scales and guidelines.....	154
Appendix B:	Variance components and intraclass correlations for all scales.....	173
Appendix C:	Models of alternative health outcomes.....	181

CHAPTER 1

INTRODUCTION

Problem Statement

Work-related musculoskeletal disorders (WRMSDs) of the upper extremity are a major occupational safety and health concern in much of the world (Armstrong et al., 1993b). These disorders, which include tendinitis, epicondylitis, and carpal tunnel syndrome, are a major cause of disability, lost productivity, and worker pain and impairment. In the past decade, the total number of cases of occupational illnesses associated with repeated trauma rose nearly tenfold, from 34,700 cases in 1984 to 332,000 cases in 1994 (Bureau of Labor Statistics, 1995). In 1989, the total direct compensable cost for these disorders in the United States was estimated at \$563 million (Webster and Snook, 1994). This figure does not include indirect costs associated with lost productivity, retraining, or replacement of injured workers. When these factors are considered, the estimated yearly cost of these disorders has been placed as high as 10.8 billion (CTD News, 1995). Several physical stressors associated with manual work, including repeated, sustained, and forceful exertions, localized mechanical stresses, and awkward postures have been linked to the development of these disorders (Kuorinka and Forcier, 1995; Viikari-Juntura, 1995). There is no agreement on how exposure to these stressors can best be quantified, or on acceptable exposure levels for workers. The goal of this work is to evaluate a newly developed observation exposure assessment method in terms of its reliability, validity, and ability to predict risk of upper extremity musculoskeletal disorders in active workers.

Background

Various physical stresses, including repetitive, sustained, and forceful exertions, awkward postures, localized mechanical stress, highly dynamic

movements, exposure to low temperatures, and vibration have been linked to increased risk of work-related musculoskeletal disorders (WRMSD) (Table 1.1). There is widespread evidence to indicate that an exposure-response relationship exists between work-related exposure to repetitive work and development of disorders (Armstrong et al., 1993b) (Figure 1.1). Numerous studies indicate that risk increases as exposure increases from very low levels to very high levels. For example, in the nine cross-sectional studies cited by Kuorinka and Forcier (1994) as evidence of the work-relatedness of hand-wrist tendinitis, prevalence rates were at least three times higher for the exposed groups than the control groups, with odds ratios ranging from 1.4 to 29 (mean = 7.4). Less is known about the exact shape of the dose-response curve, i.e. the critical exposure levels at which a worker's risk significantly increases. A major reason for this knowledge gap is the difficulty in quantifying the level of exposure to these stressors in manual work.

Difficulties in exposure definition, assessment, and evaluation have been cited among the greatest challenges in ergonomic epidemiology (Hagberg, 1992). In a meta-analysis of studies concerning workplace ergonomic factors and the development of neck and upper extremity musculoskeletal disorders, Stock (1991) eliminated 18 of 43 studies (42%) because of inadequate exposure measures. There are several competing constraints involved in selecting an exposure assessment method. Winkel et al. (1992) suggest four characteristics of exposure assessment methods that must be considered in the selection of a method for a particular application:

- cost (in terms of equipment, time, and expertise)
- capacity (size of population that can be evaluated)
- specificity/versatility (degree to which method addresses only one or a number of stressors)
- generality (ability to estimate exposures in general rather than at a particular point in time).

In addition, the issues of reliability and validity are critical in determining the adequacy of an exposure assessment method (Kelsey et al., 1986).

Numerous methods have been utilized to assess exposure to these stressors. These can generally be categorized as instrumentation based, worker reported, or observational. Each type of method has its own strengths and limitations, and is applicable in certain situations (Winkel et al., 1992; Kilbom, 1994).

Table 1.1: Work-related factors which have been identified as risk factors for musculoskeletal disorders and selected references.

factor	explanation	references
repeated and sustained exertions	cyclical use of the same tissues, either as a repeated motion or a repeated muscular effort without movement (Kuorinka and Forcier, 1994)	Luopajarvi et al., 1979; Cannon et al. , 1981; Silverstein et al., 1986; Armstrong et al., 1987
forceful exertions	high magnitude or intensity of muscular load	Silverstein, 1986; Armstrong, 1987; Goldstein et al., 1987; Bystrom and Fransson-Hall, 1994
localized mechanical stresses	mechanical stresses produced on tissues due to contact with external objects	Lundborg et al., 1982; Sauter et al., 1987; Fransson-Hall and Kilbom, 1993; Hoffman and Hoffman, 1995
posture	extreme postures close to the range of motion of the joint, or non-extreme postures which change the musculoskeletal geometry and increase stress on tendons, muscles, or other tissues (Kuorinka and Forcier, 1994)	Tichauer, 1966; Phalen, 1966; Armstrong and Chaffin, 1978; Gelberman et al., 1981; deKrom et al., 1990
joint kinematics	high rates of angular velocity and acceleration	Marras and Schoenmarklin, 1993; Schoenmarklin et al., 1994
recovery time	insufficient periods when the muscle is not contracting	Viersted et al. 1990; Sundelin and Hagberg, 1989; Bystrom and Fransson-Hall, 1994
exposure to low temperatures	working in cold environments or handling cold objects	Hammar skjold et al. 1992; Clarke, 1961; Williamson et al., 1984
vibration	exposure to hand-arm vibration	Lundstrom and Johansson, 1986; Nilsson et al., 1989; Hagberg et al., 1992
jerk/impulse	rapid eccentric contraction of muscles, as when resisting the torque of a tool	Armstrong et al, 1993a; Edwards, 1988

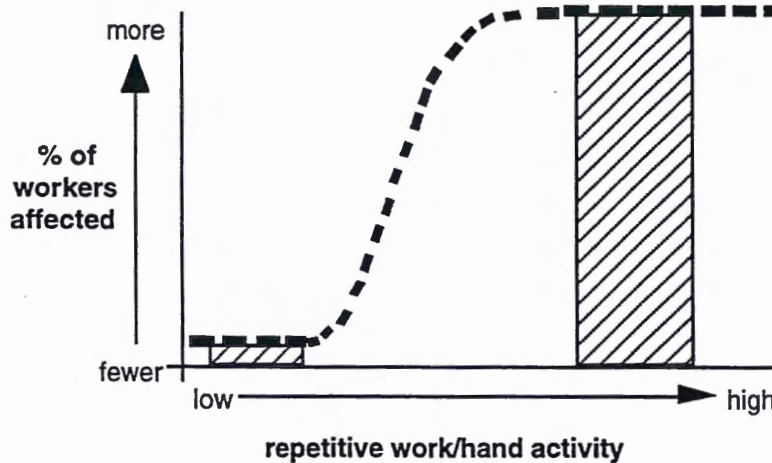


Figure 1.1: Proposed exposure-response curve for work-related upper extremity musculoskeletal disorders.

Instrumentation such as electromyography (EMG), electrogoniometry, and digitized video motion analysis systems are perceived as having high validity because they provide quantitative data. While these measurements are objective in one sense, critical metrics have not yet been determined, thus a subjective component is necessary in the analysis (Muckler and Seven, 1992). For example, with EMG, there is not agreement on which aspects of the signal are most critical: peak levels, mean/median levels, baseline levels, or a complex model involving the distribution of exertion levels over time (Mathiassen and Winkel, 1991). Instrumentation also has high requirements in terms of equipment and training, thus limiting sample sizes and feasibility for field studies (low capacity and high cost). Table 1.2 lists several instrumentation techniques and the categories of stressors that they have been used to quantify. While an instrumentation technique may provide detailed information on the occurrence of one or more of the stressors listed above, there is no single method of instrumentation that addresses all the major stressors (thus they have high specificity/low versatility).

Worker interviews, questionnaires, or ratings are often used as exposure assessment methodologies (Putz-Anderson and Grant, 1995). These methods have the benefit of being readily available and relatively inexpensive to collect (low cost, high capacity). Another benefit is that workers can be asked about their exposure to a wide range of stressors (high versatility) over a period of time (high generality). Although it is clear that workers can provide valuable information about the activities and demands of their jobs, there is debate about

the accuracy of worker reports as the sole source of exposure data (Baty et al., 1986; Wiktorin, 1993).

Table 1.2: Instrumental techniques commonly used to assess exposure to physical stressors in manual work and examples of applications.

	repeated/ sustained exertions	forceful exertions	local mech. stress	awkward postures	joint kinematics	recovery time	low temp	vibration
electro- myography (EMG)	Mathiassen and Winkel, 1991	Armstrong and Chaffin, 1979; Jonsson, 1988, Aaras and Westgaard, 1987				Viersted, 1990; Mathiassen and Winkel, 1991		
electro- goniometry	Radwin and Lin, 1993			Marras and Schoen- marklin, 1993	Marras and Schoen- marklin, 1993			
computerized video motion tracking systems	Leskinen et al., 1997			Leskinen et al., 1997				
accelero- meters	Andersson et al., 1996				Andersson et al., 1996			ISO, 1986; NIOSH 1989
force sensors/ transducers		Armstrong et al., 1994	Jensen et al., 1991					

Observational systems are perhaps the most widely used methods to assess exposure to physical stressors. Table 1.3 lists several observational systems that have been proposed to evaluate upper limb stress, and indicates which stressors are addressed in each. Typically, observational methods utilize one of three general procedures for recording stresses: 1) continuous, real-time observation, 2) sampling strategies, or 3) summary methods. In the continuous, real-time observational techniques, the parameter of interest (e.g. posture) is recorded every time it changes (e.g. Keyserling, 1986; Kilbom et al., 1986; Fransson-Hall et al., 1995). This allows both the frequency and duration of exposures over the course of a work task to be determined. These methods often employ computers to assist in data recording and summarizing. Sampling methods involve recording the parameters of interest at a series of points in time. Sampling rates may range from occasionally throughout the work task to several times per second (e.g. Armstrong et al, 1981). Summary methods generalize exposures over the entire work task. These methods often take the form of binary or categorical checklists (e.g. Karhu et al, 1977; Keyserling et al.

1993; McAtamney and Corlett, 1993), or counts of certain types of activities (e.g. Stetson et al., 1991, Drury, 1987). Equipment requirements for observational methods are low, and these methods allow analysts to obtain information on a wide variety of stressors fairly rapidly (low cost, high capacity, high versatility). The amount of training required for the different techniques varies widely, however, and there are often questions concerning the reliability and validity of such tools (Keyserling and Wittig, 1988; Baty et al., 1986; Kilbom, 1994).

The observational methods vary greatly as to the level of information provided about each stressor (Table 1.3). For example, in assessing wrist posture, the checklist described by Stetson et al. (1991) requires the user to count the number of times the wrist is in involuntary flexion/extension and ulnar deviation. In contrast, the method presented by Armstrong et al. (1982) involves the classification of wrist posture several times per second into one of five categories, yielding an exposure/time profile. The methods also differ in the number of body parts they consider. For example, the method described by Armstrong et al. (1979) only addresses hand posture, while the method discussed by Armstrong et al (1982) considers the shoulder, elbow, wrist, and hand. There is no consistent way in which these methods define or quantify the exposures. Schierhout et al. (1994) defines repetition in terms of cycle time and basic cycles, while Moore and Garg (1995) address repetition from the standpoint of efforts per minute and speed of work. The methods also differ in how specifically they address the exposures, and often reflect the purpose for which they were designed. For example, the checklist described by Keyserling et al. (1993), which was designed for the automotive industry, identifies exposures to low temperatures only if they are related to exhaust from an air tool, whereas the checklist discussed by Lifshitz and Armstrong considers whether the worker's hands are "exposed to moderate temperature". It is clear that certain stresses (e.g. exposure to low temperature for a worker handling frozen packages in a food warehouse) would not be identified with the Keyserling checklist.

Along with providing varying amounts of information, the methods differ in application time. Regardless of the specific system requirements, however, all methods require an initial amount of time for the analyst to document and understand the job. The time required for documentation is related to the complexity of the job task, with highly complex jobs encompassing many

discrete operations requiring more time to understand than short cycle time, highly stereotyped jobs.

Full understanding of the dose-response relationship between worker exposure to the physical stressors listed above and musculoskeletal disorders is dependent on the availability of valid, reliable methods for assessing those exposures. In order to be feasible for application in large scale epidemiological studies, methods must have low cost, high capacity, high versatility, and high generality. Once these demands are met, the ultimate test of any exposure assessment method is its external validity, or its ability to identify exposures associated with increased risk of WRMSDs (Kilbom, 1994).

Table 1.3: Several observational job analysis methods and the upper limb stressors they consider.

reference (type and name of method, if given)	repeated / sustained exertions	forceful exertions	localized mechanical stress	awkward postures	joint kinematics	recovery time	low temp	vibration / jerk
Armstrong et al., 1979 sampling				hand: 4 positions				
Armstrong et al., 1981 sampling				shoulder: 4 levels F/E 4 levels ab./ad. 4 levels lat./med. elbow: 4 levels F/E forearm: 3 levels P/S wrist: 3 levels U/R 5 levels F/E hand: 6 levels				
Corlett et al., 1979 (Posture Targetting) sampling				upper arm, lower arm: continuous position in horizontal and vertical plane				
Drury, 1987 summary	number of daily damaging wrist motions = number of exertions with grip or force plus non-neutral wrist posture	force estimated by operator: continuous value		Estimated angle categorized into 4 levels based on % range of motion: shoulder: rot., ab./ad, F/E elbow: F/E forearm: P/S wrist: U/R, F/E				vibration and "shock": yes/no

Table 1.3 Continued

reference (type and name of method, if given)	repeated / sustained exertions	forceful exertions	localized mechanical stress	awkward postures	joint kinematics	recovery time	low temp	vibration / jerk
Fransson-Hall et al., 1995 (PEO) real time recording	repetitive movements of forearms, wrist, or fingers: yes/no	manual handling: 4 levels of load		shoulder: 4 levels				
Holzman, 1982 (ARBAN) sampling	static muscle load: rating 0-10 based on Borg scale	force due to dynamic muscle effort: rating 0-10 based on Borg scale:		effort due to posture (shoulder-arm unit): rating 0-10 based on Borg scale:				vibration and shock: rating 0-10 based on Borg scale:
Karhu, 1977 (OWAS) sampling		strength req'd: 3 levels		overall arm position: 3 levels				
Keyserling, 1986 real time recording				shoulder position: 3 levels				
Keyserling et al., 1993 summary	cycle time <30 s or hands repeating same motions more than 1/2 of work cycle: yes/no	several items based on weight of object, posture, use of gloves: yes/no weight of part or tool >2.7kg: 3 categories based on % cycle	palm, base of hand, fingers, forearm, elbow, armpit: yes/no	finger wrist forearm shoulder 3 categories based on % cycle in certain positions	finger used in rapid triggering motion 3 categories based on % cycle		cold exhaust from air tool: 3 categories based on % cycle	vibration and jerk: 3 categories based on % cycle
Kilbom et al., 1986 (VIRA) real time recording	cycle time and number of cycle/hour			shoulder F/E: 5 levels ab/ad: 4 levels shoulder elev: yes/no	frequency of changes in posture and duration of each posture	frequency and duration of arm at rest		

Table 1.3 Continued

reference (type and name of method, if given)	repeated / sustained exertions	forceful exertions	localized mechanical stress	awkward postures	joint kinematics	recovery time	low temp	vibration / jerk
Lifshitz & Armstrong, 1986 summary	cycle time >30 s: yes/no	gloves, wt. < 10 lbs, tool suspended: yes/no	fingers/wrist: yes/no	pinch grip, wrist F/E, wrist R/U, clothes-wringing motion, adjustable worksurface and tool, grip 5-8 cm: yes/no			moderate temp.: yes/no	vibrating tool: yes/no
McAtamney & Corlett, 1993 (RULA) summary		discount overall rating if weight of arm is supported		upper arm: 4 levels plus shoulder elevation & abduction lower arm: 2 levels plus lat/med wrist: 3 levels F/E plus R/U forearm: 2 levels				
Moore & Garg, 1995 (Strain Index) summary	efforts/min: 5 levels duration of exertion: 5 levels	intensity of exertion: 5 levels		hand/wrist posture: 5 levels		speed of work: 5 levels		
Nathan, 1993 (Occupational Hand Use Categories) summary	frequency of operations: 5 levels	resistance: 5 levels						

Table 1.3 Continued

reference (type and name of method, if given)	repeated / sustained exertions	forceful exertions	localized mechanical stress	awkward postures	joint kinematics	recovery time	low temp	vibration / jerk
Schierhout et al., 1994 summary	based on cycle time and % time in fundamental cycle: 4 levels	0-10 based on Borg scale		4 categories based on % cycle in certain positions: shoulder ext. elbow elev. reach shoulder abd.				
Stetson et al., 1991 summary	number of hand exertions	number of hand exertions using high force	number of times palm used as striking tool	number of times wrist in F/E and R/U number of hand exertions using pinch grip				number of hand exertions using power tools

Notes: F/E = flexion/extension; R/U = radial/ulnar; ab/ad = abduction/adduction; lat./med. = lateral/medial; P/S = pronation/supination

Research Objectives

The goal of this work is to advance the understanding of the risk factors for WRMSDs, particularly repetition, and to progress toward determining acceptable exposure levels. The work focuses on the development of an method for defining and assessing exposure to repetition, evaluating the validity and reliability of the method, and determining workers' risk of WRMSDs relative to the repetition level of their jobs. A secondary goal included extension of the above objectives to other stressors, including force, localized mechanical stress, and awkward postures.

This effort was undertaken in four distinct phases:

- 1) Development of an operational definition of "repetition" based on previous literature and observable characteristics of manual work, and development of decision criteria for rating repetition and other stressors.
- 2) Investigation of the inter-rater and test-retest reliability of the method
- 3) Comparison of the observational method for rating repetition, force, wrist posture, and wrist movement velocity with instrumental measurements of the same stressors.
- 4) Completion of a cross-sectional epidemiological study comparing the repetition level of jobs with the prevalence of WRMSDs in workers active in those jobs.

Dissertation Organization

The dissertation is organized into five chapters and an appendix. Chapter 1 serves as an introduction to the research problem and details the significance of the dissertation. Chapters 2 through 5 are presented as a series of self contained manuscripts, discussing each phase of the research. Chapter 6 summarizes the studies and their findings, and presents recommendations for further research. Appendix A contains the current version of the job analysis method, including the written guidelines.

Chapter 2 presents the development process for the repetition rating scale. It reviews the existing literature on repetition as a risk factor for WRMSDs, proposes a new rating metric, and compares this metric to previous methods.

Chapter 3 is a study addressing how consistently different raters can apply the rating method, and how consistent results are at different points in time. Inter-rater reliability of four groups of raters (17 individuals) is evaluated for twelve jobs. Inter-rater reliability of two groups of raters (7 individuals) is evaluated for a total of forty-five jobs. Test-retest reliability was evaluated for twelve jobs over a seven week period for one group of raters. Comparisons are made of the consistency of initial (individual) and final (group) ratings.

Chapter 4 presents a laboratory study in which subjects performed a simulated task. Finger force was measured using electromyography, and wrist posture and angular velocity were measured using electrogoniometry. The instrumental measures were compared to observational ratings of the corresponding parameters.

Chapter 5 details an epidemiologic study evaluating the relationship between repetitive work and prevalence of WRMSDs. It is a cross-sectional, double blind study of 352 workers in three industries.

Chapter 6 concludes the dissertation by summarizing each of the four studies, discussing the findings, and suggesting areas in which further work is needed.

Appendix A includes a written version of the complete rating scales, as well as the written guidelines, which include verbal decision criteria and benchmark examples.

References

- Aaras, A. & Westgaard, R. H. (1987). Further studies of postural load and musculoskeletal injuries of workers at an electro-mechanical assembly plant. *Applied Ergonomics*, 18(3), 211-219.
- Andersson, A.; Nordgren, B.; & Hall, J. (1996). Measurements of movements during highly repetitive industrial work. *Applied Ergonomics*, 27(5), 343-344.
- Anonymous (1995). CTDs taking bigger bite of corporate bottom line. *CTD News*, 4(6),
- Armstrong, T. J.; Bir, C., Finsen, L., Foulke, J., Martin, B., Sjogaard, G., & Tseng, K. (1993a). Muscle responses to torques in hand held power tools. *Proceedings of the International Society of Biomechanics*, 114-115.
- Armstrong, T. J., Buckle, P., Fine, L. J., Hagberg, M., Jonsson, B., Kilbom, A., Kuorinka, I. A., Silverstein, B. A., Sjogaard, G., & Viikari-Juntura, E. R. (1993b). A conceptual model for work-related neck and upper-limb musculoskeletal disorders. *Scandinavian Journal of Work & Environmental Health*, 19, 73-84.
- Armstrong, T. J., & Chaffin, D. B. (1978). An Investigation of the relationship between displacements of the finger and the wrist joints and the extrinsic finger flexor tendons. *Journal of Biomechanics*, (11), 119-128.
- Armstrong, T. J., Chaffin, D. B., & Foulke, J. A. (1979). A methodology for documenting hand positions and forces during manual work. *Journal of Biomechanics*, 12, 131-133.
- Armstrong, T. J., Fine, L. J., Goldstein, S. A., Lifshitz, Y. R., & Silverstein, B. A. (1987). Ergonomics considerations in hand and wrist tendinitis. *The Journal of Hand Surgery*, 12a(5, part 2), 830-837.

- Armstrong, T. J., Foulke, J. A., Joseph, B. S., & Goldstein, S. A. (1982). Investigation of cumulative trauma disorders in a poultry processing plant. *American Industrial Hygiene Association Journal*, 43, 103-116.
- Armstrong, T. J., Foulke, B. J., Martin, B. J., Gerson, J., & Rempel, D. M. (1994). Investigation of applied forces in alphanumeric keyboard work. *American Industrial Hygiene Association Journal*, 55(1), 30-35.
- Baty, D., Buckle, P. W., & Stubbs, D. A. (1986). *Posture Recording by Direct Observation, Questionnaire Assessment and Instrumentation: a Comparison Based on a Recent Field Study*. In N. Corlett, J. Wilson, & I. Manenica, eds., *The Ergonomics of Working Postures, Models, Methods, and Cases*. London: Taylor and Francis.
- Bureau of Labor Statistics (1995). *Workplace injuries and illnesses in 1994*. Washington, DC: U.S. Government Printing Office.
- Bystrom, S., & Fransson-Hall, C. (1994). Acceptability of intermittent handgrip contractions based on physiological response. *Human Factors*, 36(1), 158-171.
- Cannon, L. J., Bernacki, E. J., & Walter, S. D. (1981). Personal and occupational factors associated with carpal tunnel syndrome. *Journal of Occupational Medicine*, 23(4), 255-258.
- Clarke, R. G. (1961). The limiting hand skin temperature for unaffected manual performance in the cold. *Journal of Applied Psychology*, 45, 193-194.
- Corlett, E. N., Madeley, S. J., & Manenica, I. (1979). Posture Targetting: A technique for recording working postures. *Ergonomics*, 22(3), 357-366.
- deKrom, M., Kester, A., Knipschild, P., & Spaans, F. (1990). Risk Factors for carpal tunnel syndrome. *American Journal of Epidemiology*, 132(6), 1102-1110.
- Drury, C. G. (1987). A biomechanical evaluation of the repetitive motion injury potential of industrial jobs. *Seminars in Occupational Medicine*, 2(1), 41-49.

- Edwards, R. (1988). Hypotheses of peripheral and central mechanisms underlying occupational muscle pain and injury. *European Journal of Applied Physiology*, 57(3), 275-281.
- Fransson-Hall, C., Gloria, R., Winkel, J., & Kilbom, A. (1995). A portable ergonomic observation method (PEO) for computerized on-line recording of postures and manual handling. *Applied Ergonomics*, 26(2), 93-100.
- Fransson-Hall, C., & Kilbom, A. (1993). Sensitivity of the hand to surface pressure. *Applied Ergonomics*, 24, 3.
- Gelberman, R., Hergenroeder, P., Hargens, A., Lundborg, G., & Akeson, W. (1981). The carpal tunnel syndrome: a study of carpal canal pressures. *The Journal of Bone and Joint Surgery*, 63A(3), 380-383.
- Goldstein, S. A., Armstrong, T. J., & Chaffin, D. B. (1987). Analysis of cumulative strain in tendons and tendon sheaths. *Journal of Biomechanics*, 20(1), 1-6.
- Hagberg, M. (1992). Exposure variables in ergonomic epidemiology. *American Journal of Industrial Medicine*, 21, 91-100.
- Hagberg, M., Morgenstern, H., & Kelsh, M. (1992). Impact of occupations and job tasks on the prevalence of carpal tunnel syndrome: a review. *Scandinavian Journal of Work, Environment, and Health*, 12, 277-279.
- Hammariskjold, E., Harms-Ringdahl, K., & Ekholm, J. (1992). Reproducibility of carpenters work after cold exposure. *International Journal of Epidemiology*, 9, 195-204.
- Hoffman, J., & Hoffman, P. (1995). Staple gun carpal tunnel syndrome. *Journal of Occupational Medicine*, 27(11), 848-849.
- Holzmann, P. (1982). ARBAN-A new method for analysis of ergonomic effort. *Applied Ergonomics*, 13(2), 82-86.
- International Standards Organization (ISO). (1986). Mechanical vibration--guidelines for the measurement and the assessment of human exposure to

- hand-transmitted vibration. Geneva: International Organization for Standardization. (ISO 5349-1986(E)).
- Jensen, T. R., Radwin, R. G., & Webster, J. G. (1991). A conductive polymer sensor for measuring external finger forces. *Journal of Biomechanics*, 24(9), 851-858.
- Jonsson, B. (1988). The static load component in muscle work. *European Journal of Applied Physiology*, 57, 305-310.
- Karhu, O., Kansi, P., & Kuorinka, I. (1977). Correcting working postures in industry: a practical method for analysis. *Applied Ergonomics*, 8(4), 199-201.
- Kelsey, J. L., Thompson, W. D., & Evans, A. S. (1986). *Methods in Observational Epidemiology*. New York: Oxford University Press.
- Keyserling, W. M. (1986). Postural analysis of the trunk and shoulders in simulated real time. *Ergonomics*, 29(4), 569-583.
- Keyserling, W., Stetson, D., Silverstein, B., & Brouwer, M. (1993). A checklist for evaluating ergonomic risk factors associated with upper extremity cumulative trauma disorders. *Ergonomics*, 31(5), 735-747.
- Keyserling, W. M., & Wittig, S. J. (1988). An analysis of experts' ratings of ergonomic stress. *International Journal of Industrial Ergonomics*, 2, 291-304.
- Kilbom, A. (1994). Assessment of physical exposure in relation to work-related musculoskeletal disorders-what information can be obtained from systematic observations? *Scandinavian Journal of Work, Environment, and Health*, 20 (special issue), 30-45.
- Kilbom, A., Persson, J., & Jonsson, B. (1986). Disorders of the cervicobrachial region among female workers in the electronics industry. *International Journal of Industrial Ergonomics*, 1, 37-47.
- Kuorinka, I., & Forcier, L. (1995). *Work related musculoskeletal disorders (WMSDs): a reference book for prevention*. London: Taylor & Francis.

- Leskinen, T., Hall, C., Rauas, S., Ulin, S., Tonnes, M., Viikari-Juntura, E., & Takala, E. P. (1997). Validation of portable ergonomic observation (PEO) method using optoelectronic and video recordings. *Applied Ergonomics*, 28(2), 75-83.
- Lifshitz, Y., & Armstrong, T. J. (1986). *A design checklist for control and prediction of cumulative trauma disorder in intensive manual jobs. Proceedings of the Human Factors Society 30th Annual Meeting. Vol. 2* Dayton, OH, September 29 (pp. 837-841). : Human Factors Society.
- Lundborg, G., Gelberman, G., Minter-Convery, M., Lee, Y., & Hargens, A. (1982). Median nerve compression in the carpal tunnel-functional response to experimentally induced controlled pressure. *The Journal of Hand Surgery*, 7(3), 252-259.
- Lundstrom, R., & Johansson, R. S. (1986). Acute impairment of the sensitivity of skin mechanoreceptive units caused by vibration exposure of the hand. *Ergonomics*, 29(5), 687-698.
- Luopajarvi, T., Kuorinka, I., Virolainen, M., & Holmberg, M. (1979). Prevalence of tenosynovitis and other injuries of the upper extremities in repetitive work. *Scandinavian Journal of Work, Environment, and Health*, 5(suppl. 3), 48-55.
- Marras, W. S., & Schoenmarklin, R. W. (1993). Wrist motions in industry. *Ergonomics*, 36(4), 341-351.
- Mathiassen, S., & Winkel, J. (1991). Quantifying variation in physical load using exposure-vs-time data. *Ergonomics*, 34(12), 1455-1468.
- McAtamney, L., & Corlett, E. N. (1993). RULA: a survey method for the investigation of work related upper limb disorders. *Applied Ergonomics*, 24(2), 91-99.
- Moore, J. S., & Garg, A. (1995). The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. *American Industrial Hygiene Association Journal*, 56, 443-458.

- Muckler, F. A., & Seven, S. A. (1992). Selecting performance measures: "objective" versus "subjective" measurement. *Human Factors*, 34(4), 441-455.
- Nathan, P. A., Keniston, R. C., Meadows, K. D., & Lockwood, R. S. (1993). Validation of occupational hand use categories. *Journal of Occupational Medicine*, 35(10), 1034-1042.
- National Institute for Occupational Safety and Health (NIOSH). (1989). Criteria for a recommended standard: Occupational exposure to hand-arm vibration. Washington, DC: U.S. Department of Health and Human Services, Public Health Service/Centers for Disease Control. (DHHS (NIOSH) publication No. 89-106).
- Nilsson, T., Burstrom, L., & Hagberg, M. (1989). Risk assessment of vibration exposure and white fingers among plasterers. *International Archives of Occupational and Environmental Health*, 61, 473-481.
- Phalen, G. (1966). The carpal tunnel syndrome: Seventeen years' experience in diagnosis and treatment of six hundred fifty-four hands. *The Journal of Bone and Joint Surgery*, 48A(2), 211-228.
- Fransson-Hall, C., & Kilbom, A. (1993). Sensitivity of the hand to surface pressure. *Applied Ergonomics*, 24, 3.
- Radwin, R. G., & Lin, M. L. (1993). An analytical method for characterizing repetition motion and postural stress using spectral analysis. *Ergonomics*, 36(4), 379-389.
- Sauter, S. L., Chapman, L. J., Knutson, S. J., & Anderson, H. A. (1987). Case example of wrist trauma in keyboard use. *Applied Ergonomics*, 18(3), 183-186.
- Schierhout, G. H., Bridger, R. S., & Myers, J. E. (1994). Development of observational methods for estimation of exposure to workplace postural stress. *Occupational Medicine*, 44, 262-266.

- Schoenmarklin, R. W., Marras, W. S., & Leurgans, S. E. (1994). Industrial Wrist Motions and Incidence of Hand/Wrist Cumulative Trauma Disorders. *Ergonomics*, 37(9), 1449-1459.
- Silverstein, B. A., Fine, L. J., & Armstrong, T. J. (1986). Hand Wrist Cumulative Trauma Disorders in Industry. *British Journal of Industrial Medicine*, 43, 779-784.
- Stetson, D. S., Keyserling, W. M., Silverstein, B. A., & Leonard, J. A. (1991). Observational Analysis of the Hand and Wrist: A Pilot Study. *Appl. Occup. Environ. Hyg.*, 6(11), 927-937.
- Stock, S. R. (1991). Workplace Ergonomic Factors and the Development of Musculoskeletal Disorders of the neck and Upper Limb: A Meta Analysis. *American Journal of Industrial Medicine*, 19, 87-107.
- Sundelin, G., & Hagberg, M. (1989). The effects of different pause types on neck and shoulder EMG activity during VDU work. *Ergonomics*, 32(5), 527-537.
- Tichauer, E. R. (1966). Some aspects of stress on forearm and hand in industry. *Journal of Occupational Medicine*, 8(2), 63-71.
- van der Beek, A. J., van Gaalen, L. C., & Frings-Dresen, M. H. W. (1992). Working postures and activities of lorry drivers: a reliability study of on-site observation and recording on a pocket computer. *Applied Ergonomics*, 23(5), 331-336.
- Viersted, K. B., Westgaard, R. H., & Andersen, P. (1990). Pattern of muscle activity during stereotyped work and its relation to muscle pain. *International Archives of Occupational and Environmental Health*, 62, 31-41.
- Viikari-Juntura, E. (1995). *The role of physical stressors in the development of hand/wrist and elbow disorders*. In S. L. Gordon, S. J. Blair, & L. J. Fine, eds., *Repetitive Motion Disorders of the Upper Extremity*. (pp. 7-30). Rosemont, IL: American Academy of Orthopaedic Surgeons.

- Webster, B. S., & Snook, S. H. (1994). The cost of compensable upper extremity cumulative trauma disorders. *Journal of Occupational Medicine*, 36(7), 713-717.
- Wiktorin, C., Karlqvist, L., & Winkel, J. (1993). Validity of self-reported exposures to work postures and manual materials handling. *Scandinavian Journal of Work, Environment, and Health*, 19, 208-214.
- Williamson, D., Chrenko, F., & Hamley, E. (1994, March). A study of exposure to cold in cold stores. *Applied Ergonomics*, 25-30.

CHAPTER 2

DEVELOPMENT AND EVALUATION OF AN OBSERVATIONAL METHOD FOR ASSESSING REPETITION IN HAND TASKS*

Abstract

Several physical stressors, including repetitive, sustained, and forceful exertions, awkward postures, localized mechanical stress, highly dynamic movements, exposures to low temperatures, and vibration have been linked to increased risk of work-related musculoskeletal disorders. Repetitive exertions have been among the most widely studied of these stressors, but there is no single metric for assessing exposure to repetitive work. The development of a new methodology enables repetitive hand activity to be rated based on observable characteristics of manual work. This method utilizes a series of 10 cm visual-analog scales with verbal anchors and benchmark examples. In this system, ratings for repetition reflect both the dynamic aspect of hand movements and the amount of recovery or idle hand time. Trained job analysis experts rate the jobs individually and then discuss ratings until consensus is achieved. For a group of 33 jobs, repetition ratings using this system were compared to measurements of recovery time within the cycle, exertion counts, and cycle time. Amount of recovery time within the job cycle was found to be significantly correlated with the analysis ratings ($r^2=0.58$), as were the number of exertions per second ($r^2=0.53$). Cycle time was not related to the analyst ratings. Repeated analyses using the new method were performed 1 1/2 to 2 years apart on the same jobs with the same group of raters. Ratings for repetition differed less than 1 unit (on the ten cm scale), on average, between the different sessions. These results indicate that the method is sensitive to exertion level and recovery time, and that the decision criteria and benchmark

* Copyright 1997, American Industrial Hygiene Association, reprinted with permission.

examples allow for a consistent application of these methods over a period of time. This method of rating repetition can be combined with similar scales for other physical stressors.

Background

Various physical stresses, including repetitive, sustained, and forceful exertions, awkward postures, localized mechanical stress, highly dynamic movements, exposure to low temperatures, and vibration have been linked to increased risk of work-related musculoskeletal disorders (WRMSD) (Andersson, 1995; Armstrong, et al., 1993; Kuorinka and Forcier, 1995; NIOSH, 1992; Stock, 1991;- Viikari-Juntura, 1995). Specifically, several studies have implicated repetitive exertions as a significant contributor to workers' risk of developing these disorders (Table 2.1).

There is widespread evidence to indicate that an exposure-response relationship exists between exposure to repetitive work and development of disorders (Armstrong, et al., 1993) (Figure 2.1). Although it is clear that risk increases as exposure increases from very low levels to very high levels, less is known about the exact shape of the dose-response curve, i.e. the critical exposure levels at which a worker's risk significantly increases.

Table 2.1: Selected epidemiological studies which showed a relationship between repetitive work and upper extremity musculoskeletal disorders.

Study	Study design	Population	Health outcome	Factor(s) considered as measure of repetition
Kuorinka and Koskinen (1979)	cross-sectional	workers in manual jobs in scissors manufacturing	tension-neck and muscle-tendon syndrome	number of pieces handled per year
Luopajarvi et al.(1979)	cross-sectional	assembly line packers and shop assistants	tenosynovitis	number of repetitive exertions
Cannon et al.(1981)	case-control	workers in aircraft engine manufacturing	carpal tunnel syndrome	observation of repetitive motion tasks of the wrists

Punnett et al. (1985)	cross-sectional	garment workers and hospital employees	persistent shoulder, wrist, and hand pain	observation of repetitive upper extremity exertions
Silverstein et al. (1986)	cross-sectional	workers from six industrial sites	hand/wrist cumulative trauma disorders (specific diagnoses and non-specific symptoms)	cycle time < 30 s or 50% of cycle in same fundamental movements
Armstrong et al. (1987)	cross-sectional	workers from seven worksites in six industries	hand and wrist tendinitis	cycle time < 30 s or 50% of cycle in same fundamental movements
Wieslander et al. (1988)	case-control	surgical carpal tunnel syndrome patients, other surgical cases, and general population	carpal tunnel syndrome	presence of repetitive wrist movements identified by worker and occupational hygienist
Barnhart et al. (1991)	cross-sectional	workers at a ski assembly plant	carpal tunnel syndrome	observed repeated and/or sustained activities involving wrist motion and/or pinch grip
Kurppa et al. (1991)	cross-sectional	workers in a meat-processing factory	tenosynovitis, peritendinitis, and epicondylitis	observed work "strenuous to the muscle-tendon structures of the upper limbs"
Chiang et al. (1993)	cross-sectional	fish-processing workers	shoulder girdle pain, epicondylitis, and carpal tunnel syndrome	cycle time < 30 s or 50% of cycle in same fundamental movements
Schoenmarklin et al. (1994)	cross-sectional	workers in 20 jobs in 8 industrial plants	"hand/wrist CTDs" as reported on OSHA 200 logs	velocity and acceleration of wrist movements
English et al. (1995)	case-control	patients attending 3 orthopedic clinics	various soft-tissue diseases of the upper limb	frequency and duration of various upper extremity motions
Schierhout et al. (1995)	cross-sectional	workers in 46 jobs in 11 factories	self-reported regional musculo-skeletal pain	cycle time

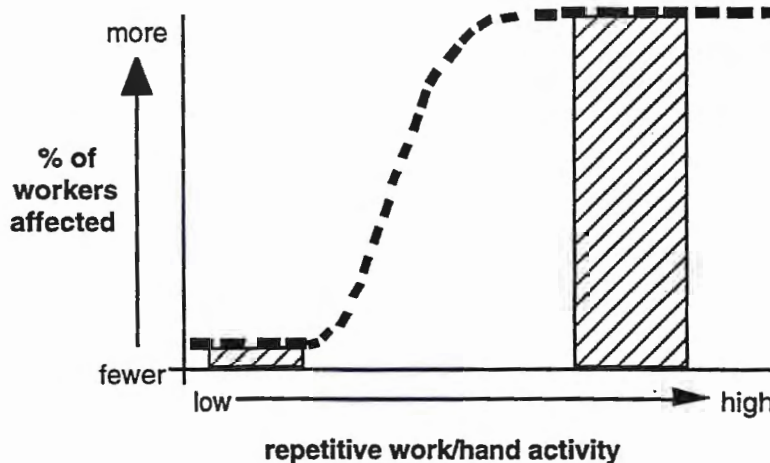


Figure 2.1: Proposed exposure-response curve for work-related upper extremity musculoskeletal disorders. Studies have shown higher risk at higher exposure levels (Andersson, 1995, Stock, 1991). The exact shape of the curve is not known.

A major reason for this knowledge gap is the difficulty in quantifying the level of repetitiveness (i.e. exposure) in manual work. There is no single metric that has been used to assess repetition. Previous studies of repetitive work have examined repetitiveness in terms of repeated cycles or amount of hand activity (Table 2.1). For example, some investigators have attempted to quantify repetition based on production standard data, e.g. cycle time or parts produced (see Table 2.1). However, such methods are difficult to apply to certain tasks, especially those with long cycle times or complex operations. Also, a measure of cycle time alone does not account for how busy the hands are during the cycle. For example, a 30 second cycle time used by some investigators (Silverstein, et al., 1986, Armstrong, et al., 1987, Chiang, et al., 1993) would not be suitable for certain jobs, such as those involving computer keyboard use or some assembly operations in which the worker performs a brief exertion and is idle for the rest of the cycle. In situations such as these, production standard-based methods may inaccurately estimate the true activity level of the hands. Other methods consider hand activity more directly, by defining repetition according to number of hand exertions per cycle or unit time (Kuorinka, and Koskinen, 1979, Luopajarvi, et al., 1979), the velocity and acceleration of motions (Schoenmarklin, et al., 1994), or the duration of micropauses (Henning, et al., 1989.). Potential difficulties with these methods include difficulty in defining and identifying "exertions", complications arising from exertion length,

and the technical requirements of instrumentation. Other investigators have relied on observation or worker reports to classify the repetitiveness of jobs (Cannon, et al., 1981, Punnett, et al., 1985, Wieslander, et al., 1989, Barnhart, et al., 1991, English, et al., 1995, Schierhout, et al., 1995), but these studies provide few details of the specific criteria used in making the classifications.

In general, existing methods for exposure assessment fall into one of the following categories (see Table 2.1): 1) production standard data, 2) observational methods, including checklists, and 3) instrumentation. The system described in this chapter is observational, in which repetition or hand activity is characterized using a visual-analog scale ranging from the lowest to the highest amount imaginable. The rating system consists of a 10 cm visual-analog scale which ranges from 0, which corresponds to no hand activity, to 10, the most possible hand activity. Rating scales such as these are common subjective assessment techniques which are often used to elicit workers' perception of job attributes (Sinclair, 1990). The system includes decision rules and benchmark examples to aid in determining the magnitude of the stress. This paper will focus on the rating scale for repetition/hand activity, although similar scales can be developed for the other stressors such as force, contact stress, and awkward postures.

The verbal anchors and decision criteria that form the foundation of the rating scale have been developed through extensive observation of a large number of jobs. These anchors and criteria are intended to provide benchmarks against which the raters can compare the job being observed, promoting consistency in ratings between analysts and jobs. The methods employed in these ratings are similar to those used to determine performance ratings in work measurement (Niebel, 1988), and in scoring many competitive sporting events at both the recreational and professional level, notably diving (Leas, 1993), figure skating (United States Figure Skating Association, 1995), and gymnastics (Loken and Willoughby, 1971). In all these cases, the outcome of interest is not an easily quantifiable entity. The ratings are based on professional judgment, applied by highly-trained individuals and aided by a set of well-defined decision criteria.

System development

The development and initial application of this system consisted of three subtasks: 1) Proposal of initial scale based on factors considered in previous studies (see Table 2.1), 2) preliminary job selection/classification and 3) ratings/rating system application and enhancement (Figure 2.2). In the preliminary job selection and classification stage, a team of 2 or more researchers conducted a plant walk-through, during which available jobs were observed. The goal of this job selection was to obtain examples of jobs encompassing the full range of the possible repetition rates. Jobs were selected based on an initial subjective assessment of repetition/hand activity. Representative workers, equipment, and job cycles were identified, and the selected jobs were videotaped and documented. Written documentation included production standards, job tasks, workstation layout and nominal dimensions, and materials, tools, and equipment. The job information was then taken back to the university laboratory where it was further analyzed.

The ratings/rating system enhancement was conducted at the university by a team of 4 - 6 faculty and staff who were experienced in ergonomic analysis in general, and this technique in particular. The videotaped jobs and documentation were presented to the team members, who independently rated the jobs for repetition/hand activity using the 10 cm scale and written guidelines (decision criteria and benchmark examples). Raters selected and rated the busiest hand during the cycle. Approximately 3 to 5 minutes of videotape was recorded for each job; these segments were repeated as many times as necessary for all raters to complete their ratings. When all team members had completed their individual ratings, the ratings were discussed with the goal of reaching consensus. Consensus was defined as: 1) a difference of no more than 1 unit (i.e. cm) on the 10 cm scale between the lowest score the highest score, and 2) the bases of all differences had been addressed. If the individual ratings for any stressor initially met consensus, no further discussion was necessary for that stressor. If consensus was not initially met, the outlying raters were given the opportunity to briefly explain the rationale for their ratings. In some cases, the discrepancies were due to observational differences: one rater simply noticed something that no one else did, or an individual overlooked something that everyone else saw. In these cases, the affected rater(s) adjusted their ratings accordingly and consensus was achieved (Figure 2.2).

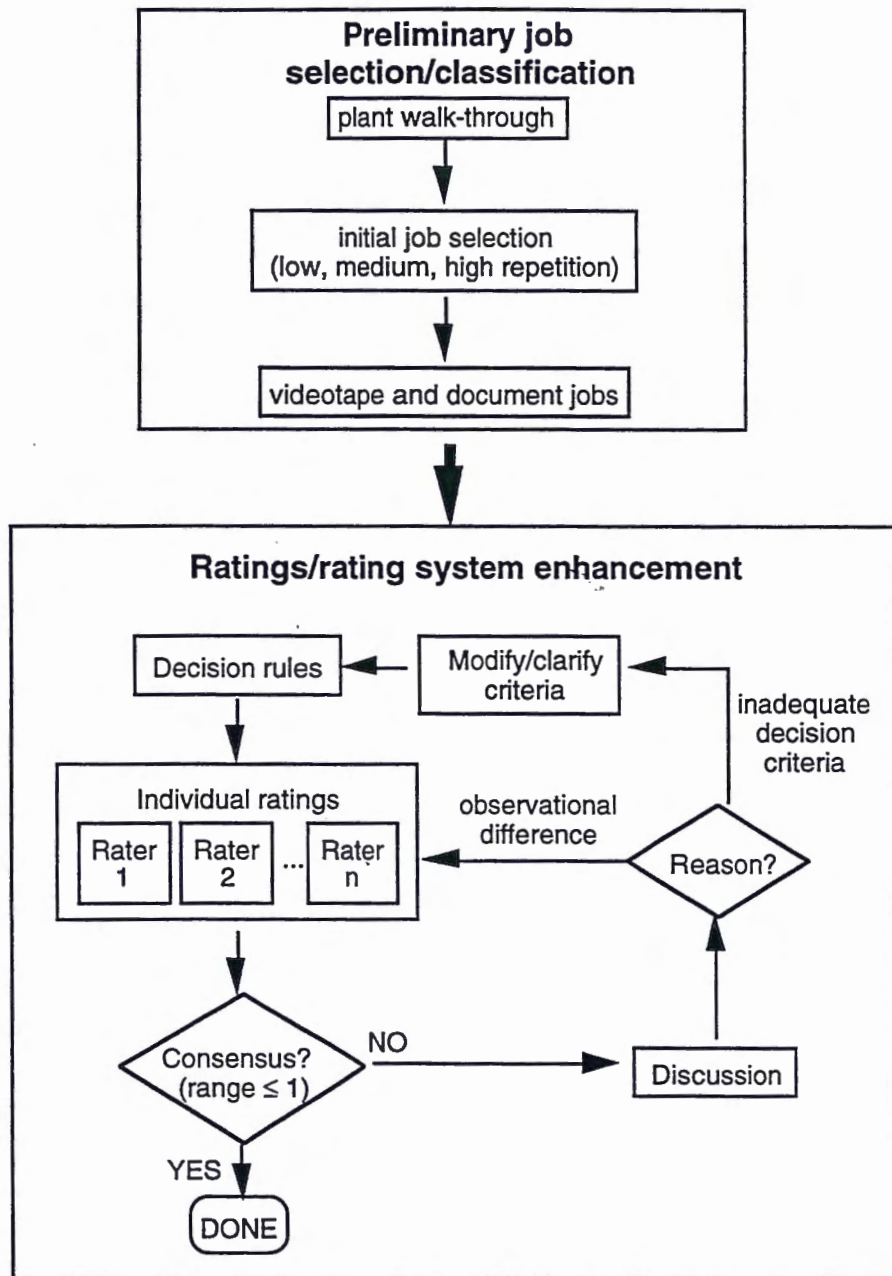


Figure 2.2: Flow chart illustrating the steps involved in job selection and system development and application.

Occasionally, jobs were encountered for which the team was not able to quickly reach consensus. These cases signaled an inadequacy in the decision criteria. In these situations, the team members discussed the repetitive activity observed relative to the existing criteria, and proposed modifications to clarify

the decision criteria. In some cases, the rating guidelines were then updated to include these enhancements (Figure 2.2).

Sometimes jobs were composed of two or more different tasks, e.g. building up subassemblies and installing subassemblies, or the workers rotated between multiple jobs on a production line or in a work cell. In most of these cases, the raters mentally integrated the two tasks and produced a single rating; however, each task could be rated separately and a time weighted average (TWA) calculated.

This rating system for repetition has evolved from the assessment of over 185 jobs. These jobs encompass a wide variety of industries, including furniture manufacturing, automotive components, appliance manufacturing, and paper products, and include both traditional, short-cycle assembly jobs as well as long-cycle time jobs with more task variability. The decision criteria and benchmark examples draw from this diverse group of jobs, providing a comprehensive set of guidelines that can be generalized to a wide variety of manufacturing jobs. Videotaped examples of the various rating levels have been assembled from these jobs and are maintained in a library; they are used for calibration and training purposes. Table 2.2 lists representative jobs from the full range of the rating scale.

The result of this iterative process is illustrated in Figure 2.3. The final observational rating method addresses repetition from the standpoint of hand activity rather than repeated work elements. This rating scheme draws from the studies cited in Table 2.1 and integrates, either directly or indirectly, several of the factors considered in those studies into a single scale. The factors directly considered in the repetition/hand activity rating include (Figure 2.3):

- 1) Duration and frequency of observed rest pauses (i.e. the amount of recovery time within the task),
- 2) The speed of hand movements (i.e. how fast the fingers and wrist move).

Table 2.2: Thirty-three jobs included in this study. Ratings are on a 10-cm scale, and are expressed as the average of five raters, rounded to the nearest 0.5 cm. In all cases, the range in ratings between the five raters was less than or equal to 1 cm.

job	industry	rating
inspector, visual observation of passing parts, no handling	appliance manuf.	1
unload plastic drum lids from molding machine; long pauses between cycles	fiber drum manuf.	1
load/operate press to form head liners; wait for machine to cycle	auto components	1
loading automatic fabric wrapping machine; wait for machine to cycle	auto components	2
load/operate trim machine for head liners; wait for machine to cycle	auto components	2
1 m transfer task, 6 s cycle	laboratory	2.5
stacking large drums; wait for next drum	fiber drum manuf.	2.5
connecting ground wire on dryer; attach 1 clip and drive 2 bolts	appliance manuf.	3.5
glass inspection, manipulating glass and visually inspecting for defects	glass/mirror manuf.	4
visor silkscreen (load, activate, unload)	auto components	4
feeding drum cover through glue machine and loading onto forming machine	fiber drum manuf.	4.5
hanging back panels of dryers on racks (2 jobs)	appliance manuf.	4.5
panel upholstery (office cubicle panels)	office furniture manuf.	5
overhead console upholstery; load fabric and frame into press, smooth fabric by hand	auto components	5
1 m transfer task, 3 s cycle	laboratory	5
securing fan on dryer; position housing, drive 3 bolts and position fan	appliance manuf.	6
wiring heater box on dryer; position wire harness and insert 4 connectors	appliance manuf.	6
securing upper back panel on dryer; position back panel and drive 3 bolts	appliance manuf.	6
securing back of console on dryer; insert 4 connectors	appliance manuf.	6.5
securing top panel on dryer; drive 3 bolts	appliance manuf.	6.5
mirror cutting; load etching machine and separate cut mirrors	glass/mirror manuf.	6.5
spray painting racks of visors and armrests	auto components	7
assembling lid for armrest/center storage unit	auto components	7
sewing armrest covers (2 jobs)	auto components	7.5
assembling cupholder	auto components	7.5
office chair upholstery; cover seats and backs	office furniture manuf.	8
press loading, small parts	fiber drum manuf.	8
1 m transfer task, 1.5 s cycle	laboratory	8
handle assembly; riveting (2 jobs)	fiber drum manuf.	9

As the repetition rating increases, the duration and frequency of rest pauses decreases until a rating of 10 is achieved, where rest pauses are virtually non-existent. Higher levels of repetition are also characterized by faster movements of the hands. This definition of repetition does not consider cycle time; a very long cycle job could have a high repetition rating if the hands are in constant, rapid motion with no recovery time. Similarly, short cycle time jobs may be rated low on the scale if the worker's hands are idle most of the cycle. Within this framework, ratings are discounted if there is a wide variety of different hand motions in the task.

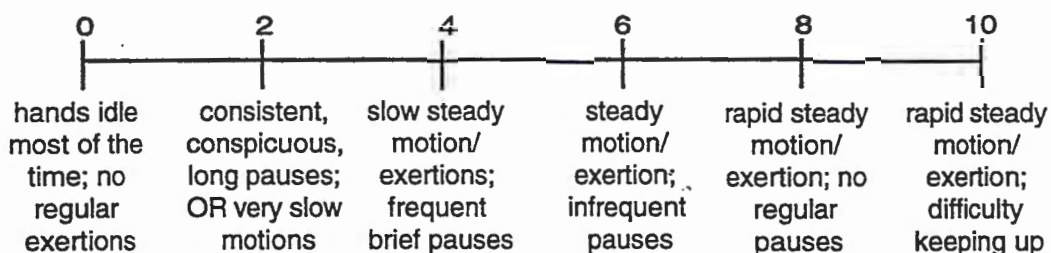


Figure 2.3: Visual-analog scale for rating repetition/hand activity, with verbal anchors.

Evaluation (Methods/results)

The evaluation entailed examining 1) how well the method assessed those factors it was designed to assess, and 2) test-retest results (Kelsey, et al., 1986). Because there is no single, agreed upon definition of repetition, the ratings were compared with characteristics of repetitive work that have been identified by previous researchers (Table 2.1). The three methods chosen for comparison are common observational metrics: hand exertions per second, amount of recovery time, and cycle time.

A subset of 33 jobs (Table 2.2) was randomly selected for this further analysis from those which previously had been rated by the team. This subset represented five different industries and one laboratory simulation. The team repetition ratings on the jobs ranged from 1 to 9 on the 10-cm scale, with an approximately uniform distribution over that range. A researcher, blinded to the

team ratings, performed the hand exertion analysis, recovery time analysis, and cycle time calculation on each of the 33 jobs. The same jobs were re-rated by the team 1 1/2 - 2 years after the initial ratings to assess the consistency of repetition/hand activity ratings by the same raters over time. All ratings and data collection were performed from the same segments of videotape.

Hand exertions:

Methods: The hand exertion analysis consisted of counting the number of exertions by the most active hand during several cycles of the job and averaging over the cycles. An exertion was defined as a conspicuous application of force by the hand, and included using the hand to hold, manipulate, trigger, push, pull, or otherwise handle an object (Stetson, et al., 1991). For each task, at least 5 cycles were observed and the exertion counts were averaged. The exertion values were divided by cycle time (yielding exertions per second) to facilitate comparative analysis.

Results: The results of this analysis are presented in Figure 2.4. The agreement (r^2) between the team ratings of repetition and the number of exertions per second was 0.53 ($p < 0.0001$).

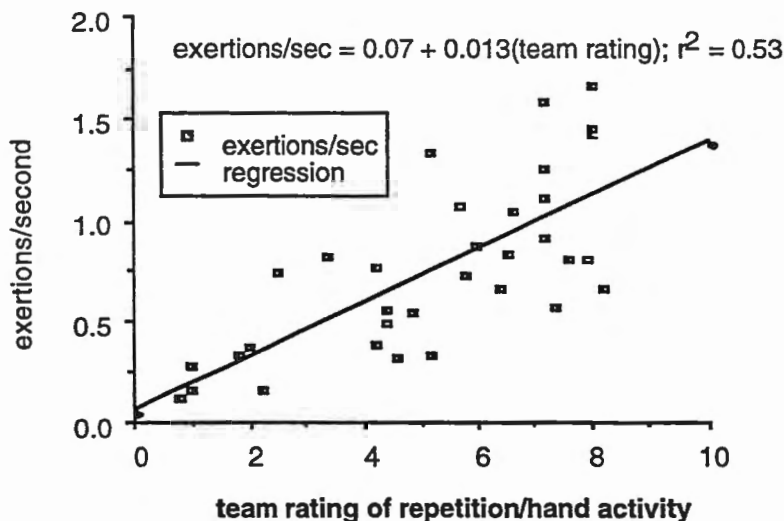


Figure 2.4: Relationship between team ratings of repetition/hand activity and number of exertions per second in 33 jobs.

Recovery time:

Methods: Recovery or rest time was operationally defined as periods when the hand was not holding, manipulating, triggering, pushing, pulling, or otherwise handling an object, and included times when the hand was completely idle, resting upon an object for voluntary support, moving freely, or reaching for an object. At least 5 cycles of the each task (Table 2.2) were observed and the recorded recovery time values were averaged. These values were then divided by cycle time to obtain the percent of recovery time within the cycle. Figure 2.5 illustrates typical work/recovery profiles over time for the three main levels of repetition/hand activity.

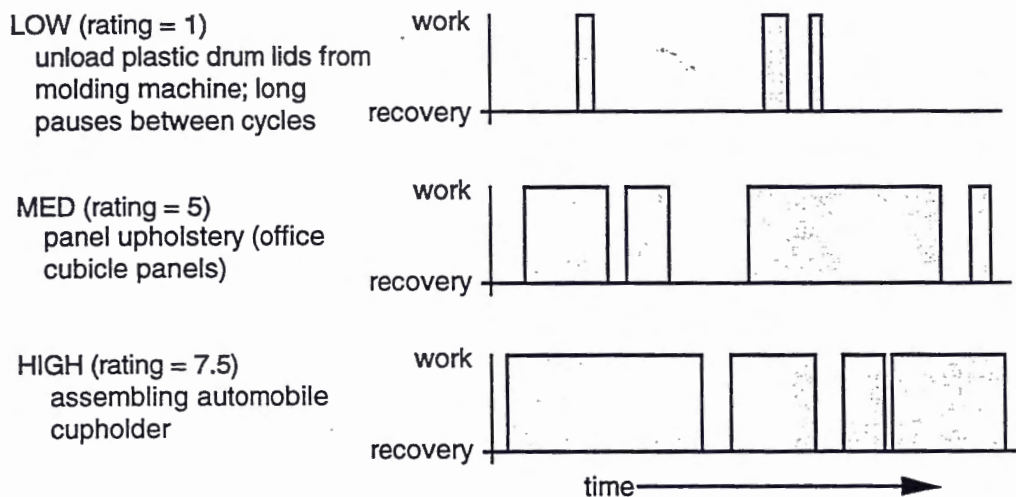


Figure 2.5: Typical work/recovery profiles for jobs at various levels of the rating scale.

Results: The results of this analysis are presented in Figure 2.6. The agreement (r^2) between the team ratings of repetition and the percent of cycle spent in recovery was 0.58 ($p < 0.0001$). Jobs in the low repetition range (rating = 0 - 3) averaged 64% ($\pm 16\%$) recovery time during the cycle. Medium repetition jobs (rating = 4 - 6) averaged 26% ($\pm 22\%$) recovery time and high repetition jobs (rating = 7 - 10) averaged 12% ($\pm 12\%$) recovery time.

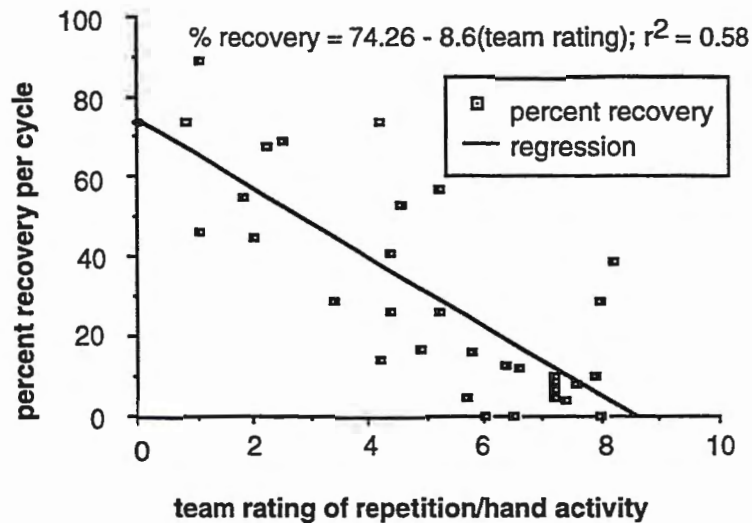


Figure 2.6: Relationship between team ratings of repetition/hand activity and % recovery time within the task cycle for 33 jobs.

Cycle time:

Methods: All 33 jobs included in this study had regular cycles, with a consistent sequence of steps being repeated throughout the work task. A cycle was defined as the time period during which the worker handled or performed operations on one unit of product. A minimum of 5 cycles of each job was observed, and average cycle time was calculated. These cycle times were compared with the team ratings of repetition.

Results: Cycle times for the 33 jobs in this study ranged from 1.2 s to 214s. Figure 2.7 illustrates the cycle time and team rating of each of the 33 jobs; there was no statistically significant relationship between cycle time and team rating. For example, the job rated lowest by the team (rating = 1) was an inspection task with a cycle time of only 8 seconds. The job with the longest cycle time (214 s) was an upholstery job which the team rated 8 on the 10 cm scale. Because a cycle time of 30 seconds has been used as a cutoff point for repetitive and non-repetitive work (Silverstein, et al., 1986, Armstrong, et al., 1987, Chiang, et al., 1993), a t-test was performed to determine if there was a significant difference in the team ratings of jobs with cycle times less than 30 s

compared to those with cycle times longer than 30 seconds. The jobs (n=11) with cycle times greater than 30 seconds had a mean team rating of 5.1 (± 2.25), while the jobs (n=22) with cycle times less than 30 seconds had a mean rating of 5.3 (± 2.3). This difference was not statistically significant ($p > 0.85$).

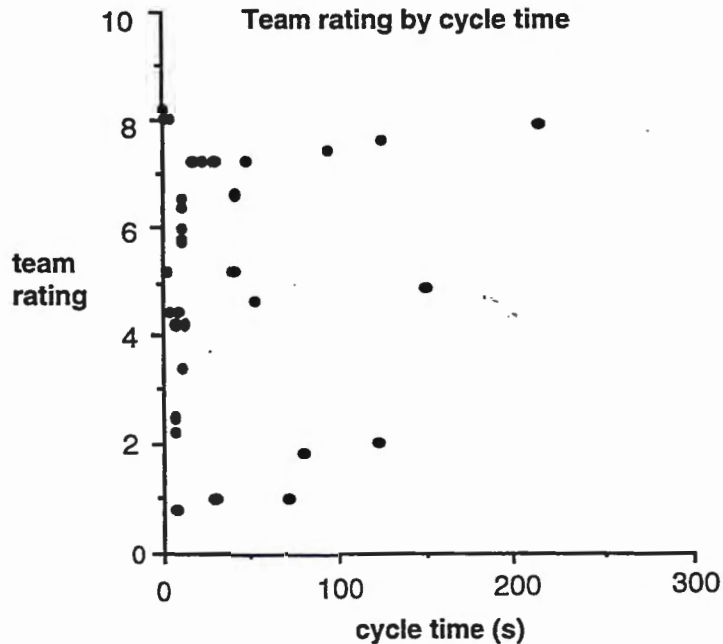


Figure 2.7: Plot of cycle time and team rating of repetition/hand activity for 33 jobs. There was no significant relationship.

Test/re-test:

Methods: The same 33 jobs were analyzed by the same team on two separate occasions to determine the consistency with which a given group is able to apply the decision criteria. Time between the two analyses ranged from 79 - 118 weeks. The team analyzed several other jobs in the interim, lowering the probability that any team members would have remembered the exact rating they had given the jobs originally. The 33 jobs were presented to the team in a random order. The rating procedure on both occasions followed that described in the Repetition Scale Development section above.

Results

The results of this evaluation are reported graphically in Figure 2.8. The two sets of ratings exhibited a good correlation ($r^2 = 0.88$). For the 33 jobs, 66% of the ratings differed by 1 unit or less between the two sessions. A paired t-test indicated a significant difference between the ratings for the two sessions ($p < 0.05$), with the ratings from the second session being approximately 0.6 units higher than the first session. Although this difference was statistically significant, it was well within the 1 unit range which was defined as consensus.

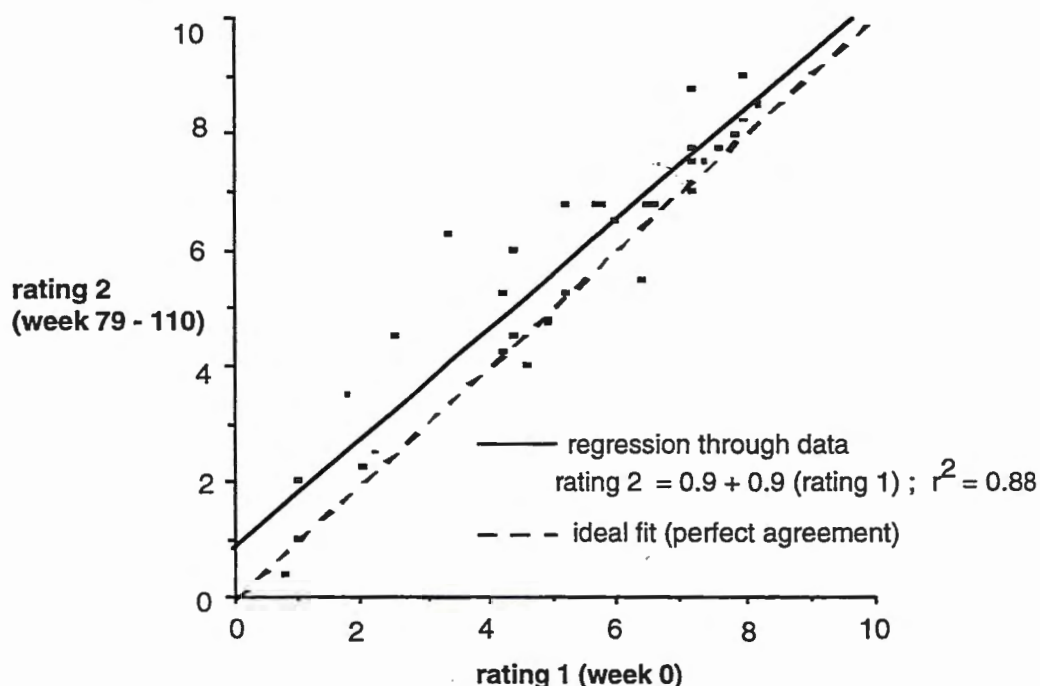


Figure 2.8: Average team ratings of repetition on two different occasions for 33 jobs ($r^2 = 0.88$).

Discussion

Hand exertion and recovery time analysis:

Both the hand exertion count (Figure 2.4) and the percent of recovery time within the cycle (Figure 2.6) were significantly correlated with the team ratings of repetition/hand activity ($r^2 = 0.53$ and 0.58 , respectively). It is expected that there would be some relationship between the measures of

exertion frequency and recovery time and the repetition/hand activity ratings. A perfect relationship would not be predicted, because the team rating for repetition/hand activity is an integrated score, which takes into account the speed of hand movements as well as the amount of recovery time within the cycle. The exertion counts and number of pieces handled, as used by Kuorinka and Koskinen (1979) and Luopajarvi et al. (1979) (Table 2.1) give an indirect indication of speed, in that a worker can accomplish more exertions per unit time if those exertions are rapid; however, the exertion counts do not directly assess movement speed. Also, the exertion counts do not account for the length of the exertions or the amount of recovery time within the cycle, both of which are factors that are considered by analysts using the rating method. The raters consider the trade off between recovery time and movement speed. For example, one worker may move quickly to load a machine, then wait while it cycles, while another may perform an assembly operation at a steady speed throughout the task, with few pauses. In such a case, both workers could receive similar ratings on the repetition/hand activity scale because of the interrelationship between movement speed and recovery time.

Cycle time:

No significant relationship was observed between the team ratings and cycle time (Figure 2.7). This is an important finding, because several epidemiological studies have used production standard based metrics to assess exposure to repetitive exertions (Table 2.1). It is possible to find high repetition, short cycle jobs, and low repetition, long cycle jobs (see Figure 2.7). This appears to be the case for the cited studies. This does not, however, mean cycle time can be used to characterize repetition in all settings. Although the criteria applied in these previous studies may have been appropriate for those particular groups of jobs, the findings of this study suggest that applying such criteria to a random group of jobs, which were not preselected with these criteria in mind, may not be appropriate.

Part of the reason for this poor agreement may be because the rating system takes into account the tradeoff between speed of movements and the amount of recovery time, rather than considering cycle time directly. Based on the rating system, for a given task, the repetition level will increase as the cycle time decreases because movements will have to be faster or recovery time will be shorter if the same amount of work must be accomplished in a shorter time.

However, merely calculating cycle time does not account for the different physiological requirements in different tasks. Similarly, counting the number of pieces handled does not consider what the worker does while handling each piece.

Some investigators have attempted to address this issue by adding additional constraints to the cycle time criterion. For example, Silverstein et al.(1986) and Armstrong et al.(1987) identified "fundamental cycles" within the work cycles as a sequence of steps that were repeated within the cycle. "High repetitive" jobs were defined as having a cycle time less than 30 seconds or performing the same type of fundamental cycles more than 50% of the cycle time, while "low repetitive" jobs had a cycle time greater than 30 seconds and less than 50% of the cycle time involving the same type of fundamental cycles. These criteria have been used in several subsequent studies by other researchers. While increasing the sensitivity of the cycle time metric, these additional conditions require more detailed analysis and analyst training. Although this definition of repetitiveness can be used to categorize certain jobs for epidemiological purposes, there are some limitations to its use. One is the difficulty in expanding it beyond a dichotomous characterization. Because the exposure is defined both in terms of cycle time and fundamental cycles, it is unclear how similar criteria could be developed to divide exposures into three or more levels. For example, three conditions of laboratory simulation were included in this study: a 1 m transfer at 1.5s, 3s, and 6 s intervals (see Table 2.2). Each transfer took approximately 1.2 second, allowing 20%, 60%, and 80% recovery time, respectively. By the Silverstein/Armstrong criteria, all of these conditions would be classified as "high repetitive". Extending the transfer task example, the same task performed once every 29 seconds would also be categorized as "high repetitive", and the worker would be able to rest nearly 28s of every cycle. Clearly the 1.5 s, 3 s, 6 s, and 29 s cycles differ in their exposure to repetitive exertions, but the cycle time metric does not allow differentiation between them.

Other investigators have provided fewer details of the criteria by which they judged work to be repetitive or non-repetitive. Criteria such as occupations requiring "repetitive motion tasks involving the wrists"(Cannon, et al., 1981), observation of repetitive upper extremity exertions (Punnett, et al., 1985), observed "repeated and/or sustained activities" involving wrist motion or pinch grip (Barnhart, et al., 1991), and observed tasks "strenuous to the muscle-

tendon structures of the upper limbs" (Kurppa, et al., 1991) have been used to describe these exposure assessments. Although observation can be a powerful tool in assessing repetition and these classification methods may have been adequate for the particular study for which they were employed, the descriptions supplied by the investigators do not provide a framework which can be used to relate the results of their studies to other jobs and types of work.

There is evidence to indicate that both movement speed and recovery time are important parameters related to risk of WRMSDs. Schoenmarklin et al. (1994) found both acceleration and velocity of the wrist to be associated with OSHA reportable disorders associated with repetitive motion among their study population, with the highest predictive power exerted by acceleration in wrist flexion/extension (OR = 5.03). Bystrom et al. (1994) suggest load and work/rest ratio acceptability limits for continuous and intermittent isometric contractions, based on a number of physiological indicators of localized muscle fatigue. In a laboratory study of cadaver hands, Goldstein et al. (1987) found that cumulative tendon strain was related to the load placed on the tendon, and increased as the ratio of tendon loading time/non-loading time increased. These findings suggest that both the contraction intensity and the exertion/recovery patterns are significant in predicting physiological signs of muscle fatigue and tendon strain. This suggests that consideration of recovery time is important in determining the physiological burden of work from the standpoint of repetition. We suspect that there is a high correlation between velocity, acceleration, and muscle force for work tasks. There is a need for further laboratory and field studies combining electromyography (EMG) and goniometers to clarify this relationship.

This observational system has some advantages over traditional methods for rating repetition in manual work. By relying on decision rules and verbal anchors, it considers the dynamic nature of the hand motion, rather than just the number of motions or the amount of idle time. Distinct advantages of this method are:

- application time is relatively short (repetition/hand activity can be rated in less than 5 minutes)
- no instrumentation is required
- team members can be trained to rate repetition/hand activity in approximately one hour

- the system can be applied by a team using videotape and written job documentation, or by a single investigator on site during a walk-through survey (although we recommend spot checks by a group of raters).

As similar scales are included for the other physical stressors (force, contact stress, and posture stress), it is predicted that training raters can be accomplished in approximately one day, while a thorough analysis of a job, evaluating all of the stressors, will be possible in between 30 and 60 minutes.

There is a need for a consistent method of exposure assessment which can be applied by different people. The ultimate goal of epidemiological research is to provide guidelines for safe exposure limits, in order to minimize workers' risk of WRMSDs while maximizing productivity. In order for this goal to be reached, it is necessary to utilize exposure assessment techniques that are applicable in a wide variety of job settings and which will eventually allow the generalization from these studies to the working world in general. The method described in this paper is one such possible technique.

Conclusions/future work:

This work indicates that the proposed rating system is sensitive to the parameters of movement frequency and recovery time in hand intensive tasks. The method can easily be applied, requiring no instrumentation and a minimum amount of training. Experienced raters are able to consistently apply this method at points in time.

The decision criteria will continue to be enhanced with the addition of benchmark examples from different types of work, e.g. office, assembly, meat processing, etc. This repetition scale can be used in combination with similar scales for the other physical stressors associated with risk of work-related musculoskeletal disorders (e.g. force, localized contact stress, and posture stress) to provide a comprehensive exposure assessment.

With regard to this method, there are several application and validity issues which are in need of further study.

- Inter- and intra-rater reliability

It is important to evaluate the consistency with which this, or any, method can be applied by various analysts. Future studies are planned which compare the ratings of the same group of jobs by several groups of analysts, thus

assessing the inter-group reliability. Studies are also planned which will examine the difference in individual vs. group ratings, to determine the amount of consistency that can be obtained without the benefit of the consensus procedure. It is predicted that the method can be satisfactorily applied by individuals, although with some increase in the variability in the ratings.

- Training/analyst experience

Experts were chosen as a starting point for evaluation because they represent a best case situation. Consistent expert ratings justify future studies of non-experts. Because it is recognized that it is not always possible or feasible to have raters of the same level of expertise as those in this study, an important future consideration is the level of background knowledge necessary to adequately apply the decision criteria. The authors foresee a tool such as this being useful to a wide variety of users, including employee-based plant ergonomic teams. Trials with users of this type of background are planned. It is necessary to determine the amount of method-specific training that is required for analysts (both expert and non-expert) to be able to effectively apply this technique. In addition, an evaluation of the "learning curve" by examining the length of time to rate and achieve consensus and the amount of variability in initial ratings, is planned.

- Comparison with instrumental methods

Although there is no singular measure that all investigators have agreed on for assessing "repetition" (see Table 2.1), instrumentation such as electromyography and electrogoniometry can provide information on patterns of muscle activity and movement. Studies comparing data from this instrumentation with ratings of the same tasks will be conducted to examine the relationships.

- Evaluation of similar scales for other physical stressors

While repetition is an important factor in the development of WRMSDs, several other factors, including forceful exertions, localized mechanical stresses, and awkward postures are also important. Similar observational scales can be used to assess exposure to these other stressors, giving a comprehensive system for exposure assessment. As these other scales are completed, reliability and validity analyses similar to those discussed above will be necessary.

- Establishment of exposure-response relationship.

One of the original goals in the development of this method was the establishment of an exposure assessment tool to be used in epidemiological research. As previously stated, the ultimate goal in epidemiological research is to provide guidelines for safe exposure limits, in order to minimize workers' risk of WRMSDs while maintaining necessary levels of production. In order to more fully determine these relationships, more epidemiological studies are needed that quantitatively compare workers' exposure with their health status.

References

- Andersson, G. B. (1995). *Epidemiology of occupational neck and shoulder disorders*. In S. L. Gordon; S. J. Blair, & L. J. Fine, eds., *Repetitive motion disorders of the upper extremity*. (pp. 31-42). Rosemont, IL: American Academy of Orthopaedic Surgeons.
- Armstrong, T. J., Buckle, P., Fine, L. J., Hagberg, M., Jonsson, B., Kilbom, A., Kuorinka, I. A., Silverstein, B. A., Sjogaard, G., & Viikari-Juntura, E. R. (1993). A conceptual model for work-related neck and upper-limb musculoskeletal disorders. *Scandinavian Journal of Work & Environmental Health*, 19, 73-84.
- Armstrong, T. J., Fine, L. J., Goldstein, S. A., Lifshitz, Y. R., & Silverstein, B. A. (1987). Ergonomics considerations in hand and wrist tendinitis. *The Journal of Hand Surgery*, 12a(5, part 2), 830-837.
- Barnhart, S., Demers, P. A., Miller, M., Longstreth, W. T., & Rosenstock, L. (1991). Carpal Tunnel Syndrome among ski manufacturing workers. *Scandinavian Journal of Work, Environment, and Health*, 17, 46-52.
- Bystrom, S., & Fransson-Hall, C. (1994). Acceptability of intermittent handgrip contractions based on physiological response. *Human Factors*, 36(1), 158-171.
- Cannon, L. J., Bernacki, E. J., & Walter, S. D. (1981). Personal and occupational factors associated with carpal tunnel syndrome. *Journal of Occupational Medicine*, 23(4), 255-258.
- Chiang, J., Ko, Y., Chen, S., Yu, H., Wu, T., & Chang, P. (1993). Prevalence of Shoulder and Upper-limb Disorders Among Workers in the Fish-processing Industry. *Scandinavian Journal of Work, Environment, and Health*, 19, 126-131.
- English, C. J., Maclaren, W. M., Court-Brown, C., Hughes, S., Porter, R. W., Wallace, W. A., Graves, R. J., Pethick, A. J., & Soutar, C. A. (1995). Relations

between upper limb soft tissue disorders and repetitive movements at work. *American Journal of Industrial Medicine*, 27, 75-90.

Goldstein, S. A., Armstrong, T. J., & Chaffin, D. B. (1987). Analysis of cumulative strain in tendons and tendon sheaths. *Journal of Biomechanics*, 20(1), 1-6.

Henning, R. A., Sauter, S. L., Salvendy, G., & Krieg, E. F. (1989). Microbreak Length, Performance, and Stress in a Data Entry Task. *Ergonomics*, (32), 855-864.

Kelsey, J. L., Thompson, W. D., & Evans, A. S. (1986). *Methods in Observational Epidemiology*. New York: Oxford University Press.

Kuorinka, I., & Forcier, L. (1995). *Work related musculoskeletal disorders (WMSDs): a reference book for prevention*. London: Taylor & Francis.

Kuorinka, I., & Koskinen, P. (1979). Occupational Rheumatic Diseases and Upper Limb Strain in Manual Jobs in a Light Mechanical Industry. *Scandinavian Journal of Work, Environment, and Health*, 5(suppl. 3), 39-47.

Kurppa, K., Viikari-Juntura, E., Kuosma, E., Huuskonen, M., & Kivi, P. (1991). Incidence of Tenosynovitis or Peritendinitis and Epicondylitis in a Meat-Processing Factory. *Scandinavian Journal of Work, Environment, and Health*, 17, 32-37.

Leas, D. E. (1993). *1993 & 1994 Official Rules of Diving and Code of Regulations*. Indianapolis, IN: United States Diving, Inc.

Loken, N. C., & Willoughby, R. J. (1971). *The Complete Book of Gymnastics*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Luopajarvi, T., Kuorinka, I., Virolainen, M., & Holmberg, M. (1979). Prevalence of tenosynovitis and other injuries of the upper extremities in repetitive work. *Scandinavian Journal of Work, Environment, and Health*, 5(suppl. 3), 48-55.

National Institute for Occupational Safety and Health (1992). *A National Strategy for Occupational Musculoskeletal Injuries: Implementation and*

Research Needs (DHHS (NIOSH) Publication 93-101). Washington, DC: Government Printing Office.

Niebel, B. W. (1988). *Motion and Time Study (8th ed.)*. Homewood, IL: Irwin.

Punnett, L., Robins, J. M., Wegman, D. H., & Keyserling, W. M. (1985). Soft Tissue Disorders in the Upper Limbs of Female Garment Workers. *Scandinavian Journal of Work, Environment, and Health*, 11, 417-425.

Schierhout, G. H., Meyers, J. E., & Bridger, R. S. (1995). Work Related Musculoskeletal Disorders and Ergonomic Stressors in the South African Workforce. *Occupational and Environmental Medicine*, 52, 46-50.

Schoenmarklin, R. W., Marras, W. S., & Leurgans, S. E. (1994). Industrial Wrist Motions and Incidence of Hand/Wrist Cumulative Trauma Disorders. *Ergonomics*, 37(9), 1449-1459.

Silverstein, B. A., Fine, L. J., & Armstrong, T. J. (1986). Hand Wrist Cumulative Trauma Disorders in Industry. *British Journal of Industrial Medicine*, 43, 779-784.

Silverstein, B. A., Fine, L. J., & Armstrong, T. J. (1987). Occupational factors and carpal tunnel syndrome. *American Journal of Industrial Medicine*, 11, 343-358.

Sinclair, M. A. (1990). *Subjective Assessment*. In J. R. Wilson, & E. N. Corlett, eds., In *Evaluation of Human Work: A Practical Ergonomics Methodology*. (pp. 69-100). London: Taylor and Francis Ltd.

Stetson, D. S., Keyserling, W. M., Silverstein, B. A., & Leonard, J. A. (1991). Observational Analysis of the Hand and Wrist: A Pilot Study. *Appl. Occup. Environ. Hyg.*, 6(11), 927-937.

Stock, S. R. (1991). Workplace Ergonomic Factors and the Development of Musculoskeletal Disorders of the neck and Upper Limb: A Meta Analysis. *American Journal of Industrial Medicine*, 19, 87-107.

United States Figure Skating Association (1995). *1995-1996 USFSA Media Guide*. Colorado Springs, CO.

Viikari-Juntura, E. (1995). *The role of physical stressors in the development of hand/wrist and elbow disorders*. In S. L. Gordon; S. J. Blair, & L. J. Fine, eds., *Repetitive Motion Disorders of the Upper Extremity*. (pp. 7-30). Rosemont, IL: American Academy of Orthopaedic Surgeons.

Wieslander, G., Norback, D., Gothe, C. J., & Juhlin, L. (1989). Carpal Tunnel Syndrome (CTS) and Exposure to Vibration, Repetitive Wrist Movements, and Heavy Manual Work: a Case-Referent Study. *British Journal of Industrial Medicine*, 46, 43-47.

CHAPTER 3

EVALUATION OF INTERRATER RELIABILITY AND TEST-RETEST RELIABILITY OF OBSERVATIONAL RATINGS OF PHYSICAL STRESSORS IN MANUAL WORK

Abstract

An observational method of quantifying exposure to physical stressors such as repetition, forceful exertions, localized mechanical stresses, and awkward postures has been developed. A study was conducted to determine the inter-rated reliability and the test-retest reliability of observations using the method.

Four teams of raters (3-5 raters per team, 17 raters total) rated a set of twelve jobs using a modified nominal group technique. The variance in ratings due to raters was determined for both the initial and final ratings. In general, intraclass correlations (ICCs), which measure the percent of total variance explained by differences between jobs, were very good ($ICC=0.79$) for the initial ratings of repetition, indicating that the method is sensitive to between-job differences, even with different raters. Other stressors exhibited weak ICCs after the initial ratings ($0.10 < ICC < 0.39$). The consensus process greatly improved the ICCs. The ICC for repetition increased to 0.86, indicating that 86% of the variability in ratings was attributable to job differences. Ratings of peak and average force, peak and average contact stress on the fingers, and peak wrist posture improved into a range that is considered to be adequate. Intraclass correlations of the final ratings of these stressors ranged from 0.41 to 0.69. Only the ICCs of average wrist posture ratings remained weak.

Two of the teams ($n=3$ and $n=4$) rated an additional 33 jobs. ICCs ranged from 0.29 to 0.90 (median = 0.48) for the initial ratings and 0.21 to 0.94 (median = 0.69) for the final ratings. With these groups of raters and jobs, only the average force ratings continued to have an unacceptable level of rater variability after the consensus process. One team re-rated a subset of twelve

jobs seven weeks after the initial rating session. None of the seven scales evaluated had a statistically significant day effect.

For all phases of this study, the ratings of repetition had the highest level of agreement between raters (i.e. the strongest ICC and the highest test/retest correlation). Repetition was also the most broadly represented stressor in the jobs studied. The other stressors occurred over a relatively narrow range of potential exposures. It is unclear whether the relatively low agreement for the other stressors is due to this narrow range of exposures or to weaknesses in the method.

Background

Several physical aspects of work have been linked to increased risk of musculoskeletal disorders (WRMSDs). These factors include repeated and sustained exertions, forceful exertions, localized mechanical stress, posture, joint kinematics, recovery time, temperature, and vibration (Table 3.1).

Individually, these stressors have been associated with the development of WRMSDs. There is also evidence to suggest that stressors which occur in combination increase risk more than a single stressor alone (Silverstein et al., 1986; Armstrong et al., 1987). Because of this interaction, it is important to have exposure assessment techniques that allow quantification of exposure to all of these stressors.

It is generally agreed that a dose-response relationship exists between these stressors and musculoskeletal disorders in workers (Armstrong et al, 1993b). Numerous studies indicate that risk increases as exposure increases from very low levels to very high levels. For example, in the nine cross-sectional studies cited in Kuorinka and Forcier (1995) as evidence of the work-relatedness of hand-wrist tendinitis, prevalence rates were at least three times higher for the exposed groups than the control groups, with odds ratios ranging from 1.4 to 29 (mean = 7.4). Less is known about the degree of increased risk associated with moderate exposure levels, and the critical exposure levels at which a worker's risk significantly increases. This is in part due to the difficulty in quantifying workers' exposure to these stresses.

Numerous methods have been utilized to assess exposure to these stressors. These methods can generally be categorized as instrumentation based, worker reported, or observational. Each type of method has its own

strengths and limitations, and is applicable in certain situations (Kilbom, 1994; Winkel et al. 1991).

Table 3.1: Work-related factors which have been identified as risk factors for musculoskeletal disorders and selected references.

factor	explanation	references
repeated and sustained exertions	cyclical use of the same tissues, either as a repeated motion or a repeated muscular effort without movement (Kuorinka and Forcier, 1994)	Luopajarvi et al. ,1979; Cannon et al. ,1981; Silverstein et al., 1986, 1987; Armstrong et al., 1987
forceful exertions	high magnitude or intensity of muscular load	Silverstein, 1986; Armstrong, 1987; Goldstein et al., 1987; Bystrom and Fransson-Hall, 1994
localized mechanical stresses	mechanical stresses produced on tissues due to contact with external objects	Lundborg et al., 1982; Sauter et al., 1987; Fransson-Hall and Kilbom, 1993; Hoffman and Hoffman, 1995
posture	extreme postures close to the range of motion of the joint, or non-extreme postures which change the musculoskeletal geometry and increase stress on tendons, muscles, or other tissues (Kuorinka and Forcier, 1994)	Tichauer, 1966; Phalen, 1966; Armstrong and Chaffin, 1978; Gelberman et al., 1981; deKrom et al., 1990
joint kinematics	high rates of angular velocity and acceleration	Schoenmarklin et al., 1994; Marras and Schoenmarklin, 1993.
recovery time	insufficient periods when the muscle is not contracting	Viersted et al. 1990; Sundelin and Hagberg, 1989; Bystrom and Fransson-Hall, 1994
exposure to low temperatures	working in cold environments or handling cold objects	Hammarskjold et al. 1992; Clarke, 1961; Williamson et al., 1984
vibration	exposure to hand-arm vibration	Lundstrom and Johansson, 1986; Nilsson et al., 1990; Hagberg et al., 1992
jerk/impulse	rapid eccentric contraction of muscles, as when resisting the torque of a tool	Armstrong et al, 1993a; Edwards, 1988

Instrumental techniques are often favored because they provide quantitative data. While these measurements are obtained objectively, in the absence of corresponding health data their interpretation is subjective (Muckler

and Seven, 1992). Instrumentation also has high requirements in terms of equipment and training, and thus is not feasible in all applications. This limitation generally restricts the use of instrumentation to short periods of time and small numbers of workers; extrapolation of these data to larger groups introduces an additional subjective component. While an instrumentation technique may provide detailed information on the occurrence of one or more of the stressors listed above, there is no single method of instrumentation that addresses all the stressors.

Worker interviews, questionnaires, or ratings are often used as exposure assessment methodologies. These methods have the benefit of being readily available and relatively inexpensive to collect. Another benefit is that workers can be asked about their exposure to a wide range of stressors. Although it is clear that workers can provide valuable information about the activities and demands of their jobs, there is debate about the accuracy of worker reports as the sole source of exposure data (Wiktorin et al., 1993).

Observational methods have been used widely to assess exposure to physical stressors. These methods often take the form of binary or categorical checklists (e.g. Karhu et al, 1977; Keyserling et al., 1993; McAtamney and Corlett, 1993), or counts of certain types of activities (e.g. Stetson, 1991, Drury, 1987). Equipment requirements for observational methods are low, and these methods allow analysts to obtain information on a wide variety of stressors fairly rapidly. The amount of training required for the different techniques varies widely, however, and there are often questions concerning the reliability and validity of such tools (Kilbom, 1993).

The goal of this research was to develop and evaluate an observational rating system for assessing exposure to physical stressors in manual work. Specifically, the issues of interrater reliability and test-retest reliability were examined.

Methods

An observational rating system was developed to quantify exposure to various physical stressors in manual work. Four teams of analysts rated a group of jobs using this method. The consistency of ratings among the individual raters was evaluated. One team rated a subset of the jobs on two different occasions, to assess test-retest reliability.

I. Rating system

The observational rating system consists of a series of 10 cm visual-analog scales corresponding to the different stressors and segments of the upper limb. In general, the scales range from 0, which corresponds to no stress, to 10, corresponding to the most possible stress. Rating scales such as these are common subjective techniques, and are used in a variety of applications, including worker ratings of job attributes (Sinclair, 1995), work measurement performance ratings (Niebel, 1988), and scoring in competitive sporting events such as diving (Leas, 1993), figure skating (United States Figure Skating Association, 1995), and gymnastics (Loken and Willoughby, 1971).

The system makes use of a series of decision rules and benchmark examples to aid in determining the magnitude of each stressor. The development of this method was an iterative process, in which a team of four ergonomics experts analyzed a large number of jobs and identified the observable characteristics of each stressor. From these analyses, decision rules and benchmark example videotapes were developed (Latko et al., 1997). These verbal anchors and decision criteria are the foundation of the rating system. By including examples from a wide variety of industries and work tasks, the system allows ratings to be compared not only for jobs that are similar to each other, but also for different types of jobs. The benchmark examples have been rated by a reference group, and are used for training new users of the system.

The complete rating system includes a total of fifty-two attributes corresponding to the stressors described above (Table 3.2). The rating scales and corresponding guidelines for selected stressors are described in the Appendix.

II. Subjects

Teams of subjects were recruited from health and safety professionals at various workplaces. Requirements for selection included: 1) at least 3 team members were available from each organization, 2) each member had a fundamental understanding of basic ergonomic principles, and 3) the entire team could spend at least two full days learning and applying the method. Four teams participated (Table 3.3), with 3 - 5 members per team. The total number of subjects was 17. Educational backgrounds included ergonomics, mechanical engineering, industrial hygiene, and occupational medicine. Field

experience levels in ergonomic evaluation ranged from novice to expert, with ten of the subjects having more than five years of experience evaluating jobs for physical stressors.

Each rating group performed their ratings independently. Three of the groups performed their ratings in the university laboratory, while the fourth group performed their ratings in a computer laboratory near their place of employment.

Table 3.2: Physical stress parameters included in complete rating system.

stressor	aspects rated
repetition	overall level of hand activity
forceful exertions	peak and average hand force
localized mechanical stress	peak and average mechanical stress on: • fingers • wrist/palm • forearm • elbow
posture	peak and average posture, peak and average angular velocity, and % exertion time for: • fingers • wrist • forearm • elbow • shoulder • neck • back
low temperature	peak and average hand exposure to low temperature
vibration	peak and average hand-arm vibration exposure
jerk/impulse	peak and average jerk/impulse loading of the hand

III. Process

Each team received approximately 3 hours of training in the observational rating technique. Each rater received written guidelines describing the rating criteria. These criteria were explained by an investigator, and supplemented by a series of videotaped example jobs designed to be used as benchmarks for the ratings. The example jobs had been previously rated by the group who developed this observational rating technique. Three example jobs were shown for each stressor, generally illustrating "low", "medium", and "high" stress for each. During the videotape training, subjects were encouraged to rate the jobs using the written criteria. An investigator gave the raters feedback on their numerical ratings compared to the reference ratings.

Subjects received no feedback on ratings after the training period was complete.

Table 3.3: Composition of the four rating teams, including work organization, primary educational background, and years of experience in ergonomic job analysis.

Team	n	organization	rater background	years exp.	# jobs rated
1	5	state occupational safety and health research program	• ergonomics • industrial hygiene • ergonomics • industrial hygiene • ergonomics	• >5 • <1 • 2 • <1 • >5	12
2	3	national occupational safety and health research institute	• mechanical engineering • ergonomics • ergonomics	• <1 • >5 • >5	45
3	4	university research staff	• ergonomics • ergonomics • ergonomics • ergonomics	• >5 • >5 • >5 • >5	45
4	5	national occupational safety and health research institute	• ergonomics • medicine • medicine • industrial hygiene • ergonomics	• >5 • 0 • >5 • <1 • >5	12

After the training was complete, the group rated a set of twelve jobs. The jobs rated were from three industries: industrial container, automotive components, and office furniture manufacturing, and had been previously selected for inclusion in an epidemiologic study assessing the relationship between exposure to repetitive work and prevalence of upper limb musculoskeletal disorders in workers. The twelve jobs included assembly tasks, machine loading/monitoring, and material handling operations (Table 3.4). Subjects were shown looped videotapes which included a 3-5 minute segments of each job. These tapes had been recorded in real-time during normal working conditions. The videotaped segments were repeated as many times as necessary during the rating session.

In order to examine the rating variability over a larger set of jobs, two of the groups (groups 2 and 3) rated an additional 33 jobs. These additional jobs were from the same three companies as the initial 12 jobs, and included a wide variety of operations. In addition, group 2 returned to the laboratory seven

weeks after their initial rating session to re-rate a subset of twelve jobs in order for test-retest reliability to be assessed.

Table 3.4: Industry and job descriptions of the 12 jobs which were rated by all four teams..

code	industry	job description	production rate
a	industrial container	get small fiber drum, position and hold on crimping machine, transfer to conveyor	600 / hour
b	office furniture	secure padding and fabric to seat of office chair using hand held staple gun and scissors	18 / hour
c	automotive components	place and weld center wire in spark plug (using tweezers); activate welder; remove assembly	1000 / hour
d	industrial container	get small metal handle, position and hold in rivet machine, transfer to conveyor	1500 / hour
e	industrial container	get large fiber drum, position and hold horizontally on sealing machine, remove from machine	150 / hour
f	automotive components	transfer spark plugs from box to conveyor	5000 / hour
g	industrial container	transfer large fiber drum from conveyor to floor	500 / hour
h	industrial container	remove plastic lid from injection molding machine, trim small piece of flash as needed	120 / hour
i	industrial container	unload small fiber drum from forming machine and transfer to conveyor	600 / hour
j	industrial container	feed paper drum lid through gluing machine then place on small drum	600 / hour
k	automotive components	place spark plug shell in welding fixture, activate; transfer to press, activate; remove assembly	3400 / hour
l	office furniture	load desk top panel into edging machine	107 / hour

All groups followed a modified nominal group technique (Gustafson, 1973), which is a consensus rating process (Latko et al., 1997). Each team member rated the job independently for all stressors included in the system. After the initial ratings were complete, the group discussed their ratings with the goal of reaching consensus, which was defined as a range of no more than one unit on the 10 cm scale.

All ratings were performed on networked Macintosh™ computers. The rating program was written in Hypercard™ (Apple Computer, 1995), and consisted of a series of 10 cm sliders. The written guidelines were embedded in help screens accessible from each scale. Each rater computer transmitted the ratings to a central "collector" computer running a Labview™ program (National Instruments, 1995). The collector recorded all initial ratings, and

displayed all raters' scores for each scale when the initial rating was complete. The display also indicated whether the raters were in consensus. If a scale did not initially meet the consensus criteria, the raters discussed the rationale for their ratings, and adjusted their ratings accordingly. These changes were displayed in real time by the collector, and the consensus criterion was continuously re-evaluated. Raters were not allowed to continue to the next scale until consensus was achieved on the current scale. This procedure was followed for each of the 52 scales. When consensus had been achieved on all scales, the final ratings were recorded. The time required for the initial ratings and consensus discussion was also recorded.

IV. Analysis

Comparisons were done on the initial and final ratings of the 17 raters to determine the degree to which they were able to apply the criteria consistently. The ratings of the 45 jobs by groups 2 and 3 were also compared. Because all ratings were done using continuous 10 cm sliding scales, parametric statistics were used for analysis. Data analyses were performed using JMP™ statistical software (SAS Institute Inc., 1995). For brevity, only seven of the scales pertaining to the hand and wrist are presented in this chapter. These seven scales are representative of the remaining scales, which are presented in Appendix B.

Results

Inter-rater reliability:

a) 12 jobs, 17 raters

Figures 3.1 through 3.4 show the mean initial and final ratings for each of the twelve jobs. There were no statistically significant differences between the initial and final ratings for repetition, peak and average force, peak and average contact stress, or average wrist posture. There was a statistically significant difference between the initial and final ratings of peak wrist posture ($p < 0.02$). On average, the final ratings of peak wrist posture were 0.4 units higher than the initial ratings. Although this difference was statistically significant, the absolute magnitude of the difference is relatively small and within the 1 unit range defined as consensus for the rating process. For all jobs and stressors, the

standard deviation of the final ratings was less than the standard deviation of the initial ratings. On average, the standard deviation among the 17 raters in the final ratings was 0.4 units lower than the standard deviation of the initial ratings.

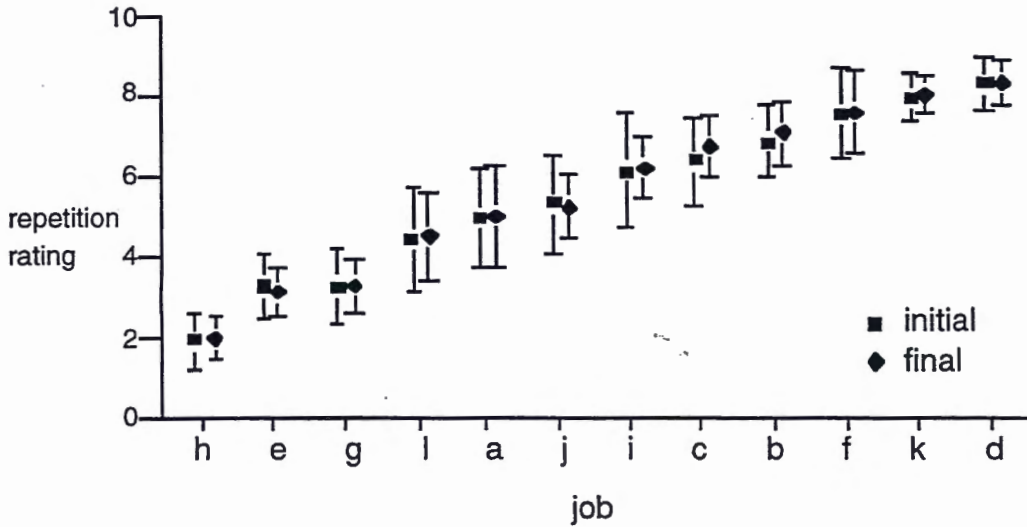


Figure 3.1: Average initial and final repetition ratings (± 1 SD), for 17 raters and 12 jobs. Job labels correspond to the jobs listed in Table 3.4, and are shown in order of increasing mean magnitude.

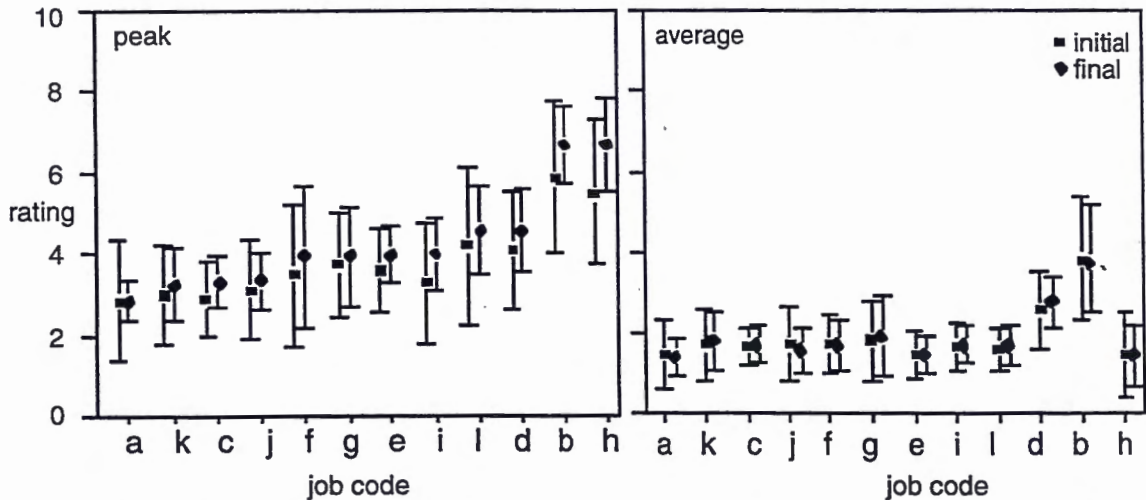


Figure 3.2: Average initial and final force ratings (± 1 SD), for 17 raters and 12 jobs.

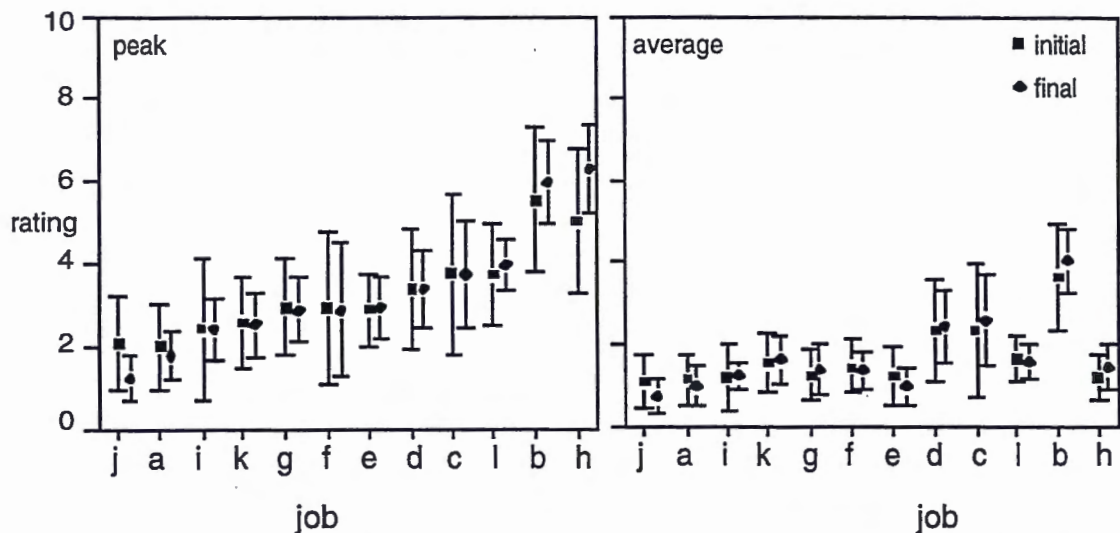


Figure 3.3: Average initial and final finger contact stress ratings (± 1 SD), for 17 raters and 12 jobs.

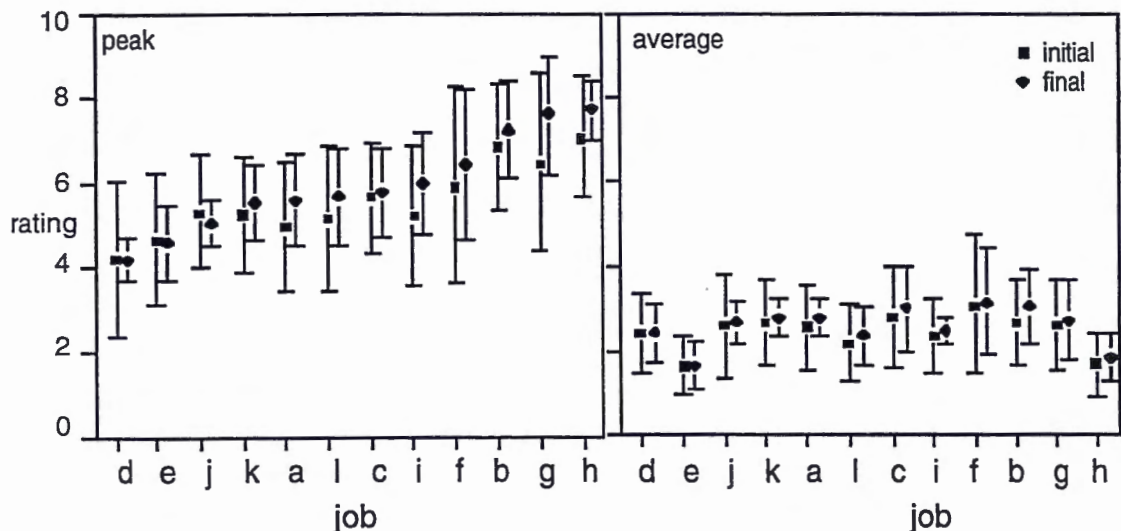


Figure 3.4: Average initial and final wrist posture ratings (± 1 SD), for 17 raters and 12 jobs.

The variance in the ratings due to differences between raters and differences between jobs was calculated for the initial (before consensus) ratings. Because the ratings of the group members were forced to converge during the consensus process, final ratings were compared only as group averages, rather than between individuals. Table 3.5 lists the estimates of variance due to jobs, raters (groups), and model error for both the initial (individual) and final (group) ratings.

To test significance of the agreement between raters, intra-class correlations (ICC) were calculated. This statistic is the ratio of between-jobs variance to total model variance (Shrout and Fleiss, 1979). For example, in the initial repetition ratings, job-to-job variance accounted for 79% of the total rating variance (Table 3.5). For the initial ratings, the data were modeled as ICC(2,1), where a random sample of 12 targets was rated by 17 judges; judges were considered random effects (Shrout and Fleiss, 1979). Final ratings were modeled as ICC(2,1) with 12 random targets and 4 groups. The ICCs for initial and final ratings are listed in Table 3.5. The initial rating correlations were all statistically significant ($F_{11, 176}$) at $\alpha=0.01$. All final group rating correlations were significant at $\alpha=0.05$ ($F_{11,33}$). Care must be taken when viewing ICCs with regard to statistical significance; significance level is dependent on sample size, and the number of jobs and raters was relatively large in this study, especially for the analysis of individual ratings. Morrow and Jackson (1993) state that significance of ICCs is not meaningful and that acceptable reliability should be determined from the standpoint of "practical significance". They state that the actual value of "practical significance" will vary with the application, and suggest ICC values in the range of 0.7 - 0.8 as indicating practical significance. Fleiss (1986) suggests that ICC values below 0.4 indicate poor reliability, values between 0.4 and 0.75 represent fair to good reliability, and values above 0.75 represent excellent reliability. Using these criteria, the initial ratings of repetition exhibit "excellent" reliability, while all other initial ratings fall into the "poor" category. According to these criteria, the final group ratings are markedly better than the initial, individual ratings. For the final ratings, the ICC for repetition still is above the "excellent" threshold, while peak and average force, peak and average contact stress on the fingers, and peak wrist posture are in the "fair to good" category. Only average wrist posture ratings exhibit poor reliability.

Table 3.5: Variance components and intra-class correlations for initial and final ratings of seven stressors (12 jobs, 17 raters in 4 groups).

initial ratings				intra-class correlation (raters)
stressor	var(jobs)	var(rater)	var(error)	
repetition	4.1	0.24	0.83	0.79**
pk force	0.86	0.24	1.99	0.28**
avg force	0.41	0.08	0.69	0.35**
peak contact stress	1.10	0.05	2.10	0.34**
avg contact stress	0.53	0.10	0.73	0.39**
peak wrist posture	0.59	0.08	2.69	0.18**
avg wrist posture	0.13	0.19	0.94	0.10**

final ratings				intra-class correlation (groups)
stressor	var(jobs)	var(groups)	var(error)	
repetition	4.24	0.34	0.35	0.86**
pk force	1.42	0.47	0.73	0.54**
avg force	0.36	0.16	0.31	0.43**
peak contact stress	2.08	0.09	0.85	0.69**
avg contact stress	0.69	0.05	0.32	0.65**
peak wrist posture	0.88	0.05	1.19	0.41**
avg wrist posture	0.11	0.16	0.38	0.16*

* ICC significant at $\alpha = 0.05$

** ICC significant at $\alpha = 0.01$

Inter-rater reliability

b) 45 jobs, 2 groups:

The estimates of variance due to jobs and raters (groups) and the ICCs for the 45 jobs and 7 raters are listed in Table 3.6. In the initial ratings, the repetition ratings had an ICC of 0.90. The ratings for peak force, peak and average contact stress, and average wrist posture were within the 0.4 to 0.75 range, indicating fair to good reliability. Average force and peak wrist posture ratings exhibited poor reliability. After consensus, ICCs for ratings of repetition and peak force were both above 0.75. All other stressors except average force were in the fair to good range. The intraclass correlations of all stressors except average force improved from the individual to group final ratings. Intra-class

correlations after the initial ratings ranged from 0.29 to 0.90 (median = 0.44). The range of ICCs after consensus was 0.21 to 0.95 (median = 0.69). These correlations were all statistically significant ($F_{44, 264}$) at $\alpha=0.05$ or better.

Table 3.6: Variance components and intra-class correlations for initial and final ratings of seven stressors (45 jobs, 7 raters).

stressor	var(jobs)	var(rater)	var(error)	intra-class correlation (raters)
initial ratings				
repetition	6.21	<0.01	0.71	0.90*
pk force	1.82	0.10	1.37	0.55*
avg force	0.31	0.15	0.56	0.29*
peak contact stress	1.31	0.06	1.47	0.48*
avg contact stress	0.46	0.10	0.44	0.44*
peak wrist posture	0.94	<0.01	1.88	0.33*
avg wrist posture	0.65	0.07	0.67	0.44*
final ratings				
	var(jobs)	var(groups)	var(error)	intra-class correlation (groups)
repetition	4.02	<0.01	0.69	0.94**
pk force	2.35	0.04	0.62	0.79**
avg force	0.18	0.35	0.38	0.21*
peak contact stress	1.99	<0.01	0.79	0.73**
avg contact stress	0.49	0.19	0.24	0.55**
peak wrist posture	1.07	<0.01	0.86	0.57**
avg wrist posture	0.69	0.08	0.24	0.69**

* ICC significant at $\alpha = 0.025$ ** ICC significant at $\alpha = 0.01$

The group final ratings of the 45 jobs by Group 2 and Group 3 were compared. Regression formulas are indicated in Figure 3.5. The relationship between each pair of ratings was significant at $p < 0.05$. Repetition, average force, average contact stress for the fingers, and average wrist posture all had intercepts which were statistically significant ($p < 0.05$). These intercepts ranged from 0.72 to 1.23. Paired t-tests were performed on each pair of ratings. Of the

seven parameters evaluated in this study, significant systematic differences were present for average force (mean difference = 0.84) and average contact stress-fingers (mean difference = 0.62). Although these differences are statistically significant, they are within the one unit range defined as consensus.

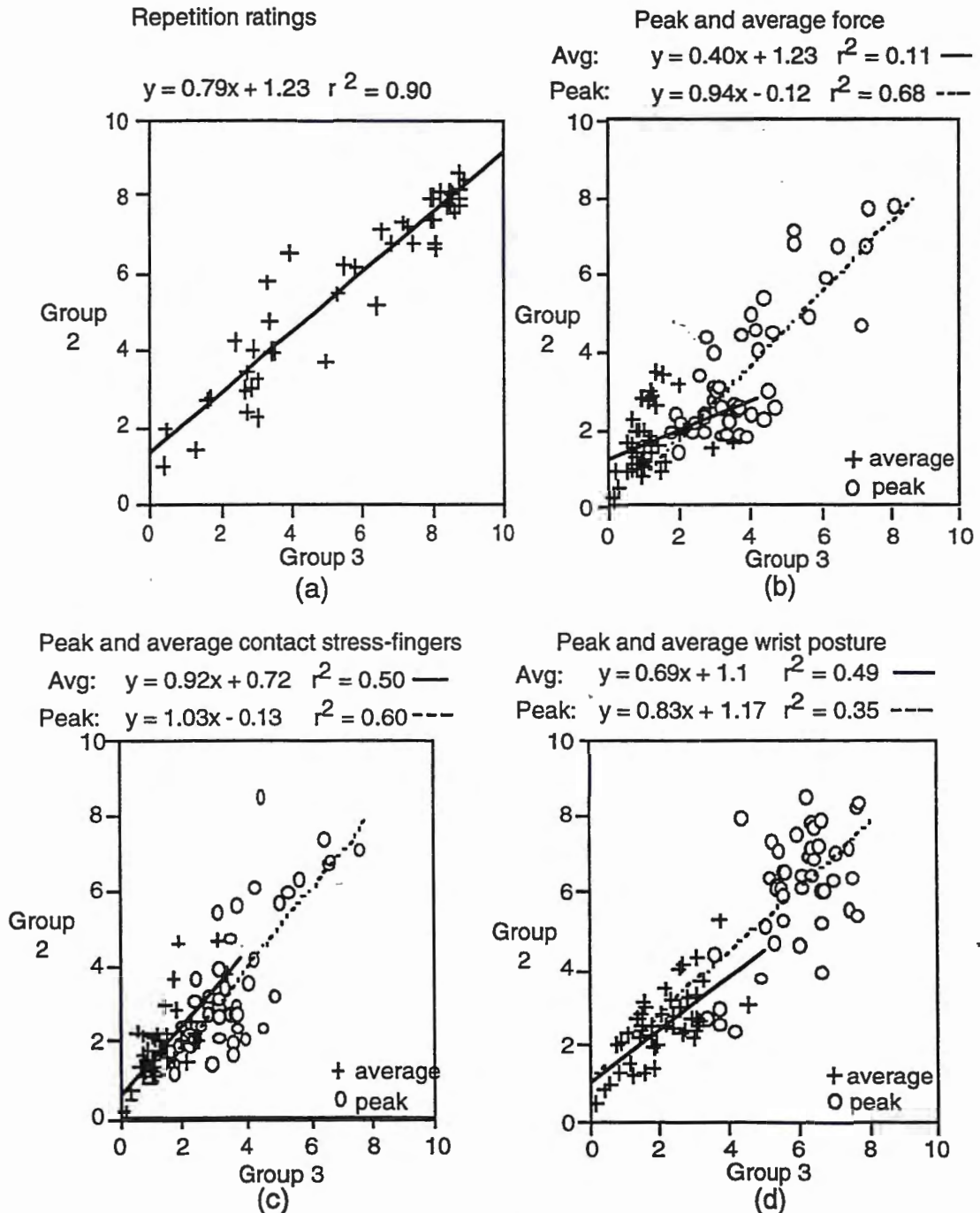


Figure 3.5: Agreement in average group ratings between Groups 2 and 3 for 45 jobs.

Test/retest reliability:

Group 2 re-rated a subset of twelve jobs seven weeks after their initial rating session. Ratings for each stressor were modeled as a function of job, rater, and day (session 1 or session 2). Table 3.7 lists the variance component estimates and indicates the significance of the terms for each of the stressors. Rater effects were included in the model of the individual ratings, but not in the final ratings. In the initial ratings, job was a significant effect in predicting the rating of each stressor. Rater was significant in the initial ratings of average contact stress only. For the final ratings, job was the only significant term in any of the stressor models. There were no significant day effects for either initial or final ratings. In general, the amount of variability due to jobs increased from initial to final ratings, while the amount of variability due to day remained low (Table 3.7).

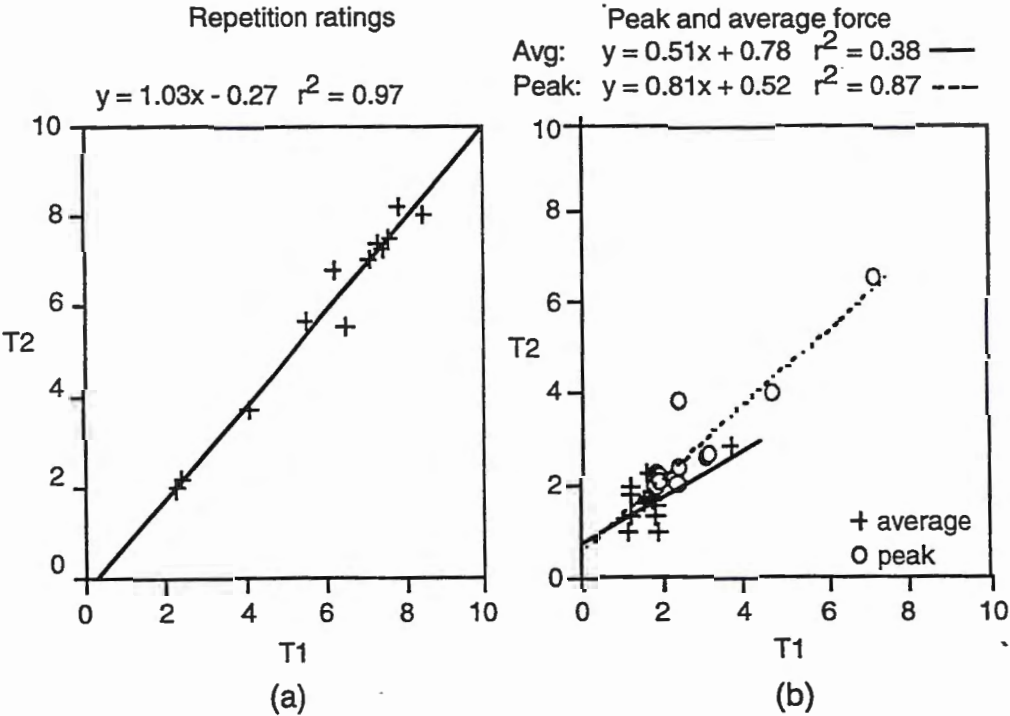
Table 3.7: Significance of job, rater, and day in test-retest study.

stressor	variance component estimate					model r^2
initial ratings	job	rater	day	error	%var (job)	
repetition	4.30*	0.03	<0.01	0.42	90.3%	0.92
pk force	1.04*	0.05	<0.01	0.76	55.9%	0.66
avg force	0.21*	0.01	<0.01	0.20	48.8%	0.59
peak contact stress	2.06*	0.03	0.04	1.05	64.8%	0.72
avg contact stress	0.48*	0.03*	<0.01	0.15	71.6%	0.81
peak wrist posture	0.66*	0.01	0.07	1.87	25.3%	0.40
avg wrist posture	0.14*	<0.01	<0.01	0.44	23.3%	0.36

final ratings	job	day	error	%var (job)	
repetition	4.50*	<0.01	0.08	98.0%	0.99
pk force	1.92*	<0.01	0.14	92.8%	0.96
avg force	0.21*	<0.01	0.13	60.0%	0.80
peak contact stress	2.99*	<0.01	0.34	89.5%	0.95
avg contact stress	0.48*	<0.01	0.07	85.7%	0.93
peak wrist posture	0.82*	0.11	0.63	52.6%	0.79
avg wrist posture	0.18*	<0.01	0.09	64.3%	0.84

* effect significant at $p < 0.05$

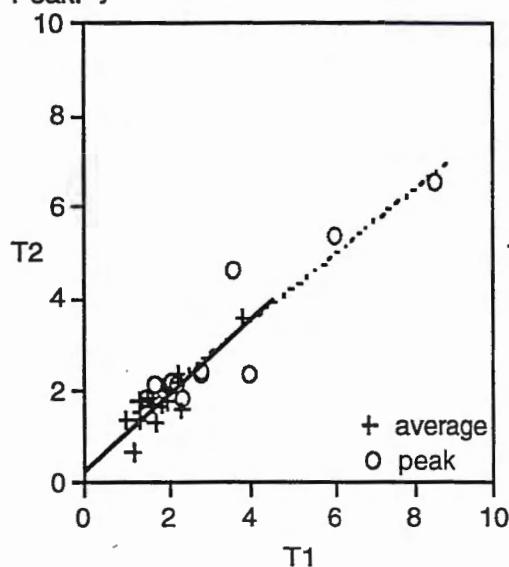
Paired t-tests revealed no systematic differences in the final group ratings between the two rating sessions for any of the scales. When modeled without the day effect, R^2 values for average force and peak and average wrist posture at time 1 vs. time 2 were relatively low (0.38, 0.32 and 0.45, respectively) (Figure 3.6). Average force and wrist posture ratings covered a range of only about 2 units on the ten cm scale, which probably contributes to the small amount of variance explained by the model. While this variability is statistically significant, it is likely not meaningful in a practical sense, since all ratings fell within a narrow range. Peak wrist posture ratings represented a range of about 5 units on the scale; the low r^2 in this case indicates a potentially more meaningful level of variability in the ratings.



Peak and average contact stress-fingers

Avg: $y = 0.83x + 0.26$ $r^2 = 0.76$ —

Peak: $y = 0.71x + 0.69$ $r^2 = 0.86$ ---

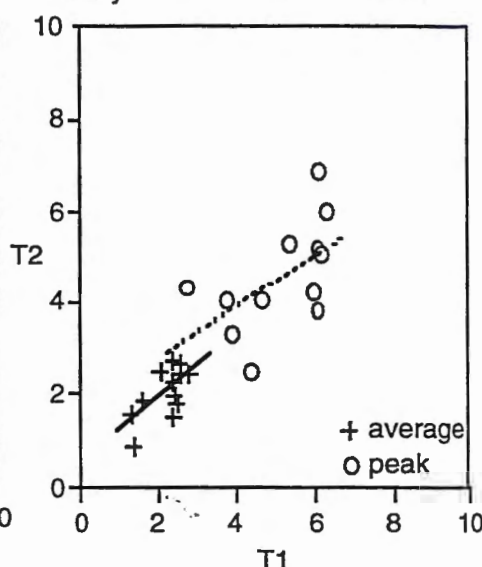


(c)

Peak and average wrist posture

Avg: $y = 0.74x + 0.44$ $r^2 = 0.45$ —

Peak: $y = 0.57x + 1.65$ $r^2 = 0.32$ ---



(d)

Figure 3.6: Agreement in Group 2 average ratings at times 1 and 2 for 12 jobs.

Discussion

1. Comparison to other methods

Although observational techniques are often used to characterize physical stressors in the workplace, relatively few studies have specifically examined the variability between observers. Keyserling and Wittig (1988) compared expert ratings on three-point scales for four categories of ergonomic stress (materials handling, upper extremity activities, posture, and energy expenditure) for ten jobs. They found perfect agreement among the five experts on 30% of the job/stressor combinations. Three of the five experts were in agreement on 87.5% of the job/stressor combinations. Among the four categories of stressors characterized, the posture ratings achieved the lowest level of agreement, although the difference was not statistically significant. The five experts did not achieve perfect agreement on any of the ten jobs evaluated for posture stresses.

Stetson et al. (1991) evaluated an observational technique for quantifying hand exertions in manual work. Four analysts counted and classified exertions in two jobs based on duration and posture or force characteristics. No statistically significant differences in exertion counts were found, although the coefficient of variation ranged up to 40% (mean = 12%). Baty et al. (1986) did not find acceptable inter-rater reliability for categorical estimates of trunk posture during nursing duties. They found high levels of agreement (80% - 90%) among the seven raters for the frequently occurring categories of bend, stand, walk, and sit. Agreement for the infrequent behaviors, which included squat, kneel, push, pull, carry, reach, and twist, ranged from 20% to 96%. These results indicate that some postural behaviors are more difficult to define than others.

Van Der Beek et al. (1992) examined the inter-observer reliabilities of posture and load categorization of lorry drivers. Several posture and load parameters were examined, including gross body posture, trunk position, general arm posture, and loads lifted, carried, pushed, or pulled. Postures and loads were classified into four or five categories, depending on body part. Percent agreement between the two raters ranged from 78.7% to 96.9% ($0.24 \leq \kappa \leq 0.89$). Using similar posture and load classifications to those of van der Beek, DeLooze et al. (1994) conducted a study during a simulated case packing and stacking task. The two raters exhibited strong agreement for all body parts ($\kappa \geq 0.66$), although the observer ratings did not agree as well with the direct measures that were recorded. This suggests that the observers tended to make similar mistakes.

In a study of test-retest reliability of the "posture targetting" technique, Corlett et al. (1979) had ten subjects make observational recordings of ten body segments for six different overall body postures. For the six postures, correlations ranged from 0.67 to 0.84. In comparison, the correlation between wrist posture ratings in the test-retest reliability segment of the current study was 0.67 for peak and 0.56 for average. Corlett et al. only present overall correlations for each body posture, and do not report correlations for each body segment individually. From the available data, it appears that the method presented in this chapter exhibits test-retest correlations at the low end of those found for posture targetting. Although not presented in this paper, the test-retest correlations for the elbow and shoulder joints were higher than those reported for the wrist. Correlations for average elbow and shoulder postures were 0.78

and 0.83, respectively, while correlations for the peak elbow and shoulder postures were 0.78 and 0.72. These findings are in closer agreement with the Corlett data. The lower agreement for wrist posture may be due to the inherent difficulty in observing wrist posture. The wrist is often obscured by the rest of the body or by work objects. In addition, because the wrist location is affected by the position of the more proximal joints of the arm and the torso, it is often difficult to maintain an orthogonal view of the joint.

There is no information given on the range of postures represented in the Corlett study. In the current study, there was a relatively narrow range in the identified levels of average wrist posture, with average postures ratings ranging from approximately 1 to 3 (see Figure 3.6(d)). This tends to make a low correlation coefficient likely even if the absolute magnitude of the differences is relatively small. In fact, the ratings of average posture in the second session were all within ± 0.5 unit of the first session ratings. Although it is believed that this low range is due to particular jobs included in the study, the study design does not rule out lack of sensitivity of the method.

The technique presented in this chapter is unique among most other observational techniques in that it is continuous rather than categorical. The expert raters in Keyserling's study (1988) reported difficulty in strictly classifying stresses into only three categories. The raters indicated that the observed stresses sometimes did not fall obviously into one of the three defined categories, and thus elected to qualify their recorded scores with a plus or a minus. Several of the raters expressed a preference for a continuous rating scale. Using a continuous rating scale also reduces the opportunity for misclassification of stresses near the end points of the categories. For example, RULA, an observational technique for assessing upper limb posture, assigns points to upper limb position (shoulder posture) based on inclusion in one of four categories: -15° to 15° , 15° to 45° , 45° to 90° , or $>90^{\circ}$ (McAtamney and Corlett, 1993). Although such a categorical classification appears straightforward, it is easy to see how an error in classification can be critical for postures near the categorical boundaries. For example, a 10° underestimation for a posture with 20° of flexion will have no effect on the rating, while the same 10° error will result in misclassification if the true posture is 50° . Although a few such misclassifications may not be critical for a given job, if such errors are systematic across stressors, a biased assessment may result.

Because the method evaluated in this paper is continuous rather than categorical in nature, it is difficult to directly compare the findings of the above studies to the current findings. In order to attempt a comparison, arbitrary boundaries would have to be set to determine the categorization of the ratings (e.g. five categories, each representing a 2 unit range on the scale, or three categories, with the scale equally divided into thirds). Examining the variance due to raters gives an indication of the ranges that can be expected for a given stressor. In the current study with 17 raters and 12 jobs (Table 3.5), if the rater variance is converted to a 95% confidence interval (± 1.96 SD), mean expected variability for all stressors due to raters alone is 1.4 for the initial ratings and 1.6 for the final ratings. This would suggest that a categorization into approximately 5 categories would yield a high level of agreement between raters. The study with 7 raters and 45 jobs had a lower level of rater variability; the maximum expected 95% confidence intervals (± 1.96 SD) were 1.0 for both the initial individual ratings and group final averages over the seven stressors. Based on this evaluation of more jobs with a smaller set of raters, categorization into more than five levels appears reasonable. Based on these ranges, it appears that this method exhibits resolution and repeatability similar to those described above.

The rater variance was slightly lower for the study with the 2 groups and 45 jobs than it was for the study with 4 groups and 12 jobs. Similarly, the ICCs were generally higher for the 2 group, 45 job study. There are several possible reasons for this. The two groups rated more jobs, which increased the range of potential exposures evaluated. A smaller number of raters also decreases the opportunity for multiple "outliers". Finally, six of the seven raters in groups 2 and 3 were highly experienced (Table 3.3), while only five of the ten raters in groups 1 and 4 had more than five years of experience. It is possible that the experience level among these seven raters contributed to the consistency in their ratings.

For all analyses performed in this study, the repetition scale had high intraclass correlations (above 0.75), indicating a high degree of inter-rater reliability. Because these jobs were initially selected from a pool of jobs used in a study of repetition, this stressor had the widest range of possible exposures in the job sample. It is encouraging that the level of agreement was high over nearly the entire range of the scale (from approximately 2 to 8 on the 10 cm scale), even though this scale requires the most integrating and judgment by

the analysts. Based on the definitions of the stressors and rating benchmarks (see Appendix), repetition is the most qualitative of the stressors. For example, ratings of posture are based on observations of joint position relative to maximum range of motion, parameters that could theoretically be measured from the videotape. Repetition, however, requires analysts to subjectively assess movement speed and exertion time, using descriptors such as "slow", "steady", "rapid", and "brief". The high levels of agreement among different raters and groups when assessing repetition suggest that the combination of written guidelines and videotaped benchmark examples used for training adequately promote rating consistency.

In contrast, because of the nature of the jobs available for this study, there were relatively small ranges of average force, contact stress, and wrist posture represented. These other stressors were not specifically controlled in the original job pool, thus their distribution was due merely to chance. As a percent of total variance, inter-rater variability was highest in the ratings of average force, contact stress, and wrist posture. This was due mainly to the relatively low range of ratings given to these stressors. Given the selection of jobs included in this study, there was no expectation that a wide range of exposures to these stressors would be identified.

In some cases, the range of potential exposures in terms of average values is limited by physiological tolerances of workers. For example, it has been shown that it is difficult for workers to maintain average exertion levels above 15% - 20% of maximum for extended periods (Bystrom and Fransson-Hall, 1994). Similarly, certain extreme postures may be difficult to sustain. Given the current definition of the endpoints on the ten cm scales, it is unlikely that a worker would ever receive a rating higher than two (corresponding to 20% of maximum) for average force. In light of the inter-rater variability identified in this study, further investigation is warranted to determine the degree to which the method is able to discriminate among jobs over the widest possible range of average force, contact stress, and posture levels.

In addition to the variability due to differences in perceived exposure by the groups, there is also variability due to the groups choosing to rate different activities viewed on the videotape. Groups were instructed to rate "regularly occurring" parts of the job, and groups sometimes interpreted this differently. For example, for job "f", which was a spark plug transfer operation, group 2 elected to include in their ratings a stock handling task which occurred several

times throughout the day, but was not part of the main job cycle. Their peak ratings of force, contact stress, and wrist posture are significantly higher than the other groups', because they represent stressors encountered in this second task which the other groups did not rate. This situation highlights the need to define which parts of a job should be analyzed

The issue of variability in job tasks is an important issue in analyzing the physical stresses present in a job. Sufficient data currently are not available to determine how frequently an irregular element must occur, or how severe that exposure must be, for it to have a significant impact on worker health. Towards this end, we recommend that irregular elements (i.e. tasks that occur several times throughout the work day, but are not part of the main job cycle) be analyzed separately. Total exposure ratings can then be calculated as a time weighted average (TWA). Utilizing this process will provide an overall score for the job, while also providing information about the relative contribution of the various activities to the overall stress.

II. Individual vs. group ratings

It was hypothesized that using a group consensus technique would provide more consistent results than individual ratings. In nearly all cases, the intraclass correlations were higher for the group final ratings than for the initial, individual ratings. For example, when 4 groups (17 total raters) rated twelve jobs each, intraclass correlations for the initial ratings ranged from 0.10 to 0.79 (median=0.34), while the intraclass correlations of the final group ratings were 0.16 - 0.86 (median = 0.54). Similar increases were seen in the 45 jobs rated by the two groups (Table 3.6) and in the test/retest reliability study (Table 3.7).

From a mathematical standpoint, decreased variability after forcing group consensus may seem to be an obvious expectation, since any method of grouping and averaging the individual ratings would have decreased the variability. In this study, however, raters were forced to adjust their initial ratings in order to reach consensus. Thus, theoretically, the rating groups could have converged at one of the extremes rather than at some central value. The consistency in the group ratings among the four groups indicate that the discussion and adjustment process assists groups in converging at a consistent rating.

The quality of group decisions has been widely explored in the organizational behavior literature (Lorge et al., 1958). Although it was not

possible in the current study to determine the quality of ratings as compared to a correct or true value, reliable application by different users is an important issue in itself, and many of the factors applying to quality in an absolute sense are also applicable in the sense of reliability. One reason that is commonly cited for the superiority of group decisions over those of an individual is the error-correcting property of social interaction (Blau and Scott, 1962). Holloman and Hendrick (1971) suggest that this social interaction allows the information to be viewed from "varied perspectives and backgrounds of knowledge and experience", thus reducing the probability of the final decision being tainted by any individual's bias. With the job analysis method described in this paper, this interaction not only can reduce the potential effect of bias on the observations, but also may help to remove bias in the interpretation of the decision guidelines.

Another factor which has been proposed to account for the improved performance of groups over individuals is that, merely by chance, groups have a higher probability of including a member who is able to solve the problem (Lorge and Solomon, 1955). Relating to this study, it is possible that some individuals are simply "better", by virtue of their knowledge or experience, at rating these stressors. In order for this theory to hold, however, the other members of the group would have to consistently converge on the ratings of this one member. An analysis of the change in individual ratings before and after consensus discussions did not reveal an individual member of any group who changed his/her ratings appreciably less than any other.

A third potential explanation for the superiority of group ratings over individual ratings is due to the complex nature of work tasks. Many jobs require a complex series of movements, often occurring at high speeds, thus making it difficult for an observer to notice everything. Systematic observation (e.g. following a checklist or other structured method) can reduce the chance of overlooking critical elements, but can not guarantee a comprehensive analysis. Having multiple observers increases the probability that some analyst will notice any important aspects of the job.

These three factors, error-correcting social interaction, chance for more highly skilled raters in groups, and the advantages of having multiple observers to capture a complex task support the findings in this study which indicate that groups exhibit more consistency in ratings than do individuals. This observational technique can be used to quantify exposures to physical

stressors in manual work, without forcing analysts to strictly categorize exposures.

The current consensus method sometimes becomes labored when one person adopts an extreme view. In future applications with this method, it may be desirable to perform the initial ratings, discuss differences, and allow raters to change their ratings but not force consensus. In this case, the mean group rating would be used, and the range or standard deviation would be a measure of rater agreement. In the twelve jobs which were rated by 17 raters, there were no statistically significant differences in the initial and final overall average ratings (Figures 3.1 through 3.4). This suggests that if the goal of an analysis is merely to obtain a group average stress rating, the consensus process may not be necessary. However, because the comparison and discussion of ratings force the raters to share their thought processes and observations of the job, we feel it greatly enhances the overall understanding of the stresses.

Although this method has been mainly applied by groups using the consensus procedure, it can also be used by individuals, with the recognition that there may be an increase in variability in the ratings. The increased variability in individual ratings can be attributed to two factors: 1) differing interpretations of the magnitude of the stresses observed, and 2) raters "missing" something while watching the job. Many jobs involve a complex series of operations, and having two or more analysts helps insure that all aspects of the job are considered in the ratings.

III. Limitations and future work

Despite the fact that this study was performed in the laboratory, the data illustrated above reflect conditions similar to those that might be encountered in actual workplaces. The rating groups were presented with videotapes of real jobs, which had not been edited to remove extraneous work elements. Thus each rater was responsible for determining which elements were "regularly occurring" parts of the work cycle, which should be rated, as opposed to irregular elements that should be ignored. Decisions to include or not include such elements have a large impact on the "peak" ratings which are reported above. It is important, not only for this method but for any job analysis technique, to develop criteria for determining which, if any, aspects of a job should be excluded from analysis.

Another consideration is the ability of the system to be used in real-time, on the plant floor. The analysts in this study viewed all jobs on videotape; it was possible to view the same segments multiple times, and tasks could be viewed in slow-motion or freeze-frame. Further examination is needed to determine the level of increased variability that may be introduced by performing the ratings in real-time.

Further studies are planned to examine the impact of additional training on rating consistency. The raters in this study received approximately 3 hours of training in the method. This training consisted of a review of the written rating guidelines and viewing the benchmark jobs on videotape. Since the method consists of 51 rating scales when all stressors and body parts are considered, this is a relatively short amount of training for each scale. Rating groups also did not receive any feedback on their ratings after the initial training period. It is hypothesized that more extensive training and feedback would improve the consistency of individual and group ratings.

Clarification of the written decision rules and more extensive use of videotaped benchmark examples may also improve the consistency of ratings. More extensive development of these rating aids may provide raters with a more precise framework on which to base their ratings.

It is important to note that measures of inter-rater reliability do not indicate the accuracy, or validity, of ratings. Future studies will compare the ratings of various stressors using this method to data obtained from instrumental methods, such as EMG and electrogoniometry. Studies are also under way to evaluate the relationship between exposure to physical stresses as quantified by this method and health outcomes in workers

References

- Armstrong, T. A.; Buckle, P.; Fine, L. J.; Hagberg, M.; Jonsson, B.; Kilbom, A.; Kuorinka, I. A.; Silverstein, B. A.; Sjøgaard, G.; & Viikari-Juntura, E. R. (1993). A conceptual model for work-related neck and upper-limb musculoskeletal disorders. *Scandinavian Journal of Work & Environmental Health*, 19, 73-84.
- Armstrong, T. A.; Foulke, B. J.; Martin, B. J.; Gerson, J.; & Rempel, D. M. (1994). Investigation of applied forces in alphanumeric keyboard work. *American Industrial Hygiene Association Journal*, 55(1), 30-35.
- Armstrong, T. J.; Bir, C.; Finsen, L.; Foulke, J.; Martin, B.; Sjøgaard, G.; & Tseng, K. (1993). Muscle responses to torques in hand held power tools. *Proceedings of the International Society of Biomechanics*, 114-115.
- Armstrong, T. J.; & Chaffin, D. B. (1978). An Investigation of the relationship between displacements of the finger and the wrist joints and the extrinsic finger flexor tendons. *Journal of Biomechanics*, (11), 119-128.
- Armstrong, T. J.; Fine, L. J.; Goldstein, S. A.; Lifshitz, Y. R.; & Silverstein, B. A. (1987). Ergonomics considerations in hand and wrist tendinitis. *The Journal of Hand Surgery*, 12a(5, part 2), 830-837.
- Baty, D.; Buckle, P. W.; & Stubbs, D. A. (1986). *Posture Recording by Direct Observation, Questionnaire Assessment and Instrumentation: a Comparison Based on a Recent Field Study*. In N. Corlett; J. Wilson; & I. Manenica, eds., *The Ergonomics of Working Postures, Models, Methods, and Cases*. London: Taylor and Francis.
- Blau, P. M.; & Scott, W. R. (1962). *Formal organizations*. San Francisco: Chandler.
- Bystrom, S.; & Fransson-Hall, C. (1994). Acceptability of intermittent handgrip contractions based on physiological response. *Human Factors*, 36(1), 158-171.

- Cannon, L. J.; Bernacki, E. J.; & Walter, S. D. (1981). Personal and occupational factors associated with carpal tunnel syndrome. *Journal of Occupational Medicine*, 23(4), 255-258.
- Clarke, R. G. (1961). The limiting hand skin temperature for unaffected manual performance in the cold. *Journal of Applied Psychology*, 45, 193-194.
- Corlett, E. N.; Madeley, S. J.; & Manenica, I. (1979). Posture Targetting: A technique for recording working postures. *Ergonomics*, 22(3), 357-366.
- deKrom, M.; Kester, A.; Knipschild, P.; & Spaans, F. (1990). Risk Factors for carpal tunnel syndrome. *American Journal of Epidemiology*, 132(6), 1102-1110.
- DeLooze, M. P.; Toussaint, H. M.; Ensink, J.; & Mangnus, C. (1994). The validity of visual observation to assess posture in a laboratory-simulated, manual material handling task. *Ergonomics*, 37(8), 1335-1343.
- Drury, C. G. (1987). A biomechanical evaluation of the repetitive motion injury potential of industrial jobs. *Seminars in Occupational Medicine*, 2(1), 41-49.
- Edwards, R. (1988). Hypotheses of peripheral and central mechanisms underlying occupational muscle pain and injury. *European Journal of Applied Physiology*, 57(3), 275-281.
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York: John Wiley and Sons.
- Fransson-Hall, C.; & Kilbom, A. (1993). Sensitivity of the hand to surface pressure. *Applied Ergonomics*, 24, 3.
- Gelberman, R.; Hergenroeder, P.; Hargens, A.; Lundborg, G.; & Akeson, W. (1981). The carpal tunnel syndrome: a study of carpal canal pressures. *The Journal of Bone and Joint Surgery*, 63A(3), 380-383.
- Goldstein, S. A.; Armstrong, T. J.; & Chaffin, D. B. (1987). Analysis of cumulative strain in tendons and tendon sheaths. *Journal of Biomechanics*, 20(1), 1-6.

- Gustafson, D. H., Shulka, R. K., Delbecq, A., & Walster, G. W. (1973). A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, delphi groups, and nominal groups. *Organizational Behavior and Human Performance*, 9, 280-291.
- Hagberg, M.; Morgenstern, H.; & Kelsh, M. (1992). Impact of occupations and job tasks on the prevalence of carpal tunnel syndrome: a review. *Scandinavian Journal of Work, Environment, and Health*, 12, 277-279.
- Hammarskjold, E.; Harms-Ringdahl, K.; & Ekholm, J. (1992). Reproducibility of carpenters work after cold exposure. *International Journal of Epidemiology*, 9, 195-204.
- Hoffman, J.; & Hoffman, P. (1995). Staple gun carpal tunnel syndrome. *Journal of Occupational Medicine*, 27(11), 848-849.
- Holloman, C. R.; & Hendrick, H. W. (1971). Problem solving in different sized groups. *Personnel Psychology*, 24, 489-500.
- Hypercard (v. 2.3). (1995). Cupertino, CA: Apple Computer, Inc.
- JMP (v. 3.1). (1995). Cary, NC: SAS Institute, Inc.
- Karhu, O.; Kansil, P.; & Kuorinka, I. (1977). Correcting working postures in industry: a practical method for analysis. *Applied Ergonomics*, 8(4), 199-201.
- Keyserling, W.; Stetson, D.; Silverstein, B.; & Brouwer, M. (1993). A checklist for evaluating ergonomic risk factors associated with upper extremity cumulative trauma disorders. *Ergonomics*, 31(5), 735-747.
- Keyserling, W. M.; & Wittig, S. J. (1988). An analysis of experts' ratings of ergonomic stress. *International Journal of Industrial Ergonomics*, 2, 291-304.
- Kilbom, A. (1994). Assessment of physical exposure in relation to work-related musculoskeletal disorders-what information can be obtained from systematic observations? *Scandinavian Journal of Work, Environment, and Health*, 20 (special issue), 30-45.

- Kuorinka, I.; & Forcier, L. (1995). *Work related musculoskeletal disorders (WMSDs): a reference book for prevention*. London: Taylor & Francis.
- LabVIEW for Macintosh/Power Macintosh (v. 3.1.1). (1995). Austin, TX: National Instruments Corporation.
- Latko, W. A.; Armstrong, T. J.; Foulke, J. A.; Herrin, G. D.; Raboum, R. A.; & Ulin, S. S. (1997). Development and evaluation of an observational method for assessing repetition in hand tasks. *American Industrial Hygiene Association Journal*, 58(4), 278-285.
- Leas, D. E. (1993). *1993 & 1994 Official Rules of Diving and Code of Regulations*. (pp. 115 - 120). Indianapolis, IN: United States Diving, Inc.
- Loken, N. C. & Willoughby, R. J. (1971). *The Complete Book of Gymnastics*. (pp. 309-31). Englewood Cliffs, NJ: Prentice Hall, Inc.
- Lorge, I.; Fox, D.; Davitz, J.; & Brenner, M. (1958). A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. *Psychological Bulletin*, 55, 337-372.
- Lorge, I.; & Solomon, H. I. (1955). Two models of group behavior in the solution of eureka-type problems. *Psychometrika*, 20(139-148),
- Lundborg, G.; Gelberman, G.; Minter-Convery, M.; Lee, Y.; & Hargens, A. (1982). Median nerve compression in the carpal tunnel-functional response to experimentally induced controlled pressure. *The Journal of Hand Surgery*, 7(3), 252-259.
- Lundstrom, R.; & Johansson, R. S. (1986). Acute impairment of the sensitivity of skin mechanoreceptive units caused by vibration exposure of the hand. *Ergonomics*, 29(5), 687-698.
- Luopajarvi, T.; Kuorinka, I.; Virolainen, M.; & Holmberg, M. (1979). Prevalence of tenosynovitis and other injuries of the upper extremities in repetitive work. *Scandinavian Journal of Work, Environment, and Health*, 5(suppl. 3), 48-55.

- McAtamney, L.; & Corlett, E. N. (1993). RULA: a survey method for the investigation of work related upper limb disorders. *Applied Ergonomics*, 24(2), 91-99.
- Morrow, J. R.; & Jackson, A. W. (1993). How "significant" is your reliability. *Research Quarterly for Exercise and Sport*, 63(3), 352-355.
- Niebel, B. W. (1988). *Motion and Time Study (8th ed.)*. Homewood, IL: Irwin.
- Nilsson, T.; Burstrom, L.; & Hagberg, M. (1989). Risk assessment of vibration exposure and white fingers among platers. *International Archives of Occupational and Environmental Health*, 61, 473-481.
- Phalen, G. (1966). The carpal tunnel syndrome: Seventeen years' experience in diagnosis and treatment of six hundred fifty-four hands. *The Journal of Bone and Joint Surgery*, 48A(2), 211-228.
- Sauter, S. L.; Chapman, L. J.; Knutson, S. J.; & Anderson, H. A. (1987). Case example of wrist trauma in keyboard use. *Applied Ergonomics*, 18(3), 183-186.
- Schoenmarklin, R. W.; Marras, W. S.; & Leurgans, S. E. (1994). Industrial Wrist Motions and Incidence of Hand/Wrist Cumulative Trauma Disorders. *Ergonomics*, 37(9), 1449-1459.
- Shrout, P. E.; & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428.
- Silverstein, B. A.; Fine, L. J.; & Armstrong, T. J. (1986). Hand Wrist Cumulative Trauma Disorders in Industry. *British Journal of Industrial Medicine*, 43, 779-784.
- Silverstein, B. A.; Fine, L. J.; & Armstrong, T. J. (1987). Occupational factors and carpal tunnel syndrome. *American Journal of Industrial Medicine*, 11, 343-358.

- Sinclair, M. A. (1990). *Subjective Assessment*. In J. R. Wilson; & E. N. Corlett, eds., *Evaluation of Human Work: A Practical Ergonomics Methodology*. (pp. 58-88). London: Taylor and Francis Ltd.
- Stetson, D. S.; Keyserling, W. M.; Silverstein, B. A.; & Leonard, J. A. (1991). Observational Analysis of the Hand and Wrist: A Pilot Study. *Appl. Occup. Environ. Hyg.*, 6(11), 927-937.
- Sundelin, G.; & Hagberg, M. (1989). The effects of different pause types on neck and shoulder EMG activity during VDU work. *Ergonomics*, 32(5), 527-537.
- Tichauer, E. R. (1966). Some aspects of stress on forearm and hand in industry. *Journal of Occupational Medicine*, 8(2), 63-71.
- United States Figure Skating Association (1995). *1995-1996 USFSA Media Guide*. Colorado Springs, CO.
- van der Beek, A. J.; van Gaalen, L. C.; & Frings-Dresen, M. H. W. (1992). Working postures and activities of lorry drivers: a reliability study of on-site observation and recording on a pocket computer. *Applied Ergonomics*, 23(5), 331-336.
- Viersted, K. B.; Westgaard, R. H.; & Andersen, P. (1990). Pattern of muscle activity during stereotyped work and its relation to muscle pain. *International Archives of Occupational and Environmental Health*, 62, 31-41.
- Wiktorin, C.; Karlqvist, L.; & Winkel, J. (1993). Validity of self-reported exposures to work postures and manual materials handling. *Scandinavian Journal of Work, Environment, and Health*, 19, 208-214.
- Williamson, D.; Chrenko, F.; & Hamley, E. (March 1994). A study of exposure to cold in cold stores. *Applied Ergonomics*, 25-30.
- Winkel, J.; Dallner, M.; Ericson, M.; Fransson, C.; Karlqvist, L.; Nygard, C. -.; Selin, K.; Wigaeus Hjelm, E.; & Wiktorin, C. (1991). *Evaluation of a questionnaire for the estimation of physical load in epidemiologic studies: study design*. In Y. Queinnec; & F. Daniellou, eds., *Designing for Everyone*:

Proceedings of the 11th Congress of the International Ergonomics Association. Vol. 1 Paris (pp. 227-229). London: Taylor and Francis.

CHAPTER 4

COMPARISON OF OBSERVATIONAL AND INSTRUMENTAL BASED MEASUREMENTS OF REPETITION, FORCE, AND WRIST POSTURE IN MANUAL WORK

Abstract

The primary goal of this work was to compare direct measurements (EMG and electrogoniometry) of repetition, force, posture, and movement speed with observer ratings of the same parameters. A laboratory experiment was conducted in which six subjects performed a hand transfer task. A container was repeatedly transferred a distance of 30 cm at frequencies of 12.5 and 50 transfers/minute, with masses of 1, 3, and 6 kg, and horizontal and vertical handle positions. EMG from the finger flexors and extensors and electrogoniometer data (in flexion/extension and radial/ulnar deviation) from the wrist were collected. For a representative subject, repetition, force, wrist posture, and wrist movement speed were rated using an observational method. All stressors were rated on 10 cm visual-analog scales, where 0 generally corresponded to no stress, and 10 corresponded to the most possible stress.

In the observational method, repetition is defined as a composite index which integrates movement speed and exertion time. When the repetition rating was modeled as a function of average wrist movement speed and amount of finger recovery time (defined as exertion below 2.5% MVC), both terms were significant and the model explained 88% of the variance.

Peak and average EMG were significant terms in the model of peak and average force ratings ($r^2 = .54$ and 0.49 , respectively). Peak and average wrist posture as measured with the goniometer were also significant effects in the model of peak and average wrist posture ($r^2 = .28$ and 0.53 , respectively). There was no significant relationship between the peak movement speed calculated from the goniometer data and ratings of peak movement speed.

Calculated average movement speed was significantly related to ratings of average movement speed ($r^2 = .30$).

These results suggest that observers are able to identify repetition, force, movement speed, and average rates of wrist velocity. Strongest agreement between the observations and instrumentation was for the repetition scale, with peak and average force, and average wrist posture next in terms of model strength.

Introduction

Upper extremity musculoskeletal disorders are a major occupational safety and health concern in much of the world (Armstrong et al., 1993). These disorders, which include tendinitis, epicondylitis, and carpal tunnel syndrome, are a major cause of disability, lost productivity, and worker pain and impairment. Several work-related physical stressors are associated with the development of these disorders, including repeated and sustained exertions, forceful exertions, and awkward postures (Kuorinka and Forcier, 1995; Viikari-Juntura, 1995). In addition, high rates of velocity and acceleration of the wrist (Schoenmarklin et al., 1994) have been linked to increased incidence rates of these disorders. Numerous methods have been employed in the quantification of exposure to these stressors.

Difficulties in exposure definition, assessment, and evaluation have been cited among the greatest challenges in ergonomic epidemiology (Hagberg, 1992). In a meta-analysis of studies concerning workplace ergonomic factors and the development of neck and upper extremity musculoskeletal disorders, Stock (1991) eliminated 18 of 43 studies (42%) because of inadequate exposure measures. There are several competing constraints involved in selecting an exposure assessment method, and different methods meet the various constraints to different degrees. Winkel et al. (1991) suggest four characteristics of exposure assessment methods that must be considered in the selection of a method for a particular application:

- cost (in terms of equipment, time, and expertise)
- capacity (size of population that can be evaluated)
- versatility/specificity (degree to which method addresses only one or a number of stressors)

- generality (ability to estimate exposures in general rather than at a particular point in time)

In addition, the issues of reliability and validity are critical in determining the adequacy of an exposure assessment method (Kelsey et al., 1986; Kilbom, 1994).

Subject-based questionnaires, observer checklists, and other observational methods tend to have the lowest cost, highest capacity, greatest versatility, and most generality; however, the reliability and validity of such measures are often suspect (Kilbom, 1994). Instrumentation such as electromyography (EMG), electrogoniometry, and digitized video motion analysis systems are perceived as having higher validity, but usually only address one stressor (high specificity) and are more costly to perform and thus have lower capacity.

Several investigators have evaluated the degree to which observers' categorization of posture stresses agrees with instrumentation (Baty et al., 1986; Ericson et al., 1991; DeLooze et al., 1994; Baluyut et al., 1995). Seemingly little work has been done on validating observations of other stressors (e.g. repetition, force, and movement speed).

The primary goal of this work was to compare instrumentation information (EMG and electrogoniometry) about repetition, force, posture, and movement speed with observer ratings of the same parameters.

Methods

The comparison of the observational and instrumental methods was based on the analysis of a hand transfer task performed in the laboratory. Mass, frequency, and posture were varied in each of twelve trials. Electromyography (EMG), electrogoniometry, and observational ratings were recorded for each trial.

Subjects

Six healthy subjects (3 males and 3 females) participated in the experiment. Subjects were recruited from the student population. Mean age was 27.7 ± 3.4 years, and all subjects were right handed.

Equipment

A programmable conveyor was utilized to facilitate the transfer task. The conveyor was programmed to index a distance of 30 cm at specified intervals. A container was placed on the conveyor to the left of the subject. Each time the conveyor indexed, it moved the container 30 cm to the right. Thus the subject repeatedly moved the container 30 cm from right to left. A hollow container with an adjustable handle was used as the transfer object. The handle of the container was instrumented with a strain gauge to measure grip force exerted during the transfer. Various amounts of steel shot were placed inside the same container to vary the mass.

Two transfer frequencies were used: 50 transfers/minute and 12.5 transfers/minute (1.2 s and 4.8 s cycles, respectively). Total container mass was set to 1 kg, 3 kg, and 6 kg. The handle orientation was varied between horizontal and vertical, in order to alter wrist posture. All combinations of frequency/mass/posture were performed by all subjects, yielding twelve experimental conditions.

The EMG electrodes were previously developed and have been extensively used in past studies (Armstrong et al., 1994, Gerard et al., 1996). Each surface electrode consisted of a bipolar electrode on which a preamplifier was mounted to minimize the effects of external noise. The EMG signals were then amplified and the raw EMG signal was RMS converted with a 55 millisecond time constant. The signals were digitized by a National Instruments Lab NB board at a rate of 40 Hz. Electrodes were placed on the finger flexor and extensor muscles of the subjects' right arms.

Data on wrist position and movement were collected using a Greenleaf™ electrogoniometer (Greenleaf Medical Systems, 1995). The goniometer was mounted in a glove worn on the right hand. The goniometer recorded wrist position in flexion/extension and radial/ulnar deviation. The signals were sampled at 40 Hz. The system defines flexion and radial deviation as negative values, while extension and ulnar deviation are positive.

Procedure

Subjects reported to the lab where the experiment was explained and their written consent was obtained. EMG electrodes were attached to each subjects' right forearm to record the muscle activity of the underlying finger flexors (*Flexor digitorum superficialis*) and finger extensors (*Extensor digitorum communis*). Efforts were made to minimize signal contamination from adjacent wrist flexor and extensor muscles. Muscle activity for the finger flexors was calibrated to percent of maximum voluntary contraction (%MVC) by having each subject grip the experimental container handle as hard as possible. Wrist extensors were calibrated to %MVC by having each subject loosely grip the handle, then attempt to extend their fingers outward while resistance was applied by an experimenter. The goniometer glove was then placed on each subject's right hand. The goniometer was calibrated by placing the wrist in a neutral position, then performing measured excursions to 60° of flexion, 60° of extension, 30° of radial deviation, and 20° of ulnar deviation.

The conveyor height was adjusted so that the handle was at elbow height for each trial. Each trial lasted 2 1/2 minutes; subjects were allowed to become accustomed to the task for the first 30 seconds; data were collected during the last 2 minutes of each trial. Trials were randomized among subjects. Subjects were not told which condition they were performing; however, frequency and handle position parameters were obvious. Each trial was videotaped for later observational analysis.

Observational ratings of repetition, force, wrist posture, and wrist movement speed were made for all trials for one representative subject. Ratings were performed using a series of 10 cm visual analog scales, which generally ranged from 0, which corresponded to "least possible exposure" to the stressor, to 10, corresponding to "most possible exposure" (Latko et al., 1997). A series of videotaped examples were used for training analysts in this method, and written decision rules and benchmark examples were used to facilitate the ratings.

The repetition rating is an integrated score for the hand and fingers, which takes into account movement speed and exertion time (Figure 4.1a). Hand forces were rated in terms of percent maximum effort, where a rating of 0

represented no force exertion and 10 represented maximum effort (Figure 4.1b). Wrist posture was rated relative to the range of motion of that joint; 0 corresponded to neutral, and 10 corresponded to the extreme of the range of motion in either flexion/extension or radial/ulnar deviation (Figure 4.1c). Movement speed of the wrist was rated with respect to the verbal anchor points at 0 of "zero velocity, no movement" and 10 of "maximum velocity" (Figure 4.1d). The maximum instantaneous (peak) and time-weighted average values of force, wrist posture, and wrist movement speed were rated.

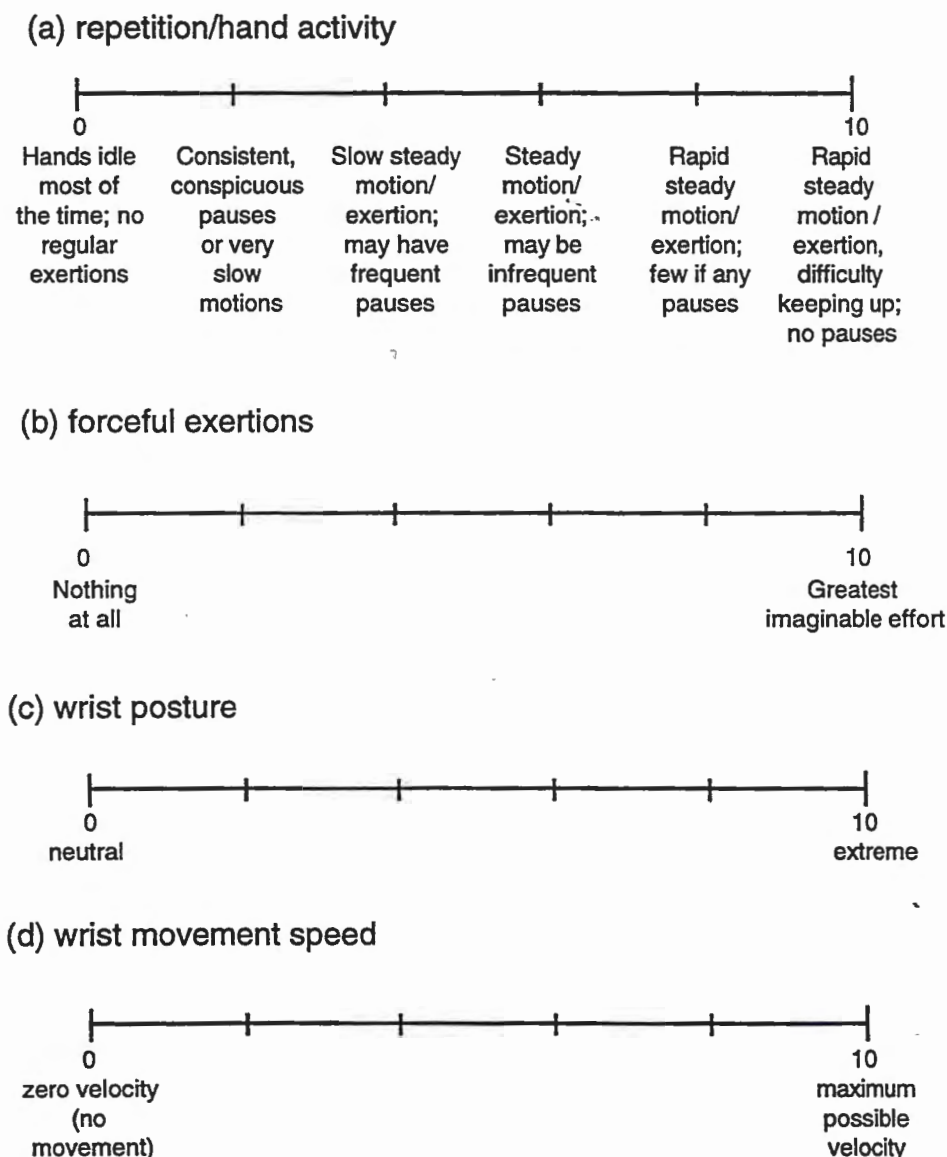


Figure 4.1: Ten cm visual-analog scales for rating repetition/hand activity (a), forceful exertions (b), wrist posture (c), and wrist movement speed (d).

A team of four analysts performed the observational ratings. Analysts were given no information concerning cycle time, mass, or handle position. Ratings were made solely on the basis of observation. The ratings followed a modified nominal group procedure (Gustafson et al., 1973). The raters were presented with a looped videotape of each condition, allowing them to view the trail as many times as necessary. Each team member rated the job independently for each of the stressors included in the study. After the initial ratings were complete, the group members discussed their ratings with the goal of reaching consensus, which was defined as a range of no more than one unit on the 10 cm scale.

All ratings were performed on networked Macintosh™ computers. The rating program was written in Hypercard™ (Apple Computer, 1995), and consisted of a series of 10 cm sliders. The written guidelines were embedded in help screens accessible from each scale. Each rater computer wrote the ratings to a central "collector" computer running a Labview™ program (National Instruments, 1995). The collector recorded all initial ratings, and displayed all raters' scores for each scale when the initial rating was complete. The display also indicated whether the raters were in consensus. If a scale did not initially meet the consensus criteria, the raters discussed the rationale for their ratings, and adjusted their ratings accordingly. These changes were displayed in real time by the collector, and the consensus criterion was continuously re-evaluated. Raters were not allowed to continue to the next scale until consensus was achieved on the current scale. This procedure was followed for each of the 7 scales in this study. When consensus had been achieved on all scales, the final ratings were recorded and the raters moved on to the next experimental condition. The twelve conditions were rated over a period of three days.

Data analysis

Peak and average calibrated EMG values, wrist postures, and wrist movement speeds were calculated from the corresponding signals. These values were compared to the observational ratings of the same parameters. From the EMG data, the percent of time that subjects' finger flexors and extensors were active at or above 2.5% MVC was calculated. This threshold was selected based on the work of Jonsson (1978) which suggests that the

static load level in continuous work should not exceed 2 to 5% MVC. The peak and average of the absolute value of the angular velocity in each plane (flexion/extension and radial/ulnar deviation) was calculated from the goniometer data.

Results

The pooled main effects are listed in Table 1. With the exception of peak velocity, all of the dependent measures differed significantly between the two cycle times. The magnitude of these differences were relatively large (Table 4.1). Handle position was significant for extensor recovery time and average wrist velocity in flexion extension, and there was a significant difference in peak radial/ulnar wrist velocity between the 1 kg and 3 kg load.

Table 4.1: Pooled main effects of each experimental parameter.
Different superscripts indicate significant effects at $p < 0.05$.

	handle position		mass			cycle time	
	hor.	vert.	1 kg	3 kg	6 kg	4.8 s	1.2 s
repetition rating (0-10)	5.2	5.4	5.2	5.7	5.1	2.7 ^A	7.9 ^B
%time $\leq$ 2.5% MVC-flx	55 $\pm$ 25	62 $\pm$ 23	65 $\pm$ 24	56 $\pm$ 23	54 $\pm$ 24	78 $\pm$ 7 ^A	38 $\pm$ 16 ^B
%time $\leq$ 2.5% MVC-ext	46 $\pm$ 25 ^A	30 $\pm$ 27 ^B	39 $\pm$ 27	39 $\pm$ 28	36 $\pm$ 27	62 $\pm$ 12 ^A	13 $\pm$ 11 ^B
Peak velocity-F/E* (deg/s)	145 $\pm$ 61 ^A	194 $\pm$ 66 ^B	151 $\pm$ 55	190 $\pm$ 73	168 $\pm$ 71	184 $\pm$ 64	155 $\pm$ 69
Peak velocity-R/U* (deg/s)	80 $\pm$ 34	89 $\pm$ 31	71 $\pm$ 25 ^A	95 $\pm$ 33 ^B	88 $\pm$ 36 ^{A,B}	86 $\pm$ 28	83 $\pm$ 38
Avg. velocity-F/E* (deg/s)	35 $\pm$ 15 ^A	46 $\pm$ 18 ^B	39 $\pm$ 18	43 $\pm$ 18	40 $\pm$ 16	28 $\pm$ 10 ^A	53 $\pm$ 14 ^B
Avg. velocity-R/U* (deg/s)	17 $\pm$ 9	20 $\pm$ 10	16 $\pm$ 8	20 $\pm$ 10	20 $\pm$ 10	12 $\pm$ 4 ^A	29 $\pm$ 9 ^B

*absolute value

Repetition:

The observational rating method defines repetition in terms of speed of movements and amount of recovery time within the job cycle (Figure 4.1a). The most active of the two muscle groups and most dynamic wrist movement plane for each condition were selected for further analysis. In 65 of the 72 trials, the extensors had less recovery time than did the flexors. Average movement speed in flexion/extension was higher than average movement speed in radial/ulnar deviation for 70 of the 72 trials. A model was generated to determine the degree to which the rating of repetition reflected movement speed and recovery time as measured by the instrumentation.

These two parameters (percent time below 2.5% MVC for the most active finger muscle and the average wrist velocity in flexion/extension or radial/ulnar deviation) were both statistically significant terms and explain 88% of the variance in repetition ratings. There were no significant subject effects. The final model is shown in Table 4.2.

Table 4.2: Model coefficients and p values for repetition rating as a function of movement speed and exertion time.

term	coefficient	F ratio	p-value
intercept	6.9		<0.0001
finger muscle recovery time (% time below 2.5% MVC)	-0.08	101.9	<0.0001
avg angular wrist velocity	0.03	5.3	0.025

Figure 4.2 illustrates the individual relationships between the modeled movement speed and recovery time parameters and the repetition ratings for the six subjects. Individually, average angular velocity is moderately correlated with repetition rating ($r^2=0.57$), while percent recovery time is strongly correlated with repetition rating ($r^2=0.84$).

Further analysis was performed to determine the correlation between the average velocity and the exertion time. The relationship between the velocity and exertion time parameters was significant ($p<0.0001$); however, the models only explained 26% of the variance between the parameters. A correlation between average velocity and exertion time would be expected, since exertion is generally required for movement (i.e. non-zero velocity). However, the relatively low r^2 value between the two sets of parameters

indicates that they reflect different aspects of the activity pattern, and thus may both be important in determining hand activity level.

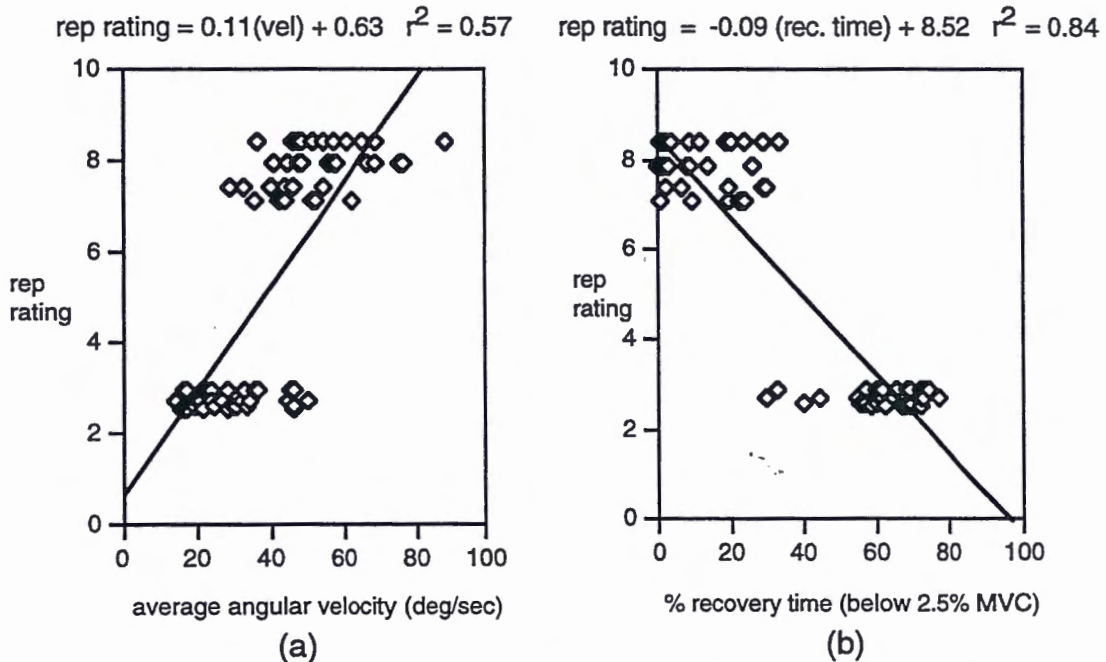


Figure 4.2: Repetition ratings as a function of average angular velocity (a), and repetition ratings as a function of % time muscles were active below 2.5% MVC (b).

Force

The peak RMS EMG values for the finger flexor muscles and peak handle grip force for each transfer was determined. These peaks were averaged over each trial to obtain the mean peak finger flexor EMG and mean peak grip force. Both EMG and grip force values were normalized to percent of maximum: percent maximum voluntary contraction (%MVC) for EMG and percent maximum grip force (%MGF) for grip. Instrumental data were averaged for the six subjects. For the twelve conditions, ratings of peak force using the observational method ranged from 3 to 6 on the ten cm scale (Figure 4.3). Peak force rating was modeled as a function of mean peak EMG; the relationship was significant at $p < 0.0001$ ($r^2 = 0.78$). Similar results were obtained for the model using mean peak grip force levels and subject as independent variables ($p < 0.0001$, $r^2 = 0.81$). In the model containing peak EMG, the intercept was not significant. The intercept on the grip force model was significant at $p < 0.001$.

The slopes in both the EMG and grip force models were close to 0.10 (0.13 and 0.12, respectively).

Average finger flexor EMG and handle grip force were determined by taking the arithmetic mean of all finger flexor EMG data for each trial. Observational ratings of average force ranged from 0.5 to 2.0 on the ten cm scale. In their respective models, average EMG and average grip force were each significant factors ($p < 0.02$, $r^2 = 0.57$ and 0.49 , respectively). In both models, the intercept was statistically significant ($p < 0.02$), and the slopes were at or near 0.1, which indicates direct correspondence between % MVC and rating. The data for peak and average EMG and peak and average grip force vs. peak and average force ratings are shown in Figure 4.3.

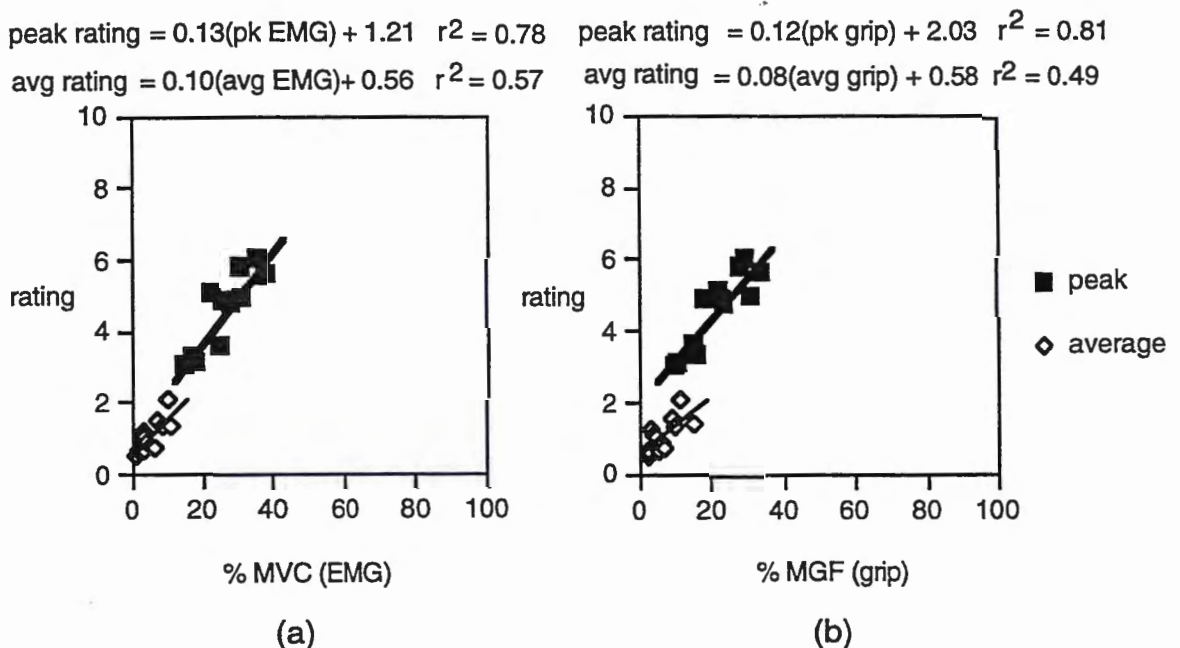


Figure 4.3: Observational rating of peak and average force as a function of (a) peak and average finger flexor EMG and (b) peak and average grip force.

Wrist posture

The most extreme wrist position in both flexion/extension and radial/ulnar deviation recorded during each transfer was determined using a peak detection routine. These peaks were averaged over each trial to determine the mean peak wrist posture in each plane. Wrist postures were normalized with respect

to population maximum ranges of motion (American Academy of Orthopaedic Surgeons, 1965) in each plane, and the plane with the most extreme normalized posture was compared to the ratings of peak wrist posture. In general, for the conditions where the handle position was vertical, flexion/extension was the most severe plane. When the handle position was horizontal, deviation in the radial/ulnar plane was more severe than flexion/extension. Because the subjects' actual wrist angles were compared to population norms rather than their own maximum range of motion, some subjects exhibited wrist postures in excess of 100% of the normalized maximum. Instrumental data were averaged for the six subjects. Observational ratings of peak wrist posture ranged from 3 to 8 (Figure 4.4). An ANOVA was performed with peak wrist posture rating as the dependent variable and mean peak posture (as a percent of maximum range of motion) as the independent variable. The wrist posture variable was significant at $p < 0.001$, and the model explained 77% of the variance in the peak wrist posture ratings. The intercept was also significant ($p < 0.02$).

The median normalized wrist postures in flexion/extension and radial/ulnar deviation were used in the comparison with average wrist posture ratings. Because position in both the flexion/extension and radial/ulnar planes were considered in the ratings, the plane with the higher average position was selected for the comparison. The ratings of average wrist posture ranged from 1 to 7 (Figure 4.4). In the model with median normalized wrist posture as the independent variable, the term was significant at $p < 0.0001$ ($r^2 = 0.89$). The intercept was not significant.

Wrist movement speed

The absolute value of the angular velocity of the wrist in both flexion/extension and radial/ulnar deviation was calculated from the goniometer data. The peak wrist velocity during each transfer cycle was determined and averaged over each trial. Observational ratings of peak movement speed ranged from 4.5 to 6 on the ten cm scale. There was no statistically significant relationship between the recorded angular velocity of the wrist and ratings of peak wrist movement speed.

The ratings of average wrist movement speed were compared to the median absolute value of angular wrist velocity. Ratings of average wrist movement speed ranged from 0.9 to 2.0 on the ten cm scale (Figure 4.5). When

ratings of average wrist movement speed were modeled as a function of maximum average wrist velocity (in either flexion/extension or radial/ulnar deviation), maximum average wrist velocity was significant ($p < 0.02$, $r^2 = 0.44$). The intercept was significant as well.

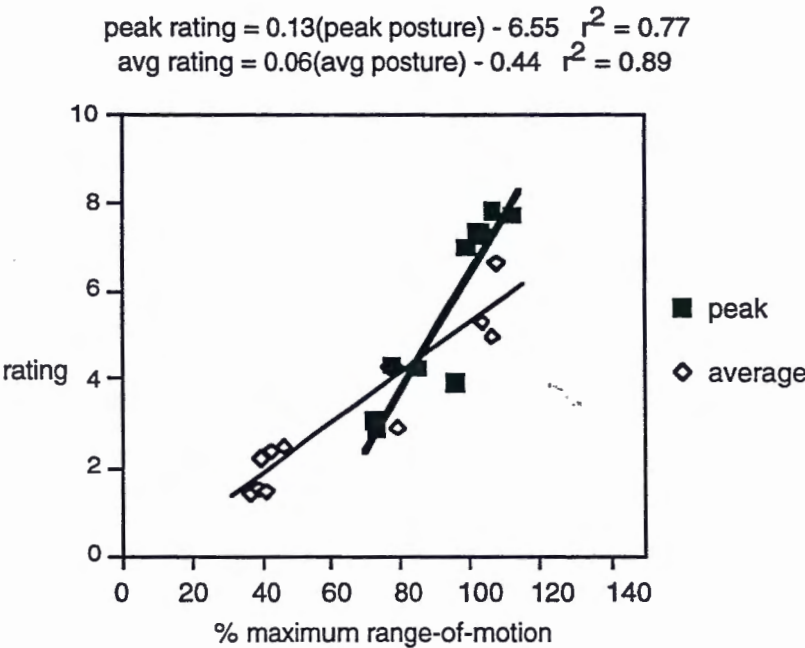


Figure 4.4: Observational ratings of peak and average wrist posture as a function of peak and average wrist angle.

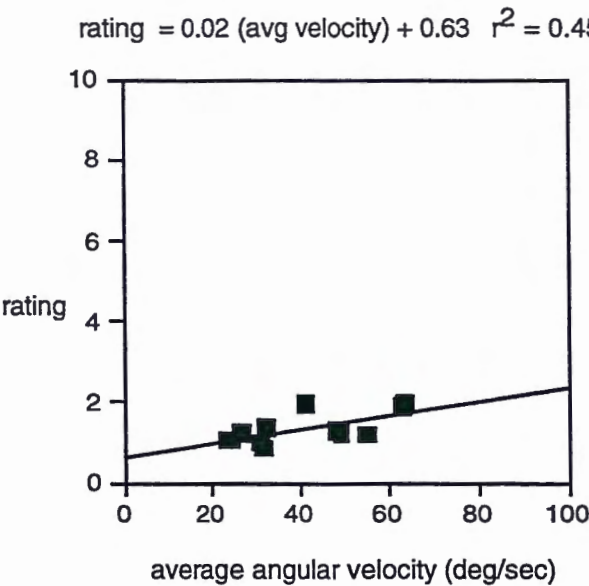


Figure 4.5: Observational ratings of average wrist movement speed as a function of average wrist angular velocity.

Discussion

For all stressors except peak movement speed (velocity), the instrumental measurements of the stressor were significantly related to the ratings of the stressor. There was a wide range in the amount of variability in these models, however.

The model of repetition ratings as a function of movement speed and exertion time was the strongest of the seven parameters evaluated ($r^2 = 0.88$). It can be argued that this rating is the most complex, because it requires the rater to integrate two distinct aspects of the work task (movement speed and exertion time). Repetition has been defined in a variety of manners by other investigators (see Latko et al., 1997 for review). A recent study by Leskinen et al. (1997) compared opto-electronic and video measurements of hand repetitive motions to observed repetitive activities recorded using the Portable Ergonomic Observation (PEO) method (Fransson-Hall et al, 1995). For the two subjects evaluated, agreement between the video and observer was 0.96 and 0.98. In the PEO method, repetitive activities are defined as "pronation/supination, volar/dorsal wrist flexion, radial/ulnar wrist deviation, finger flexion/extension performed several times/min" (Fransson-Hall et al., 1995). No details are given as to the method for determining repetitive activities from the instrumentation. Since only video-based instrumentation was used, however, it is clear that the instrumental metric did not include a measure of muscle exertion level or duration.

Relatively strong agreement was also found between the peak and average finger flexor EMG levels and grip forces, and the ratings of peak and average force ($r^2 = 0.49$ to 0.81). By definition, the peak and average force ratings should correspond to %maximum effort, which in this experiment was approximated by EMG normalized to %MVC. Theoretically, a peak exertion of 10% percent MVC would correspond to a rating of 1, a 20%MVC exertion would correspond to a rating of 2, etc. Perfect estimation of percent effort would be reflected in a slope of 0.1 and intercept of 0 for the regression line. For the two peak force models, the slopes were close to 0.1 (0.12 and 0.13). In this study, the raters tended to overestimate the percent effort being exerted. Ratings of peak force ranged from approximately 3-6 on the ten cm scale, while mean peak muscle activity levels for the same conditions ranged from approximately 15% MVC to 37% MVC (Figure 4.3). The intercept was not significant in the

peak EMG model, but in the model of peak grip force, the intercept was approximately 2.0 and was statistically significant. Ratings of average force tended to agree more closely with the actual percent effort exerted in an absolute sense, although the model r^2 values were lower. The lower r^2 values were most likely due to the relatively small range in exertion levels (approximately 1%MVC to 11%MVC) (Figure 4.3). For the average EMG and average grip force models, the slope of the relationships were close to 0.1 (0.1 and 0.08, respectively). For both models the intercept was significant, but relatively small (0.56 and 0.58, respectively).

Relatively few studies have examined the validity of observer ratings of force. In a study evaluating the validity of observations, DeLooze et al. (1994) had subjects perform a task involving handling crates and transferring bottles between crates. Observers were asked to record whether the subject was handling a full crate, empty crate, bottles alone, or no load. The investigators found modest agreement ($\kappa = 0.50$) between observations and actual load. These findings can not be directly compared to the current study because of the difference in the observational methods in the two studies (categorical versus numerical). However, in the current study, correlation coefficients for the peak force variables are 0.88 and 0.90, which indicates strong agreement. Correlation coefficients for the average force parameters were lower than the peaks, but still relatively high (0.70 and 0.75). The relatively high r^2 values between the ratings and the instrumentally measured force exertion levels indicates that the observers were able to detect differences in force exertion levels; however, the absolute agreement between the measured force levels and the ratings was lower. The raters were able to correctly determine the relative magnitude of the exertion (in comparison to the other conditions observed), but these ratings did not directly correspond to the actual percent of maximum effort exerted. The strong linear relationship observed does indicate that efforts to further train or calibrate the raters may produce better absolute estimates.

In the current study, a high percentage of the variability in the wrist posture ratings was explained by the measured wrist posture ($r^2 = 0.77$ for peak, and 0.89 for average). As in the force ratings, however, the absolute agreement was lower than the raters' ability to consistently detect proportional differences in wrist position.

Several other investigators have compared observational ratings or estimates of wrist posture with instrumental recordings, with varying results. Most of these studies utilized categorical posture classifications. In one of the earliest reliability studies of an observational technique, Corlett et al. (1979) compared observations of postures with measures taken directly from film. They found relatively strong agreement between head and trunk postures ($0.65 \leq r \leq 0.82$), but much lower correlations for upper extremity postures (shoulder and elbow) ($0.00 \leq r \leq 0.49$). In the crate handling study mentioned above (DeLooze et al., 1994), observers also categorized arm position as being greater than or less than 60° with arm angles recorded from an opto-electronic system. Fair agreement was found between rater categorization and measured values ($\kappa = 0.43$). Baluyut et al. (1995) had observers estimate non-neutral static postures around the wrist, elbow, shoulder, neck, and lower back. Wrist postures were classified into eight categories: five categories of flexion/extension, two categories of radial/ulnar deviation, and an overall category. The percentage of subjects giving the correct response ranged from approximately 60% to 82% (mean = 72%). Burdorf et al. (1992) and Baty et al. (1986) compared observer ratings of trunk position with inclinometer recordings; both studies found large differences between the two techniques. Ericson et al. (1991) compared observational estimates of upper arm postures with "true" segment angles measured from video. The median error in estimation was 5 degrees for static postures and 10-13 degrees during dynamic activities. Leskinen et al. (1997) also compared opto-electronic and video measurements of hand position to PEO ratings; proportion of agreement for percent of time with the hand above shoulder level was 0.69 and 0.76 for the two subjects, while agreement for number of occurrences of the hand above shoulder height was much higher (0.92 and 0.94). Of the above studies, only the Ericson work utilizes angular estimation rather than categorization. It appears that the analysts in that study were more accurate in their observations than were the analysts in the current study. In the current study, the raters tended to underestimate both peak and average wrist posture by as much as 40% of the range of motion in any one direction (Figure 4.4). For example, peak wrist postures which were measured as approximately 70% of the total range motion were rated approximately 3 on the ten cm scale (corresponding to 30% of maximum range of motion). Similarly, peak joint angles which were recorded as being at the limits of the range of motion for the wrist (i.e. 100%) were rated

approximately 8 (80%) on the ten cm scale. Translated to degrees, this would be an error of up to 24° in the flexion/extension plane, and 12° in the radial/ulnar deviation plane. Again, since the raters appear to be able to discriminate between different postures (as evidenced by the strong linear relationship observed), it is possible that increased training to overcome these systematic errors may be beneficial.

The study by Ericson et al. (1991) only considered upper arm posture, whereas the current study examined wrist posture. It is possible that the position of the shoulder is easier to observe, because of the large size of the two body segments it connects (i.e. torso and upper arm), and because of the large range of motion. The wrist has a smaller range of motion, especially in radial/ulnar deviation, which may make a given percentage change in angle more difficult to detect. Wrist movements tend to be more highly dynamic than shoulder movements, as well. Not only can the speed make postures more difficult to discern, but the highly dynamic nature of wrist movements often makes it impossible to maintain an orthogonal view of the joint. Close up video recordings and multiple recording angles can improve this situation, although observation remains difficult.

The observational ratings of peak movement speed were not statistically related to the calculated peak movement speed. Interestingly, peak velocity was also not related to transfer frequency in the tested frequency range (Table 4.1). Although there was no statistically significant relationship between frequency and movement speed, the general trend in the data is opposite what might be expected: the shorter cycle time tasks have generally lower peak velocities than the longer cycle time conditions. A visual analysis of the data indicates that this is due to the movement patterns during the task. In the slower condition (4.8 s cycle), the subjects had time to relax their hands between transfers. Plots of the goniometer data from this condition show periods of a few seconds where the wrist is nearly static, with brief, quick (i.e. high velocity) movements when the container is picked up. During the 1.2 s cycle conditions, the subject has no time to relax between transfers. The subject's wrist steadily oscillates as it completes each transfer cycle. In this study observers were not able to discern these differences in peak velocities. The peak velocities recorded during this task did not represent a wide range of exposures. In a study of forty industrial workers, Marras and Schoenmarklin (1993) recorded peak angular wrist velocities in the range of 120°/s-185°/s in flexion/extension

and 77°/s-115°/s in radial/ulnar deviation. In the current study, peak wrist velocities ranged from 145°/s to 194°/s for flexion/extension (i.e. toward the higher end of Marras and Schoenmarklin's range) and from 71°/s to 95°/s in radial/ulnar deviation (i.e. at the lower end of Marras and Schoenmarklin's range) for the various conditions (Table 4.1). This suggests that raters may not be able to perceive differences of this magnitude. It is possible that in other types of tasks (e.g. that have larger differences between conditions) differences in peak movement velocities could be detected. The raters were, however, able to effectively differentiate between different average levels of movement speed. The ranges of average velocities in this study (28-53°/s in flexion/extension, 12-29°/s in radial/ulnar deviation --Table 4.1) were within $\pm 11^\circ/\text{s}$ of the ranges identified by Marras and Schoenmarklin (28.7-42.2°/s in flexion/extension, 17-26 °/s in radial/ulnar deviation).

In general, the study described in this paper indicated varied amounts of agreement between observational ratings of repetition, force, wrist posture, and wrist movement speed and instrumental measures of the same parameters. Repetition exhibited the strongest agreement with the instrumental model, while peak wrist movement speed (velocity) showed no statistical relationship. In general, for the six parameters that did have statistically significant relationships between the ratings and instrumentation, the instrumental measures explained a large amount of the variation in the ratings. However, the absolute agreement between the measured parameters and the raters was not as strong. Raters tended to overestimate force and underestimate wrist postures.

There were several limitations to the design of this study. First, it was a laboratory simulation with a stereotyped task. The motions required for this task were less complex than those in many typical manufacturing operations. It is unclear whether the strength of the observations would be as high in a more complex task. The second major limitation is that the ratings were based on the videotape of only one representative subject, and therefore the instrumentation data are presented as averages over the six subjects. Although ideally each of the six subjects would have been rated, the methods followed in this study reflect the way exposure assessment is often performed in the workplace and in epidemiological studies. Exposures are often determined for the job as a whole, rather than for each individual worker. In general, data collection using instrumentation is much more time and equipment intensive than are observations. In the laboratory it is possible to collect instrumentation data on a

relatively large number of subjects. In the workplace, however, instrumentation is usually not feasible, especially on a large number of subjects. In this study, efforts were made to standardize conditions as much as possible between subjects (e.g. the handle of the container was adjusted to elbow height for each subject).

This results of this study highlight the need for further work in several areas. It was shown that the raters were able to differentiate between different forces, postures, and movement speeds from a proportional standpoint but not in terms of absolute estimation. Further studies should examine whether training raters to adjust for these systematic errors may reduce their occurrence. The tasks used in this study were simple, short cycle time tasks. Future studies are needed to examine the validity of the observations when more complex tasks are performed. In the current study, ratings of peak movement speed were not related to the measured peak angular wrist velocities. However, in this study, a relatively small range of wrist velocities were recorded. Further examination of whether larger differences in movement speeds can be detected by observation is necessary. Also, ratings should be performed for individual subjects, to determine the degree to which differences in individuals, rather than gross differences between jobs, can be detected.

References

- American Academy of Orthopaedic Surgeons (1965). *Joint Motion: Methods of Measuring and Recording*. Chicago: American Academy of Orthopaedic Surgeons.
- Armstrong, T. J., Buckle, P., Fine, L. J., Hagberg, M., Jonsson, B., Kilbom, A., Kuorinka, I. A., Silverstein, B. A., Sjøgaard, G., & Viikari-Juntura, E. R. (1993). A conceptual model for work-related neck and upper-limb musculoskeletal disorders. *Scandinavian Journal of Work & Environmental Health*, 19, 73-84.
- Armstrong, T. J., Foulke, B. J., Martin, B. J., Gerson, J., & Rempel, D. M. (1994). Investigation of applied forces in alphanumeric keyboard work. *American Industrial Hygiene Association Journal*, 55(1), 30-35.
- Baluyut, R., Genaidy, A. M., Davis, L. S., Shell, R. L., & Simmons, R. J. (1995). Use of visual perception in estimating static postural stresses: magnitudes and sources of errors. *Ergonomics*, 38(9), 1841-1850.
- Baty, D., Buckle, P. W., & Stubbs, D. A. (1986). *Posture Recording by Direct Observation, Questionnaire Assessment and Instrumentation: a Comparison Based on a Recent Field Study*. In N. Corlett, J. Wilson, & I. Manenica, eds. *The Ergonomics of Working Postures, Models, Methods, and Cases*. London: Taylor and Francis.
- Burdorf, A., Derksen, J., Naaktgeboren, B., & vanRiel, M. (1992). Measurement of trunk bending during work by direct observation and continuous measurement. *Applied Ergonomics*, 23(4), 263-267.
- Corlett, E. N., Madeley, S. J., & Manenica, I. (1979). Posture Targetting: A technique for recording working postures. *Ergonomics*, 22(3), 357-366.
- DeLooze, M. P., Toussaint, H. M., Ensink, J., & Magnus, C. (1994). The validity of visual observation to assess posture in a laboratory-simulated, manual material handling task. *Ergonomics*, 37(8), 1335-1343.

- Ericson, M., Kilbom, A., Wiktorin, C., & Winkel, J. (1991). *Validity and reliability in the estimation of trunk, arm and neck inclination by observation*. Y. Queinnec, & F. Daniellou, *Designing for Everyone: Proceedings of the 11th Congress of the International Ergonomics Association*. Vol. 1 Paris (pp. 245-247). London: Taylor and Francis.
- Fransson-Hall, C., Gloria, R., Winkel, J., & Kilbom, A. (1995). A portable ergonomic observation method (PEO) for computerized on-line recording of postures and manual handling. *Applied Ergonomics*, 26(2), 93-100.
- Gerard, M. J., Armstrong, T. J., Foulke, J. A., & Martin, B. J. (1996). Effects of key stiffness on force and the development of fatigue while typing. *American Industrial Hygiene Association Journal*, 57, 849-854.
- Greenleaf Medical Systems, I. Wrist system [Computer Program]. Palo Alto, CA.
- Gustafson, D. H., Shulka, R. K., Delbecq, A., & Walster, G. W. (1973). A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, delphi groups, and nominal groups. *Organizational Behavior and Human Performance*, 9, 280-291.
- Hypercard [Computer Program] (2.3). Cupertino, CA: Apple Computer, Inc.
- Jonsson, B. (1978). *Kinesiology (with special reference to electromyographic kinesiology)*. In W. A. Cobb and H. Van Duijn, eds. *Contemporary Clinical Neurophysiology (EEG Suppl. No. 34)*. Amsterdam: Elsevier Scientific Publishing Company.
- Kelsey, J. L., Thompson, W. D., & Evans, A. S. (1986). *Methods in Observational Epidemiology*. New York: Oxford University Press.
- Kilbom, A. (1994). Assessment of physical exposure in relation to work-related musculoskeletal disorders-what information can be obtained from systematic observations? *Scandinavian Journal of Work, Environment, and Health*, 20 (special issue), 30-45.

- Kuorinka, I., & Forcier, L. (1995). *Work related musculoskeletal disorders (WMSDs): a reference book for prevention*. London: Taylor & Francis.
- LabVIEW for Macintosh/Power Macintosh [Computer Program] (3.1.1). Austin, TX: National Instruments Corporation.
- Latko, W. A., Armstrong, T. J., Foulke, J. A., Herrin, G. D., Raboum, R. A., & Ulin, S. S. (1997). Development and evaluation of an observational method for assessing repetition in hand tasks. *American Industrial Hygiene Association Journal*, 58(4), 278-285.
- Leskinen, T., Hall, C., Rauas, S., Ulin, S., Tonnes, M., Viikari-Juntura, E., & Takala, E. P. (1997). Validation of portable ergonomic observation (PEO) method using optoelectronic and video recordings: *Applied Ergonomics*, 28(2), 75-83.
- Marras, W. S., & Schoenmarklin, R. W. (1993). Wrist motions in industry. *Ergonomics*, 36(4), 341-351.
- National Instruments. (1992). LabNB Manual. Austin, Texas.
- Schoenmarklin, R. W., Marras, W. S., & Leurgans, S. E. (1994). Industrial Wrist Motions and Incidence of Hand/Wrist Cumulative Trauma Disorders. *Ergonomics*, 37(9), 1449-1459.
- Stock, S. R. (1991). Workplace ergonomic factors and the development of musculoskeletal disorders of the neck and upper limb: A meta analysis. *American Journal of Industrial Medicine*, 19, 87-107.
- Viikari-Juntura, E. (1995). *The role of physical stressors in the development of hand/wrist and elbow disorders*. S. L. Gordon, S. J. Blair, & L. J. Fine, *Repetitive Motion Disorders of the Upper Extremity*. (pp. 7-30). Rosemont, IL: American Academy of Orthopaedic Surgeons.
- Winkel, J., Dallner, M., Ericson, M., Fransson, C., Karlqvist, L., Nygard, C. -, Selin, K., Wigaeus Hjelm, E., & Wiktorin, C. (1991). *Evaluation of a questionnaire for the estimation of physical load in epidemiologic studies:*

study design. Y. Queinnec, & F. Daniellou, *Designing for Everyone: Proceedings of the 11th Congress of the International Ergonomics Association. Vol. 1* Paris (pp. 227-229). London: Taylor and Francis.

CHAPTER 5

A CROSS-SECTIONAL STUDY OF THE RELATIONSHIP BETWEEN REPETITIVE WORK AND THE PREVALENCE OF UPPER LIMB MUSCULOSKELETAL DISORDERS

Abstract

A cross sectional study was conducted in which workers' exposure to repetition and other physical stressors was quantified using an observational method, and compared to the prevalence rates of various upper limb disorders. Three hundred fifty-two workers from three companies participated. Jobs were selected based on their repetition levels, to insure that a wide range of repetition exposures were represented. Only companies in which 30 or more workers in low, medium, and high repetition categories were selected in order to control for inter-plant effects. Other stressors, including force, posture, and contact stress were evaluated, but did not have a wide range of exposures in the sample jobs. Physical evaluations on all participating workers were performed by medical professionals, and included a self-administered questionnaire, physical exam, and limited electrodiagnostic testing.

Repetitiveness of work was found to be significantly associated with prevalence of reported discomfort in the wrist, hand, or fingers, (OR=2.45 for change for low to high repetition) tendinitis in the wrist, hand, or fingers, (OR=2.90 for change for low to high repetition) and symptoms consistent with carpal tunnel syndrome as indicated on a hand diagram (OR=2.32 for change for low to high repetition). The relationship between repetitiveness of work and median mononeuropathy indicated by the combination of positive electrodiagnostic results and symptoms consistent with carpal tunnel syndrome achieved borderline statistical significance (OR=3.11).

Background

Several physical stressors encountered in industrial work, including repeated, sustained and forceful exertions, localized mechanical stress, awkward postures, highly dynamic movements, insufficient recovery time, exposures to low temperatures, vibration, and impulse loads have been linked to increased risk of work-related musculoskeletal disorders (WRMSDs) (Table 5.1). These disorders are also referred to as "cumulative trauma disorders", "repetitive motion injuries", or "repetitive stress injuries" (Armstrong et al., 1993b). WRMSDs represent a class of related disorders, which have in common that they develop over an extended period of time, rather than from an instantaneous exposure, and they affect the soft tissues of the upper limb, namely the muscles, tendons, and/or nerves (Armstrong et al., 1993b). Specific diagnoses include carpal tunnel syndrome (CTS), tendinitis, DeQuervain's Disease, and epicondylitis (Putz-Anderson, 1988).

Numerous studies have shown a relationship between these stressors and WRMSDs (Table 5.1). It is generally agreed that a dose-response relationship exists between exposure to these stressors and the prevalence or incidence rates of these disorders (Armstrong et al., 1993b). Most of these epidemiological studies, however, have only examined exposures in a binary classification, either present/absent or low/high (Figure 5.1). Thus there is evidence to support the endpoints of this curve for the various stressors, but there is relatively little information concerning the increased risk associated with intermediate exposure levels. Part of the reason for this lack of information is the difficulty in quantifying exposure to these stressors.

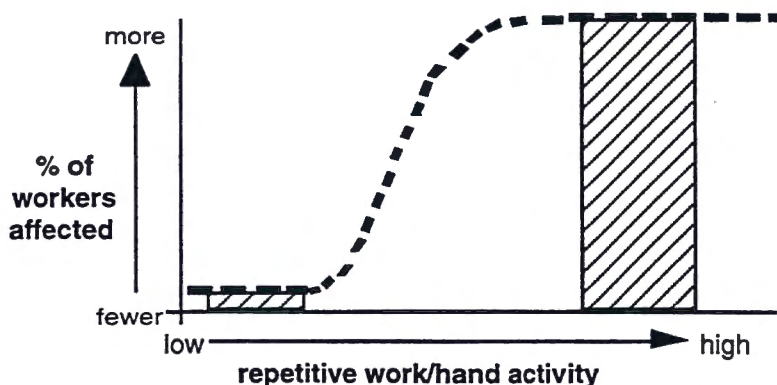


Figure 5.1: Proposed exposure-response curve for work-related upper extremity musculoskeletal disorders.

Table 5.1: Work-related factors which have been identified as risk factors for musculoskeletal disorders and selected references.

factor	explanation	references
repeated and sustained exertions	cyclical use of the same tissues, either as a repeated motion or a repeated muscular effort without movement (Kuorinka and Forcier, 1994)	Luopajarvi et al. ,1979; Cannon et al. ,1981; Silverstein et al., 1986, 1987; Armstrong et al., 1987
forceful exertions	high magnitude or intensity of muscular load	Silverstein et al., 1986, 1987; Armstrong et al., 1987; Goldstein et al., 1987; Bystrom and Fransson-Hall, 1994
localized mechanical stresses	mechanical stresses produced on tissues due to contact with external objects	Lundborg et al., 1982; Sauter et al., 1987; Fransson-Hall and Kilbom, 1993; Hoffman and Hoffman, 1995
posture	extreme postures close to the range of motion of the joint, or non-extreme postures which change the musculoskeletal geometry and increase stress on tendons, muscles, or other tissues (Kuorinka and Forcier, 1994)	Tichauer, 1966; Phalen, 1966; Armstrong and Chaffin, 1978; Gelberman et al., 1981; deKrom et al., 1990
joint kinematics	high rates of angular velocity and acceleration	Schoenmarklin et al., 1994; Marras and Schoenmarklin, 1993.
recovery time	insufficient periods when the muscle is not contracting	Viersted 1990; Sundelin and Hagberg, 1989; Bystrom and Fransson-Hall, 1994
exposure to low temperatures	working in cold environments or handling cold objects	Hammar skjold et al. 1992; Clarke, 1961; Williamson et al., 1984
vibration	exposure to hand-arm vibration	Lundstrom and Johansson, 1986; Nilsson et al., 1990; Hagberg et al., 1992
jerk/impulse	rapid eccentric contraction of muscles, as when resisting the torque of a tool	Armstrong et al, 1993a; Edwards, 1988

Repetition has been among the most widely studied of these stressors, yet no universal definition or quantification technique exists for it (Latko et al., 1997c). Table 5.2 lists several of the means by which repetition has been assessed in various epidemiological studies.

Table 5.2: Definitions of "repetition" that have been used in previous epidemiological studies.

definition of "repetition"	specific criteria	references
number of pieces handled	number of pieces handled per year (3 categories): <200,000; 200,000-300,000; >300,000	Kuorinka & Koskinen, 1979
"observed repetitive activities"	qualitative analysis; presence/absence of high-speed repetitive motions, static load, and extreme hand positions presence/absence of highly repetitive, low force wrist and fine finger motions presence/absence of subject-reported and observed performance of repetitive wrist movements presence/absence of repeated and/or sustained activities involving flexion/extension, or ulnar/radial deviation of the wrist, or use of pinch grip presence/absence of performance of repetitive motion tasks	Luopajarvi et al., 1979 Punnett et al., 1985 Wieslander et al., 1988 Barnhart et al., 1991 Cannon et al., 1991
cycle time	2 classes: short cycle (2.0-9.5s) & long cycle (7.3-26.1s) 2 classes: low = cycle time >30 s; high = cycle time <30 s or at least 50% of cycle in same fundamental movements 4 classes: 1=no cycles; 2=cycle time>30s, same fundamental cycles < 50% of cycle time; 3=10s< cycle time <30s, or 50-80% of cycle time in same basic cycles; 3=cycle time < 10s, or >80% of cycle time in same basic cycles.	Kuorinka & Koskinen, 1979 Silverstein et al., 1986, 1987; Armstrong et al., 1987; Chiang et al., 1993; Schierhout et al., 1995
velocity and acceleration	angular velocity and acceleration of the wrist	Marras & Schoenmarklin, 1993

The work presented in this paper was designed with the primary goal of providing more information on the shape of the dose-response curve for

repetition. Other stressors (e.g. force, localized mechanical stress, posture, etc.) were also considered.

Methods

A double-blind, cross-sectional epidemiological study was conducted to determine the relationship between exposure to physical stressors and prevalence of WRMSDs in industrial workers. Repetition was the stressor of primary interest, although other stressors were treated as covariates. Study participants included 352 workers at three manufacturing facilities. Exposure to physical stressors was quantified for each job by a team of experienced ergonomic analysts. Participating workers in each selected job underwent a medical evaluation conducted by a team of health professionals.

Job selection

The selection of jobs for inclusion in this study consisted of two subtasks: 1) preliminary job selection/classification and 2) formal analysis and final job classification. In the preliminary job selection and classification stage, a team of 2 or more researchers conducted a plant walk-through, during which available jobs were observed. The goal of this job selection was to obtain examples of jobs encompassing three distinct levels of repetition: low, medium, and high. Jobs were selected based on an initial subjective assessment of repetition/hand activity, based on inspection and supervisor interviews. Representative workers, equipment, and job cycles were identified, and the selected jobs were videotaped and documented. Written documentation included production standards, job tasks, workstation layout and nominal dimensions, and materials, tools, and equipment. The job information was then taken back to the university laboratory for further analysis.

In order for a site to be eligible for the study, it was necessary that the three levels of repetition were present at the plant, with at least thirty eligible workers per repetition level. Workers were eligible if they had been in their current job for at least 6 months prior to the study date. It was not possible in all cases to find single job classifications with the necessary number of workers; jobs with similar repetition rates were sometimes combined to achieve the requisite number of workers. In addition, plant management had to agree to

allow the medical evaluations to occur on company time, during normal work hours.

Job analysis

Quantification of the exposure levels for repetition and the other stressors was performed using an observational rating method developed for this study (Latko et al., 1997a, 1997b, 1997c). The rating method utilizes a series of 10 cm visual-analog scales which range from 0, corresponding to no stress, to 10, corresponding to the most possible stress. A series of verbal anchors, written decision criteria, and videotaped benchmark examples form the foundation of the system. These decision criteria and benchmark examples draw from a diverse group of jobs, providing a comprehensive set of guidelines that can be generalized to a wide variety of manufacturing jobs. These guidelines are intended to provide benchmarks against which the raters can compare the job being observed, promoting consistency in ratings between analysts and jobs.

In this method, repetition is defined in terms of hand activity, or how busy the hands are during the work cycle. Ratings of repetition take into account two factors: 1) amount of recovery time within the cycle, and 2) how fast the hands are moving (Figure 5.2).

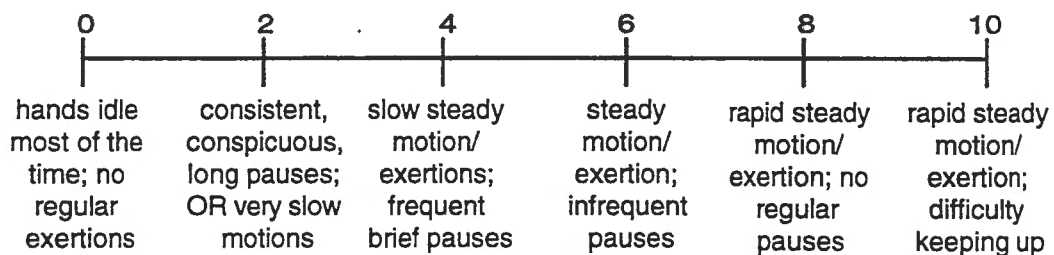


Figure 5.2: Visual-analog scale for rating repetition/hand activity, with verbal anchors.

A total of fifty-two ergonomic variables were quantified for each of the jobs (Table 5.3). The formal ratings were performed using a modified nominal group technique (Gustafson, 1973). The rating team consisted of 4 university faculty and research staff members who were experienced in ergonomic analysis in general, and this technique in particular. The videotaped jobs and documentation were presented to the team members, who independently rated the jobs for each stressor using the 10 cm scale and written guidelines. Each

hand was analyzed independently. When all team members had completed their individual ratings, the ratings were discussed with the goal of reaching consensus. Consensus was defined as: 1) a difference of no more than 1 unit on the 10 cm scale between the lowest score the highest score, and 2) the bases of all differences had been addressed. If the individual ratings for any stressor initially met consensus, no further discussion was necessary for that stressor. If consensus was not initially met, the outlying raters were given the opportunity to briefly explain the rationale for their ratings. In some cases, the discrepancies were due to observational differences: one rater simply noticed something that no one else did, or an individual overlooked something that everyone else saw. In these cases, the affected rater(s) adjusted their ratings accordingly and consensus was achieved.

Table 5.3: Physical stress parameters included in complete rating system (Latko, 1997b).

stressor	aspects rated
repetition	overall level of hand activity
forceful exertions	peak and average hand force
localized mechanical stress	peak and average mechanical stress on: • fingers • wrist/palm • forearm • elbow
posture	peak and average posture, peak and average angular velocity, and % exertion time for: • fingers • wrist • forearm • elbow • shoulder • neck • back
low temperature	peak and average hand exposure to low temperature
vibration	peak and average hand-arm vibration exposure
jerk/impulse	peak and average jerk/impulse loading of the hand

After the group agreed on the ratings for each job, the jobs from each site were again examined with regard to the inclusion criteria (i.e. availability of adequate numbers of workers in all three repetition categories). For the purposes of stratification in this study, three ranges of repetition were defined: low, medium, and high. These categories were defined by a strict division of the scale into thirds: 0 - 3.3 = low, 3.3 - 6.6 = medium, and above 6.6 = high. The average of the four raters was used for all analyses. Three manufacturing

facilities were ultimately included in the study; they represented office furniture (OF), spark plug (SP), and industrial container (IC) manufacturing. The included jobs and their repetition ratings are listed in Table 5.4.

Table 5.4: Jobs included in study, stratified on repetition level. Average repetition ratings are in parentheses.

repetition category	plant and job	rating
low (0 - 3.3) mean rating = 2.4	<i>office furniture manufacturing (n=25)</i>	
	machine loading	(1.9)
	NC machine operation	(2.9)
	<i>industrial container manufacturing (n=18)</i>	
	injection molding machine operator	(1.6)
	material handling	(1.7)
	drum forming machine operator	(2.4)
	drum sealing machine operator	(3.0)
	<i>spark plug manufacturing (n=75)</i>	
	shipping/receiving clerk	(1.2)
	machine setup and monitor A	(1.4)
	printer setup	(2.1)
	inspector A	(2.5)
	machine setup and monitor B	(2.7)
	machine setup and monitor C	(2.8)
	machine setup and monitor D	(3.0)
	machine setup and monitor E	(3.1)
	inspector B	(3.3)
medium (3.4 - 6.6) mean rating = 5.4	<i>office furniture manufacturing (n=24)</i>	
	office cubicle panel upholstery	(6.4)
	<i>industrial container manufacturing (n=25)</i>	
	stacker-large drums A	(3.5)
	ironing machine operator	(3.5)
	molding machine operator	(4.8)
	molding machine helper	(5.2)
	stacker-large drums B	(5.3)
	small drum cover glue	(5.3)
	small drum crimping mach. operator	(5.5)
	large drum slitter operator	(6.1)
	punch press operator	(6.5)
	small drum unload	(6.5)
	<i>spark plug manufacturing (n=13)</i>	
	spark plug transfer	(5.8)

high (6.7 - 10) mean rating =8.0	<i>office furniture manufacturing (n=36)</i>	
	office chair upholstery A	(7.9)
	office chair upholstery B	(8.1)
	<i>industrial container manufacturing (n=25)</i>	
	rubber gasket machine operator	(6.7)
	large drum label stamping	(6.8)
	handle riveting	(8.4)
	band forming press	(8.4)
	band welding	(8.5)
	handle assembly	(8.8)
	<i>spark plug manufacturing (n=111)</i>	
	spark plug transfer A	(7.2)
	platinum spark plug weld	(7.9)
	spark plug transfer B	(8.1)
	checker/packer	(8.6)

Table 5.5 lists the mean ratings of selected other stressors by repetition category. The nine stressors listed in the table are representative of the entire set. Of the 51 other ergonomic stressors quantified, there were statistically significant differences between exposure levels for several. For example, mean "peak force" ratings ranged from 4.0 to 4.4 and mean "average force" ratings ranged from 0.8 to 1.3 for the three repetition levels (see Table 5.5). The difference in force ratings between the repetition categories is statistically significant. However, while these differences are significant statistically, their absolute magnitude is relatively small, thus it is probable that they are not meaningful from a practical standpoint. Because the jobs were stratified and selected solely based on repetition level, the similarity in exposure to these other stressors is not surprising. The exceptions among the parameters listed in Table 5.5 are average wrist velocity and wrist exertion time, which had ranges of 3.4 and 5.7 units, respectively. Because repetition was defined in this study as a combination of movement speed and exertion time (Figure 5.2), these parameters are not independent of the repetition ratings. Previous work has shown that repetition rating can be closely modeled as a linear function of average wrist velocity and finger exertion time (Latko, 1997b).

Table 5.5: Average exposures to selected stressors stratified by repetition (mean $\pm$ std. dev.).

stressor	low (0<rep<3.3)	medium (3.3<rep<6.6)	high (6.6<rep<10)	p-value
force (peak)	4.4 $\pm$ 1.8	4.4 $\pm$ 1.3	4.0 $\pm$ 1.7	0.02
force (avg)	0.8 $\pm$ 0.4	1.3 $\pm$ 0.2	1.3 $\pm$ 0.5	<0.0001
contact stress- fingers (peak)	3.8 $\pm$ 1.7	4.6 $\pm$ 1.6	3.4 $\pm$ 1.3	<0.0001
contact stress- fingers (avg)	0.8 $\pm$ 0.3	1.6 $\pm$ 0.4	1.8 $\pm$ 0.9	<0.0001
wrist posture (peak)	5.2 $\pm$ 1.1	6.4 $\pm$ 0.5	6.4 $\pm$ 1.0	<0.0001
wrist posture (avg)	1.1 $\pm$ 0.4	2.2 $\pm$ 0.5	2.9 $\pm$ 0.5	<0.0001
wrist velocity (peak)	5.2 $\pm$ 1.3	5.3 $\pm$ 0.5	6.3 $\pm$ 0.5	<0.0001
wrist velocity (avg)	0.9 $\pm$ 0.3	2.1 $\pm$ 0.3	3.3 $\pm$ 0.3	<0.0001
wrist exertion time	3.5 $\pm$ 1.0	7.8 $\pm$ 1.3	9.2 $\pm$ 0.2	<0.0001

Medical survey

After the jobs were selected based on the above criteria, subjects were recruited for the medical survey from among the active workers performing those jobs. The only constraint on worker participation was that they had to have been performing the job of interest for at least the six consecutive months prior to the study date. All participants provided written informed consent, and all medical evaluations were performed on company time, during normal work hours.

The medical survey consisted of four types of evaluations: a questionnaire, limited physical examination of the upper extremities, limited electrodiagnostic testing at both wrists, and general anthropometry measurements. All clinical procedures were performed by appropriately trained health professionals. Clinicians were blinded to the medical and job-related data collected by other members of the study team.

The self-administered questionnaire focused on a variety of types of information, and has been described previously by Franzblau et al. (1993). Demographic information, including age, race, gender, educational level, history of tobacco use, work history, job tenure, and job title was obtained. A medical history section requested information concerning the presence of diseases or conditions thought to be related to WRMSDs, including diabetes, thyroid disease, rheumatoid arthritis, gynecological factors, previous surgeries, and previous acute injuries. Information was also obtained on current health

status and symptoms potentially related to WRMSDs, including burning, stiffness, pain, cramping, tightness, aching, soreness, tingling, and numbness in each of fifteen body locations. Subjects were instructed to report a symptom if it had been present in at least three separate episodes, or if one episode had lasted more than one week, during the previous twelve months. This portion of the questionnaire did not ask subjects to distinguish the exact location of the symptoms, e.g. symptoms in the distribution of the median nerve from symptoms elsewhere in the fingers, hands, or wrists. If symptoms in the wrist, hand, or fingers were reported, subjects were asked to indicate if they had experienced nocturnal occurrence of the symptoms. The test-retest reliability of all aspects of the questionnaire (excluding psychosocial questions) has been shown to be good to excellent (Franzblau et al., in press)

Each subject also completed a hand diagram (Katz et al., 1990a and 1990b; Franzblau et al., 1994) indicating whether he/she had experienced numbness, tingling, burning, or pain (NTBP) in the fingers, hands or wrists in at least three separate episodes, or in one episode lasting more than one week, during the previous twelve months. If any symptoms meeting these criteria had been present, he/she was instructed to shade in the distribution of the symptoms on the hand diagram. The hand diagrams were scored independently for likelihood of underlying Carpal Tunnel Syndrome by two physicians blinded to other test results. Differences were resolved by consensus. The rating system used has been described previously by Franzblau et al. (1994) (Table 5.6).

Table 5.6: Hand diagram scoring definitions.

Score	Criteria
Classic (score = 3):	Tingling, numbness, burning or pain in at least 2 of digits 1,2, or 3. Palm and dorsum of the hand excluded; wrist pain or radiation proximal to the wrist allowed.
Probable (score = 2):	Same as classic, except palmar symptoms allowed unless confined solely to ulnar aspect.
Possible (score = 1):	Tingling, numbness, burning, or pain in at least one of digits 1, 2, or 3.
Unlikely (score = 0):	No symptoms in digits 1, 2, or 3.

Subjects completed a series of psychosocial evaluations, which included selected scales from the Karasek Job Content Questionnaire (Karasek, 1985), Cohen's Perceived Stress Scale (Cohen, 1983) and a social network questionnaire.

Anthropometric data, including height, weight, finger circumference and length (digit 2), and wrist width and depth, and triceps skinfold thickness were collected.

The screening physical examination was adapted from that described by Fine and Silverstein (Fine, 1988; Fine and Silverstein, 1995). The physical examination included inspection, palpation, active and passive range of motion, and specific screening tests, including Tinel's, Phalen's, and Finkelstein's tests. Two-point discrimination was performed and was considered normal if a subject could correctly perceive 2 points which were 4 mm apart at the tip of the index finger. Physical exams were performed by physicians trained in occupational medicine.

Bilateral limited electrophysiologic testing of each subject was performed by a board certified electromyographer or a certified electrodiagnostic technician working under direct supervision of a board certified electromyographer. This testing consisted of assessing sensory function of the median and ulnar nerves at the wrists using surface electrodes and 14 cm antidromic stimulation. Sensory amplitude, peak latency, and takeoff latency were measured in each nerve tested. Hand temperature was monitored; if hand temperature was below 32°C, the hands were warmed to 32°C before testing commenced.

Because WRMSDs encompass a variety of specific diagnoses, several specific health outcome measures were modeled in this study: tendinitis of the wrist, carpal tunnel syndrome (CTS), and subject-reported non-specific discomfort (Table 5.7). Since there are no universally accepted criteria for the clinical or laboratory diagnosis of CTS (Stock, 1991; 1992), three different case definitions of CTS are presented in this paper; several other definitions were evaluated (Franzblau et al., 1993).

Table 5.7: Case definitions of health outcome measures presented in this paper.

Outcome measure	Definition
Non-specific discomfort	burning, stiffness, pain, cramping, tightness, aching, soreness, tingling, or numbness in the fingers, hand, or wrist
Tendinitis	Symptoms (pain, stiffness, burning, tightness, aching, or soreness) plus physical exam findings consistent with tendinitis in the elbow, forearm, wrist, hand, or fingers (pain with resisted motion, tenderness, or positive finding on appropriate test maneuver-e.g. Finkelstein's maneuver)
CTS-symptoms alone	symptoms consistent with CTS indicated by the hand diagram ("Classic" or "Probable"--see Table 5.6)
CTS-electrophysiology alone	median mononeuropathy ("MM5"- difference in peak latency of 0.5 ms between ulnar and median nerve)
CTS- symptoms and electrophysiology combined	positive findings on both hand diagram and electrophysiologic testing, as defined above

Data analysis

Data were analyzed following the approach used by Hales et al. (1994).

There were a total of 109 exposure variables analyzed: 10 anthropometry parameters, 25 medical history parameters, 5 demographic parameters, 13 psychosocial parameters, 4 tobacco use parameters, and 52 ergonomic parameters. Separate analyses were performed for each of the five health outcomes presented. For each health outcome, the following three stage process was followed:

- Stage 1) Univariate analyses were performed with each of the 109 exposure parameters as independent variables. Fisher's exact test, the Pearson chi-squared test, or logistic regression was used, depending on the nature of each exposure variable. Non-significant variables ($p > 0.10$) were excluded from further analyses.
- Stage 2) The independent variables which were not excluded in step 1 were grouped into six general categories: anthropometry, medical history, demographic, psychosocial, tobacco, and ergonomic. Multiple variable logistic analyses were conducted for each of the six categories using these remaining variables. Variables which were

not significant after this stage ($p>0.05$) were excluded from further analyses.

Stage 3) Multivariate logistic analyses were performed using all variables which were not eliminated in Stages 1 or 2. If a subject was missing data for any variables included in the final model, that subject was excluded from the analysis. Non significant variables were iteratively eliminated until the final model was achieved.

All health parameters and ergonomic stress assessments were analyzed for the dominant hand only. Because individuals with diabetes are likely to develop neuropathy, and because the diabetics in the study population exhibited significantly different electrophysiological results than non-diabetics, subjects who reported diabetes on the health history questionnaire were excluded from all analyses which included electrophysiologic parameters ($n=16$). All analyses were adjusted for age and gender, regardless of significance of these covariates. Ergonomic parameters were modeled as continuous effects in the range 0-10.

Results

A total of 438 workers were employed in the selected jobs at the time of the study and met the inclusion criteria (i.e. 6 month job tenure). Three hundred fifty-two (80%) participated in the study. Participation rates for the individual sites were 88% (Office Furniture), 84% (Industrial Container), and 76% (Spark Plug). Reasons for non-participation included absenteeism, scheduling conflicts, and refusal to participate. Table 5.8 shows the major demographic characteristics by plant and repetition category.

There was a statistically significant difference in worker age between the three repetition categories. Workers in the medium repetition jobs were the youngest (mean = 37.9), while those in the low repetition jobs were the oldest (mean = 43.0) (Table 5.8). The gender distribution was not equal over the three categories. The low repetition jobs were predominately male (90%), while the high repetition jobs had a slightly higher percentage of females (55%). There were no differences in BMI between the groups.

Table 5.8: Age, gender, and BMI distribution by plant and repetition category.

	plant	total	low	medium	high
n	COMBINED	352	118	62	172
	OF	85	25	24	36
	IC	68	18	25	25
	SP	199	75	13	111
age *	COMBINED	41.3±10.5	43.0±9.9	37.9±9.4	41.4±10.9
mean ±SD	OF	37.2±9.0			
	IC	37.0±9.4			
	SP	44.5±10.3			
gender*	COMBINED	206/146	98/20	30/32	78/94
M/F	OF	54/31			
	IC	50/18			
	SP	102/97			
BMI	COMBINED	28.6±5.8	28±4.8	29.7±7.1	28.6±5.9
(kg/m ²)	OF	26.6±4.6			
	IC	29.6±6.7			
	SP	29.1±5.8			

* statistically significant difference between levels at p<0.05.

Table 5.9 shows the general linear trends for five health outcome measures. Subjects were stratified into three repetition categories based on the repetition rating of their jobs. A strict division of the scale into thirds was used for this analysis (low = 0 - 3.3 , medium = 3.3 - 6.6, high = above 6.6). In general, the linear trend was significant for discomfort, tendinitis, and the hand diagrams, but non significant for MM5. The linear trend for CTS (defined by hand diagram and MM5) was borderline.

Table 5.9: Linear trend of symptoms, tendinitis, and CTS in the dominant extremity.

symptom	total	low (mean = 2.4)	medium (mean = 5.4)	high (mean = 8.0)	Prob> χ^2 overall	Prob> χ^2 linear trend
n	352	118	62	172		
wrist, hand, finger discomfort	129 (36.7%)	26 (22.0%)	23 (37.1%)	80 (46.5%)	0.0001	<0.0001
tendinitis	35 (9.9%)	5 (4.2%)	5 (8.1%)	25 (14.5%)	0.01	0.004
CTS						
hand diagram	47 (13.4%)	8 (6.8%)	9 (14.5%)	30 (17.4%)	0.03	0.01
MM5	81 (24.0%)	30 (26.8%)	10 (16.4%)	41 (25.0%)	0.29	0.83
hand diag. + MM5	19 (5.6%)	3 (2.7%)	3 (4.9%)	13 (7.9%)	0.17	0.06

Non-specific discomfort

Thirty-seven percent of the workers reported symptoms of burning, stiffness, pain, cramping, tightness, aching, soreness, tingling, or numbness in the fingers, hand, or wrist (Table 5.9). The prevalence of self-reported symptoms increased from 22% for workers in the low repetition jobs to 46.5% for workers in the high repetition jobs. The χ^2 test for linear trend was significant ($p<0.0001$) (Table 5.9). The final stage logistic regression model using the presence/absence of these symptoms as the outcome measure is shown in Table 5.10. History of physician-diagnosed carpal tunnel syndrome or tendinitis and repetition level were significant terms in the final model.

Table 5.10: Predictors of discomfort in the wrist, hand, or fingers; N=351. Log likelihood = -208.02

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-1.86 (0.59)	0.002		
history of CTS (Y=1, N=0)	0.75 (0.38)	0.05	2.11	1.00 - 4.47
history of tendinitis (Y=1, N=0)	0.68 (0.31)	0.03	1.98	1.07 - 3.63
sex (F=1, M=0)	0.72 (0.25)	0.003	2.07	1.27 - 3.35
age (yrs)	-0.004 (0.01)	0.71	1.00	0.97 - 1.02
repetition rating (0-10)	0.16 (0.05)	0.001	1.17*	1.06 - 1.29

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(\text{coefficient} \times \Delta \text{rating})}$

Tendinitis

Ten percent of the total subject population had symptoms and physical exam findings consistent with tendinitis in the dominant elbow, forearm, wrist, hand, or fingers (Table 5.9). No subjects, however, exhibited the classical signs of tendinitis, i.e. swelling and redness. Prevalence rates increased from 4.2% for workers in low repetition jobs to 14.5% for workers in the high repetition jobs. There was a significant linear trend for the three levels of repetition ($p<0.01$) (Table 5.9). History of soft tissue disease (i.e. tendinitis, epicondylitis, or rotator cuff syndrome), triceps skinfold thickness, and repetition rating were significant in the final regression model (Table 5.11). Gender was also significant, with females having approximately twice the odds of males, while age was not significant. Four subjects were excluded from this analysis due to missing data.

Table 5.11: Predictors of tendinitis in the dominant elbow, forearm, wrist, hand, or fingers; N=348. Log likelihood = -100.85

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-5.93 (1.09)	<0.001		
history of soft-tissue disease (Y=1, N=0)	0.98 (0.38)	0.01	2.66	1.27 - 5.55
sex (F=1, M=0)	-0.35 (0.50)	0.48	0.70	0.27 - 1.85
age (yrs)	0.03 (0.02)	0.15	1.03	0.99 - 1.07
triceps skinfold thickness	0.04 (0.02)	0.02	1.04	1.01 - 1.07
repetition rating (0-10)	0.19 (0.09)	0.02	1.21*	1.03 - 1.44

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR = e^{(\text{coefficient} \times \Delta \text{rating})}$

Carpal Tunnel Syndrome

a) Hand diagram score only (symptoms consistent with CTS)

Based on the hand diagram scores, 13.4% of the subjects exhibited "classic" or "probable" CTS (i.e. hand diagram score of 3 or 2, see Table 5.6). There was a significant linear trend with respect to repetition level, with prevalence increasing from 6.8% for workers in the low repetition jobs to 17.4% for workers in the high repetition jobs (Table 5.9). Wrist ratio and repetition level were significant in the final model (Table 5.12). Wrist ratio was defined as the ratio of wrist depth to width and was modeled as a binary variable, with ratios at or below the 75th percentile (0.73) of the study population modeled as 0 and ratios above the 75th percentile (>0.73) modeled as 1. Gender achieved borderline significance, with females having elevated odds, while age was not significant.

Table 5.12: Predictors of CTS based on hand diagram score of 2 or 3; N=351. Log likelihood = -127.14

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-3.52 (0.85)	<0.001		
sex (F=1, M=0)	0.63 (0.34)	0.07	1.88	0.96 - 3.70
age (yrs.)	0.003 (0.02)	0.85	1.00	0.97 - 1.03
wrist ratio--depth/width (1=>0.73, 0=<0.73)	0.95 (0.33)	0.004	2.59	1.35 - 4.96
repetition rating (0-10)	0.15 (0.07)	0.05	1.16*	1.00 - 1.34

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR = e^{(\text{coefficient} \times \Delta \text{rating})}$

b) Median mononeuropathy (0.5 ms threshold)

Based on the diagnostic criterion of a difference in peak latency of 0.5ms between the ulnar and median nerve, 24% of the subjects were classified as having median mononeuropathy in the dominant wrist. Sixteen subjects were excluded from the analyses of electrodiagnostic results due to reported diabetes. There was no significant linear trend with repetition level (Table 5.9); prevalence rates were similar for workers in the low and high repetition jobs (26.8% and 25%, respectively), but lower for the medium repetition jobs (16.4%). Significant terms in the logistic model were age, gender, body mass index (BMI) and wrist ratio (Table 5.13). No ergonomic parameters were statistically associated with the health outcome. Males were more likely than females to have median mononeuropathy, and risk also increased with age.

Table 5.13: Predictors of CTS based on median mononeuropathy (0.5 ms threshold); N=336. Log likelihood = -163.88

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-5.35 (0.88)			
sex (F=1, M=0)	-0.61 (0.29)	0.04	0.54	0.31 - 0.97
age (yrs)	0.03 (0.01)	0.05	1.03	1.00 - 1.06
BMI (kg/m ²)	0.10 (0.02)	<0.001	1.11	1.06 - 1.16
wrist ratio --depth/width (1=>0.73, 0=<0.73)	1.02 (0.29)	0.001	2.77	1.56 - 4.92

- c) Hand diagram score of 2 or 3 (symptoms consistent with CTS) and median mononeuropathy (0.5 ms threshold)

When the strictest definition of CTS was used, requiring both hand diagram indications consistent with CTS and positive electrodiagnostic findings, 5.6% (n=19) of the study group met the criteria. There was an observable increasing trend in prevalence with increasing repetition. 2.7% of the workers in low repetition jobs met the diagnostic criteria, while 7.9% of the workers in the high repetition jobs met the criteria (Table 5.9). This trend achieved borderline statistical significance ($p<0.06$). For this health outcome, the three stage logistic regression procedure did not indicate any significant predictors at $p=0.05$; wrist ratio and repetition rating were close to achieving statistical significance ($p<0.06$ and 0.08 , respectively, Table 5.14). Gender and age were non-significant. A negative log likelihood test was performed for each of the two borderline terms, wrist ratio and repetition rating. Both terms were again borderline significant (repetition $\chi^2=3.6$, $p<0.06$ and wrist ratio $\chi^2=3.5$, $p<0.07$).

Table 5.14: Predictors of CTS based on median mononeuropathy (0.5 ms threshold) combined with hand diagram scores of at least 2; N=336. Log likelihood = -68.87

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-5.02 (1.30)	<0.001		
sex (F=1, M=0)	-0.15 (0.50)	0.77	0.86	0.32 - 2.31
age	0.02 (0.02)	0.49	1.02	0.97 - 1.06
wrist ratio (1=>0.73, 0=<0.73)	0.93 (0.49)	0.06	2.53	0.97 - 6.57
repetition rating (0-10)	0.20 (0.11)	0.08	1.22*	0.98 - 1.53

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(coefficient \times \Delta rating)}$

In the above analyses, repetition rating was modeled as a continuous variable, in the range 0-10. The original job selection criteria relied on a 3 category classification of repetition (low, medium, and high), based on a division of the 10 cm scale into thirds (Table 5.4). Table 5.15 shows the odds ratios for changes in repetition between the mean values in the three categories: low, medium, and high, based on the above analyses for the five health outcomes. The range between mean ratings for the "low" and "medium"

categories was 3.0 (2.4 - 5.4), the range between "low" and "high" category means was 5.6 (2.4 - 8.0), and the range between "medium" and "high" category means was 2.6 (5.4 - 8.0). For the three health outcome measures which were significantly related to repetition level, the odds ratios were similar. A change from low to medium repetition resulted in an odds ratio of 1.57 to 1.77 depending on the outcome measure, a change from low to high yielded an OR of 2.32 to 2.90, and changes from medium to high resulted in OR of 1.48 to 1.64 (Table 5.15).

Table 5.15: Odds ratios based on changes in repetition level for the five health outcome variables presented.

outcome measure	low-medium (Δ rating = 3.0) OR (95%CI)	low-high (Δ rating = 5.6) OR (95%CI)	med-high (Δ rating = 2.6) OR (95%CI)
non-specific discomfort	1.62 (1.20 - 2.17)	2.45 (1.42 - 4.24)	1.52 (1.17 - 1.96)
tendinitis	1.77 (1.04 - 3.00)	2.90 (1.08 - 7.78)	1.64 (1.04 - 2.59)
hand diagram ("classic" or "probable")	1.57 (1.04 - 2.37)	2.32 (1.07 - 4.99)	1.48 (1.03 - 2.11)
MM5	NS	NS	NS
hand diagram + MM5	1.84 (0.94 - 3.59)	3.11 (0.89 - 10.87)	1.69 (0.95 - 3.03)

Discussion

In this study, repetition was found to be associated with worker-reported discomfort in the upper limb, tendinitis, and symptoms consistent with CTS as reported on a hand diagram. Workers in high repetition jobs (average repetition rating of 8) had 2 to 3 times higher risk of these health outcomes than did workers in low repetition jobs (average repetition rating of 2.4).

Electrophysiology alone did not show any statistically significant relationship to repetitiveness of work. When the health outcome was defined as the combination of median mononeuropathy and hand diagram symptoms, repetition was close to achieving significance ($p < 0.08$ and $p < 0.06$ for logistic model and linear trend, respectively). There were a relatively small number of subjects ($n=19$) who tested positive for CTS using this case definition (Table 5.9), making it more difficult to show a statistical association. A clear linear trend is observable in the data, which is borderline significant. It is possible that this trend would achieve significance with a larger sample size.

For the health outcome parameters where a statistically significant association was found, the odds ratios for repetition in this study were similar to those determined in previous studies linking repetitive work to adverse health outcomes in workers. In this study, the increased odds of discomfort associated with an increase of 1 unit on the 10 cm repetition scale was approximately 1.16 (Table 5.10). Extending this to the average ratings of jobs in the high versus low repetition categories results in an odds ratio of 2.45. In a study of garment industry workers, Punnett et al. (1985) found odds ratios of 2.7 and 3.9 for persistent pain, numbness, or tingling in the hand or wrist, respectively, for workers whose jobs required repetitive hand movements versus those whose jobs did not require repetitive motions. In a comparison of ski manufacturing workers in "clearly highly repetitive" and "not repetitive" jobs, Barnhart et al. (1991) reported odds ratios of 1.22 for hand pain in the right hand, and 1.17 for any symptoms in the right hand.

Symptoms consistent with carpal tunnel syndrome as identified on the hand diagram had an odds ratio of 2.32 for high vs. low repetition in the current study. In a study of CTS as determined by physical exam and interview among a large industrial population, Silverstein et al. (1987) found an OR of 2.7 for high repetitive-low force jobs compared to low repetitive-low force jobs. When Punnett et al. (1985) examined symptoms specifically related to CTS, the odds ratios for the two groups was 3.0. Cannon (1981), in a case-control study of CTS patients vs. matched controls, found an odds ratio of 2.1 related to performance of repetitive motion tasks. These three studies relied only on symptoms and/or physical exam findings consistent with CTS, and did not include any electrodiagnostic testing.

In the current study, an association was found between repetitive work and diagnosis of CTS when positive electrophysiologic findings and symptoms were included in the diagnostic criteria, although the statistical significance of this association was borderline. Previous investigators have had differing results. Barnhart et al. (1991) found a significant relationship between repetitive work and difference in peak latency of 0.5ms (OR=2.32), but no significant relationship when the stricter diagnostic criteria of the 0.5ms latency plus positive test results (Phalen's or Tinel's) or symptoms consistent with CTS were used. Wieslander (1989) found a positive relationship between the performance of repetitive wrist movements at work and CTS defined by both diagnosis by a surgeon and positive electrophysiologic results for workers who

had the work exposure for more than 20 years ($OR = 4.6$), but not for workers with less job tenure.

Twenty-four percent of all the workers in the current study had abnormal electrophysiologic test results (Table 5.9). This rate is comparable to rates determined by other investigators among working populations. In a previous study of industrial workers using methods similar to those described here, Franzblau et al. (1993) found 23.8% of subjects with median mononeuropathy in the wrist(s). Barnhart (1991) reported that 31% of the workers with repetitive jobs and 19% of the workers in non-repetitive jobs had peak sensory latency differences between the ulnar and median nerve of at least 0.5ms. Nathan et al. (1988) reported that 39% of 471 randomly selected workers in four industries had peak sensory latency differences between the ulnar and median nerve of at least 0.4ms.

Further analyses were performed using slight variations in the definitions of both exposure and health outcome criteria in order to test the robustness of the associations found. The complete models for these analyses are presented in Appendix C. The analysis presented in Table 5.8 considers a hand diagram score of 2 ("probable") or 3 ("classic") to indicate presence of CTS. Changing the analysis to also include scores of 1 ("possible") does not change the significant parameters. Restricting the criteria to include only scores of 3 ("classic") yields similar results. When presence of subjects reporting "numbness, tingling, burning, or pain" is substituted into the model for the hand diagram score, repetition still is a significant parameter, although wrist ratio is not. In this study, median mononeuropathy was defined as a difference in peak latencies (ulnar-median) of at least 0.5ms. While this is a commonly used threshold for diagnosis of CTS, a threshold of 0.8ms is also sometimes used (Franzblau et al., 1994; Stetson et al., 1993). Using this more restrictive criterion does not change the significant terms in the model containing median mononeuropathy alone as the health outcome. When these alternative definitions are combined in the strictest CTS model (presence of symptoms plus positive electrodiagnostic findings), there are minor changes in the magnitudes of the model terms. When the definition of CTS is changed to reflect a hand diagram score of 2 or 3 and median mononeuropathy (0.8ms threshold), repetition becomes significant, while none of the other parameters change. When CTS is modeled as reported numbness, tingling, burning, or pain and median mononeuropathy (0.5 ms threshold), repetition is again significant,

while wrist ratio is not. The high level of similarity between the models when alternative diagnostic parameters are substituted indicates that the findings are robust to alternative definitions of disease. The significance of repetition in these alternative models of CTS provides further evidence of an association between repetitive work and CTS, despite the borderline statistical significance reported in Table 5.14.

Previous studies examining the relationship between repetitive work and adverse health outcomes in workers have modeled exposure as a binary variable. The definition of repetition used in this study is unique in that it quantifies repetition on a continuous scale from 0 to 10. As indicated by Hagberg (1992), modeling more levels of an exposure tends to decrease the likelihood of significance. Modeling as two or three categories (low/high or low/medium/high) based on the repetition ratings yields similar odds ratios to those reported in Table 5.15. These models are included in Appendix C. The repetition rating method used in this study has undergone both reliability and validity testing (Latko 1997a and 1997b), with satisfactory results for both. Similarly, the questionnaire used in this study has been shown to have good to excellent test-retest reliability (Franzblau et al., in press).

No other ergonomic stresses (e.g. force, posture) were found to be associated with any health outcome measure. This is not surprising, considering the range of exposures to these other stressors which were encountered. Because the job selection was stratified only on repetition, no effort was made to insure wide representation of the other stressors. In general, most of the exposure levels for the other stressors fell within a 2 or 3 unit range on the ten cm scale (Table 5.5). It is likely that this is not enough variability in exposure to produce significant results.

Several of the studies cited above (e.g. Punnett et al. (1985) and Barnhart et al. (1991)) defined repetition solely in qualitative terms, identifying the presence or absence of repetitive movements. This classification of repetition does not isolate exposure to repetition from exposure to other stressors. It is possible that exposure to repetitive hand activities may be confounded with exposure to other stressors, such as forceful exertions or awkward postures. In the study reported in this chapter, exposure to these other ergonomic stressors was also quantified. Exposure levels were found to be similar between the different levels of repetition (see Table 5.5). This strengthens the conclusion that the increased risk observed is due to increased

repetitiveness, and not due to another factor (such as force or posture) that coincidentally varied with repetition.

There were no consistent gender effects found in this study. Females were approximately 2 times more likely to report discomfort than were males (see Table 5.10). In contrast, females were approximately half as likely as males to exhibit median mononeuropathy (see Table 5.13). None of the other 3 health outcome parameters addressed in this chapter had significant gender effects. Age was significant only for median mononeuropathy, for which a one year increase in age increased risk 3%. Plant effects were modeled in the initial univariate analysis of each health outcome. These effects were not significant.

There are several limitations to this study. This was a cross-sectional study, which does not allow any temporal association to be determined (Hennekens and Buring, 1987), although study participants were limited to those who had been on the job at least six months in an attempt to lessen the chance that disorders observed were related to previous work experience. There is also the possibility of a survivor effect, i.e. that workers who have had problems have left the workplace or changed jobs because of them, and that the population remaining is particularly resistant to such problems. The survivor effect would, however, tend to reduce the associations found between work and health outcomes.

Worker exposure to ergonomic stressors other than repetition was relatively low and did not vary much between jobs, making it unlikely to determine an association between these stressors and the various health outcomes. Exposure to ergonomic stressors was evaluated for one representative worker in each job classification. Although the investigators attempted to ascertain that the worker evaluated was representative, it is possible that exposures varied between workers. It is unlikely that repetition rates would vary significantly, because most of the jobs included in this study were machine paced, performed in a workcell (i.e. group paced), or had strict production standards. Stresses such as posture, however, may have had variability due to individual work style and anthropometry. Finally, the subjects in this study were limited to employees at companies which were receptive to participation in the study and agreed to allow the medical evaluations to be performed on company time. There is a possibility that the management attitude and culture at these companies may be different than at other companies, thus biasing the results.

This study clearly indicates a link between repetitive work and upper-limb musculoskeletal disorders in workers. No clear link was found between repetitive work alone and purely electrophysiologic signs of Carpal Tunnel Syndrome. However, the number of cases for some outcomes was modest or small, which makes it more difficult to detect significant effects. The results of this study indicate the need for future studies examining the relationship between a wide range of exposures to combinations of stressors and health effects on workers. Similar studies examining a wider variety of workplaces (e.g. offices), with more variability in exposures to other stressors, are also needed.

References

- Armstrong, T. J.; Bir, C.; Finsen, L.; Foulke, J.; Martin, B.; Sjogaard, G.; & Tseng, K. (1993a). Muscle responses to torques in hand held power tools. *Proceedings of the International Society of Biomechanics*, 114-115.
- Armstrong, T. J.; Buckle, P.; Fine, L. J.; Hagberg, M.; Jonsson, B.; Kilbom, A.; Kuorinka, I. A.; Silverstein, B. A.; Sjogaard, G.; & Viikari-Juntura, E. R. (1993b). A conceptual model for work-related neck and upper-limb musculoskeletal disorders. *Scandinavian Journal of Work & Environmental Health*, 19, 73-84.
- Armstrong, T. J.; & Chaffin, D. B. (1978). An Investigation of the relationship between displacements of the finger and the wrist joints and the extrinsic finger flexor tendons. *Journal of Biomechanics*, (11), 119-128.
- Armstrong, T. J.; Fine, L. J.; Goldstein, S. A.; Lifshitz, Y. R.; & Silverstein, B. A. (1987). Ergonomics considerations in hand and wrist tendinitis. *The Journal of Hand Surgery*, 12a(5, part 2), 830-837.
- Barnhart, S.; Demers, P. A.; Miller, M.; Longstreth, W. T.; & Rosenstock, L. (1991). Carpal Tunnel Syndrome among ski manufacturing workers. *Scandinavian Journal of Work, Environment, and Health*, 17, 46-52.
- Bystrom, S.; & Fransson-Hall, C. (1994). Acceptability of intermittent handgrip contractions based on physiological response. *Human Factors*, 36(1), 158-171.
- Cannon, L. J.; Bernacki, E. J.; & Walter, S. D. (1981). Personal and occupational factors associated with carpal tunnel syndrome. *Journal of Occupational Medicine*, 23(4), 255-258.
- Chiang, J.; Ko, Y.; Chen, S.; Yu, H.; Wu, T.; & Chang, P. (1993). Prevalence of Shoulder and Upper-limb Disorders Among Workers in the Fish-processing

- Industry. *Scandinavian Journal of Work, Environment, and Health*, 19, 126-131.
- Clarke, R. G. (1961). The limiting hand skin temperature for unaffected manual performance in the cold. *Journal of Applied Psychology*, 45, 193-194.
- Cohen S; Kamarck T.; and Mermelstein R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4):385-96.
- deKrom, M.; Kester, A.; Knipschild, P.; & Spaans, F. (1990). Risk Factors for carpal tunnel syndrome. *American Journal of Epidemiology*, 132(6), 1102-1110.
- Edwards, R. (1988). Hypotheses of peripheral and central mechanisms underlying occupational muscle pain and injury. *European Journal of Applied Physiology*, 57(3), 275-281.
- Fine, L. J. (1988). *Appendix B: Medical Treatment and Diagnosis*. In V. Putz-Anderson, eds., *Cumulative trauma disorders: a manual for musculoskeletal diseases of the upper limbs*. (pp. 119-127). Bristol, PA: Taylor and Francis, Inc.
- Fine, L. J.; & Silverstein, B. A. (1995). *Work-related disorders of the neck and upper extremity*. In B. S. Levy; & D. H. Wegman, eds., *Occupational Health: Recognizing and Preventing Work-Related Disease*. (3) (pp. 470-490). Boston: Little, Brown and Company.
- Fransson-Hall, C.; & Kilbom, A. (1993). Sensitivity of the hand to surface pressure. *Applied Ergonomics*, 24, 3.
- Franzblau, A.; Salemo, D.F.; Armstrong, T.J.; Werner, R.A. (in press). Test-Retest Reliability of an Upper Extremity Discomfort Questionnaire in an Industrial Population. *Scandinavian Journal of Work, Environment, and Health*.

- Franzblau, A.; Werner, R. A.; Albers, J. W.; Grant, C. L.; Olinski, D.; & Johnston, E. (1994). Workplace Surveillance for Carpal Tunnel Syndrome Using Hand Diagrams. *Journal of Occupational Rehabilitation*, 4(4), 185-198.
- Franzblau, A.; Werner, R.; Valle, J.; & Johnston, E. (1993). Workplace surveillance for Carpal Tunnel Syndrome: A comparison of methods. *Journal of Occupational Rehabilitation*, 3(1), 1-13.
- Gelberman, R.; Hergenroeder, P.; Hargens, A.; Lundborg, G.; & Akeson, W. (1981). The carpal tunnel syndrome: a study of carpal canal pressures. *The Journal of Bone and Joint Surgery*, 63A(3), 380-383.
- Goldstein, S. A.; Armstrong, T. J.; & Chaffin, D. B. (1987). Analysis of cumulative strain in tendons and tendon sheaths. *Journal of Biomechanics*, 20(1), 1-6.
- Gustafson, D. H.; Shulka, R. K.; Delbecq, A.; & Walster, G. W. (1973). A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, delphi groups, and nominal groups. *Organizational Behavior and Human Performance*, 9, 280-291.
- Hagberg, M. (1992). Exposure variables in ergonomic epidemiology. *American Journal of Industrial Medicine*, 21, 91-100.
- Hagberg, M.; Morgenstern, H.; & Kelsh, M. (1992). Impact of occupations and job tasks on the prevalence of carpal tunnel syndrome: a review. *Scandinavian Journal of Work, Environment, and Health*, 12, 277-279.
- Hales, T. R.; Sauter, S. L.; Peterson, M. R.; Fine, L. J.; Putz-Anderson, V.; Schleifer, L. R.; Ochs, T. T.; & Bernard, B. P. (1994). Musculoskeletal disorders among visual display terminal users in a telecommunications company. *Ergonomics*, 37(10), 1603-1621.
- Hammariskjold, E.; Harms-Ringdahl, K.; & Ekholm, J. (1992). Reproducibility of carpenters work after cold exposure. *International Journal of Epidemiology*, 9, 195-204.

- Hennekens, C. H. & Buring, J. E. (1987). *Epidemiology in Medicine*. Boston: Little, Brown, and Company.
- Hoffman, J.; & Hoffman, P. (1995). Staple gun carpal tunnel syndrome. *Journal of Occupational Medicine*, 27(11), 848-849.
- Karasek, R. (1985). *Job content questionnaire and user's guide, rev. 1.1*. Lowell, MA: University of Massachusetts.
- Katz, J. N. & Stirrat, C. R. (1990a). A self-administered hand diagram for the diagnosis of carpal tunnel syndrome. *Journal of Hand Surgery*, 15A, 360-363.
- Katz, J. N.; Stirrat, C. R.; Larson, M. G.; Fossel, A. H.; Eaton, H. M.; & Liang, M. H. (1990b). A self-administered hand diagram for the diagnosis and epidemiologic study of carpal tunnel syndrome. *Journal of Rheumatology*, 17, 1495-1498.
- Kuorinka, I.; & Koskinen, P. (1979). Occupational Rheumatic Diseases and Upper Limb Strain in Manual Jobs in a Light Mechanical Industry. *Scandinavian Journal of Work, Environment, and Health*, 5(suppl. 3), 39-47.
- Latko, W. A. (1997a). Comparison of observational and instrumental based measurements of repetition, force, and wrist posture in manual work (Chapter 4). Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI.
- Latko, W. A. (1997b). Evaluation of interrater reliability and test-retest reliability of observational ratings of physical stressors in manual work (Chapter 3). Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI.
- Latko, W. A.; Armstrong, T. J.; Foulke, J. A.; Herrin, G. D.; Rabourn, R. A.; & Ulin, S. S. (1997c). Development and evaluation of an observational method for assessing repetition in hand tasks. *American Industrial Hygiene Association Journal*, 58(4), 278-285.

- Lundborg, G.; Gelberman, G.; Minter-Convery, M.; Lee, Y.; & Hargens, A. (1982). Median nerve compression in the carpal tunnel-functional response to experimentally induced controlled pressure. *The Journal of Hand Surgery*, 7(3), 252-259.
- Lundstrom, R.; & Johansson, R. S. (1986). Acute impairment of the sensitivity of skin mechanoreceptive units caused by vibration exposure of the hand. *Ergonomics*, 29(5), 687-698.
- Luopajarvi, T.; Kuorinka, I.; Virolainen, M.; & Holmberg, M. (1979). Prevalence of tenosynovitis and other injuries of the upper extremities in repetitive work. *Scandinavian Journal of Work, Environment, and Health*, 5(suppl. 3), 48-55.
- Marras, W. S.; & Schoenmarklin, R. W. (1993). Wrist motions in industry. *Ergonomics*, 36(4), 341-351.
- Nathan, P.A.; Meadows, K.D.; and Doyle, L.S. (1988). Occupation as a risk factor for impaired sensory conduction of the median nerve at the carpal tunnel. *Journal of Hand Surgery*, 13B(2), 167-70.
- Nilsson, T.; Burstrom, L.; & Hagberg, M. (1989). Risk assessment of vibration exposure and white fingers among platers. *International Archives of Occupational and Environmental Health*, 61, 473-481.
- Phalen, G. (1966). The carpal tunnel syndrome: Seventeen years' experience in diagnosis and treatment of six hundred fifty-four hands. *The Journal of Bone and Joint Surgery*, 48A(2), 211-228.
- Punnett, L.; Robins, J. M.; Wegman, D. H.; & Keyserling, W. M. (1985). Soft Tissue Disorders in the Upper Limbs of Female Garment Workers. *Scandinavian Journal of Work, Environment, and Health*, 11, 417-425.
- Putz-Anderson, V. (1988). *Cumulative trauma disorders: a manual for musculoskeletal diseases of the upper limbs*. Bristol, PA: Taylor and Francis, Inc.

- Fransson-Hall, C.; & Kilbom, A. (1993). Sensitivity of the hand to surface pressure. *Applied Ergonomics*, 24, 3.
- Sauter, S. L.; Chapman, L. J.; Knutson, S. J.; & Anderson, H. A. (1987). Case example of wrist trauma in keyboard use. *Applied Ergonomics*, 18(3), 183-186.
- Schierhout, G. H.; Meyers, J. E.; & Bridger, R. S. (1995). Work Related Musculoskeletal Disorders and Ergonomic Stressors in the South African Workforce. *Occupational and Environmental Medicine*, 52, 46-50.
- Schoenmarklin, R. W.; Marras, W. S.; & Leurgans, S. E. (1994). Industrial Wrist Motions and Incidence of Hand/Wrist Cumulative Trauma Disorders. *Ergonomics*, 37(9), 1449-1459.
- Silverstein, B. A.; & Fine, L. J. (1984). *Evaluation of upper extremity and low back cumulative trauma disorders*. Ann Arbor, MI: The University of Michigan School of Public Health.
- Silverstein, B. A.; Fine, L. J.; & Armstrong, T. J. (1986). Hand Wrist Cumulative Trauma Disorders in Industry. *British Journal of Industrial Medicine*, 43, 779-784.
- Silverstein, B. A.; Fine, L. J.; & Armstrong, T. J. (1987). Occupational factors and carpal tunnel syndrome. *American Journal of Industrial Medicine*, 11, 343-358.
- Stetson, D. S.; Silverstein, B. A.; Keyserling, W. M.; Wolfe, R. A.; and Albers, J. W. (1993). Median sensory distal amplitude and latency: comparisons between non-exposed managerial/professional employees and industrial workers. *American Journal of Industrial Medicine*, 24(2), 175-189.
- Stock, S. R. (1992). Epidemiology of work-related musculoskeletal disorders of the neck and upper limbs: a response (letter to the editor). *American Journal of Industrial Medicine*, 21, 899-901.

- Stock, S. R. (1991). Workplace ergonomic factors and the development of musculoskeletal disorders of the neck and upper limb: A meta analysis. *American Journal of Industrial Medicine*, 19, 87-107.
- Sundelin, G.; & Hagberg, M. (1989). The effects of different pause types on neck and shoulder EMG activity during VDU work. *Ergonomics*, 32(5), 527-537.
- Tichauer, E. R. (1966). Some aspects of stress on forearm and hand in industry. *Journal of Occupational Medicine*, 8(2), 63-71.
- Viersted, K. B.; Westgaard, R. H.; & Andersen, P. (1990). Pattern of muscle activity during stereotyped work and its relation to muscle pain. *International Archives of Occupational and Environmental Health*, 62, 31-41.
- Viikari-Juntura, E. (1995). *The role of physical stressors in the development of hand/wrist and elbow disorders*. In S. L. Gordon; S. J. Blair; & L. J. Fine, eds., *Repetitive Motion Disorders of the Upper Extremity*. (pp. 7-30). Rosemont, IL: American Academy of Orthopaedic Surgeons.
- Wieslander, G.; Norback, D.; Gothe, C. J.; & Juhlin, L. (1989). Carpal Tunnel Syndrome (CTS) and Exposure to Vibration, Repetitive Wrist Movements, and Heavy Manual Work: a Case-Referent Study. *British Journal of Industrial Medicine*, 46, 43-47.

CHAPTER 6

DISCUSSION

The purpose of this dissertation was to evaluate a newly-developed observational method for assessing upper limb exposures to physical stressors in manual work. The stressor of primary concern was hand activity (repetition), although several other stressors (including forceful exertions, localized mechanical stress, and awkward postures) were also considered. The evaluation of this method was conducted in four phases. Each of the phases is summarized below, followed by sections discussing these findings and directions for future work.

Summary of findings

1) Ratings of hand activity/repetition exhibit varying levels of agreement with previous metrics for assessing repetitive work.

The purpose of this study was to compare a method for assessing repetition to other methods which has been previously applied by other investigators. A single team of analysts rated 33 jobs for repetition using the new method. The repetition ratings using this system were compared to measurements of recovery time within the cycle, exertion counts, and cycle time. Amount of recovery time within the job cycle was found to be significantly correlated with the analyst ratings ($r^2=0.58$, Figure 2.6), as were the number of exertions per second ($r^2=0.53$, Figure 2.4). Cycle time was not related to the analyst ratings (Figure 2.7). Repeated analyses using the new method were performed 1 1/2 to 2 years apart on the same jobs with the same group of raters. Ratings for repetition differed less than 1 unit (on the ten cm scale), on average, between the different sessions (Figure 2.8). The ratings from the two sessions were highly correlated ($r^2=0.88$). There was a statistically significant difference between the two sets of ratings, with the ratings from the second session

approximately 1/2 unit higher than the first session ratings. Although this difference was statistically significant, it is less than the one unit range defined as "consensus" in the rating process, and is therefore not large enough to be of practical concern.

2) Inter- and intra-rater reliability are strong for hand activity ratings, low to moderate for other stressors. The use of the group consensus process generally increases the consistency of ratings.

This study had three primary goals: 1) to determine the consistency with which different analysts can apply the rating method; 2) to evaluate the benefit (in terms of reducing rating variability) of using the group consensus process when rating, and 3) to determine the consistency with which a group of raters can apply the method at different points in time. Four teams of raters (3-5 raters per team, 17 raters total) rated a set of twelve jobs using a modified nominal group technique. The variance in ratings due to raters was determined for both the initial and final ratings. In general, intraclass correlations (ICCs), which measure the percent of total variance explained by differences between jobs, were very good ($ICC=0.79$, Table 3.5) for the initial ratings of repetition, indicating that the method is sensitive to between-job differences, even with different individual raters. Other stressors exhibited weak ICCs after the initial ratings ($0.10 < ICC < 0.39$, Table 3.5). The total variance in these models was low, however, which results in underestimation of inter-individual rating consistency. The consensus process greatly improved the ICCs. The ICC for repetition increased to 0.86, indicating that 86% of the variability in ratings was attributable to job differences. Ratings of peak and average force, peak and average contact stress on the fingers, and peak wrist posture improved into a range that is considered to be "good" (Fleiss, 1986). Intraclass correlations of the final ratings of these stressors ranged from 0.41 to 0.69. Only the ICCs of average wrist posture ratings remained weak.

Two of the teams ($n=3$ and $n=4$) rated an additional 33 jobs. ICCs ranged from 0.29 to 0.90 (median = 0.48) for the initial ratings and 0.21 to 0.94 (median = 0.69) for the final ratings (Table 3.7). With these groups of raters and jobs, only the average force ratings continued to have an unacceptable level of rater variability after the consensus process. One team re-rated a subset of twelve jobs seven weeks after the initial rating session. None of the seven scales evaluated had a statistically significant day effect.

3) Observational ratings of repetition, force, and wrist posture are consistent with direct measurements of the same parameters.

The primary goal of this study was to compare instrumentation information (EMG and electrogoniometry) about repetition, force, posture, and movement speed with observer ratings of the same parameters. In the observational method, repetition is defined as a composite index which integrates movement speed and exertion time. When the repetition rating was modeled as a function of average wrist movement speed and amount of finger recovery time (defined as exertion below 2.5% MVC), both terms were significant and the model explained 88% of the variance (Table 4.2).

Peak and average EMG were significant terms in the model of peak and average force ratings ($r^2 = .54$ and 0.49 , respectively, Figure 4.3). Peak and average wrist posture as measured with the goniometer were also significant effects in the model of peak and average wrist posture ($r^2 = .28$ and 0.53 , respectively, Figure 4.4). There was no significant relationship between the peak movement speed calculated from the goniometer data and ratings of peak movement speed. Calculated average movement speed was significantly related to ratings of average movement speed ($r^2 = .30$, Figure 4.5).

These results suggest that observers are able to identify repetition, force, movement speed, and average rates of wrist velocity. Strongest agreement between the observations and instrumentation was for the repetition scale, with peak and average force, and average wrist posture next in terms of model strength.

4) Increased levels of hand activity is associated with increased prevalence of upper limb musculoskeletal disorders in workers.

A cross sectional study was conducted in which workers' exposure to repetition and other physical stressors was quantified using the new method, and compared to the prevalence rates of various upper limb disorders. Three hundred fifty-two workers from three companies participated. Jobs were selected based on their repetition levels, to insure that a wide range of repetition exposures were represented. Other stressors, including force, posture, and contact stress were evaluated, but did not have a wide range of exposures in the sample jobs and were not controlled. Physical evaluations on

all participating workers were performed by medical professionals, and included a self-administered questionnaire, physical exam, and limited electrodiagnostic testing.

Repetitiveness of work was found to be significantly associated with prevalence of reported discomfort in the wrist, hand, or fingers, (OR=2.45 for change for low to high repetition) tendinitis in the wrist, hand, or fingers, (OR=2.90 for change for low to high repetition) and symptoms consistent with carpal tunnel syndrome as indicated on a hand diagram (OR=2.32 for change for low to high repetition). The relationship between repetitiveness of work and median mononeuropathy indicated by the combination of positive electrodiagnostic results and symptoms consistent with carpal tunnel syndrome achieved borderline statistical significance (OR=3.11 for change for low to high repetition).

Discussion of Findings

i. Comparison with other measures of "repetition"

When the ratings of repetition were compared to metrics that had been used by other investigators (i.e. number of exertions and amount of recovery time), there were moderate levels of agreement between the pairs of methods. It is expected that there would be some relationship between the measures of exertion frequency and recovery time and the repetition/hand activity ratings. A perfect relationship would not be predicted, because the team rating for repetition/hand activity is an integrated score, which takes into account the speed of hand movements as well as the amount of recovery time within the cycle. The exertion counts and number of pieces handled give an indirect indication of speed, in that a worker can accomplish more exertions per unit time if those exertions are rapid; however, the exertion counts do not directly assess movement speed. Also, the exertion counts do not account for the length of the exertions or the amount of recovery time within the cycle, both of which are factors that are considered by analysts using the rating method. The raters consider the trade off between recovery time and movement speed. No significant relationship was observed between the team ratings and cycle time (Figure 2.7). This is an important finding, because several epidemiological studies have used production standard based metrics to assess exposure to

repetitive exertions (Table 2.1). It is possible to find high repetition, short cycle jobs, and low repetition, long cycle jobs (see Figure 2.7). This appears to be the case for the cited studies. This does not, however, mean cycle time can be used to characterize repetition in all settings. Although the criteria applied in these previous studies may have been appropriate for those particular groups of jobs, the findings of this study suggest that applying such criteria to a random group of jobs, which were not preselected with these criteria in mind, may not be appropriate.

Part of the reason for this poor agreement may be because the rating system takes into account the tradeoff between speed of movements and the amount of recovery time, rather than considering cycle time directly. Based on the rating system, for a given task, the repetition level will increase as the cycle time decreases because movements will have to be faster or recovery time will be shorter if the same amount of work must be accomplished in a shorter time. However, merely calculating cycle time does not account for the different physiological requirements in different tasks. Similarly, counting the number of pieces handled does not consider what the worker does while handling each piece. The new method directly takes into account the level of activity of the hands, or how busy the hands are during the work cycle. Cycle time, number of exertions, and amount of recovery time are considered only for the impact that they have on the overall level of hand activity.

ii. Consistency of different raters, benefit of consensus, and consistency over time

In spite of the fact that the repetition ratings require the analysts to integrate several aspects of the job, users appear to be able to perform these ratings with a high level of consistency. For all analyses performed in this study, the repetition scale had high intraclass correlations (above 0.75), indicating a high degree of inter-rater reliability. The other stressors had varying degrees of inter-rater reliability. Among the lowest were average force and average wrist posture. It is not clear whether these weak ICCs were due to an inadequacy in the method, or to the narrow range of exposures to these stressors in the jobs evaluated.

Because these jobs were initially selected from a pool of jobs used in a study of repetition, repetition had the widest range of possible exposures in the job sample. It is encouraging that the level of agreement was high over nearly

the entire range of the scale (from approximately 2 to 8 on the 10 cm scale), even though this scale requires the most integrating and judgment by the analysts. Based the definitions of the stressors and rating benchmarks (see Appendix), repetition is the most qualitative of the stressors. For example, ratings of posture are based on observations of joint position relative to maximum range of motion, parameters that could theoretically be measured from the videotape. Repetition, however, requires analysts to subjectively assess movement speed and exertion time, using descriptors such as "slow", "steady", "rapid", and "brief". The high levels of agreement among different raters and groups when assessing repetition suggest that the combination of written guidelines and videotaped benchmark examples used for training adequately promote rating consistency.

In contrast, because of the nature of the jobs available for this study, there were relatively small ranges of average force, contact stress, and wrist posture represented. These other stressors were not specifically controlled in the original job pool, thus their distribution was due merely to chance. As a percent of total variance, inter-rater variability was highest in the ratings of average force, contact stress, and wrist posture. This was due mainly to the relatively low range of ratings given to these stressors. Given the selection of jobs included in this study, there was no expectation that a wide range of exposures to these stressors would be identified.

In some cases, the range of potential exposures in terms of average values is limited by physiological tolerances of workers. For example, it has been shown that it is difficult for workers to maintain average exertion levels above 15% - 20% of maximum for extended periods (Bystrom and Fransson-Hall, 1994). Similarly, certain extreme postures may be difficult to sustain. Given the current definition of the endpoints on the ten cm scales, it is unlikely that a worker would ever receive a rating higher than two (corresponding to 20% of maximum) for average force. In light of the inter-rater variability identified in this study, further investigation is warranted to determine the degree to which the method is able to discriminate among jobs over the widest possible range of average force, contact stress, and posture levels.

In addition to the variability due to differences in perceived exposure by the groups, there is also variability due to the groups choosing to rate different activities viewed on the videotape. Analysts were instructed to rate "regularly occurring work elements", and there was some discrepancy as to how this was

interpreted by the different groups. This is an important issue in analyzing the physical stresses present in a job. Sufficient data currently are not available to determine how frequently an irregular element must occur, or how severe that exposure must be, for it to have a significant impact on worker health. Towards this end, it is recommended that irregular elements (i.e. tasks that occur several times throughout the work day, but are not part of the main job cycle) be analyzed separately. Total exposure ratings can then be calculated as a time weighted average (TWA). Utilizing this process will provide an overall score for the job, while also providing information about the relative contribution of the various activities to the overall stress.

It was hypothesized that using a group consensus technique would provide more consistent results than individual ratings. In nearly all cases, the intraclass correlations were higher for the group final ratings than for the initial, individual ratings. For example, when 4 groups (17 total raters) rated twelve jobs each, intraclass correlations for the initial ratings ranged from 0.10 to 0.79 (median=0.34), while the intraclass correlations of the final group ratings were 0.16 - 0.86 (median = 0.54). Similar increases were seen in the 45 jobs rated by the two groups (Table 3.6) and in the test/retest reliability study (Table 3.7).

From a mathematical standpoint, decreased variability after forcing group consensus may seem to be an obvious expectation, since any method of grouping and averaging the individual ratings would have decreased the variability. In this study, however, raters were forced to adjust their initial ratings in order to reach consensus. Thus, theoretically, the rating groups could have converged at one of the extremes rather than at some central value. The consistency in the group ratings among the four groups indicate that the discussion and adjustment process assists groups in converging at a consistent rating. It is likely that the increased consistency in ratings of groups over individuals is due to several factors, most notably that the group process reduces the chance of any one observer's bias influencing the final ratings, and that having multiple observers better enables a complex task to be captured.

The current consensus method sometimes becomes labored when one person adopts an extreme view. In future applications with this method, it may be desirable to perform the initial ratings, discuss differences, and allow raters to change their ratings but not force consensus. In this case, the mean group rating would be used, and the range would be a measure of rating quality. In the twelve jobs which were rated by 17 raters, there were no statistically

significant differences in the initial and final overall average ratings (Figures 3.1 through 3.4). This suggests that if the goal of an analysis is merely to obtain a group average stress rating, the consensus process may not be necessary. However, because the comparison and discussion of ratings force the raters to share their thought processes and observations of the job, it greatly enhances the overall understanding of the stresses and is preferable to individual ratings.

iii. Accuracy (internal validity) of ratings

The fact that multiple observers interpret a job the same way with regard to the rating criteria does not mean that these interpretations are necessarily accurate. This is illustrated in Figure 6.1. The center of the target can be thought of as the "true", or "correct" value. In case (a), although all raters have similar values (thus high reliability), they are still far from the center. In case, (b), the raters all have similar values and are at the center of the target, indicating both high reliability and high accuracy.

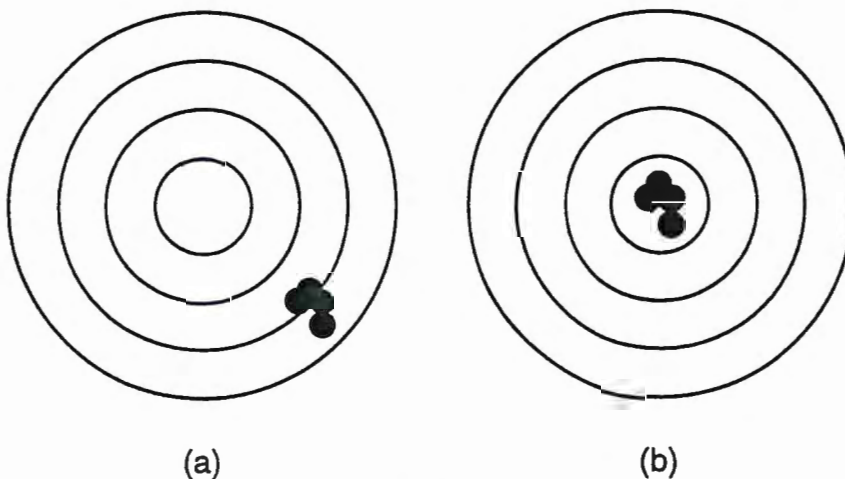


Figure 6.1: Illustration of high reliability but low accuracy (a), and high reliability/high accuracy (b) (adapted from Montgomery, 1996).

For all stressors except peak movement speed (velocity), the instrumental measurements of the stressor were significantly related to the ratings of the stressor. There was a wide range in the amount of variability in these models, however. The model of repetition ratings as a function of movement speed and exertion time was the strongest of the seven parameters evaluated ($r^2 = 0.88$). Relatively strong agreement was also found between the peak and average finger flexor EMG levels and grip forces, and the ratings of

peak and average force ($r^2 = 0.49$ to 0.81). By definition, the peak and average force ratings should correspond to %maximum effort, which in this experiment was approximated by EMG normalized to %MVC. Theoretically, a peak exertion of 10% percent MVC would correspond to a rating of 1, a 20%MVC exertion would correspond to a rating of 2, etc. Ratings of peak force ranged from approximately 3-6 on the ten cm scale, while mean peak muscle activity levels for the same conditions ranged from approximately 15% MVC to 37% MVC (Figure 4.3). Ratings of average force tended to agree more closely with the actual percent effort exerted in an absolute sense, although the model r^2 values were lower. The lower r^2 values were most likely due to the relatively small range in exertion levels (approximately 1%MVC to 11%MVC) (Figure 4.3). For the average EMG and average grip force models, the slope of the relationships were close to 0.1 (0.1 and 0.08, respectively). For both models the intercept was significant, but relatively small (0.56 and 0.58, respectively).

Correlation coefficients for the average force parameters were lower than the peaks, but still relatively high (0.70 and 0.75). The relatively high r^2 values between the ratings and the instrumentally measured force exertion levels indicates that the observers were able detect differences in force exertion levels; however, the absolute agreement between the measured force levels and the ratings was lower. The raters were able to correctly determine the relative magnitude of the exertion (in comparison to the other conditions observed), but these ratings did not directly correspond to the actual percent of maximum effort exerted. The strong linear relationship observed, however, indicates that efforts to further train or calibrate the raters may produce better absolute estimates.

In the study presented in Chapter 4, a high percentage of the variability in the wrist posture ratings was explained by the measured wrist posture ($r^2 = 0.77$ for peak, and 0.89 for average). As in the force ratings, however, the absolute agreement (in terms of estimating the actual percent range of motion) was lower than the raters' ability to consistently detect proportional differences in wrist position.

Due to instrumentation availability, the wrist was the only joint for which posture was analyzed. It is possible that other joints, such as the shoulder, would have higher levels of agreement between the true and observed values. The position of the shoulder may be easier to observe, because of the large size of the two body segments it connects (i.e. torso and upper arm), and

because of the large range of motion. The wrist has a smaller range of motion, especially in radial/ulnar deviation, which may make a given percentage change in angle more difficult to detect. Wrist movements tend to be more highly dynamic than shoulder movements, as well. Not only can the speed make postures more difficult to discern, but the highly dynamic nature of wrist movements often makes it impossible to maintain an orthogonal view of the joint. In addition, the wrist may become obscured by work objects and equipment.

The observational ratings of peak movement speed were not statistically related to the calculated peak movement speed. Interestingly, peak velocity was also not related to transfer frequency (Table 4.1). Although there was no statistically significant relationship between frequency and movement speed, the general trend in the data is opposite what might be expected: the shorter cycle time tasks have generally lower peak velocities than the slower cycle time conditions. A visual analysis of the data indicates that this is due to the movement patterns during the task. In the slower condition (4.8 s cycle), the subjects had time to relax their hands between transfers. Plots of the goniometer data from this condition show periods of a few seconds where the wrist is nearly static, with brief, quick (i.e. high velocity) movements when the container is picked up. During the 1.2 s cycle conditions, the subject has no time to relax between transfers. The subject's wrist steadily oscillates as it completes each transfer cycle. In this study observers were not able to discern these differences in peak velocities. The peak velocities recorded during this task did not represent a wide range of exposures, as indicated by other investigators (e.g. Marras and Schoenmarklin 1993), and representation of a wider range may improve the agreement.

iv. relationship to health outcomes in workers (external validity)

Finally, in this study, repetition was found to be associated with worker-reported discomfort in the upper limb, tendinitis, and symptoms consistent with CTS as reported on a hand diagram. A similar trend was apparent in the model combining hand diagram scores indicating symptoms consistent with CTS and median mononeuropathy (0.5 ms threshold), in which repetition was close to achieving statistical significance ($p < 0.08$ and $p < 0.06$ for logistic model and linear trend, respectively). Figure 6.2 illustrates the exposure-response

relationship between level of hand activity and the five health outcomes presented in this dissertation.

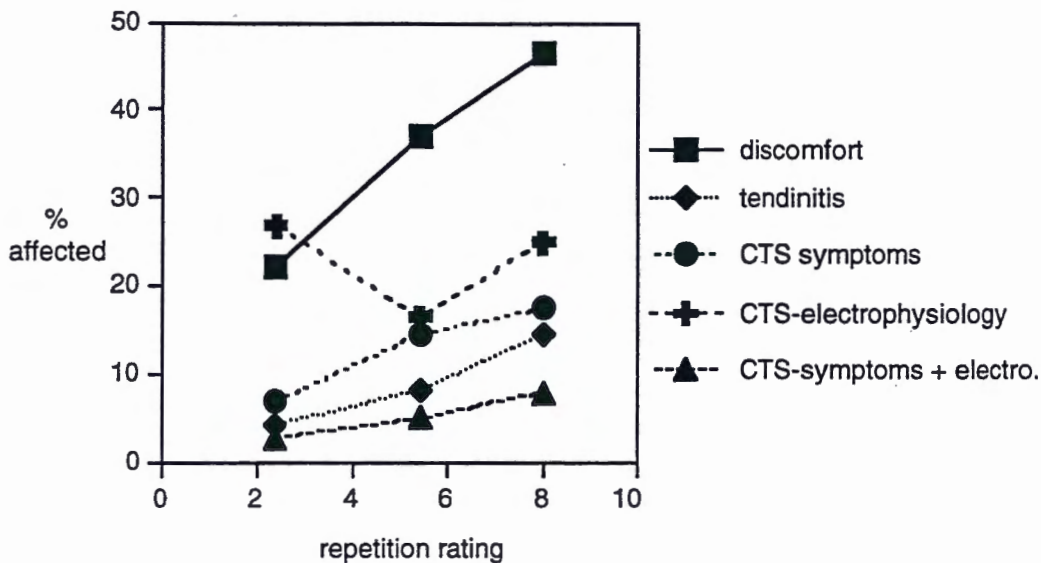


Figure 6.2: Prevalence of five WRMSDs by repetition/hand activity category.

For the health outcome parameters where a statistically significant association was found, the odds ratios for repetition in this study were similar to those determined in previous studies linking repetitive work to adverse health outcomes in workers. In this study, the increased odds of discomfort associated with an increase of 1 unit on the 10 point repetition scale was approximately 16% (Table 5.8). Extending this to the average ratings of jobs in the high versus low repetition categories results in an odds ratio of 2.45. Symptoms consistent with carpal tunnel syndrome as identified on the hand diagram had an odds ratio of 2.32 for high vs. low repetition in the current study.

In the current study, no statistically significant association was found between repetitive work and diagnosis of CTS when positive electrophysiologic findings were the only diagnostic criterion considered. Twenty-four percent of all the workers in the current study had abnormal electrophysiologic test results (Table 5.9). This rate is comparable to rates determined by other investigators among working populations (e.g. Franzblau et al., 1993; Barnhart et al., 1991; and Nathan et al., 1988).

Further analyses were performed using slight variations in the definitions of both exposure and health outcome criteria in order to test the robustness of

the associations found. There was a high level of similarity between the models when alternative diagnostic parameters are substituted, indicating that the findings are robust to alternative definitions of health outcomes.

No other ergonomic stresses (e.g. force, posture) were found to be associated with any health outcome measure. This is not surprising, considering the range of exposures to these other stressors which were encountered. Because the job selection was stratified only on repetition, no effort was made to insure wide representation of the other stressors. In general, most of the exposure levels for the other stressors fell within a 2 or 3 point range on the ten point scale (Table 5.5). It is likely that this is not enough variability in exposure to produce significant results.

v. Advantages of method

This observational system has some advantages over traditional methods for assessing stressors in manual work. Distinct advantages of this method are:

- application time is relatively short. A single stressor can be rated in a matter of minutes. Rating all 52 stressors currently included in the system can be completed in 30-45 minutes (after the job has been videotaped and documented), including the group consensus process.
- no instrumentation is required.
- team members can be trained to rate all the stressors in a few hours.
- the system can be applied by a team using videotape and written job documentation, or by a single investigator on site during a walk-through survey (although a lower level of precision would be expected with individual raters).

There is a need for a consistent method of exposure assessment which is applicable to a wide variety of jobs and can be applied by different people. The ultimate goal of epidemiological research is to provide guidelines for safe exposure limits, in order to minimize workers' risk of WRMSDs while maximizing productivity. In order for this goal to be reached, it is necessary to utilize exposure assessment techniques that are applicable in a wide variety of job settings and which will eventually allow the generalization from these studies to the working world in general. The method described and evaluated in this dissertation is one such possible technique.

Recommendations for future work

i. Enhancement of decision criteria.

The results of the studies presented in this dissertation indicate that the method for rating repetition can be consistently applied by different users, and demonstrated internal and external validity. The ratings for the other stressors (e.g. force, posture, contact stress) exhibited lower levels of both reliability and validity. More extensive development of the rating aids (i.e. decision criteria and videotaped benchmark examples) may provide raters with a more precise framework on which to base their ratings.

The results of the study in Chapter 4 (in which ratings were compared with direct measurements) indicated that raters were able to detect incremental changes in force and wrist posture, but were not able to accurately estimate the magnitude of those exposures in an absolute sense. Further studies should examine whether training raters to adjust for these systematic errors may reduce their occurrence. The work presented here is the first attempt at validation of these scales. The current benchmarks and written guidelines have been developed from the observation of many jobs, and are based on the agreement of a group of experts. There was no corroboration of these ratings with the "true" or "correct" values of the parameters of interest. Because direct measurement techniques exist for force, posture, movement speed, and localized mechanical stress in some situations (Table 1.2) it is possible to collect both observational and instrumental data on these parameters. An iterative process similar to that outlined for the development of the repetition guidelines (Figure 2.2) could be undertaken to refine the written decision criteria and videotape benchmarks. Figure 6.3 illustrates the proposed process for refining the rating system to better reflect the "true" values of those parameters that can be measured.

ii. Further comparisons between ratings and direct measurements

The work presented in Chapter 4 (comparison of ratings and instrumental methods) utilized a laboratory task. This task was a short cycle, highly stereotyped activity. Additional studies which instrumental data and

observational ratings are compared in industrial jobs, which have a wide variety of activities, are warranted.

iii. Evaluation of training requirements

Further evaluation of the training requirements of the method are also warranted. The raters who were studied in Chapter 3 were all experienced ergonomists and received less than 3 hours of training on the complete method. Since there are currently 52 scales which comprise the complete method, this is an average of approximately 3 minutes per scale. The training consisted of reviewing the written decision criteria and videotaped benchmark examples. Each group then rated one complete job, and received feedback. No feedback was given after the initial job. It is likely that improved training and increased feedback as the new raters analyze complete jobs would be beneficial.

iv. Investigation of effects of analyst experience

Experienced analysts were chosen as a starting point for evaluation because they represent a best case situation. Consistent expert ratings justify future studies of non-experts. Because it is recognized that it is not always possible or feasible to have raters of the same level of expertise as those in this study, an important future consideration is the level of background knowledge necessary to adequately apply the decision criteria. A tool such as this is potentially useful to a wide variety of users, including employee-based plant ergonomic teams. Trials with users of this type of background are necessary. It is necessary to determine the amount of method-specific training that is required for analysts (both expert and non-expert) to be able to effectively apply this technique. In addition, an evaluation of the "learning curve" by examining the length of time to rate and achieve consensus and the amount of variability in initial ratings, should be performed.

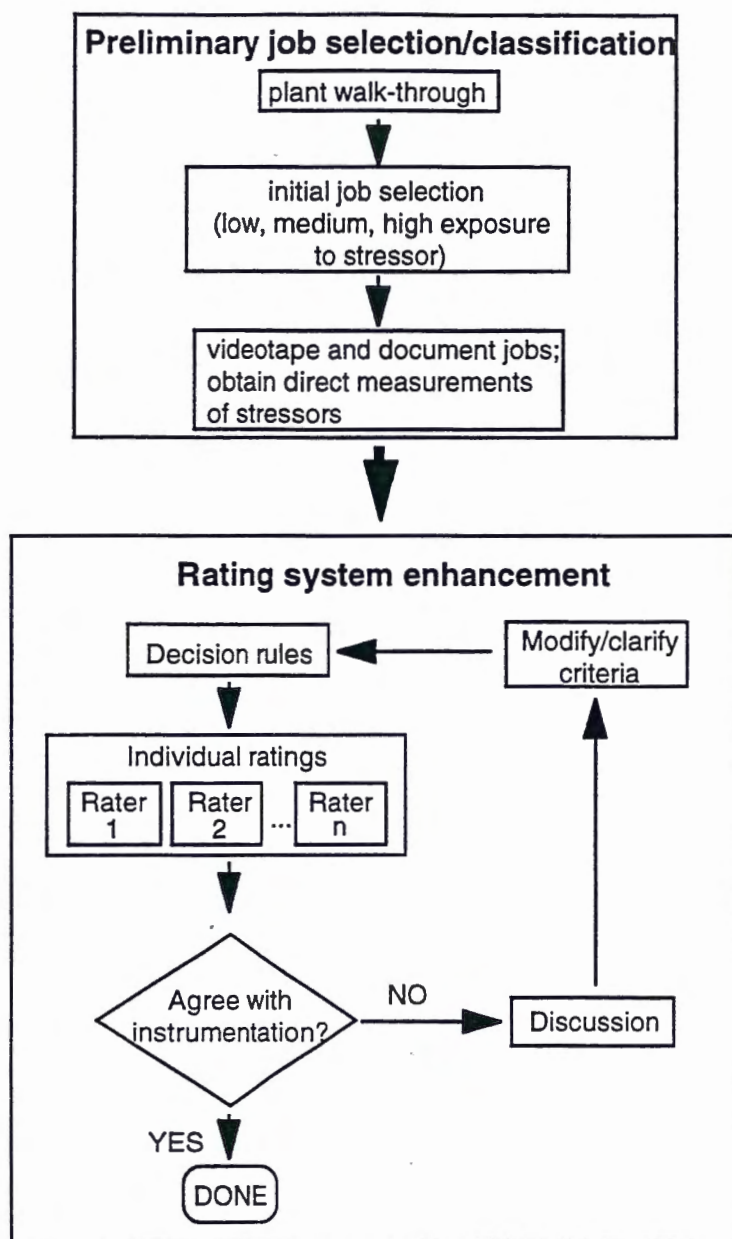


Figure 6.3: Proposed process for enhancing rating guidelines for force, posture, etc.

v. Evaluation of wider range of other stressors.

The ratings of repetition had the highest level of agreement between raters (i.e. the strongest ICC and the highest test/retest correlation), compared to the other stressors rated. Repetition was also the most broadly represented stressor in the jobs studied. The other stressors occurred over a relatively

narrow range of potential exposures. It is unclear whether the relatively low agreement for the other stressors is due to this narrow range of exposures or to inadequacies in the method.

vi. Use of system in real-time

Another consideration is the ability of the system to be used in real-time, on the plant floor. The analysts in this study viewed all jobs on videotape; it was possible to view the same segments multiple times, and tasks could be viewed in slow-motion or freeze-frame. Further examination is needed to determine the level of increased variability that may be introduced by performing the ratings in real-time. Although this would reduce the ability to view the same segment of the job repeatedly, it would allow the job to be viewed from multiple angles and allow multiple workers to be observed.

vii. Establishment of exposure-response relationship.

One of the original goals in the development of this method was the establishment of an exposure assessment tool to be used in epidemiological research. As previously stated, the ultimate goal in epidemiological research is to provide guidelines for safe exposure limits, in order to minimize workers' risk of WRMSDs while maintaining necessary levels of production. The work presented in Chapter 5 provides strong evidence of this relationship with regards to hand activity. In order to more fully determine these relationships, more epidemiological studies are needed that quantitatively compare workers' exposure with their health status, for both individual stressors and combinations of stressors.

viii. Determining the relative importance of "peak" vs. "average" exposures.

Most jobs, even the most cyclical, have occasional irregular elements that introduce different stresses. For example, an assembly line worker with a low force, short cycle time job may have high force exertions several times during the work shift to replenish stock. Currently, the relationship between the frequency with which a stressor occurs and health outcomes in workers is not understood. It is important, not only for this method but for any job analysis technique, to develop criteria for determining how to address task variability in the analysis procedure.

ix. Streamlining of method

The complete method currently consists of 52 scales, encompassing seven major stressor categories and seven body parts. These stressors are not independent of each other. For example, movement speed and exertion time are rated for every joint of the upper limb; however, the observed exposures are often the same for each joint; correlations for ratings of exertion time between the various joints are generally above 0.90. Similarly, some stressors in the current method are redundant with others. For example, repetition/hand activity is defined as a function of movement speed and exertion time. It may be possible to eliminate these redundant scales without losing power of the method. As more knowledge is gained of the relative importance of these stressors and interaction between them, it may be possible to remove several scales from the method while preserving the amount of useful information obtained.

References

- Barnhart, S.; Demers, P. A.; Miller, M.; Longstreth, W. T.; & Rosenstock, L. (1991). Carpal Tunnel Syndrome among ski manufacturing workers. *Scandinavian Journal of Work, Environment, and Health*, 17, 46-52.
- Bystrom, S. & Fransson-Hall, C. (1994). Acceptability of intermittent handgrip contractions based on physiological response. *Human Factors*, 36(1), 158-171.
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York: John Wiley and Sons.
- Franzblau, A.; Werner, R.; Valle, J.; & Johnston, E. (1993). Workplace surveillance for Carpal Tunnel Syndrome: A comparison of methods. *Journal of Occupational Rehabilitation*, 3(1), 1-13.
- Marras, W. S. & Schoenmarklin, R. W. (1993). Wrist motions in industry. *Ergonomics*, 36(4), 341-351.
- Montgomery, D. C. (1996). *Introduction to Statistical Quality Control*. (pp. 455-461). New York: John Wiley and Sons, Inc.
- Nathan, P.A.; Meadows, K.D.; and Doyle, L.S. (1988). Occupation as a risk factor for impaired sensory conduction of the median nerve at the carpal tunnel. *Journal of Hand Surgery*, 13B(2), 167-70.

APPENDICES

APPENDIX A

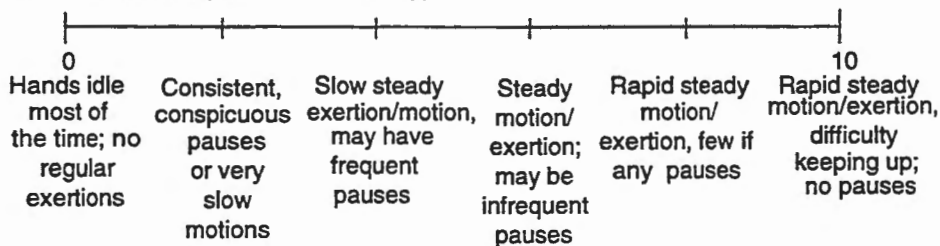
JOB ANALYSIS RATING SCALES

AND GUIDELINES

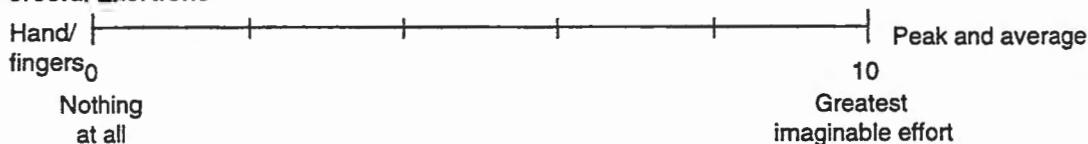
Plant _____ Date: _____ Rater: _____

_____ Job _____ work standard _____ hand rated _____

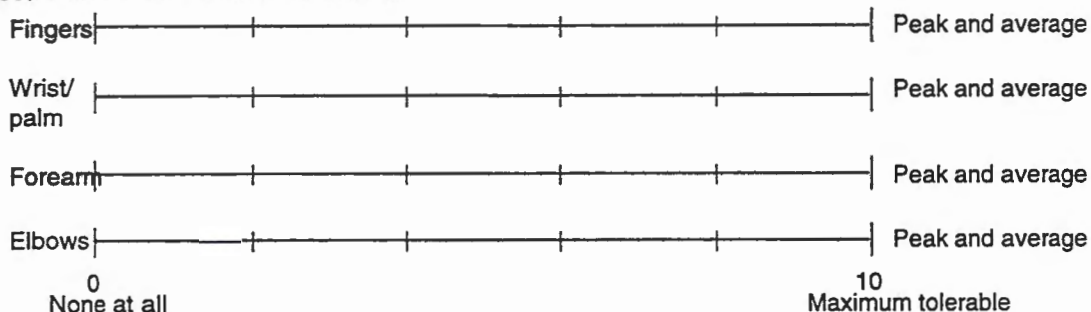
Repetitive Exertions (hand/wrist activity)



Forceful Exertions

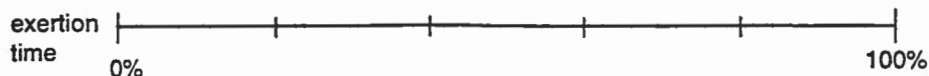
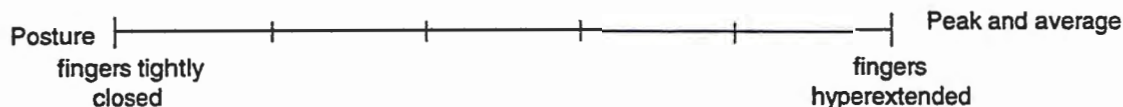


Localized contact stress: force/area



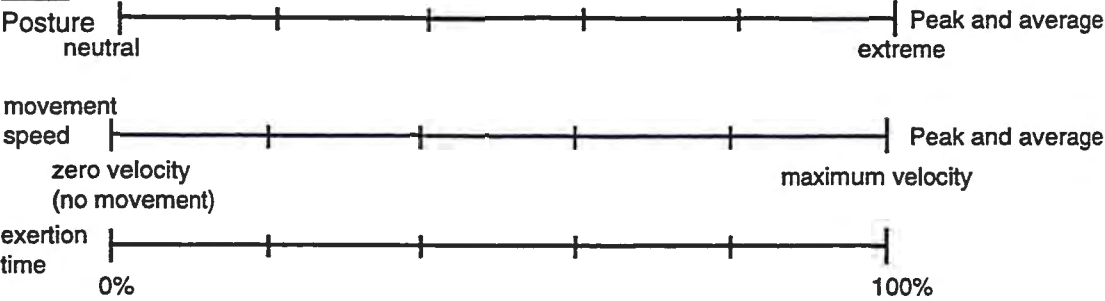
Posture/movement speed/exertion time

Hand/fingers:

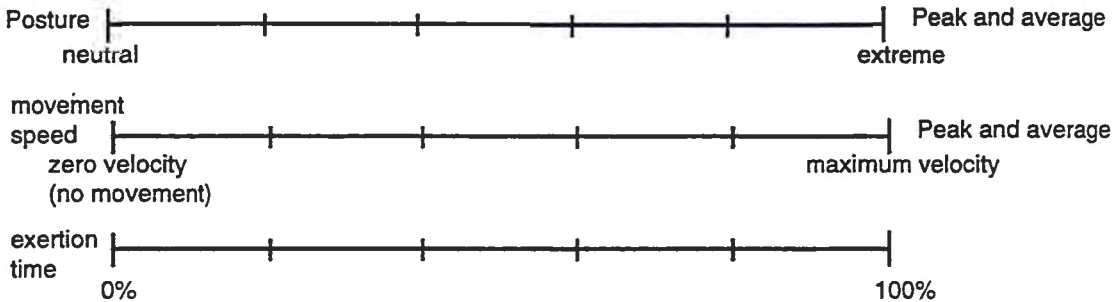


Posture/movement speed/exertion time—cont'd

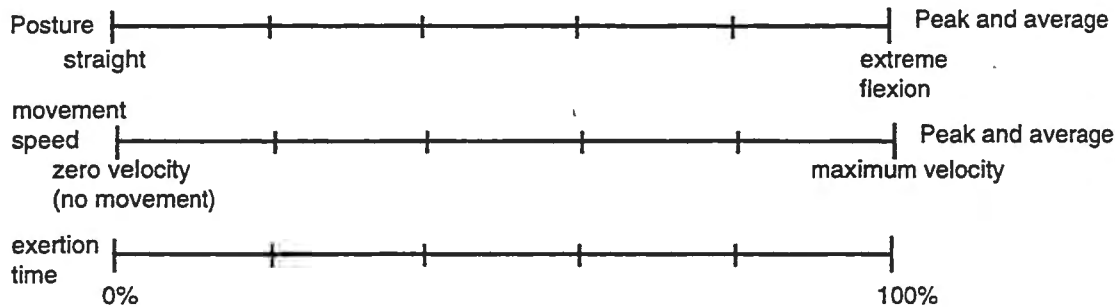
Wrist:



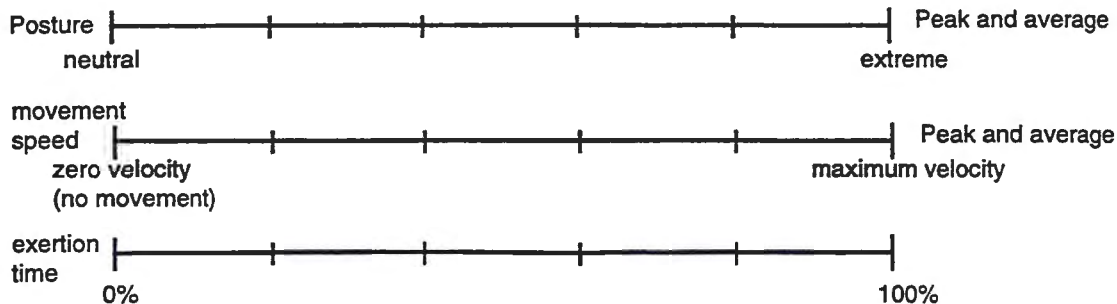
Forearm:



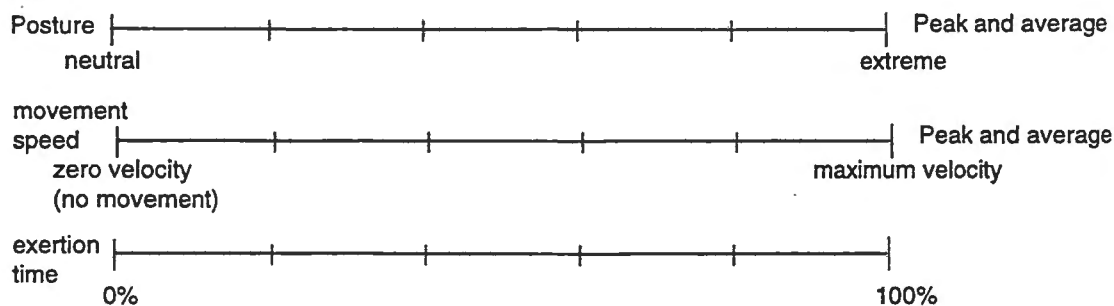
Elbow:



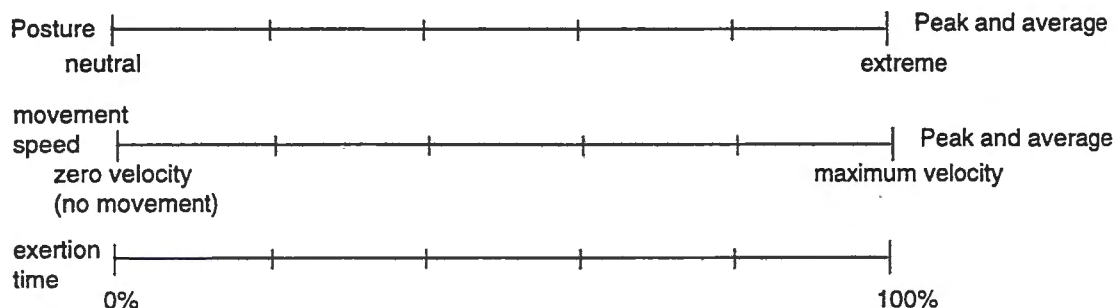
Shoulder:



Neck:



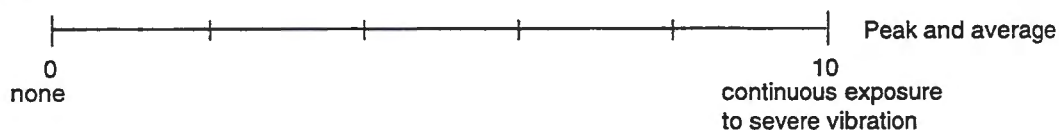
Back:



Exposure to low temp.



vibration



jerk



Guidelines for rating work-related factors of upper limb disorders

As a supplement to other methods of job documentation and analysis, the following rating categories can be used to assess physical stressors. Several cycles of the job should be observed and the various categories of risk factors rated according to the following guidelines.

General guidelines:

1. Identify and rate busiest hand in the operation. Be careful to only focus on the busiest hand when making ratings.

Repetitive Exertion Ratings (hand/wrist activity):

1. Consider recovery time within and between cycles
2. If cycle includes, for example, bursts of rapid steady motion with frequent pauses the rating should be appropriately weighted

NOTE: Repetition rating may be reduced if work is interrupted with idle time or slower tasks. In such cases, ratings for hand activity are a time-weighted average, taking into account the speed of the motions that do occur, discounted based on the percent of the cycle that is idle.

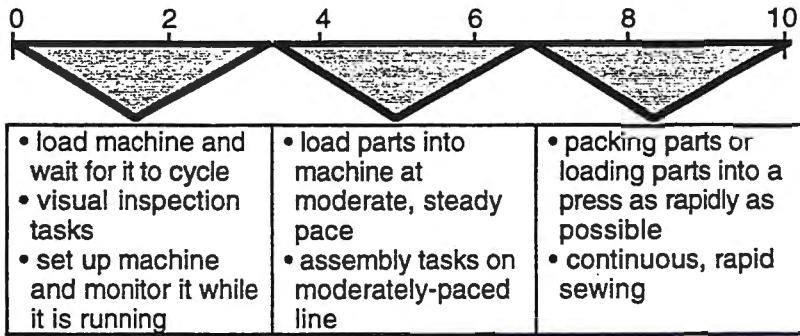
e.g. a grocery checkout job with a rating of 8 while the worker's hands are busy may be reduced to as low as 4 if half (50%) of the cycle time is spent waiting for the customer to pay, for credit card approval, etc. ($8 * 0.5 = 4$)

HINTS:

- Consider how fast the worker *is* moving vs. how fast he/she *could be* moving
- If motions are *very* different during various parts of the cycle, discount the rating.

- | | |
|----|--|
| 0 | = hands idle most of the time; no regular exertions |
| 2 | = short periods of activity/exertion separated by consistent, conspicuous, <i>long</i> pauses
<i>OR</i> continuous, very slow motions |
| 4 | = slow steady motion or exertion; may be frequent brief pauses |
| 6 | = moderate steady motion or exertion; may be infrequent, brief pauses |
| 8 | = rapid steady motion or exertion, no regular pauses |
| 10 | = rapid steady motion or exertion, difficulty keeping up; no pauses |

EXAMPLES:



Forceful Exertion Ratings (hand):

NOTE: force ratings should take into account the posture of the hand during the force exertion (e.g. pinch or power grip). The exertion of a given amount of force will receive a higher rating if the grasp is a pinch instead of a power grip.

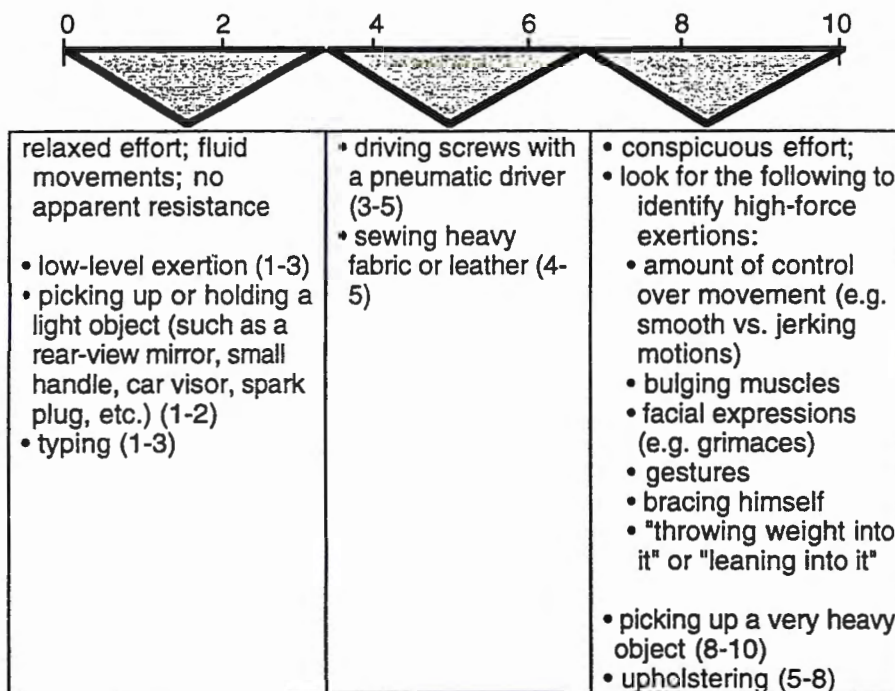
Peak force:

1. Rating should represent the *peak* hand exertion.
2. Rating is **not** dependent on length or frequency of exertions.

Factors to consider in rating force:

- Is force exerted by all fingers, one finger, palm, etc.? A given level of force (e.g. 1 kg.) will receive a higher rating if it is being exerted by a single finger than if it is being exerted with all fingers.
- Load (weight of tools/work objects, resistances, reaction forces)
- Friction (handles, gloves, finger tape)
- Balance (hand position vs. center of gravity of tool/object)
- Torque (shape of tool, length of handles, reaction bars, etc.)
- Posture (pinch vs. power grip, handle design)
- Gloves (hand posture, fit, material)
- Quality control (fit of parts)

Examples:



Note: Force ratings may be adjusted much higher than the guideline score depending on what the worker does with the object (e.g. a worker "holding a light object" may be rated in the "high" range if the worker must press the object into a fixture with a lot of force)

Average Force:

1. Rating should reflect the average amount of force exerted over the entire cycle
2. Rating is dependent on length, frequency, and magnitude of exertions.
3. Average force ratings generally can not be higher than 2, because workers can't sustain a higher average level of exertion for an entire day.

Contact Stress Ratings (force/area):

1. Mechanical stresses are produced as a result of contact between the body and hard or sharp work objects, tools, etc. Stress intensifies with increasing contact force and decreasing surface area.
2. Average rating is based on duration and frequency of contact as well as intensity of stress.
3. Ratings are reduced if gloves or other protective coverings are worn which partially protect the point of contact.
4. Guideline values may be adjusted up or down based on the characteristics of the trigger, handle, etc. (e.g. rounded or sharp edges)

Peak contact stress:

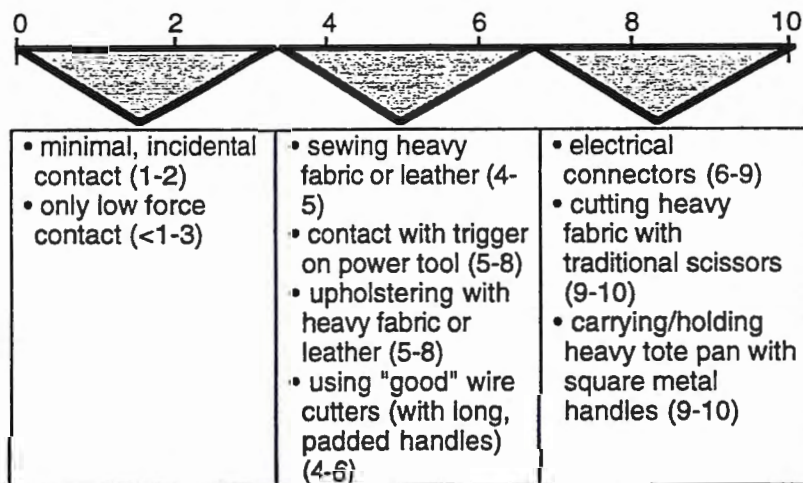
- 0 = no observed contact
10 = worst tolerable contact stress

Average contact stress:

- 0 = no observed contact
10 = worst tolerable mechanical stress 100% of time

Fingers

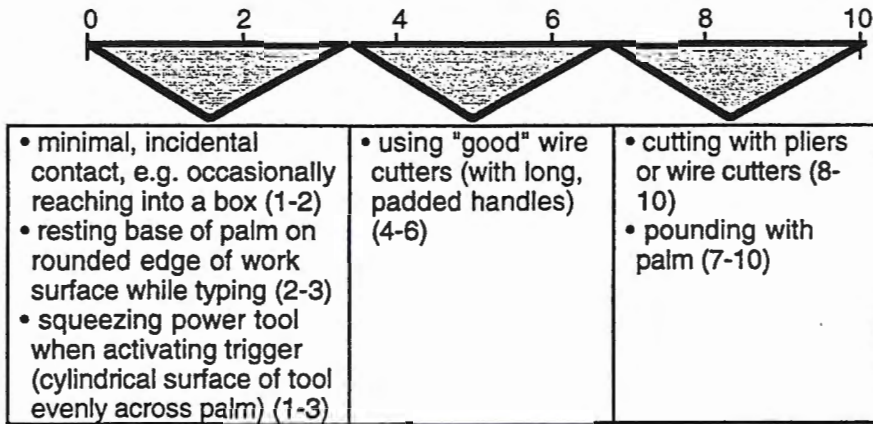
Peak: (Examples)



- notes:**
- 1) these ratings may be adjusted based on how much force is being exerted.
 - 2) discount ratings if gloves are being used
 - 3) nature of surface must be considered (e.g. sharp or rounded edges)

Wrist/palm:

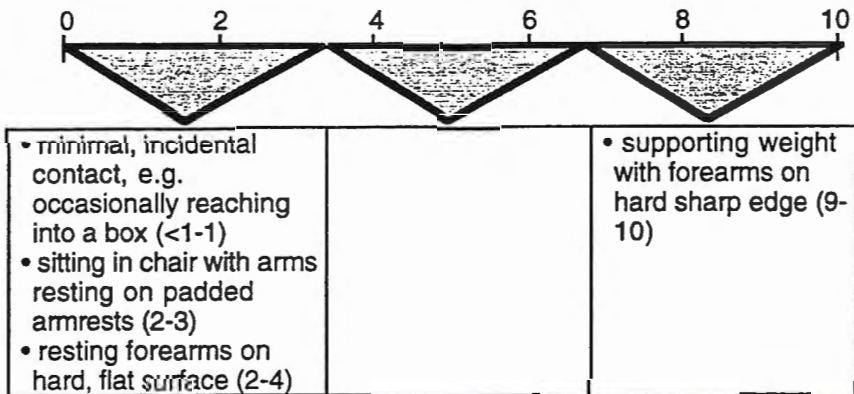
Peak: (Examples)



- notes:**
- 1) these ratings may be adjusted based on how much force is being exerted.
 - 2) discount ratings if gloves are being used
 - 3) nature of surface must be considered (e.g. sharp or rounded edges)

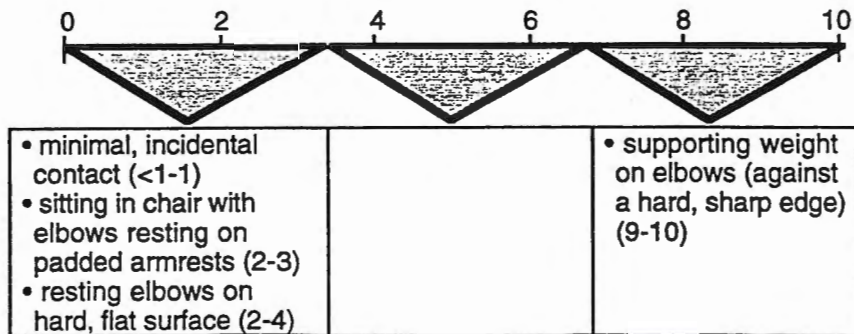
Forearm

Peak: (Examples)



Elbows

Peak: (Examples)



Posture/movement speed/exertion time ratings:

Posture stress--Peak

1. Worst posture observed during the regular cycle.
2. Rating is **not** dependent on duration or frequency of postures.
3. *Irregular tasks* --if the rater is aware of irregular tasks (e.g. replenishing stock, machine setup, cleanup, etc.) that occur as a *normal* part of the workday (as opposed to something that only happens once a week/month/year/etc.), include those tasks in the ratings.

- 0 = only neutral posture observed
1 0 = extreme posture (to the limits of the range of motion) observed

Posture stress--average:

1. Average rating takes into account both severity (relative to total range of motion) and duration of postures during the cycle

- 0 = best possible posture all the time, but not constrained in that posture
1 =
2 =
3 = "bad" posture occasionally during the cycle
4 =
5 = "bad" posture approximately half of cycle
6 =
7 = extremely "bad" posture half of the cycle *or* "bad" posture most of the cycle
8 =
9 =
1 0 = worst possible posture 100% of the time

(Note: ratings in the middle range increase if extremely severe postures are observed)

Motion (Movement speed) (peak):

1. Maximum velocity at which body part is moving
2. Rating does not consider the magnitude of displacement (i.e. how "big" the motion is)
3. Rating does not consider how often the motion occurs

- 0 = zero velocity (no movement)
1 =
2 =
3 =
4 =
5 = body part in moderate motion
6 =
7 =
8 =
9 =
1 0 = maximum possible velocity

Motion (Movement speed) (average):

1. Average velocity at which body part is moving during entire cycle (integrated score taking into account periods of both slow and fast motions)
2. rating does not consider the magnitude of displacement (i.e. how "big" the motion is)

0	= zero velocity (no movement)
1	=
2	=
3	=
4	=
5	= body part in moderate steady motion = body part in rapid motion approximately half the time
6	=
7	=
8	=
9	=
10	= maximum possible velocity

Exertion time:

1. Exertion = body part is exerting effort to produce motion or to hold a position in space.
1. Opposite of recovery time (recovery = time when body part is supported by an external object or hanging relaxed)
2. Rating = percent of cycle spent in exertion

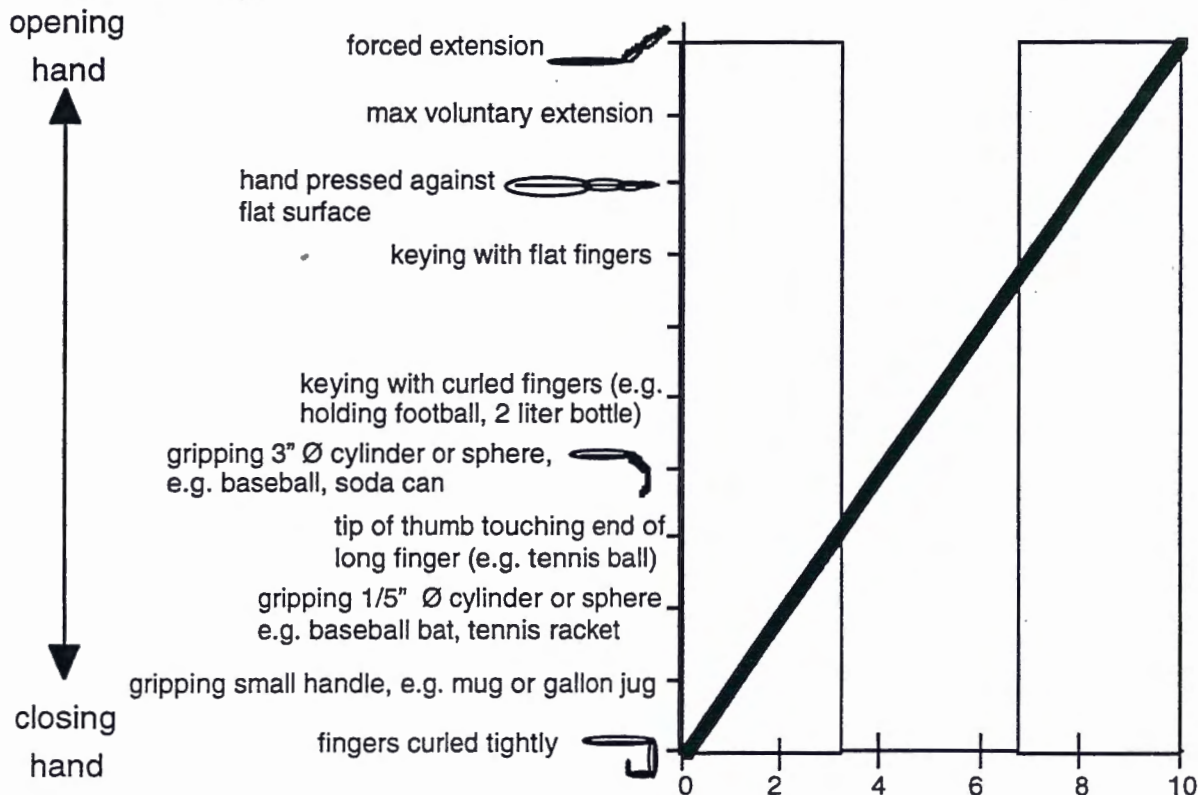
0	= no exertion (e.g. visual inspection tasks)
1	=
2	=
3	=
4	=
5	= exerting effort during 50% of cycle
6	=
7	=
8	=
9	=
10	= constant exertion of effort (no recovery time within cycle or between tasks)

hand/fingers:

(NOTE: hand posture ratings are related only to the position of the fingers, and do *not* consider what part of the fingers may be applying force.)

Posture

Peak: (Examples)



Average:

-Weighted by percent of time in different postures

Motion (movement speed):

(Examples)



- extremely slow, deliberate motions (0-2)
- holding an object (e.g. power tool) continuously with no opportunity to adjust grip other than depressing trigger (avg = 0-2)

- moderate, "normal" movement speed (peak = 5)
- waiting with hands poised over kbd, with intermittent bursts of typing (e.g. travel agent) (avg. = 2-5)
- slowly starting screws (peak = 6)

- rapid typing (peak = 8-10)
- starting loose-fitting screws (peak = 8)

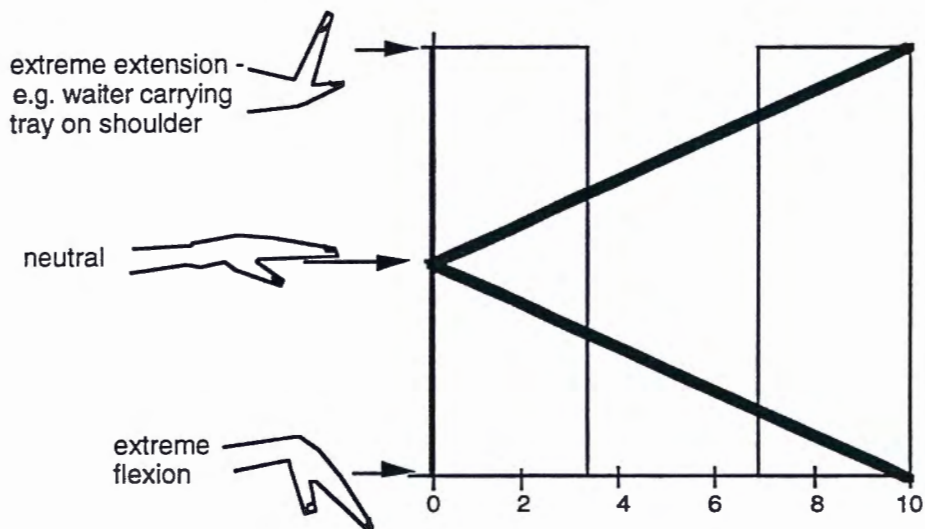
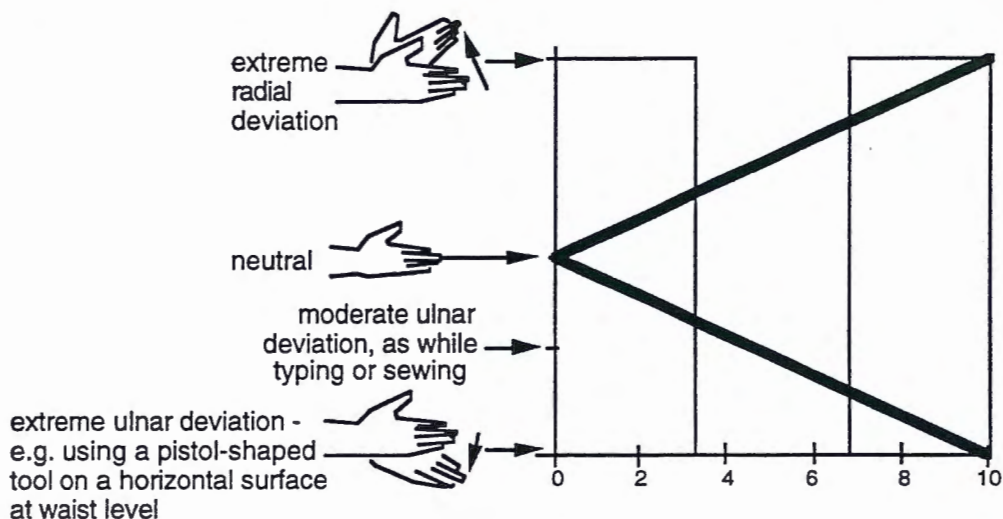
Exertion time:

Note: Don't include totally *voluntary* hand activity as part of the exertion time (e.g. the worker drumming his/her fingers against the table while waiting for the next part should not be included as exertion time)

Wrist:

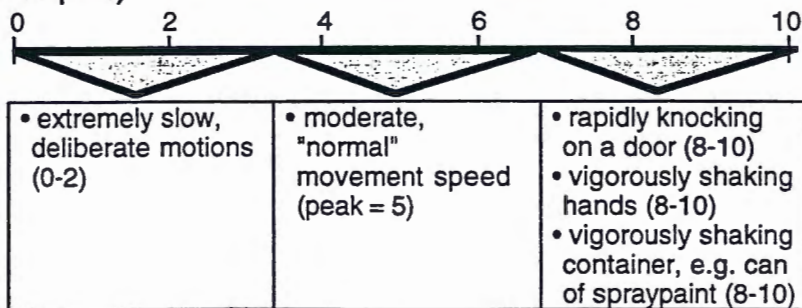
Posture

Peak: (Examples)



Motion (movement speed):

Peak: (Examples)

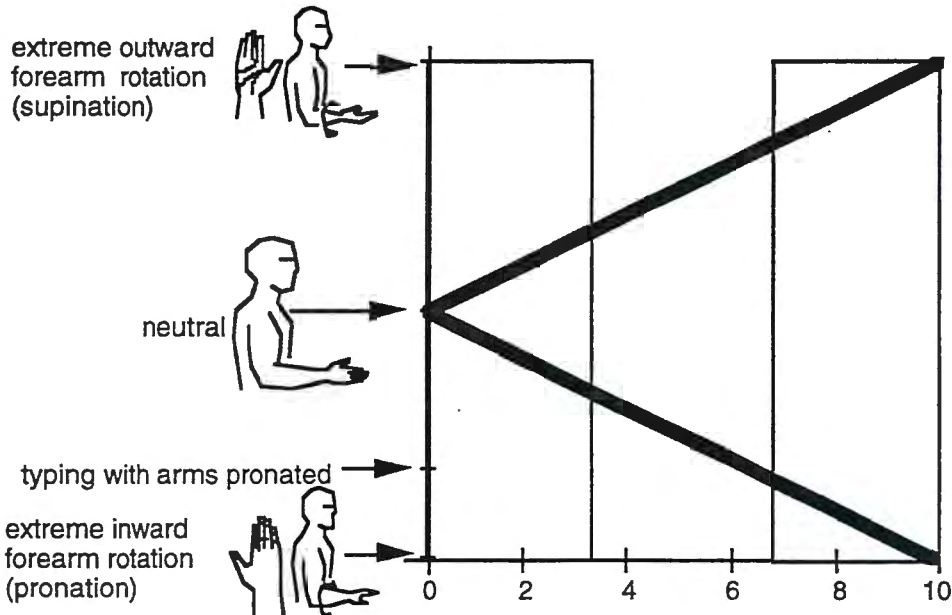


Forearm

Posture

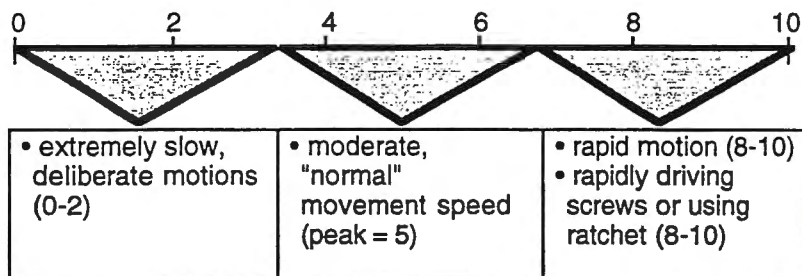
Peak: (Examples)

Look at alignment of long axis of grip with upper arm (i.e. would cylinder being held in power grip be parallel to upper arm (neutral)?)



Motion (movement speed):

Peak: (Examples)



Exertion time:

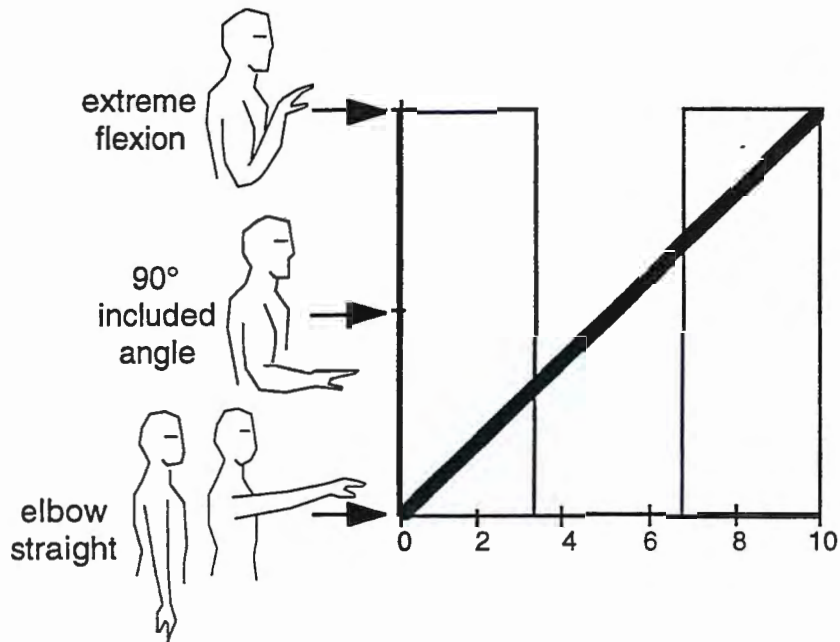
Exertion time is when the forearm is working against a load (either an external object or maintaining position).

- For example, typing with arms pronated would result in forearm exertion if the arms are held in space, but if the palms/wrists are resting on a wristrest, there would be no forearm exertion to maintain that posture.
- Holding a balanced tool upright would not be exertion time for the forearm (in general).

Elbow

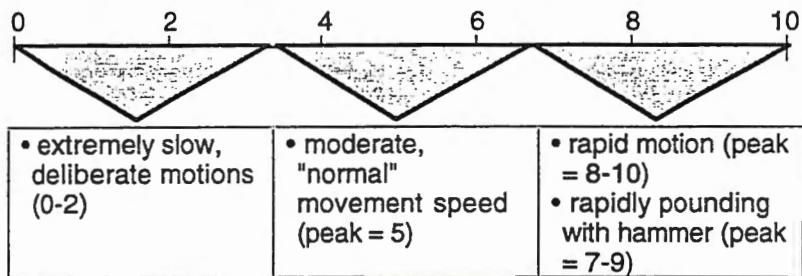
Posture

Peak: (Examples)



Motion (movement speed):

Peak: (Examples)



Exertion time:

Effort is not being exerted at the elbow when:

- 1) the elbow is straight and hanging down at the side, or,
- 2) the elbow is bent but the hands or forearms are fully supported by an external object

0.



Peak: (Examples)



Exertion time:

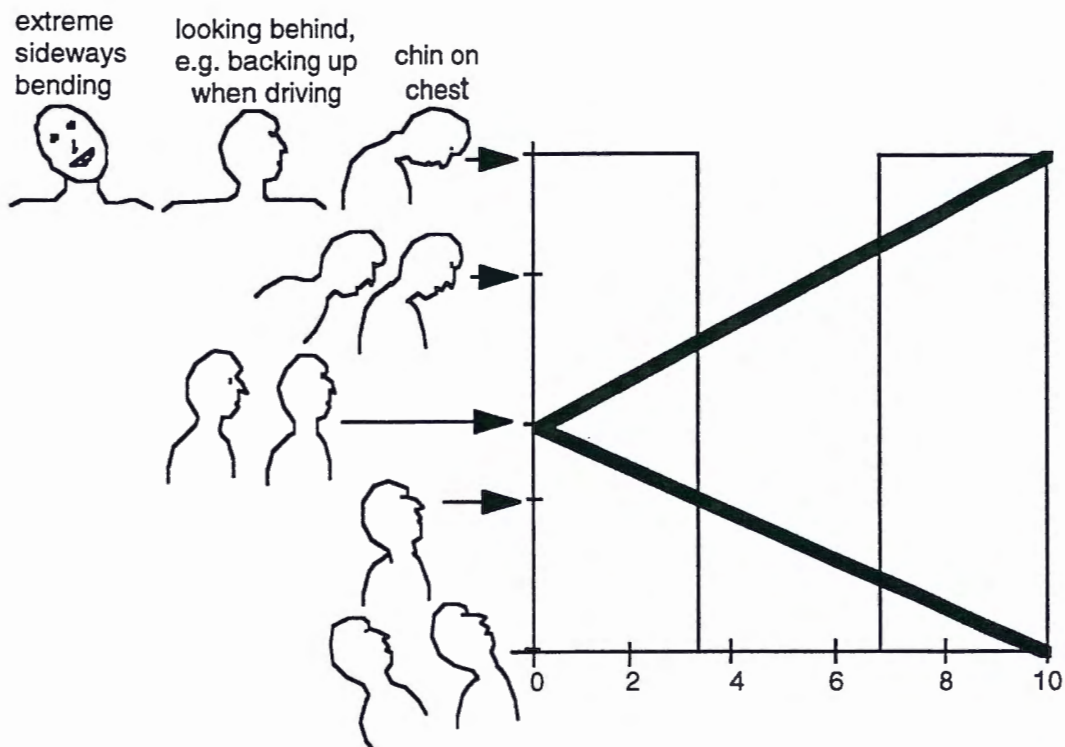
Effort is not being exerted at the shoulder when:

- 1) the arm is against the side of the body, or,
- 2) the arm is slightly extended but the hands, forearms, or elbows are fully supported by an external object

Neck

Posture

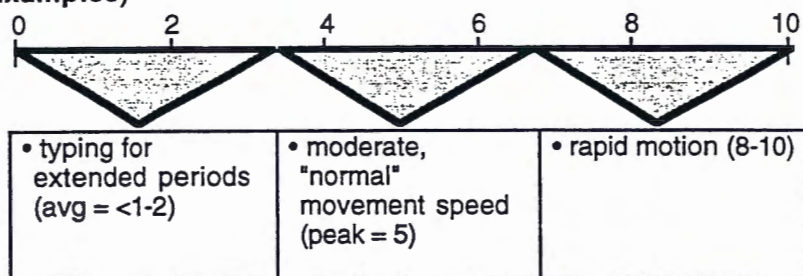
Peak: (Examples)



Note: neck posture scores increase if the torso is not upright and the neck is horizontal. Even if the neck is straight with respect to the back, if back is horizontal to the ground, the neck posture would rate near a 10.

Motion (movement speed):

Peak: (Examples)



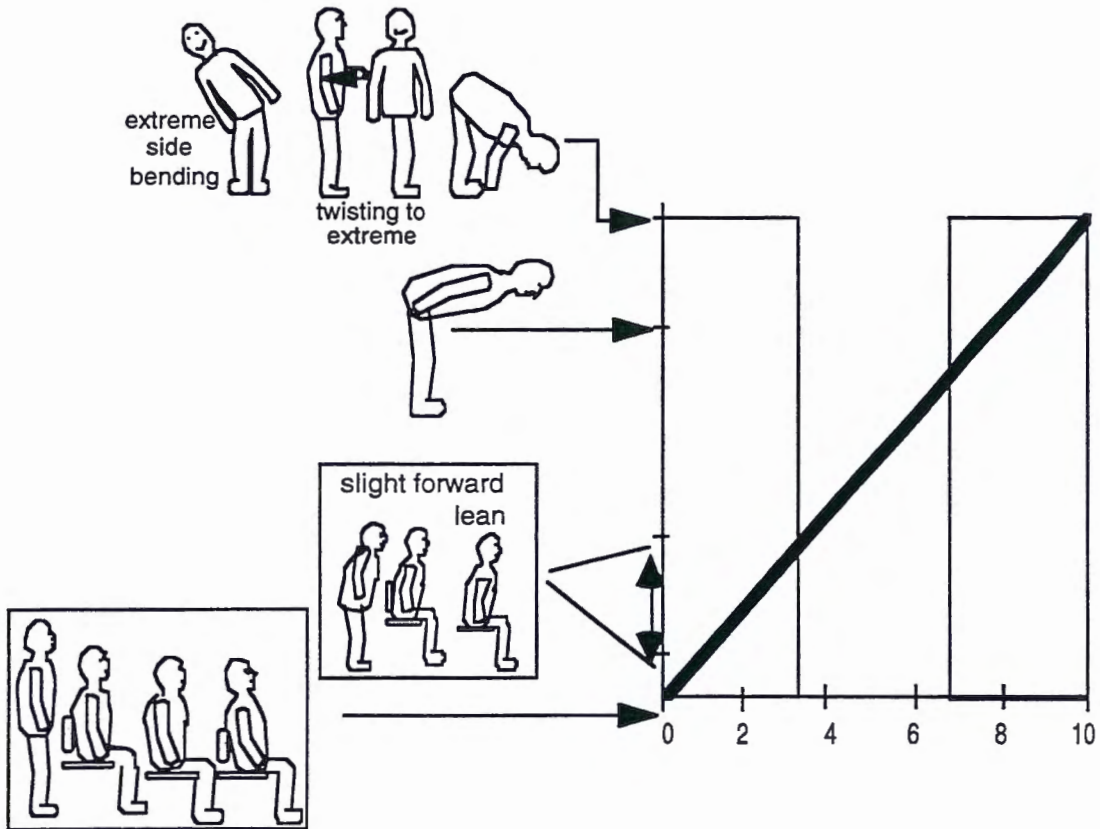
Exertion time:

This rating will be 10 unless the head/neck is supported by an external object during all or part of the cycle (e.g. a headrest, lying down with the head supported).

Back

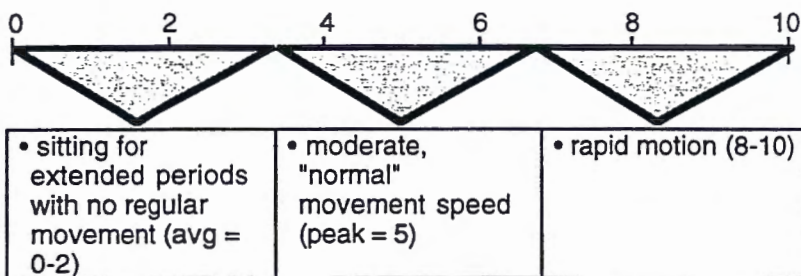
Posture

Peak: (Examples)



Motion (movement speed):

Peak: (Examples)



Exertion time:

1. This rating will be 10 unless the back is fully supported by an external object during all or part of the cycle (e.g. lying down with the back supported, sitting with back relaxed against backrest).
2. Sitting upright in a chair, even with a backrest, *may* be considered exertion time, depending on the degree to which the torso weight is being supported by the backrest.
3. Hint: Could person sleep in that posture without falling over? If not, consider it exertion.

Temperature

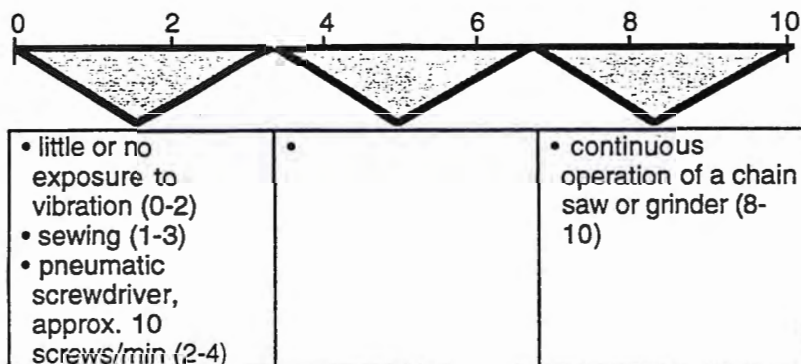
1. Based on duration of exposure and severity
2. Must consider work environment and work object
3. Weighting for thermal conductivity (e.g. wood vs. metal)
4. Consider possibility of cold exhaust from air tools.

0 = contact temperature not below 70° F
 10 = continuous exposure below freezing

Vibration Ratings:

1. Rating is based on intensity of vibration and duration/frequency
2. To help assess intensity, look at how much the worker's hands/arms/body are shaking.

Examples

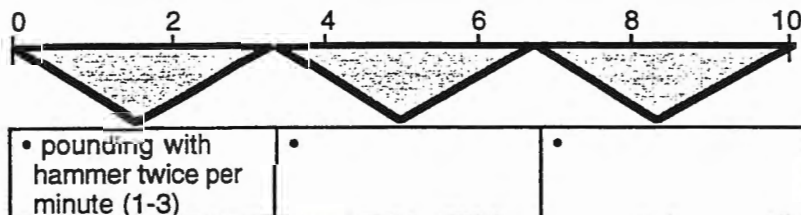


Jerk

(Reaction force)

1. Snap of tool or other jerking motion
2. Function of frequency of occurrence and severity
3. Considerations:
 - power tool torque
 - bolt size
 - hard or soft joint
 - snap when object is broken or cut

Examples



APPENDIX B

VARIANCE COMPONENTS AND INTRACLAS CORRELATIONS FOR ALL SCALES

This section contains the variance component estimates and intraclass correlations for all stressors in each of the data sets presented in Chapter 3. The following scales are not reported due to changes in the rating method over the time during which this study was conducted:

- Peak finger posture
- average finger posture
- peak finger movement speed
- average finger movement speed
- finger exertion time
- peak low temperature
- average low temperature
- peak vibration
- average low temperature
- peak jerk
- average jerk.

Table B.1: Variance components and intra-class correlations (ICC) for initial ratings of all stressors (12 jobs, 17 raters).

stressor	var (jobs)	var (raters)	var (error)	ICC (raters)
repetition	4.1	0.24	0.83	0.79
pk force	0.86	0.24	1.99	0.28
avg force	0.41	0.08	0.69	0.35
peak contact stress-fingers	1.10	0.05	2.10	0.34
avg contact stress-fingers	0.53	0.10	0.73	0.39
peak contact stress-wrist/palm	1.80	0.09	1.84	0.48
avg contact stress-wrist/palm	0.44	0.03	0.41	0.50
peak contact stress-forearm	0.31	0.05	0.67	0.30
avg contact stress-forearm	0.05	0.01	0.07	0.36
peak contact stress-elbow	<0.01	0.02	0.06	0.03
avg contact stress-elbow	<0.01	<0.01	0.01	0.01
pk posture-wrist	0.59	0.08	2.69	0.18
avg posture-wrist	0.13	0.19	0.94	0.10
pk movement speed-wrist	0.69	0.01	1.24	0.35
avg movement speed-wrist	1.38	0.17	1.19	0.50
exertion time-wrist	3.52	0.36	2.55	0.55
pk posture-forearm	0.61	0.43	1.59	0.23
avg posture-forearm	1.17	0.32	1.93	0.34
pk movement speed-forearm	0.69	0.02	1.36	0.33
avg movement speed-forearm	0.97	0.15	1.20	0.42
exertion time-forearm	2.47	0.16	3.84	0.38
pk posture-elbow	0.48	0.12	1.23	0.26
avg posture-elbow	0.39	0.11	0.98	0.26
pk movement speed-elbow	0.52	0.05	0.97	0.34
avg movement speed-elbow	1.08	0.14	1.16	0.46
exertion time-elbow	5.03	0.23	1.70	0.72
pk posture-shoulder	3.05	<0.01	1.59	0.66
avg posture-shoulder	0.58	0.11	0.91	0.36
pk movement speed-shoulder	0.32	0.09	0.70	0.29
avg movement speed-shoulder	0.63	0.27	1.16	0.30
exertion time-shoulder	3.93	0.24	2.73	0.57
pk posture-neck	1.64	0.25	2.57	0.37
avg posture-neck	1.44	0.26	1.36	0.47
pk movement speed-neck	0.86	0.11	1.51	0.35
avg movement speed-neck	0.58	0.12	0.96	0.35
exertion time-neck	0.01	0.08	0.43	0.01
pk posture-back	2.19	0.20	2.26	0.47
avg posture-back	0.31	0.10	0.89	0.24
pk movement speed-back	1.13	0.24	1.60	0.38
avg movement speed-back	0.46	0.11	0.71	0.36
exertion time-back	0.05	<0.01	2.06	0.03

Table B.2: Variance components and intra-class correlations for final ratings of all stressors (12 jobs, 4 groups).

stressor	var (jobs)	var (groups)	var (error)	ICC (groups)
repetition	4.24	0.34	0.35	0.86
pk force	1.42	0.47	0.73	0.54
avg force	0.36	0.16	0.31	0.43
peak contact stress-fingers	2.08	0.09	0.85	0.69
avg contact stress-fingers	0.69	0.05	0.32	0.65
peak contact stress-wrist/palm	2.32	0.08	0.59	0.78
avg contact stress-wrist/palm	0.42	0.01	0.12	0.75
peak contact stress-forearm	0.16	0.07	0.42	0.24
avg contact stress-forearm	0.02	0.01	0.03	0.41
peak contact stress-elbow	<0.01	0.03	0.05	<0.01
avg contact stress-elbow	<0.01	<0.01	<0.01	<0.01
pk posture-wrist	0.88	0.05	1.19	0.41
avg posture-wrist	0.11	0.16	0.38	0.16
pk movement speed-wrist	0.72	0.07	0.42	0.60
avg movement speed-wrist	1.74	0.13	0.35	0.78
exertion time-wrist	4.12	0.26	1.23	0.73
pk posture-forearm	0.50	0.37	0.39	0.39
avg posture-forearm	1.24	0.42	1.03	0.46
pk movement speed-forearm	0.66	<0.01	0.39	0.63
avg movement speed-forearm	1.22	0.19	0.30	0.71
exertion time-forearm	3.36	0.33	1.44	0.66
pk posture-elbow	0.29	0.03	0.45	0.38
avg posture-elbow	0.50	<0.01	0.39	0.56
pk movement speed-elbow	0.47	0.04	0.20	0.66
avg movement speed-elbow	1.49	0.15	0.31	0.77
exertion time-elbow	5.49	0.14	0.85	0.85
pk posture-shoulder	3.19	0.02	0.57	0.84
avg posture-shoulder	0.57	0.07	0.31	0.60
pk movement speed-shoulder	0.32	0.05	0.10	0.68
avg movement speed-shoulder	1.04	0.25	0.46	0.59
exertion time-shoulder	4.46	0.28	1.09	0.76
pk posture-neck	2.60	0.19	1.37	0.62
avg posture-neck	1.65	0.40	0.58	0.63
pk movement speed-neck	0.66	0.14	0.82	0.41
avg movement speed-neck	0.63	0.08	0.37	0.59
exertion time-neck	<0.01	0.04	0.12	<0.01
pk posture-back	2.48	0.56	1.89	0.50
avg posture-back	0.27	0.12	0.36	0.36
pk movement speed-back	0.69	0.31	1.06	0.34
avg movement speed-back	0.47	0.15	0.29	0.51
exertion time-back	<0.01	0.05	0.13	0.02

Table B.3: Variance components and intra-class correlations (ICC) for initial ratings of all stressors (45 jobs, 7 raters).

stressor	var (jobs)	var (raters)	var (error)	ICC (raters)
repetition	6.21	<0.01	0.71	0.90
pk force	1.82	0.10	1.37	0.55
avg force	0.31	0.15	0.56	0.29
peak contact stress-fingers	1.31	0.06	1.47	0.48
avg contact stress-fingers	0.46	0.10	0.44	0.44
peak contact stress-wrist/palm	1.32	0.04	1.49	0.46
avg contact stress-wrist/palm	0.30	0.07	0.41	0.39
peak contact stress-forearm	0.83	0.04	0.85	0.48
avg contact stress-forearm	0.03	<0.01	0.07	0.27
peak contact stress-elbow	0.33	0.02	0.52	0.38
avg contact stress-elbow	<0.01	<0.01	0.01	0.13
pk posture-wrist	0.94	<0.01	1.88	0.33
avg posture-wrist	0.65	0.07	0.67	0.44
pk movement speed-wrist	0.62	0.07	1.16	0.33
avg movement speed-wrist	1.64	0.21	1.16	0.54
exertion time-wrist	5.48	0.16	2.01	0.72
pk posture-forearm	0.52	0.62	1.19	0.22
avg posture-forearm	2.40	0.49	1.54	0.54
pk movement speed-forearm	0.53	0.05	0.96	0.34
avg movement speed-forearm	1.43	0.25	1.01	0.53
exertion time-forearm	5.47	0.08	2.17	0.71
pk posture-elbow	0.59	0.03	0.83	0.41
avg posture-elbow	1.13	0.06	0.75	0.58
pk movement speed-elbow	0.55	0.05	0.71	0.42
avg movement speed-elbow	1.56	0.29	0.96	0.55
exertion time-elbow	6.07	0.11	1.72	0.77
pk posture-shoulder	1.75	0.09	2.08	0.45
avg posture-shoulder	0.59	0.15	0.72	0.41
pk movement speed-shoulder	0.55	0.05	0.77	0.40
avg movement speed-shoulder	1.21	0.33	1.04	0.47
exertion time-shoulder	5.22	0.03	2.59	0.67
pk posture-neck	0.92	0.08	2.96	0.23
avg posture-neck	0.80	0.02	0.95	0.45
pk movement speed-neck	0.26	0.24	1.07	0.17
avg movement speed-neck	0.44	0.17	0.67	0.34
exertion time-neck	0.05	0.18	1.03	0.04
pk posture-back	2.66	0.10	1.69	0.60
avg posture-back	0.26	0.05	0.44	0.35
pk movement speed-back	0.41	0.13	1.16	0.24
avg movement speed-back	0.37	0.15	0.68	0.31
exertion time-back	0.11	0.21	1.70	0.05

Table B.4: Variance components and intra-class correlations for final ratings of all stressors (45 jobs, 2 groups).

stressor	var (jobs)	var (groups)	var (error)	ICC (groups)
repetition	4.02	<0.01	0.69	0.94
pk force	2.35	0.04	0.62	0.79
avg force	0.18	0.35	0.38	0.21
peak contact stress-fingers	1.99	<0.01	0.79	0.73
avg contact stress-fingers	0.49	0.19	0.24	0.55
peak contact stress-wrist/palm	1.67	<0.01	0.83	0.67
avg contact stress-wrist/palm	0.25	0.12	0.28	0.38
peak contact stress-forearm	0.93	<0.01	0.59	0.61
avg contact stress-forearm	0.02	<0.01	0.05	0.27
peak contact stress-elbow	0.80	<0.01	0.50	0.61
avg contact stress-elbow	<0.01	<0.01	0.01	0.36
pk posture-wrist	1.07	<0.01	0.86	0.57
avg posture-wrist	0.69	0.08	0.24	0.69
pk movement speed-wrist	0.59	0.05	0.51	0.51
avg movement speed-wrist	1.70	0.33	0.47	0.68
exertion time-wrist	6.17	<0.01	0.77	0.89
pk posture-forearm	0.47	0.89	0.25	0.29
avg posture-forearm	2.84	0.71	0.54	0.70
pk movement speed-forearm	0.40	<0.01	0.58	0.41
avg movement speed-forearm	1.54	0.31	0.46	0.67
exertion time-forearm	6.46	0.03	0.94	0.87
pk posture-elbow	0.61	<0.01	0.37	0.62
avg posture-elbow	0.95	0.05	0.45	0.66
pk movement speed-elbow	0.47	0.03	0.24	0.64
avg movement speed-elbow	1.56	0.51	0.50	0.61
exertion time-elbow	6.80	0.01	0.68	0.91
pk posture-shoulder	2.16	0.32	0.58	0.71
avg posture-shoulder	0.67	0.28	0.30	0.54
pk movement speed-shoulder	0.53	0.01	0.13	0.79
avg movement speed-shoulder	1.34	0.56	0.45	0.57
exertion time-shoulder	5.87	0.02	0.89	0.87
pk posture-neck	1.38	0.07	2.29	0.37
avg posture-neck	0.96	<0.01	0.29	0.77
pk movement speed-neck	0.08	0.30	0.81	0.07
avg movement speed-neck	0.44	0.26	0.41	0.40
exertion time-neck	<0.01	0.08	0.03	<0.01
pk posture-back	3.20	0.27	1.08	0.70
avg posture-back	0.28	0.07	0.14	0.57
pk movement speed-back	0.17	0.26	0.83	0.13
avg movement speed-back	0.36	0.25	0.25	0.42
exertion time-back	<0.01	0.26	0.74	<0.01

Table B.5: Job, rater, and day variance for initial ratings in test/retest reliability study.

stressor	var (jobs)	var (rater)	var (day)	var (error)	%var (job)	model r^2
repetition	4.30	0.03	<0.01	0.42	90.3%	0.92
pk force	1.04	0.05	<0.01	0.76	55.9%	0.66
avg force	0.21	0.01	<0.01	0.20	48.8%	0.59
peak contact stress-fingers	2.06	0.03	0.04	1.05	64.8%	0.72
avg contact stress-fingers	0.48	0.03	<0.01	0.15	71.6%	0.81
peak contact stress-wrist/palm	0.84	0.02	<0.01	1.26	39.7%	0.50
avg contact stress-wrist/palm	0.22	<0.01	<0.01	0.24	47.1%	0.55
peak contact stress-forearm	0.07	<0.01	<0.01	0.43	13.4%	0.29
avg contact stress-forearm	<0.01	<0.01	<0.01	0.02	15.0%	0.30
peak contact stress-elbow	0	0	0	0	0	*
avg contact stress-elbow	0	0	0	0	0	*
pk posture-wrist	0.66	0.01	0.07	1.87	25.3%	0.40
avg posture-wrist	0.14	<0.01	<0.01	0.44	23.3%	0.36
pk movement speed-wrist	0.52	0.07	<0.01	0.83	36.5%	0.51
avg movement speed-wrist	1.02	0.01	0.22	1.05	44.3%	0.60
exertion time-wrist	4.27	0.30	<0.01	1.84	66.7%	0.75
pk posture-forearm	0.22	0.07	<0.01	1.09	16.1%	0.34
avg posture-forearm	2.41	<0.01	<0.01	1.27	65.5%	0.71
pk movement speed-forearm	0.67	<0.01	<0.01	0.66	50.3%	0.58
avg movement speed-forearm	1.26	<0.01	<0.01	0.20	86.1%	0.69
exertion time-forearm	3.29	0.29	<0.01	3.45	46.8%	0.58
pk posture-elbow	0.39	<0.01	<0.01	0.66	36.9%	0.48
avg posture-elbow	0.88	0.01	<0.01	0.40	68.2%	0.75
pk movement speed-elbow	0.33	<0.01	<0.01	0.54	38.3%	0.50
avg movement speed-elbow	1.64	0.01	0.10	0.48	73.6%	0.85
exertion time-elbow	5.53	0.07	<0.01	0.90	85.0%	0.89
pk posture-shoulder	2.06	<0.01	<0.01	2.64	43.8%	0.54
avg posture-shoulder	0.61	0.13	<0.01	0.82	39.0%	0.58
pk movement speed-shoulder	0.38	0.01	<0.01	1.06	26.3%	0.43
avg movement speed-shoulder	1.34	<0.01	0.10	0.87	58.0%	0.71
exertion time-shoulder	3.85	<0.01	<0.01	2.65	59.2%	0.67
pk posture-neck	2.37	0.01	0.08	2.84	44.8%	0.58
avg posture-neck	0.67	<0.01	<0.01	0.87	43.5%	0.53
pk movement speed-neck	0.07	<0.01	0.03	1.22	5.6%	0.28
avg movement speed-neck	0.62	0.02	0.15	0.67	42.6%	0.68
exertion time-neck	0.81	0.26	<0.01	2.45	23.0%	0.47
pk posture-back	3.94	0.12	<0.01	1.55	70.2%	0.78
avg posture-back	0.19	<0.01	<0.01	0.53	26.1%	0.40
pk movement speed-back	0.51	0.02	<0.01	0.98	33.8%	0.50
avg movement speed-back	0.62	0.04	0.05	0.55	49.3%	0.67
exertion time-back	0.81	0.26	0.12	1.24	33.4%	0.63

* all jobs were rated 0.

Table B.6: Job and day variance for final ratings in test/retest reliability study.

stressor	var (jobs)	var (day)	var (error)	%var (job)	model r^2
repetition	4.50	<0.01	0.08	98.0%	0.99
pk force	1.92	<0.01	0.14	92.8%	0.96
avg force	0.21	<0.01	0.13	60.0%	0.80
peak contact stress-fingers	2.99	<0.01	0.34	89.5%	0.95
avg contact stress-fingers	0.48	<0.01	0.07	85.7%	0.93
peak contact stress-wrist/palm	0.92	<0.01	0.54	62.9%	0.80
avg contact stress-wrist/palm	0.23	<0.01	0.10	70.4%	0.85
peak contact stress-forearm	0.22	<0.01	0.14	61.2%	0.81
avg contact stress-forearm	0.01	<0.01	0.01	68.8%	0.84
peak contact stress-elbow	0	0	0	0	*
avg contact stress-elbow	0	0	0	0	*
pk posture-wrist	0.82	0.11	0.63	52.6%	0.79
avg posture-wrist	0.18	<0.01	0.09	64.3%	0.84
pk movement speed-wrist	0.33	<0.01	0.40	45.3%	0.72
avg movement speed-wrist	0.90	0.26	0.73	47.6%	0.80
exertion time-wrist	5.00	<0.01	0.15	97.0%	0.99
pk posture-forearm	0.22	<0.01	0.12	64.6%	0.83
avg posture-forearm	2.40	<0.01	0.73	76.6%	0.87
pk movement speed-forearm	0.36	0.07	0.06	74.6%	0.94
avg movement speed-forearm	1.43	0.19	0.30	74.4%	0.92
exertion time-forearm	4.61	<0.01	1.45	76.1%	0.88
pk posture-elbow	0.47	0.02	0.16	72.2%	0.88
avg posture-elbow	0.77	<0.01	0.14	84.7%	0.92
pk movement speed-elbow	0.16	<0.01	0.13	55.6%	0.76
avg movement speed-elbow	1.50	0.14	0.24	79.8%	0.93
exertion time-elbow	5.87	<0.01	0.46	92.7%	0.96
pk posture-shoulder	2.36	<0.01	0.89	72.6%	0.86
avg posture-shoulder	0.61	<0.01	0.18	77.3%	0.89
pk movement speed-shoulder	0.13	0.03	0.04	66.7%	0.90
avg movement speed-shoulder	1.44	0.17	0.32	74.9%	0.92
exertion time-shoulder	4.37	<0.01	0.44	90.9%	0.95
pk posture-neck	4.56	0.01	0.96	82.3%	0.91
avg posture-neck	1.05	<0.01	0.22	82.7%	0.91
pk movement speed-neck	0.07	<0.01	0.04	59.6%	0.55
avg movement speed-neck	0.61	0.15	0.23	61.1%	0.87
exertion time-neck	<0.01	<0.01	0.01	20.0%	0.58
pk posture-back	4.52	<0.01	0.91	83.2%	0.91
avg posture-back	0.14	<0.01	0.37	28.3%	0.64
pk movement speed-back	0.58	<0.01	0.60	48.8%	0.75
avg movement speed-back	0.65	0.09	0.32	61.4%	0.84
exertion time-back	0.15	0.24	1.27	9.2%	0.61

* all jobs were rated 0.

APPENDIX C

MODELS OF ALTERNATIVE HEALTH OUTCOMES

Table C.1: Predictors of hand diagram scores of 1,2, or 3; N=351. Log likelihood = -165.03

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-2.6 (0.70)	<0.01		
sex (F=1, M=0)	0.65 (0.29)	0.03	1.91	1.08 - 3.36
age	-0.004 (0.01)	0.74	1.00	0.97 - 1.02
wrist ratio (1=>0.73, 0=<0.73)	0.72 (0.29)	0.01	2.06	1.16 - 3.64
repetition rating (0-10)	0.15 (0.06)	0.01	1.17*	1.03 - 1.31

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(\text{coefficient} \times \Delta \text{rating})}$

Table C.2: Predictors of hand diagram scores of 3; N=351. Log likelihood = -90.57

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-4.14 (1.10)	<0.01		
sex (F=1, M=0)	0.26 (0.42)	0.54	1.29	0.56 - 2.97
age	0.002 (0.02)	0.93	1.00	0.96 - 1.04
wrist ratio (1=>0.73, 0=<0.73)	0.85 (0.41)	0.04	2.35	1.05 - 5.26
repetition rating (0-10)	0.23 (0.10)	0.02	1.26*	1.03 - 1.54

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(\text{coefficient} \times \Delta \text{rating})}$

Table C.3: Predictors of numbness, tingling, burning, or pain; N=351.
Log likelihood = -210.23

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-2.09 (0.59)	<0.01		
sex (F=1, M=0)	0.43 (0.25)	0.08	1.54	0.94 - 2.49
age	0.005 (0.01)	0.67	1.00	0.98 - 1.03
wrist ratio (1=>0.73, 0=<0.73)	0.36 (0.26)	0.17	1.44	0.86 - 2.41
repetition rating (0-10)	0.15 (0.05)	<0.01	1.16*	1.05 - 1.28

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(\text{coefficient} \times \Delta \text{rating})}$

Table C.4: Predictors of CTS based on median mononeuropathy (0.8 ms threshold); N=337. Log likelihood = -97.4

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-7.69 (1.30)	<0.01		
sex (F=1, M=0)	-0.18 (0.39)	0.65	0.84	0.39 - 1.80
age	0.05 (0.02)	<0.01	1.06	1.01 - 1.10
BMI	0.09 (0.03)	<0.01	1.10	1.04 - 1.16
wrist ratio (1=>0.73, 0=<0.73)	1.21 (0.38)	<0.01	3.34	1.59 - 7.05

Table C.5: Predictors of CTS based on median mononeuropathy (0.8 ms threshold) combined with hand diagram scores of at least 2; N=337. Log likelihood = -48.10

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-8.13 (1.96)	<0.01		
sex (F=1, M=0)	-0.05 (0.60)	0.93	0.95	0.30 - 3.06
age	0.04 (0.03)	0.14	1.04	0.99 - 1.10
wrist ratio (1=>0.73, 0=<0.73)	1.06 (0.59)	0.07	2.88	0.91 - 9.08
repetition rating (0-10)	0.41 (0.19)	0.03	1.50*	1.04 - 2.17

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(\text{coefficient} \times \Delta \text{rating})}$

Table C.6: Predictors of CTS based on median mononeuropathy (0.5 ms threshold) combined with reported numbness, tingling, burning, or pain; N=336. Log likelihood = -95.04

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-5.10 (1.08)	<0.01		
sex (F=1, M=0)	-0.18 (0.41)	0.67	0.84	0.38 - 1.87
age	0.03 (0.02)	0.13	1.03	0.99 - 1.07
wrist ratio (1=>0.73, 0=<0.73)	0.61 (0.41)	0.13	1.84	0.83 - 4.08
repetition rating (0-10)	0.23 (0.09)	0.01	1.26*	1.05 - 1.51

* OR for 1 unit increase in repetition rating (on 10 cm scale); $OR=e^{(\text{coefficient} \times \Delta \text{rating})}$

Table C.7: Predictors of non-specific discomfort with repetition modeled as 3 levels (low, medium, high); N=351. Log likelihood = -209.26

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-1.76 (0.59)	<0.01		
history of CTS	0.76 (0.38)	0.05	2.14	1.01 - 4.50
history of tendinitis	0.68 (0.31)	0.03	1.97	1.07 - 3.60
sex (F=1, M=0)	0.74 (0.25)	<0.01	2.10	1.29 - 3.41
age	-0.006 (0.01)	0.59	0.99	0.97 - 1.02
repetition category (low, med, high)	0.41 (0.14)	<0.01	1.51*	1.14 - 2.00

* OR for 1 category increase in repetition rating (e.g. low to med, med to high).

Table C.8: Predictors of tendinitis with repetition modeled as 3 levels (low, medium, high); N=351. Log likelihood = -100.24

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-6.13 (1.11)	<0.01		
history of soft tissue dis.	1.00 (0.38)	0.01	2.71	1.29 - 5.68
sex (F=1, M=0)	-0.39 (0.50)	0.43	0.67	0.25 - 1.78
age	0.02 (0.02)	0.18	1.03	0.99 - 1.78
skinfold thickness	0.04 (0.02)	0.02	1.04	1.01 - 1.07
repetition category (low, med, high)	0.63 (0.25)	0.01	1.88*	1.14 - 3.09

* OR for 1 category increase in repetition rating (e.g. low to med, med to high).

Table C.9: Predictors of hand diagram scores of at least 2 with repetition modeled as 3 levels (low, medium, high); N=351. Log likelihood = -127.99

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-3.32 (0.84)	<0.01		
sex (F=1, M=0)	0.67 (0.35)	0.05	1.95	0.99 - 3.85
age	0.001 (0.02)	0.95	1.00	0.97 - 1.03
wrist ratio (1=>0.73, 0=<0.73)	0.96 (0.33)	<0.01	2.62	1.37 - 5.02
repetition category (low, med, high)	0.32 (0.21)	0.13	1.38*	0.91 - 2.09

* OR for 1 category increase in repetition rating (e.g. low to med, med to high).

Table C.10: Predictors of CTS based on median mononeuropathy (0.5 ms threshold) combined with hand diagram scores of at least 2 with repetition modeled as 3 levels (low, medium, high); N=336. Log likelihood = -69.06

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-4.97 (1.31)	<0.01		
sex (F=1, M=0)	-0.14 (0.50)	0.78	0.87	0.33 - 2.33
age	0.01 (0.02)	0.54	1.01	0.97 - 1.06
wrist ratio (1=>0.73, 0=<0.73)	0.95 (0.49)	0.05	2.57	0.99 - 6.68
repetition category (low, med, high)	0.55 (0.32)	0.09	1.73*	0.92 - 3.25

* OR for 1 category increase in repetition rating (e.g. low to med, med to high).

Table C.11: Predictors of non-specific discomfort with repetition modeled as 2 levels (low, high); N=351. Log likelihood = -208.79

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-2.18 (0.68)	<0.01		
history of CTS	0.73 (0.38)	0.06	2.07	0.98 - 4.36
history of tendinitis	0.69 (0.31)	0.03	1.99	1.08 - 3.66
sex (F=1, M=0)	0.73 (0.25)	<0.01	2.07	1.27 - 3.37
age	-0.005 (0.01)	0.65	0.99	0.97 - 1.02
repetition category (low, high)	0.79 (0.26)	<0.01	2.21*	1.32 - 3.71

* OR for 1 category increase in repetition rating (i.e. low to high).

Table C.12: Predictors of tendinitis with repetition modeled as 2 levels (low, high); N=351. Log likelihood = -101.34

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-6.29 (1.23)	<0.01		
history of soft tissue dis.	0.96 (0.38)	0.01	2.61	1.25 - 5.44
sex (F=1, M=0)	-0.35 (0.50)	0.48	0.70	0.26 - 1.87
age	0.03 (0.02)	0.17	1.03	0.99 - 1.06
skinfold thickness	0.04 (0.02)	0.02	1.04	1.01 - 1.07
repetition category (low, high)	0.97 (0.47)	0.04	2.63*	1.05 - 6.60

* OR for 1 category increase in repetition rating (i.e. low to high).

Table C.13: Predictors of hand diagram scores of at least 2 with repetition modeled as 2 levels (low, high); N=351. Log likelihood = -127.54

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-3.78 (0.97)	<0.01		
sex (F=1, M=0)	0.64 (0.35)	0.07	1.89	0.96 - 3.74
age	0.002 (0.02)	0.90	1.00	0.97 - 1.03
wrist ratio (1=>0.73, 0=<0.73)	0.96 (0.33)	<0.01	2.61	1.36 - 4.99
repetition category (low, high)	0.71 (0.40)	0.08	2.03*	0.92 - 4.44

* OR for 1 category increase in repetition rating (i.e. low to high).

Table C.14: Predictors of CTS based on median mononeuropathy (0.5 ms threshold) combined with hand diagram scores of at least 2 with repetition modeled as 2 levels (low, high); N=336. Log likelihood = -69.31

model term	coefficient (std. err.)	prob> Z	Odds ratio (OR)	95% confidence interval (OR)
intercept	-5.31 (1.50)	<0.01		
sex (F=1, M=0)	-0.13 (0.50)	0.80	0.88	0.33 - 2.37
age	0.01 (0.02)	0.52	1.02	0.97 - 1.06
wrist ratio (1=>0.73, 0=<0.73)	0.94 (0.49)	0.05	2.56	0.99 - 6.64
repetition category (low, high)	0.94 (0.61)	0.12	2.56*	0.78 - 8.40

* OR for 1 category increase in repetition rating (i.e. low to high).