ABSTRACT

COVALLA, ELIZABETH. Visual Posture Observation Error and Training. (Under the direction of Dr. Gary A. Mirka)

The purpose of this study was to determine people's ability to visually estimate postural angles of the shoulder, trunk, and wrist. One application of these findings is to determine the effect of estimation error on common risk analysis tools that incorporate posture. Considerations are given to the effect of training, video mode, gender, body region, and subject characteristics on estimation error. Absolute error, algebraic error, and subject confidence are used to characterize visual estimation abilities. Results indicate that visual estimation error ranges between 7 and 10 degrees. Error further increased with wrist postures and female observers. Due to estimation errors, analysis tools that include posture are less accurate in predicting risk of injury. Eight, 12, and 14 percent of shoulder, trunk, and wrist postures, respectively, were misclassified causing Rapid Upper Limb Assessment (RULA) scores to shift by at least one point. For the Strain Index, forty percent of wrist postures were misclassified by participants causing as much as a two-thirds change in the final score.

VISUAL POSTURE OBSERVATION ERROR AND TRAINING

by
Elizabeth Danielle Covalla

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Masters of Science

INDUSTRIAL ENGINEERING

Raleigh

2003

APPROVED BY:

_____Dr. Nelson W. Couch_____ __Dr. Caroyln M. Sommerich__

________Dr. Gary A. Mirka________
Chair of Advisory Committee

Biography

Elizabeth Covalla graduated with a B.S. in Industrial Engineering and a B.A. in Political Science from North Carolina State University in 2001. During her undergraduate career, she was a member of the Benjamin Franklin Scholar's program and Phi Beta Kappa.

Table of Contents

List of Tables

List of Figures

1 Introduction

The term musculoskeletal disorder (MSD) has been used to describe a variety of injuries and illness affecting the musculoskeletal system, which are related to repetitive motion, cumulative trauma, or overexertion. In 1998, overexertion injuries caused by excessive lifting, pushing, pulling, holding, carrying or throwing causing five or more missed days of work were estimated to cost 9.8 billion dollars and accounted for 25.57% of workers' compensation direct costs (Liberty Mutual, 2001a). This same year repetitive motion injuries causing five or more lost work days accounted for 6.10% of injuries and cost $2.3 billion in workers' compensation direct costs.

The true cost of MSDs is thought to be even higher than the above direct cost estimates. Forty percent of US executives believe that between three and five dollars of indirect costs exist for each $1 of direct costs (Liberty Mutual, 2001b). The total cost associated with MSDs is difficult to measure. Estimates vary from 13 billion to 27 billion dollars a year (National Institute for Occupational Safety and Health, 1997, Mallory and Bradford, 1992). Despite the fact that exact costs are unknown, it is known that thousands of employees are being diagnosed with MSDs every year.

Repetitive motions, including typing or key entry, repetitive use of tools, and repetitive placing, grasping, or moving of objects other than tools, caused 92,576 injuries in 1994 (National Institute for Occupational Safety and Health, 1997). Of these repetitive motion injuries, 55% affected the wrist, 7% affected the shoulder, and 6% affected the back. This same year, overexertion accounted for approximately 661,000 injuries or illnesses.

Not every MSD is work related. Sports, hobbies, medical conditions, genetic factors, and some medications can all contribute to symptoms of MSDs. One study found that over 37% of musicians are suspected to suffer from multiple MSDs (Dawson, 2002). In particular, muscle-tendon strains are commonly diagnosed in pianists, violinists and violists, guitarists, and reed instrumentalists. Pregnant women, exposed to no other risk factors, frequently suffer from symptoms of carpal tunnel syndrome that can last long after the pregnancy is over. It can take six to twenty months for the symptoms to go away (Weimer et al., 2002).

Since MSDs can occur on or off the job, industry is eager to determine the exact causes of MSDs. Knowing what causes MSDs would help industry reduce employee exposures and would help determine work-relatedness in workers' compensation cases. The first step of determining work-relatedness is to identify the work-related physical risk factors associated with MSDs.

1.1 Risk Factors

It is unlikely a single risk factor is responsible for all MSDs. Instead, some combination of risk factors more likely leads to the development of MSDs. Force, movement, vibration, and posture are suspected of being the main work-related risk factors.

1.1.1 Force

Muscles produce the forces needed for movement. Overuse of muscles leads to muscle fatigue that causes muscle pain and cramping. Overuse may occur due to an acute, one-time event. For example, a muscle strain that occurs from lifting an object that was too heavy. These injuries are more severe than muscle fatigue and cramping resulting in torn or damaged muscle tissue. Common acute overuse injuries include strains of the low back and upper extremities.

Muscle overuse may also occur due to light exertions, if the exertions are repetitive or sustained over a significant period of time and muscles are not given sufficient rest and recovery time between muscle applications. Under these circumstances, the amount of muscle fatigue is directly related to the amount of force and the duration of force experienced by the muscles. Force has been linked to neck and neck/shoulder, elbow, and hand/wrist injuries (National Institute for Occupational Safety and Health, 1997).

In addition to the forces in the muscular tissue itself, high force exertions may also lead to damage in the joints of the body. In the typical biomechanical configuration, the joint acts as a pivot point balancing the forces of the muscles against external loads. In the spine, for example, the vertebrae and the intervertebral disc act as this pivot point. When lifting a box, the distance between the box load and the spine is far greater than the distance between the back muscles and the spine. Because of this disproportionate moment arm, the back muscles must exert many more times the external load being lifted. Meanwhile, the spinal disc must withstand the forces being applied by both the muscles and the external load. Coupled with awkward postures that can place the spinal discs outside of optimal force resistance paths, disc damage may result. Similar processes may occur in joints throughout the body.

1.1.2 Movement

Too much movement (repetition) and too little movement (static postures) are both risk factors. If a body part is held in a fixed position for too long, the continuous contracting of muscles reduces the blood flow to the body part necessary for eliminating waste products and delivering fresh nutrients to the muscles. On the other hand, too much movement does not give muscles enough recovery time. This effect was shown in a recent study using the rat model (Barbe et al., 2003). In this study, motor decrements, signs of injury, and cellular and tissue responses associated with inflammation were observed in rats after voluntary reaching for food at a rate of 4 reaches/min, 2 h/day, and 3 days/week for up to 8 weeks. Much more severe exposures to repetition may be observed in industry. Repetition has been linked to neck and neck/shoulder, shoulder, and elbow injuries in general, as well as, carpal tunnel syndrome and tendinitis of the hand/wrist in particular (National Institute for Occupational Safety and Health, 1997).

1.1.3 Vibration

Workers are often exposed to vibration through power tools and machinery. Vibration injuries are linked to the magnitude, frequency, direction, and duration of vibration (Safety Line Institute, 1998). One study found that symptoms of pain in the fingers, sensitivity to cold, numbness and pain in the fingers at night, weakness of static position, wrist-elbow pain, difficulty in bending and stretching elbow, pain in shoulder when holding up arms, lower back pain, sleeping disturbance and hearing difficulty were significantly higher among workers exposed to full body vibration from rock drills (Issever et al., 2003). In addition, strong evidence has been found to link vibration to hand-arm and back injuries (National Institute for Occupational Safety and Health, 1997).

1.1.4 Posture

Awkward postures, especially when combined with other risk factors, have also been linked to MSDs. Awkward postures are thought to contribute to MSDs by increasing the forces in the body and by reducing muscle strength. Carpal tunnel syndrome is thought to occur through the compression of the median nerve passing through the carpal tunnel region, which maybe caused by swelling of the surrounding finger flexor tendons, among other things. This swelling occurs when the tendons rub against each other and the other capral tunnel contents during repetitive activities. When the wrist is in a deviated posture, friction on the tendons within the carpal tunnel make swelling more likely. Combinations of repetitive and/or forceful movements with deviated wrist postures further increase the likelihood of swelling and nerve compression.

Awkward postures can also reduce muscle strength. Shoulder, wrist, and elbow angles have been shown to have a significant effect upon grip strength (Kattel et al., 1996). Deviations in posture place muscles at a non-optimal length causing them to work closer to their maximum force production capability. The body is under the least amount of stress when it is working within a neutral posture. Neutral postures keep the body joints within the mid-range of joint motion.

1.2 Postures and MSDs

A causal relationship between posture and MSDs has been established. In a review of epidemiological literature, National Institute for Occupational Safety and Health (1997) established moderate evidence supporting a link between posture and MSDs occurring in the shoulder, hand/wrist, and low back. National Institute for Occupational Safety and Health (NIOSH) reviewed thirteen articles showing a link between awkward postures and shoulder disorders where an awkward posture was defined as shoulder abduction or flexion past sixty degrees. Six of the studies reported a link between posture and tendinitis while the remaining seven studies reported a link between posture and non-specific shoulder disorders.

Increased deviations in hand/wrist posture were linked to carpal tunnel syndrome in five out of six studies. Three out of four studies linked awkward hand/wrist postures to hand/wrist tendinitis. However, inconsistencies in posture definitions between studies makes it difficult to define an "awkward" range of hand/wrist postures.

NIOSH reviewed thirteen studies examining awkward postures of the back during bending and twisting. Three out of four studies found an elevated risk of low-back disorder associated with tasks requiring bending, twisting, or other awkward postures. The remaining eight studies used more qualitative exposure measures that resulted in conflicting findings. As with the hand/wrist studies, inconsistency in posture definitions between studies makes it difficult to define an "awkward" range of low back postures.

1.3 Postural Analysis

1.3.1 Risk Analysis Tools

Risk analysis tools are used in industry to identify ergonomic risk factors leading to MSDs. Several attempts have been made to incorporate posture into analysis tools. Juul-Kristensen et al. (1997) identified eight observation methods used to classify posture in repetitive work. These methods include: VIRA (videofilm technique for registration and analysis of work postures and movements), PEO (portable ergonomics observation method), TRAC (task recording and analysis on computer), OWAS (Ovako working posture analyzing system), RULA (rapid upper limb assessment) and three American methods published by Keyserling (1986), Armstrong et al., (1982) and Genaidy et al. (1993). The American methods, PEO, and VIRA rely on video recordings to record duration and frequency of postures. TRAC and OWAS use computer assisted posture identification. RULA uses a checklist approach requiring subjects to visually estimate angles (McAtamney and Corlett, 1993). RULA is the most widely used tool because it is simple and does not require equipment or software.

RULA is designed to be a quick survey method for identifying risk of work-related upper limb disorders (McAtamney and Corlett, 1993). Postures of the neck, trunk, and upper limbs are grouped into ranges. Each posture range is associated with a score. Ranges are added together to reach a final RULA score. Wrist ranges are as follows: neutral posture, 0° to 15°, and greater than 15°. Upper arm ranges are as follows: neutral postures, greater than 20° extension or 20° to 45° flexion, 45° to 90° flexion, and greater than 90° flexion. Trunk ranges are as follows: 0°, 0° to 20° flexion, 20° to 60° flexion, or greater than 60° flexion. Based on RULA, The Rapid Entire Body Assessment (REBA) individually scores trunk, neck, leg, upper arms, lower arms, and wrist postures to develop a final REBA score (Hignett & McAtamney, 2000).

Other tools exist that combine posture with other risk factors to create more comprehensive risk analysis tools. The Strain Index is a semi-quantitative job analysis tool developed to identify jobs associated with distal upper extremity disorders (Moore and Garg, 1995). The Strain Index applies numerical multipliers to six task variables: intensity of exertion, duration of exertion per cycle, efforts per minute, wrist posture, speed of exertion, and duration of task per day. Verbal anchors and angle ranges, listed in Table 1.1, guide users in qualitatively grouping wrist posture. Postures that are “very good” or “good” have multipliers of 1.0 and do not effect the end Strain Index score. These postures are considered neutral. As the postures become extreme the categories change from "fair" to "bad" to "very bad" and have a multiplier of 1.5, 2.0, and 3.0, respectively. The larger the multiplier the greater the effect on the end score.

Table 1.1: Hand/Wrist posture ranges for the Strain Index

Rating Criterion	Wrist Extension	Wrist Flexion	Ulnar Deviation	Perceived Posture	Stain Index Score
Very Good	0° - 10°	0° - 3°	0° - 10°	Perfectly Neutral	1.0
Good	11° - 25°	6° - 15°	11° - 25°	Near Neutral	1.0
Fair	26° - 40°	16° - 30°	16° - 20°	Nonneutral	1.5
Bad	41° - 55°	31° - 50°	21° - 25°	Marked Deviation	2.0
Very Bad	> 60°	> 50°	> 25°	Near Extreme	3.0

Similarly, Ketola et al. (2001) developed a tool for evaluating the physical load on the upper extremities that also incorporates wrist and shoulder posture along with five other risk factors including repetitive use of the hands, use of hand force, pinch grip, and local mechanical pressure. The authors limited wrist posture to two categories: neutral or greater than 20° because they felt that wrist posture observation was not very accurate. Shoulder posture was grouped as less than or greater than 90-degrees.

Each of these tools is designed to be used by safety and health professionals in the workplace. Consequently, the most useful tools will be simple, quick, inexpensive, and accurate. Though it is easy to evaluate the tools based on the first three criteria, little is known about these tools' accuracy.

1.3.2 Validation of Risk Analysis Tools

The above tools must be validated in industry. First, these tools must be proven to accurately identify the risk associated with a job. Second, it must be proven whether users can correctly and consistently apply these tools. Some research has attempted to address the first point for the Strain Index. For single task jobs, the Strain Index has been shown to have high sensitivity and high specificity (Rucker and Moore, 2002; Moore et al., 2001). In a study of 28 jobs, Rucker and Moore (2002) showed that the Strain Index was able to correctly classify 6 out of 6 problem jobs (100% sensitivity) and 20 out of 22 safe jobs (91% specificity).

While no such validation studies were found for REBA or Ketola's method, these methods have been evaluated for overall inter-observer reliability. Data from two occupational health nurses given a 12-hour course on Ketola et al. (2001)'s method, found that inter-observer repeatability was high (Ketola et al., 2001). However, with only two subjects, limited information should be inferred from these findings. Hignett & McAtamney (2000) used 14 professional therapists, nurses, and ergonomists to assess the inter-observer reliability of REBA coding. After a workshop in which 600 examples of postures from industry were presented, the 14 participants independently assessed 144 full body postures. The participants achieved between 62 and 85 percent agreement (omitting the Upper Arm category due to coding changes). Even with extensive training of highly educated individuals, inter-observer reliability was only moderate.

1.3.3 Challenges in Estimating Postural Angles

The source of inter- and intra-observer error in job analysis tools is unknown. One aspect of these tools that may lead to user error is visual estimation of postural angles. The postural angles required by these tools are most commonly estimated visually. Though more quantitative forms of measurement exist, such as video analysis, motion tracking, and goniometers, the cost and time associated with these methods limits their use in industrial settings. In addition, many of the quantitative tools require some form of apparatus to be attached to the subject. Such equipment may alter normal subject behavior. Simple visual estimation remains the most cost-effective and commonly used means of applying these tools. However, it is unknown what effect visual estimation errors have on the accuracy of these rating systems.

1.3.3.1 Inter-observer Reliability

Research has examined inter-observer reliability and absolute accuracy of visual postural estimation among, in most cases, highly trained observers. This research is summarized in Table 1.2. In a pilot experiment, Douwes and Dul (1991) studied the inter-observer reliability of three observers trained for three days together. An inter-observer reliability of 0.97 was achieved for estimating trunk angles from photographs.

Burt and Punnett (1999) examined the inter-observer reliability of a quantitative observational method for assessing non-neutral postures. The study used two observers who placed wrist, shoulder, elbow, and back postures into a limited number of categories including: wrist flexion, wrist extension, ulnar deviation, pinch grip, elbow extension, shoulder elevation < 45°, shoulder elevation > 45°, back flexion, back extension, or twisting. The first observer, highly experienced with using the methodology, trained the second observer for one day. The observers then went on to independently evaluate 75 jobs in a stamping plant simultaneously. Inter-observer reliability between two observers ranged from 26% for right shoulder elevation to 99% for left wrist flexion. Overall, raters agreed in 96% of observations, however, the kappa value was only 0.55. Because most categories were rarely observed, the percent agreement did not surpass the chance of random agreement. In both studies, the sample size was too small to draw many conclusions.

Inter-observer reliability of visual postural estimation has also been evaluated as part of Ketola et al. (2001)'s job observation method. Ketola et al. (2001) recorded poor validity and only moderate inter-observer repeatability between two health nurses placing wrist and shoulder postures into one of two groups.

1.3.3.2 Accuracy

Ericson et al. (1991) studied the accuracy of neck, trunk, and upper arm angle estimation among eight experienced ergonomists. The ergonomists, assisted by body markers on the model, had a median estimation error of 5° during both static and dynamic conditions. However, no statistics were given for this pilot data and significance cannot be determined.

Though accuracy appears high with highly experience observers, little is known about visual estimation error among laymen observers. Job risk analysis tools are often designed for plant level practitioners that most likely will have little if any training using the tools. Therefore, to understand potential limitations of the tools, the abilities of novice tool users must be assessed. In one study, Genaidy et al. (1993) found a mean absolute error of 9.2° and a mean algebraic error of –1.3° associated with the visual estimation of the shoulder among twenty untrained engineering students. However, the authors did not determine if these findings were consistent with other body parts or if these findings were consistent with participants outside of engineering.

Baluyut et al. (1995) found that untrained engineering students were fairly accurate at placing a posture in a posture range. However, their accuracy varied depending on whether the model was standing or sitting. Thirty-two untrained engineering students were able to correctly place a wrist posture in one of 7 categories (neutral, 16 to 45° flexion, > 45°

flexion, 16 to 45° extension, >45° extension, radial deviation, or ulnar deviation) 68% of the time when the model was standing and 75% of the time when the model was sitting. Similarly, the students were 76% accurate with standing shoulder postures, 81% accurate with sitting shoulder postures, 89% accurate with standing lower back postures, and 87% accurate with sitting lower back postures.

Table 1.2: Summary of angle observation research

Research	N	Body Part	No. Different Angles	Length of Training	Static or Dynamic	Findings
Baluyut et al., 1995	63	Wrist, Elbow, Shoulder, Neck, L Back	6	none	Static	- Lower back easier than wrist and elbow angles - Lower extremity position affected estimations - > 70% for sitting, >60% for standing (except for elbow) - Flexion and extension easier to evaluate than neutral and non-neutral postures
Burt and Punnett, 1999	2	Wrist, Shoulder Back	18	8 hrs	Dynamic	- Two categories for each - Inter-observer reliability 99% left wrist flexion - Inter-observer reliability 26% shoulder
Douwes and Dul, 1991	3	Trunk	6	3 days	Static	- Intra and inter reliability 0.97 - Mean absolute error 3° (sd=2.1°) - Angle affected observation error (p=0.003) - Angles near 0° and 90° had smallest error
Ericson et al., 1991*	8	Trunk, Neck, Shoulder	10	Exper-ienced	Static & Dynamic	- Higher estimation error with dynamic video - Underestimate trunk flexion - Overestimate shoulder flexion - Mean algebraic error 5° for all static body parts - Mean algebraic error 5° for all dynamic trunk and neck - Mean algebraic error 10-13° for dynamic shoulder - Mean algebraic error remains constant over two weeks
Genaidy et al., 1993	20	Shoulder	18	None	Static	- Measured against goniometer - No difference between angle range and absolute error - Mean absolute error 9.2° - Algebraic error 2° higher, 4° lower, and 2° lower for the low, medium, and high range, respectively - Mean algebraic error –1.3° across all angles
Ketola et al., 2001	2	Wrist, Shoulder	2	12 hrs	Dynamic	- Measured against goniometers - Place into <20°/>20° categories for wrist - Place into <90°/>90° categories for shoulder - Moderate or poor Inter-observer repeatability and poor accuracy - Underestimated wrist

*no statistics given

1.3.4 Influences on Angle Estimation

An individual's ability to visually estimate postural angles may be influenced by secondary factors. Based on related research findings, training, movement, gender, field of study, and body part may all affect estimation error.

Training may improve inter-observer reliability. Though no previous research has examined the effects of training on the accuracy of postural angle estimation, Waller and Wright (1965) studied the effects of training on angle estimation among navigators. Training with flashcards for thirty minutes significantly improved the accuracy of eighteen navigators used as test subjects indicating, for angles drawn on paper, training has a positive effect.

The rate of change of the body posture being observed may also influence inter-observer reliability. Previously it has been found that postural angles are easier to estimate from static photographs than dynamic video, for some body parts (Ericson et al., 1991). Though there was no observed difference with the trunk and neck, the mean algebraic error increased from 5° to 10 - 13° when the shoulder switched from static to dynamic images. However, no statistics were given for this data and significance cannot be determined.

Researchers continue to find that males possess superior spatial visualization abilities when compared with females (Amponsah, 2000; Arthur et al., 1997; Nordivk and Amponsah, 1998). Males out perform females on the Mental Rotation Test (MRT) and the Water Level Task (WLT) (Voyer et al., 2000). It is reasonable to assume that such gender differences will carry over to the visual estimation of postural angles. Consequently, males may be better postural angle estimators than females. However, research also indicates such gender differences have environmental components. Gender differences diminish in females who have spatial toy or sport preferences (Voyer et al., 2000). Field of study has also been shown to influence visualization abilities between genders (Quaiser-Pohl and Lehmann, 2002). Gender differences are most prominent with students in the arts, humanities, and social sciences and smallest in computational visualistics majors. In addition, females with computer experience tend to have better spatial abilities than those without computer experience. Therefore, any gender differences must be looked at in the context of participant background.

Burt and Punnett (1999) found that inter-operator reliability varied based on body part. Another study found that subjects had more difficulty judging elbow and wrist postures than lower back postures (Baluyut et al., 1995). Therefore, some postural angles may inherently be easier to observe than others. At least in the case of drawn angles, characteristics of angles and their surroundings influence the perceived angle size. Angles with longer arms are judged as larger (Wenderoth and Johnson, 1984). This would suggest long limb size may influence postural angle estimates. Angles presented in a vertical orientation are underestimated compared with those presented horizontally (MacLean and Stacey, 1971). Therefore, vertical postures, such as the neck, may be misjudged. Additionally, acute angles are overestimated and obtuse angles are underestimated when presented horizontally (MacLean and Stacey, 1971). In body postures, the wrist is usually measured in the horizontal plane. Therefore, wrist posture may also be subject to error due to visual illusions. However, it is unknown whether or not these phenomenon observed with simple angles on paper may be carried over to complex postural angles. Baluyut et al. (1995) came closest at measuring the effects of visual illusions of postural angle estimation. The authors found that lower extremity positions affected the accuracy of wrist, elbow, shoulder, neck, and lower back estimation.

The preceding summary of the pertinent literature has identified a number of important issues related to postural angle estimation that should be evaluated in light of the many risk analysis tools that use these data as inputs to predict risk. The current research examines the effect of hands-on training, in which participants practice postural estimation with simple goniometers, on the ability of novice users to visually estimate postural angles of the trunk, neck, and wrist. In addition to training, the effects of still photos versus dynamic video modes as well as gender, body part, and field of study, on estimation performance is examined. The following hypotheses are asserted for this study:

- Hypothesis 1 – Hands-on training will improve performance over written training materials only
- Hypothesis 2 - Higher errors in angle estimation will be associated with dynamic conditions than static conditions
- Hypothesis 3 - Higher errors in angle estimation will be associated with females than males
- Hypothesis 4 –Higher errors in angle estimation will be associated with wrist postures than trunk and shoulder postures

2 Methods

The objectives of this study were to 1) assess people's ability to visually estimate postural angles and 2) determine the effect of training and other observation factors on an observer's ability to visually estimate the postural angles of the trunk, shoulder, and wrist.

2.1 Participants

Thirty-two volunteers participated in this study. Participants were recruited from the North Carolina State University student body and the local community. The only criterion for participant selection was that the participant had no previous instruction in visual estimation of postural angles. Participants' field of study/employment, age, and education were not taken into consideration when selecting participants. However, these characteristics were recorded as seen in Table 2.1, below. Participants studying or working in engineering, mathematics, or the sciences were considered to be in a "technical" field. All other participants were placed in the "non-technical" field category.

Table 2.1: Participant Characteristics (Standard deviations in parenthesis)

Group	All	Control	Trained
Number of Participants	32	16	16
Age (years)	28.3 (9.3)	25.6 (5.8)*	30.8 (11.0)*
Gender (male)	53 %	50 %	56 %
Education (years)	15.6 (1.6)	15.4 (2.0)*	16.0 (1.2)*
Current Field :			
Technical	72 %	81 %	63 %
Non-Technical	28 %	19 %	37 %

* means are significantly different (alpha=0.05)

2.2 Experimental Design

2.2.1 Independent Variables

Subjects were split into two groups. One group was given hands-on training (trained) while the other group received only written instructions (control). Participants were presented with both static and dynamic video showing the shoulder, trunk, or wrist. The following independent variables (levels in parenthesis) were investigated: 1) training modes (control/trained), 2) video mode (static/dynamic), 3) body part (shoulder, trunk, wrist), and 4) gender (male/female), 5) field of study (technical/non-technical). Independent variables are described in more detail below.

2.2.1.1 Training (Control/Trained)

Two training modes were studied in this experiment, control and trained. Sixteen participants made up the control group. Up to two participants were run at a time. Participants were given a standard set of written instructions (Appendix A) describing how the angle formed by the trunk, shoulder, and wrist should be estimated. Photos were included to augment the written instructions. As the instructor read the instructions to the participants, the instructor verbally defined the zero- and ninety-degree position of each body part. This completed the instruction for the control group.

Sixteen participants made up the trained group. Two participants performed the experiment at a time. Participants were given the same set of written instructions as the control group. After reading a section of the instructions, participants would take turns using a simple goniometer to measure the postural angle assumed by the instructor. As seen in Figure 2.1, one participant would observe while the other participant measured a trunk, shoulder, or wrist angle. The instructor wore markers on her wrist, hip, and shoulder to assist the participants with visualizing the postural angles. In all, six postures were measured by each participant: a shoulder angle greater than 90°, a shoulder angle less than 90°, a positive wrist angle, a negative wrist angle, a trunk angle greater than 45° and a trunk angle less than 45°. The first part of the training was created to help participants identify the important body landmarks used in the methodology.

Figure 2.1: Hands-on training methodology

A second training segment was added to help the trained participants with accuracy. During the second training segment, the pair of participants used a smaller simple goniometer to measure still photos displayed on a computer monitor as illustrated in Figure 2.1. The standard set of still photos consisted of a 26° and 63° shoulder angle, a 20° and 60° trunk angle, and a 30° flexion and 34° extension wrist angle. The instructor would correct the participant's estimate and explain why the participant's estimate was wrong after the participant had first attempted to measure the postural angle by his- or herself. The participants were also shown a photo with a drawing of the correct angle. Training lasted a total of 20 minutes.

2.2.1.2 Video Mode (Static/Dynamic)

The participants were asked to estimate the maximum postural angle that occurred during a simulated work task. The postures were either presented as a static or dynamic video clip.

2.2.1.3 Body Parts (Trunk/Shoulder/Wrist)

Participants were asked to measure postural angles formed by the trunk, shoulder, or wrist. These body parts were selected because they are commonly featured in many ergonomics tools including RULA, REBA, and the Strain Index. However, since none of these methods gives detailed instructions on which body landmarks to use in measuring postural angles a new methodology was created by the author.

In this study, the shoulder angle was defined as an angle between a line perpendicular to the ground and through the top of the shoulder at the acromium and a line drawn from the top of the shoulder through the lateral epicondyle at the elbow. The trunk angle was defined as the angle between a line perpendicular to the ground and through the hip rotation point and a line through the hip rotation point and the top of the shoulder. The wrist angle was defined as the angle between a line extending from the center of the forearm to the center of rotation of the wrist and a line from the center of rotation of the wrist through the center of the palm of the hand. While only shoulder and trunk flexion were observed, both wrist extension and flexion postures were used. Extended wrist postures were considered positive. Flexed wrist postures were considered negative.

2.2.1.4 Gender

Gender was also examined to investigate any spatial perception differences between males and females.

2.2.1.5 Field of Study

The participants' field of study (field) was also taken into account. However, field of study was not considered in the participant selection process.

2.2.2 Dependent Variables

2.2.2.1 Error

Participants were asked to visually estimate angles to the nearest degree. Absolute error was calculated by subtracting the estimate from the gold standard value and taking the absolute value of the result. The gold standard value was defined by the researcher as the actual value. Calculation of the gold standard value is discussed below. Algebraic error was calculated by subtracting the gold standard from the estimate. Algebraic error is used to assess over- and under-estimation trends.

2.2.2.2 Confidence

Participants were asked to write down a confidence interval associated with their answer (e.g. 10° +/- 3°). Participants were only told that the more confident they felt the smaller their confidence interval should be. Participants were not given a range restriction for confidence.

2.3 Creation of Standard Test Video Clips

Video was recorded directly onto an IBM ThinkPad through a DVC Dazzle Video Creator attached to a JVC camcorder. The video was edited using MGI Videowave software. The camcorder was kept perpendicular to the sagittal plane of the male volunteer used to model three simulated work tasks. The video showed a close up of only the wrist and forearm for wrist angles while the entire torso was shown for the shoulder and back angles. The volunteer's characteristics are given in Table 2.2.

Table 2.2: Volunteer's characteristics

Age (years)	28
Gender	Male
Stature (inches)	76
Weight (lbs.)	170

The volunteer lifting an empty crate at any one of ten heights was used for the trunk clips. The shoulder clips showed the volunteer placing a small wood block on a shelf at any one of ten heights. For the wrist clips, the volunteer moved his wrist from neutral to five flexed positions or one of five extended positions while holding a wood block. Examples of these postures are given in Appendix A. Static video clips lasted for 25 seconds. Dynamic video clips consisted of six looped repetitions of the same clip. Though the volunteer listened to a metronome to keep consistent pace, the nature of the action caused some variation in clip length. However, the length of time the maximum angle was held by the volunteer may be considered the same. Clip lengths are given in Table 2.3.

Table 2.3: Dynamic video clip length, in seconds

Length	Total Clip	Loop
Shoulder	26.0	4.3
Trunk	26.8	4.5
Wrist	11.8	2.0

2.4 Development of the Gold Standard

A consistent methodology for describing body part angles had to be created in order to develop a gold standard. Corlett and Manenica (1980) developed one of the first methods for recording and evaluating postures. Their method used a graphical system incorporating both posture and time. Though this method has been shown useful for reconstructing gross body activity (Corlett & Manenica, 1980), it was deemed too complex to quickly teach to general users. Instead, a method similar to Ericson et al. (1991)'s was used for the trunk and shoulder. Ericson et al. (1991) defined the trunk and shoulder segments as a straight line between two landmarks. The trunk segment was defined as a line between the hip joint and the gleno-humeral joint and the shoulder segment as a straight line between the gleno-humeral joint and the lateral humeral epicondyle. However, instead of the gleno-humeral joint the acromion was used because it is easier to identify in photographs. Ericson et al. (1991) did not study the wrist but a similar methodology was developed by the author for this body part. This final methodology is described in section 2.2.1.3.

Previous researchers have used video position analyzers to generate the gold standard body angle from video (Ericson et al., 1991; Douwes and Dul, 1991). However, such equipment was not available for this research. Attaching equipment (such as goniometers) to the model to record the postural angle has also been tried by researchers (Ketola et al., 2001). However, the equipment can obscure the view of the postural angle and may influence the visual estimation. Therefore, the postural angles were calculated using the coordinate system in Microsoft Paint. This method was deemed preferable to previous subjective methods which created the gold standard angle based solely on the visual estimations of three ergonomists/physiotherapists (Hignett & McAtamney, 2000).

Two ergonomists blinded to the target angle independently calculated each body angle from coordinates found using Microsoft Paint. If the ergonomists' answer varied by more than three degrees, the photograph was discussed until a consensus was reached. All other differences were averaged. Ten images per body part were selected. For each body part, no postural angle was within 5 degrees of another. The reference body angles used as the gold standard are listed in Table 2.4.

Table 2.4: Reference angles used as gold standard, negative wrist angles indicate wrist flexion.

Clip	Shoulder (Degrees)	Trunk (Degrees)	Wrist (Degrees)
1	14	3	52
2	20	10	42
3	32	35	31
4	43	42	20
5	51	51	5
6	65	60	- 5
7	74	67	- 18
8	83	81	- 27
9	98	88	- 43
10	102	98	- 58

2.5 Procedure

A general description of the experiment was provided to subjects after which the subjects signed a consent form approved by The North Carolina State University human subjects review board. After training, the subjects viewed the video clips at a computer workstation. A variety of monitor sizes were used, however, the screen resolution was set at 1024 x 768 pixels and sixteen-bit color. Video was viewed using Microsoft Media Player. If more than one participant was present, computer workstations were placed so that the participants could not view each other's monitors.

The subjects were told they could refer to the instructions as needed and ask questions at any time. If a subject asked a question about the methodology during testing, the experimenter would refer them to the portion of the written instructions or repeat the verbal instructions given at the start of testing. Subjects were also told that they could not use the cursor for anything but selecting files and that they could not touch the computer screen.

The subjects were presented with two folders of thirty video clips labeled "Static Video" and "Dynamic Video." Video clips were named with the target body part followed by a double letter A through J (e.g. "ShoulderAA" or "WristJJ"). Participants were allowed to watch the dynamic video clips twice and the static video clips once. The order of presentation of static or dynamic video files was randomized for each group of participants. Therefore, participants run in pairs were always presented with the same folder first. Within each set of video files, the order of clips was also randomized for each participant. After viewing each clip, subjects recorded an angle estimate and a confidence interval. Upon completion of all sixty clips, information about the participant's age, educational background, and current profession or field of study was acquired by the instructor. Participants were not compensated.

2.6 Data Processing

2.6.1 Outlier Removal

A data point was labeled an outlier if it was clear that the subject had transposed their answer. For example, one subject reported a 3° shoulder angle when the gold standard was 98°. In this case, it can be assumed that the subject was measuring the angle made by the upper arm and a vertical line extending above the shoulder instead of the angle made by the upper arm and a vertical line extending below the shoulder. Under such circumstances, the gold standard added to the estimated angle approximately equaled the range of motion for the body part in question. The same methodology was applied to both the absolute and algebraic error. All data points greater than 50° off from the gold standard were dropped. This equates to more than four standard deviations and was considered a conservative approach.

2.6.2 Error Calculations for Wrist Postures

Participants did not consistently apply negative signs to indicate wrist flexion versus wrist extension. It was assumed that participants could tell the difference between wrist flexion and wrist extension, but forgot to write down negative signs. For the absolute error, calculated by subtracting the estimate from the absolute value of the gold standard and taking the absolute value of the result, the absolute value of wrist posture was used. For algebraic error, calculated by subtracting the gold standard from the result, negative signs were artificially added to the wrist flexion postures.

2.7 Statistical Analysis

SAS JMP (Version 4.0.4) and Microsoft Excel were used to explore the relationship between the variables and the error associated with angle estimation. A further analysis was conducted to establish: 1) whether or not a relationship existed between error and the size of the body angle, 2) trends in over- or under-estimation (algebraic error), 3) the effects of the independent variables on confidence, and 4) the effects of error on RULA and the Strain Index.

2.7.1 Creating a Model

Nominal categorical variables included body part and joint angles. Video type, field, gender, and training were treated as dichotomous nominal variables, being either "true" or "false". Continuous variables included the recorded confidence, absolute error, and algebraic error. Three models were created, one model each for predicting absolute error, algebraic error, and confidence.

The models were created by first including all five main effects (training, video mode, body part, gender, field) and all two-way interactions. A fit least squares analysis was performed on the model to determine R^2 and F ratio values. Using a Pareto plot of transformed estimates the least significant effects were dropped. The model was then rerun and the R square and F ratios compared. If the values improved, more estimates were dropped using the Pareto plot as guidance. This process continued until the maximum R^2 and F ratio value was achieved for the model.

2.7.2 Testing Assumptions of ANOVA

2.7.2.1 Normality Assumption

A check of the normality assumption was performed using the Shapiro-Wilk's (Goodness of Fit) test and by plotting a histogram of the residuals. If the Shapiro-Wilks test failed to confirm the normality assumption, the histogram of the residuals was examined to determine the nature of the departure from normality. As noted in Montgomery (1997), if the departure from normality appears to come from skewness in the distribution of the residuals this represents a moderate departure from normality and the robustness of the ANOVA procedure to departures from normality will overcome this violation.

2.7.2.2 Independence Assumptions

To ensure that the independence assumption on the errors was not violated participant order was randomized. This ensured that changes in the instructor's style or variations in participants do not cause a change in the error of variance over time. Additionally, the residuals were plotted versus video clip presentation order to ensure that no learning curve effects biased the data.

2.7.2.3 Homogeneity of Variances

To insure that the assumption of homogeneity of variances was not violated, a Bartlett's test was performed on the residuals. If the Bartlett's test failed, the residuals were plotted versus the fitted values to verify that no systematic structure pattern emerged (Montgomery, 1997).

2.7.3 Analysis Process

An analysis of variance (ANOVA) was performed on the model to test the effects of the independent variables (training, video mode, body part, gender, field) on the dependent variables (absolute error, algebraic error, and confidence). An alpha of 0.05 was chosen for all analyses. To test the significance between variable levels, a Tukey-Kramer HSD test was used for pairwise mean comparisons. Least Significant Difference (LSD) means diamonds were used to visually convey the significant trends in the data when concurrent with the Tukey-Kramer HSD test results. Though visually simpler to follow, the LSD test is not as reliable a measure of means comparison as Tukey-Kramer. A means diamond illustrates a sample mean and the 95% confidence interval. The diamond points span the 95% confidence interval for each group while the central line represents the group mean. For groups with equal sample sizes, overlapping marks above or below the group mean indicate that the two group means are not significantly different at the 95% confidence level. The dashed lines above and below the diamonds illustrate the group standard deviation.

In addition to examining participant accuracy, Bartlett's test was used to examine the effects of training on participant precision by testing for significant differences between the trained and control groups' absolute error variances for each body part.

3 Results

3.1 Test of Assumptions of ANOVA

3.1.1 Test of Normality of Residuals

The assessment of the normality assumption using the Shapiro-Wilks test showed a violation of the assumption of normality of residuals for several of the dependent variables. A subsequent graphical analysis of the residuals (See Appendix B) however showed that the distribution had a skewness that was the source of the violation of the assumption of normality. As stated in Section 2.7.2.1 above, this particular situation does not create a concern with regard to the interpretability of the ANOVA results.

3.1.2 Test of Independence

The residuals were then plotted versus order to ensure that there was no correlation between residuals (See Appendix C). No trend in the residuals appears to violate the assumption of equal variance for any of the three models. No learning effects were observed.

3.1.3 Test of Homogeneity of Variances

Residuals were plotted versus the fitted values. For all three models, no obvious patterns emerged (See Appendix D). The assumption of homogeneity of variances was not violated.

3.2 Absolute Error Results

3.2.1 Absolute Error Outlier Removal

Of the original 1920 angle estimates, 27 estimates were dropped. One data point was dropped because no estimate was given. The remaining 26 data points were labeled outliers. As seen in Table 3.1, angle error ranges from 0.0° to 49.0° for the remaining 1896 points. Statistical analysis established a median score of 6.0° with an upper quartile (75^{th}%ile) of 11.0° and a lower quartile (25^{th}%ile) of 3.0°. The overall average angle error was 8.0° with a standard deviation of 7.5°.

Table 3.1: Range of angle error

Measure	All Body Parts
Sample Size	1893
Maximum	49.0
75 % Quartile	11.0
Median	6.0
25 %Quartile	3.0
Minimum	0.0
Mean	8.0
Standard Deviation	7.5

3.2.2 Absolute Error Prediction

The absolute error model included five main effects (training, video mode, body part, gender, field) and all two-way interactions. With all main effects and all two-way interactions included, the model's adjusted R^2 value was 0.028 and the model's F ratio was 3.75. Based on the results of a Pareto plot and the effects test F ratios, all effects were removed (0.05 cut off criterion) except: body part, gender, and video mode*body part. After removing the variables and rerunning the model, the adjusted R^2 value dropped to 0.027 but the F-ratio increased to 11.44. The adjusted R^2 value indicates that the model is not viable for predicting error. However, the F-ratio indicates that the components of the model significantly correlate with error.

3.2.3 Absolute Error ANOVA Results

An ANOVA for the dependent variable, absolute error, was performed. Table 3.2 lists the results of the ANOVA including the F ratio and the corresponding p-value. Figure 3.1 illustrates the significant interaction effect, video mode*body part.

Table 3.2: Significant results from ANOVA for absolute error

Grouping	Analysis of Variance	
	F-Ratio	Probability > F
Gender	5.0	0.0251
Body Part	22.9	<.0001
Video Mode*Body Part	3.1	0.0477

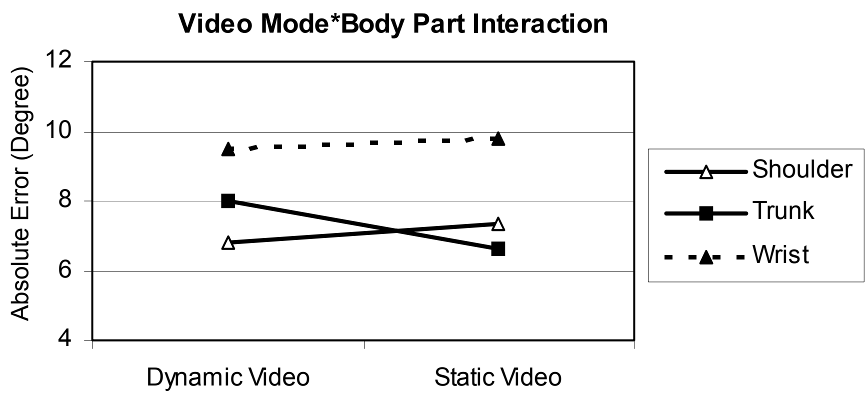


Figure 3.1: Significant interaction effect video mode*body part for absolute error

A Tukey-Kramer HSD test was used to determine the significance between effect levels. Tables 3.3 and 3.4 give the Tukey-Kramer HSD results for the two main effects, gender and body part. Positive values show pairs of means that are significantly different at an alpha of 0.05. The mean error for females, 8.4° (standard deviation 7.8°), was significantly higher then the mean error of 7.6° (7.0°) found for males. The LSD means diamonds in Figure 3.2 are used to visually illustrate the same findings found by a Tukey-Kramer HSD test.

Table 3.3: Means comparison of absolute error by body part

	Wrist	Trunk	Shoulder
Wrist	-	+	+
Trunk	+	-	-
Shoulder	+	-	-

Table 3.4: Means comparison of absolute error by gender

	Male	Female
Male	-	+
Female	+	-

Wrist was significantly different from the trunk and shoulder but the trunk was not significantly different from the shoulder. As seen in Table 3.5, a mean absolute error of 7.0°, 7.3°, and 9.6° (6.1°, 6.8°, 8.7°) resulted for the shoulder, trunk, and wrist, respectively. The mean error associated with the wrist is over 2° greater than the mean error associated with the shoulder and trunk.

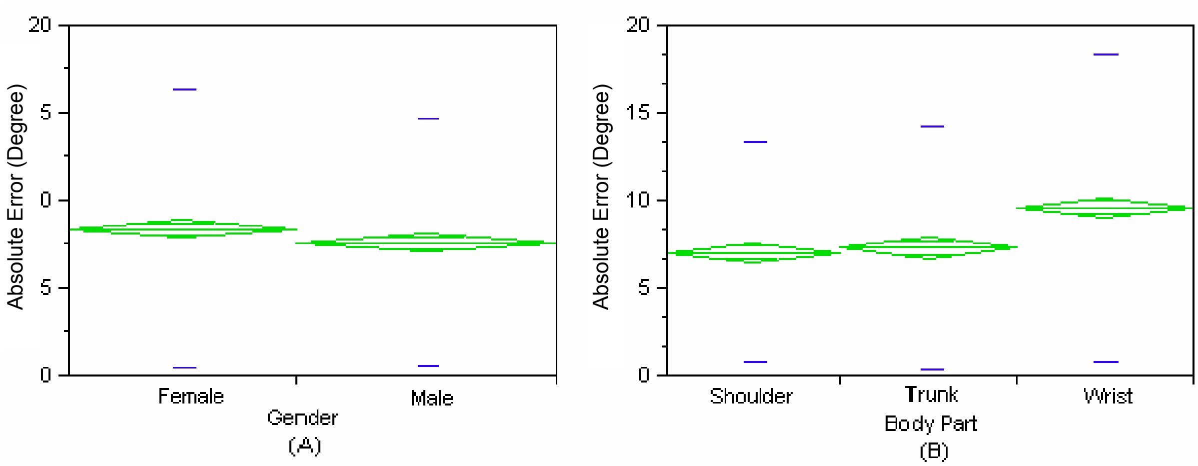


Figure 3.2: Means comparison for (a) gender and (b) body part with absolute error

Table 3.5: Means for significant main effects for absolute error

Grouping		Sample Size	Mean	Standard Deviation	Std Err Mean	Lower 95%	Upper 95%
Gender	Female	889	8.4	7.8	0.3	7.8	8.9
	Male	1004	7.6	7.0	0.2	7.2	8.0
Body Part	Shoulder	627	7.0	6.1	0.2	6.5	7.5
	Trunk	634	7.3	6.8	0.3	6.7	7.8
	Wrist	632	9.6	8.7	0.3	8.9	10.3

The interaction between body part and video mode was significant ($p = 0.047$). However, the Tukey-Kramer test (Table 3.6) indicates that only trunk means are significantly different between video modes. As seen in Table 3.7, dynamic trunk error, 8.0° (7.0°), was significantly greater than static trunk error, 6.7° (6.8°).

Table 3.6: Means comparison body part (trunk)*video mode and absolute error

For Trunk	Static	Dynamic
Static	-	+
Dynamic	+	-

Table 3.7: Means for significant interaction effect for absolute error

Grouping		Sample Size	Mean	Standard Deviation	Std Err Mean	Lower 95%	Upper 95%
Dynamic Video	Shoulder	314	6.8	6.5	0.4	6.1	7.5
	Trunk	319	8.0	7.0	0.4	7.3	8.8
	Wrist	319	9.5	8.3	0.5	8.6	10.4
Static Video	Shoulder	314	7.3	6.0	0.3	6.7	8.0
	Trunk	316	6.7	6.8	0.4	5.9	7.4
	Wrist	314	9.8	9.2	0.5	8.7	10.8

3.2.4 Gold Standard Angle and Absolute Angle Error

As seen in Table 3.8, an analysis of variance indicates significant differences between the particular angle being observed (gold standard angle) and the resulting error for the shoulder, trunk, and wrist angles. This analysis was performed separately for each body part described above because the gold standard values were different for different body regions (See Table 2.4).

Table 3.8: ANOVA results for absolute error (Gold standard value as the independent variable. Analysis by body region.)

Body Part	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Probability > F
Shoulder	626	1760.6	195.6	5.7	<.0001
Trunk	633	1648.8	183.2	4.1	<.0001
Wrist	631	11361.9	1420.2	24.5	<.0001

Mean absolute error varies based on the angle being estimated. The three LSD means diamonds shown in Figure 3.3 are used to visually illustrate the same findings found by a Tukey-Kramer HSD test shown in Table 3.9 through Table 3.11. All positive pairs (shaded cells) indicate results significantly different at an alpha of 0.05. Error peaks between 45° and 90° for the shoulder estimates and is lowest around 0° and 90°. Trunk error is also lowest around 0° and 90°. Wrist error increases as the wrist deviates from 0° peaking at the maximum flexion and extension angles. However, the peak wrist flexion is the most significant difference.

Table 3.9: Tukey-Kramer HSD results for the shoulder with absolute error

	65	102	74	51	98	43	20	32	14	83
65	-	-	-	-	-	+	+	+	+	+
102	-	-	-	-	-	-	-	-	+	+
74	-	-	-	-	-	-	-	-	-	-
51	-	-	-	-	-	-	-	-	-	-
98	-	-	-	-	-	-	-	-	-	-
43	+	-	-	-	-	-	-	-	-	-
20	+	-	-	-	-	-	-	-	-	-
32	+	-	-	-	-	-	-	-	-	-
14	+	+	-	-	-	-	-	-	-	-
83	+	+	-	-	-	-	-	-	-	-

Table 3.10: Tukey-Kramer HSD results for the trunk with absolute error

	67	10	60	51	42	35	98	81	3	88
67	-	-	-	-	-	-	-	+	+	+
10	-	-	-	-	-	-	-	-	+	+
60	-	-	-	-	-	-	-	-	-	+
51	-	-	-	-	-	-	-	-	-	-
42	-	-	-	-	-	-	-	-	-	-
35	-	-	-	-	-	-	-	-	-	-
98	-	-	-	-	-	-	-	-	-	-
81	+	-	-	-	-	-	-	-	-	-
3	+	+	-	-	-	-	-	-	-	-
88	+	+	+	-	-	-	-	-	-	-

Table 3.11: Tukey-Kramer HSD results for the wrist with absolute error

	-58	52	-43	31	20	42	-27	-18	-5	5
-58	-	+	+	+	+	+	+	+	+	+
52	+	-	-	-	-	-	+	+	+	+
-43	+	-	-	-	-	-	-	+	+	+
31	+	-	-	-	-	-	-	-	+	+
20	+	-	-	-	-	-	-	-	+	+
42	+	-	-	-	-	-	-	-	+	+
-27	+	+	-	-	-	-	-	-	-	-
-18	+	+	+	-	-	-	-	-	-	-
-5	+	+	+	+	+	+	-	-	-	-
5	+	+	+	+	+	+	-	-	-	-

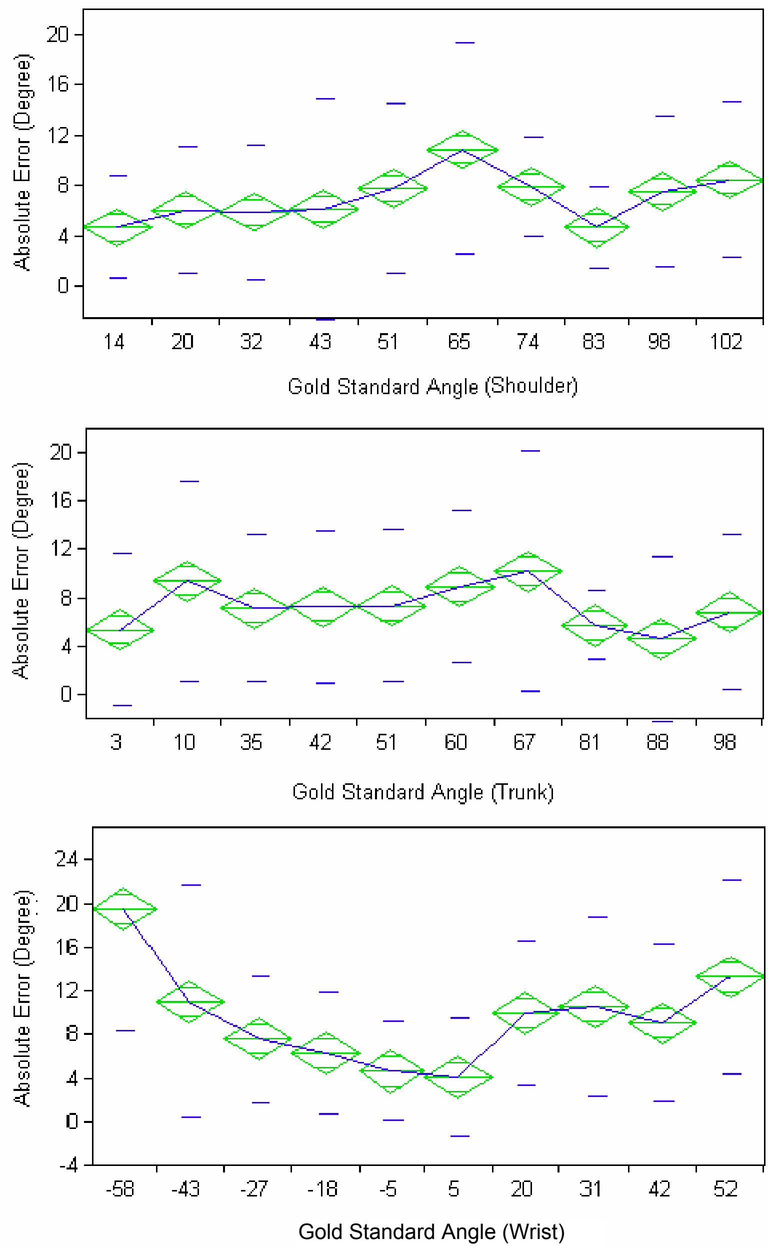


Figure 3.3: Gold standard angle by body part for absolute error

3.2.5 Absolute Error Learning Curve Effects

Comparison for all mean pairs of absolute error based on order using Tukey-Kramer HSD indicates that no mean pairs are significantly different. No learning curve effects were observed.

3.2.6 Precision and Absolute Error

A Bartlett's test was performed on the data to determine significance between the variances based on training. Overall, there was no significant difference between variances for the two groups ($p=0.1600$). However, when the data was broken down by body part, there were significant differences between variances for the shoulder ($p=0.0071$) and the trunk ($p=0.0002$). For the shoulder, the mean absolute error was 7.4° (6.7°) and 6.8° (5.8°) for the trained and control group, respectively. For the trunk, the mean absolute error was 0.7° (4.8°) and 5.8° (4.2°) for the trained and control group, respectively.

3.3 Algebraic Angle Error

3.3.1 Algebraic Error Outlier Removal

Of the original 1920 angle estimates, 23 estimates were dropped. One data point was dropped because no estimate was given. The remaining 22 data points were labeled outliers. As seen in Table 3.12, angle error ranges from -49° to 48° for the remaining 1897 points. Statistical analysis established a median score of 0° with an upper quartile (75^{th}%ile) of 5.0° and a lower quartile (25^{th}%ile) of -7.0°. The overall average angle error was –0.3° with a standard deviation of 11.0°.

Table 3.12: Range of algebraic error

Measure	All Body Parts
Sample Size	1897
Maximum	48.0
75 % Quartile	5.0
Median	0.0
25 %Quartile	-7.0
Minimum	-49.0
Mean	-0.3
Standard Deviation	11.0

3.3.2 Algebraic Error Prediction

The absolute error model included five main effects (training, video mode, body part, gender, field) and all two-way interactions. With all main effects and two-way interactions included, the model's adjusted R^2 value was 0.058 and the model's F ratio was 5.39. Based on the results of a Pareto plot and the effects test F ratios, all effects were removed except: body part and field*training. After removing the variables and rerunning the model the adjusted R^2 value dropped at 0.056 but the F-ratio increased to 38.40. The adjusted R^2 value indicates that the model is not viable for predicting error. However, the F-ratio indicates that the components of the model significantly correlate with error.

3.3.3 Algebraic Error ANOVA Results

An ANOVA for the dependent variable, algebraic error, was performed. Table 3.13 lists the results of the ANOVA, including the F ratio and the corresponding p-value. The main effect body part and the interaction effect field*training were significant.

Table 3.13: Analysis of variance with algebraic error

Grouping	Analysis of Variance	
	F-Ratio	Probability > F
Body Part	44.7	<.0001
Field*Training	25.7	<.0001

A Tukey-Kramer HSD test was used to determine the significance between effect levels. As seen from the Tukey-Kramer HSD results in Tables 3.14, all three body parts are significantly different from one another. A mean error of –3.0°, -0.5°, and 2.7° (9.0°, 10.1°, 12.8°) resulted for the shoulder, trunk, and wrist, respectively. Figure 3.4 shows the LSD means diamonds and standard deviations for body part.

Table 3.14: Means Comparison for Body Part and algebraic error

	Wrist	Trunk	Shoulder
Wrist	-	+	+
Trunk	+	-	+
Shoulder	+	+	-

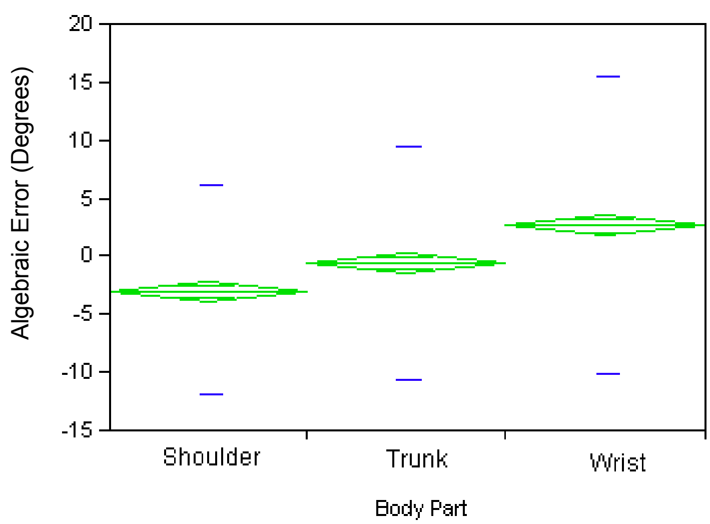


Figure 3.4: Algebraic error by body part

There was also a significant interaction between field and training (p=<0.0001). As seen by the Tukey-Kramer HSD, Table 3.15, there is only a significant difference between field for trained participants. For control participants, the average algebraic error was 1.0° (10.7°) and –0.3° (11.1°) for participants in technical and non-technical fields, respectively. For trained participants, the average algebraic error was -2.2° (10.6°) for participants in technical fields and only 0.4° (12.0°) for participants in non-technical fields. This interaction effect is illustrated in Figure 3.5.

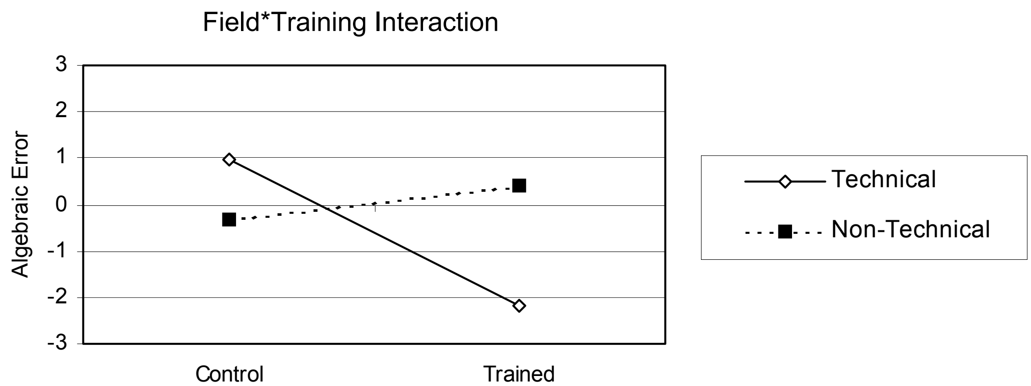


Figure 3.5: Algebraic error for field*training interaction

Table 3.15: Means Comparison for interaction effect field*training with algebraic error

Control	Non-Technical	Technical
Non-Technical	-	-
Technical	-	-

Trained	Non-Technical	Technical
Non-Technical	-	+
Technical	+	-

3.3.4 Gold Standard Angle and Algebraic Error

The mean absolute error varies based on the angle being estimated. The LSD means diamonds in Figure 3.6 are used to visually illustrate the same findings found by a Tukey-Kramer HSD test shown in Table 3.16 through 3.18. All positive pairs (shaded cells) indicate results significantly different at an alpha of 0.05. Shoulder angles greater than 50° are underestimated with the greatest mean algebraic error occurring around 74°. The mid-range trunk angles are overestimated while the end-range angles close to 0° and 90° are underestimated. The most severely underestimated angle was 10°. Wrist angles were severely overestimated at 58° flexion and slightly overestimated at 43° flexion. Wrist angles were slightly underestimated around 20° extension.

Table 3.16: Tukey-Kramer HSD results for the shoulder with algebraic error

	32	20	51	14	43	83	98	102	65	74
32	-	-	-	-	-	-	+	+	+	+
20	-	-	-	-	-	-	+	+	+	+
51	-	-	-	-	-	-	+	+	+	+
14	-	-	-	-	-	-	-	+	+	+
43	-	-	-	-	-	-	-	-	+	+
83	-	-	-	-	-	-	-	-	-	-
98	+	+	+	-	-	-	-	-	-	-
102	+	+	+	+	-	-	-	-	-	-
65	+	+	+	+	+	-	-	-	-	-
74	+	+	+	+	+	-	-	-	-	-

Table 3.17: Tukey-Kramer HSD results for the trunk with algebraic error

	35	60	51	42	67	88	98	81	3	10
35	-	-	-	-	+	+	+	+	+	+
60	-	-	-	-	+	+	+	+	+	+
51	-	-	-	-	+	+	+	+	+	+
42	-	-	-	-	+	+	+	+	+	+
67	+	+	+	+	-	-	-	-	-	+
88	+	+	+	+	-	-	-	-	-	+
98	+	+	+	+	-	-	-	-	-	-
81	+	+	+	+	-	-	-	-	-	-
3	+	+	+	+	-	-	-	-	-	-
10	+	+	+	+	+	+	-	-	-	-

Table 3.18: Tukey-Kramer HSD results for the wrist with algebraic error

	-58	-43	42	-5	52	31	-18	-27	5	20
-58	-	+	+	+	+	+	+	+	+	+
-43	+	-	-	-	-	-	-	-	+	+
42	+	-	-	-	-	-	-	-	-	-
-5	+	-	-	-	-	-	-	-	-	-
52	+	-	-	-	-	-	-	-	-	-
31	+	-	-	-	-	-	-	-	-	-
-18	+	-	-	-	-	-	-	-	-	-
-27	+	-	-	-	-	-	-	-	-	-
5	+	+	-	-	-	-	-	-	-	-
20	+	+	-	-	-	-	-	-	-	-

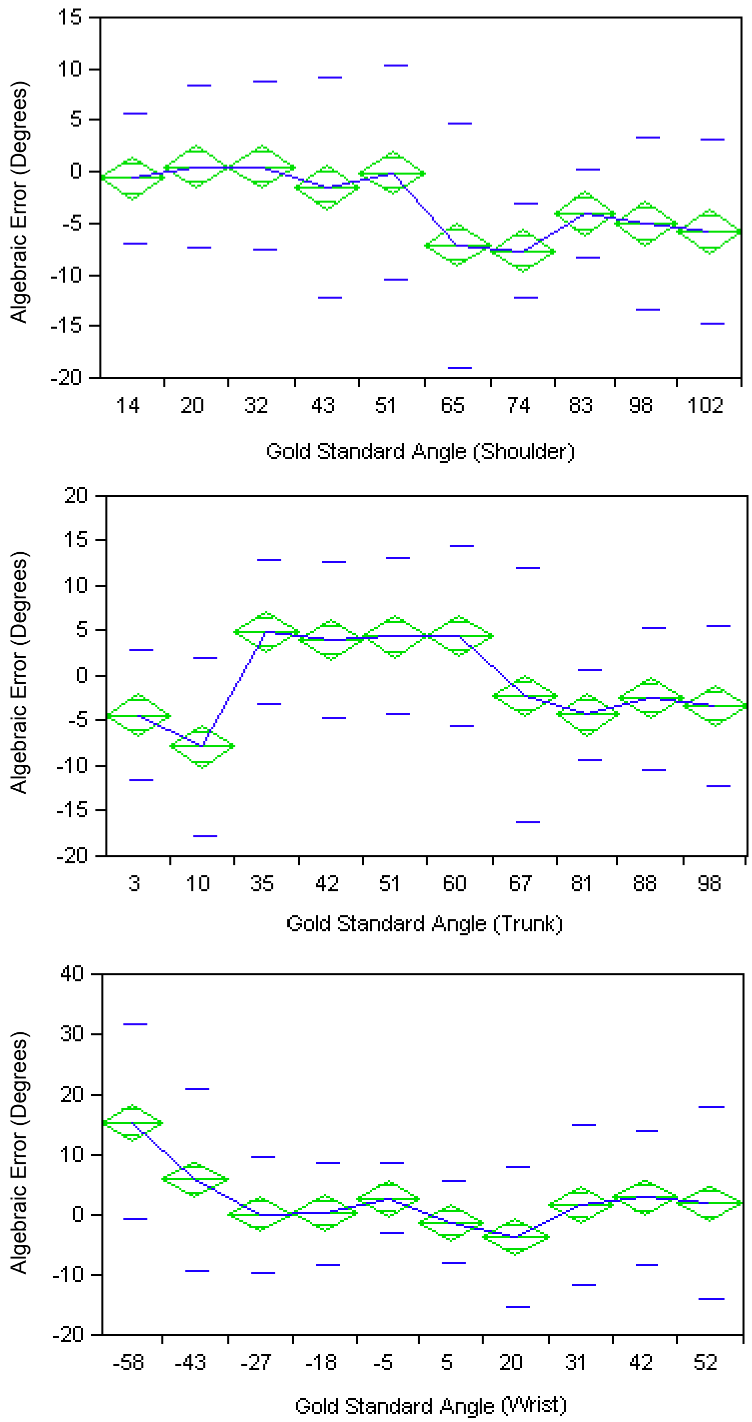


Figure 3.6: Algebraic error versus gold standard angles by body part

3.3.5 Risk Analysis Tools and Angle Error

The algebraic error was used to determine the effects of visual estimation error on the risk analysis tools, RULA and Strain Index. The number of misclassified participant answers based on RULA scoring are listed in Tables 3.19 through 3.21. In total, 8.3%, 12.4%, and 14.6% of answers for the shoulder, trunk, and wrist, respectively, would have altered the RULA score by at least one point. However, the gold standard angles used in this research are not evenly distributed across the risk analysis tool categories. This should be taken into consideration when interpreting the following tables.

Table 3.19: Misclassifications of shoulder postures with RULA

RULA Category	≤20°	20-45°	45-90°	>90°	Out of
RULA Score	1	2	3	4	
Number of Gold Standard Angles Falling in This Range	2	2	4	2	
14°		7	0	0	64
20°		0	0	0	63
32°	7		0	0	64
43°	0		15	1	64
51°	0	10		0	64
65°	1	1		1	64
74°	0	0		0	63
83°	0	0		2	59
98°	0	0	6		61
102°	0	0	1		62
Total Missclassified	**8**	**18**	**22**	**4**	**628**
% of Total Missclassified	**1.3%**	**2.9%**	**3.5%**	**0.6%**	**8.3%**

Table 3.20: Misclassifications of trunk postures with RULA

RULA Category	0°	0-20°	20-60°	>60°	Out of
RULA Score	1	2	3	4	
Number of Gold Standard Angles Falling in This Range	0	2	4	4	
3	12		5	0	64
10	1		21	0	53
35	0	1		1	64
42	0	3		1	64
51	0	0		3	64
60	0	0		13	64
67	0	2	5		64
81	0	0	0		61
88	0	0	1		64
98	0	0	0		63
Total Missclassified	**14**	**8**	**35**	**22**	**635**
% of Total Missclassified	**2.2%**	**1.3%**	**5.5%**	**3.5%**	**12.4%**

Table 3.21: Misclassifications of wrist postures with RULA

RULA Category	0°	0-15°	>15°	Out of
RULA Score	1	2	3	
Number of Gold Standard Angles Falling in This Range	0	2	8	
- 58	1	1		61
- 43	1	2		64
- 27	0	6		64
-18	0	18		64
- 5	6		5	61
5	14		2	64
20	0	14		64
31	0	8		64
42	0	0		64
52	0	0		64
Total Missclassified	**22**	**49**	**7**	**634**
% of Total Missclassified	**4.1%**	**9.2%**	**1.3%**	**14.6%**

The angle ranges are narrower for the Strain Index therefore, the percentage of misclassifications is even higher than with RULA. As seen in Table 3.22, 40.2% of wrist postures would have been given the wrong Strain Index score. The majority of misclassifications occurred for angles near the range cut-offs. For instance, almost half of wrist postures of -18° were placed into the < -15° range instead of the -16° to -30° range. The greatest number of misclassifications occurred for a Strain Index score of 1.0 and 1.5.

Table 3.22: Misclassifications of wrist postures with the Strain Index

Wrist Extension	0° - 25°	26° - 40°	41° - 55°*	> 55°*	Out of
Wrist Flexion	0° - 15°	16° - 30°	31° - 50°	> 50°	
Strain Index Score	1	1.5	2	3	
Number of Gold Standard Angles Falling in This Range	3	3	3	1	
- 58	1	1	9		61
- 43	1	2		16	64
- 27	9		2	0	64
-18	30		1	0	64
- 5		5	0	0	61
5		1	1	0	64
20		25	17	0	64
31	27		23	1	64
42	9	32		3	64
52	5	13		21	64
Total Missclassified	**82**	**79**	**53**	**41**	**634**
% of Total Missclassified	**12.9%**	**12.5%**	**8.4%**	**6.5%**	**40.2%**

* Assumed typing error in original source lists range as 41°-55° and >60°

3.4 Confidence

3.4.1 Confidence Data

Of the original 1920 angle estimates, three estimates were left blank by subjects and thus were not included in the analysis. As seen in Table 3.23, confidence ranged from 0 to 20° for the remaining 1917 points. Statistical analysis established a median score of 5° with an upper quartile (75th%ile) of 5.0° and a lower quartile (25th%ile) of 3.0°. The overall average angle error was 4.5° with a standard deviation of 2.3°.

Table 3.23: Range of confidence

Measure	All Body Parts
Sample Size	1917
Maximum	20.0
75 % Quartile	5.0
Median	5.0
25 %Quartile	3.0
Minimum	0.0
Mean	4.5
Std Deviation	2.3

3.4.2 Confidence Prediction

The absolute error model included five main effects (training, gender, field, video mode, and body part) and all two-way interactions. A Fit Least Squares test was then performed on the model. With all main effects and all two-way interactions included the adjusted R^2 value of 0.20 and an F ratio of 20.25. The Effects Tests and Pareto Diagram indicated that all factors except for the following should be eliminated from the model: field, gender, training, video mode, field*training, and field*gender. After removing the variables and rerunning the model, the adjusted R^2 value remained the same but the F-ratio increased to 76.16. The R^2 value indicates that the model is not viable for predicting error. However, the F-ratio indicates that the components of the model significantly correlate with confidence.

3.4.3 Confidence ANOVA Results

Table 3.24 lists the results on the analysis of variance for the dependent variable confidence. The main effects, field, gender, training, and video mode and the interaction effects, field*training and field*gender were significant.

Table 3.24: Trends with confidence

Grouping	Analysis of Variance	
	F-Ratio	Prob > F
Field	98.5	<.0001
Gender	117.6	<.0001
Training	20.1	<.0001
Video Mode	12.5	0.0004
Field*Gender	37.5	<.0001
Field*Training	82.5	<.0001

Confidence increased for participants in technical fields, males, trained participants, and with static video. Figure 3.7 shows the LSD means diamonds and standard deviations for the significant main effects. A Tukey-Kramer HSD test was used to determine the significance between effect levels for the interaction effects, field*gender and field*training, as seen in Table 3.14 and 3.15 and graphically show in Figure 3.8. Confidence increased with training for participants in technical fields only. Though participants in technical fields had similar confidence regardless of gender, male participants in non-technical fields were more confident that female participants in non-technical fields. In addition, the confidence did not vary based on the angle being estimated. No trends appeared in plots of the gold standard angle versus confidence.

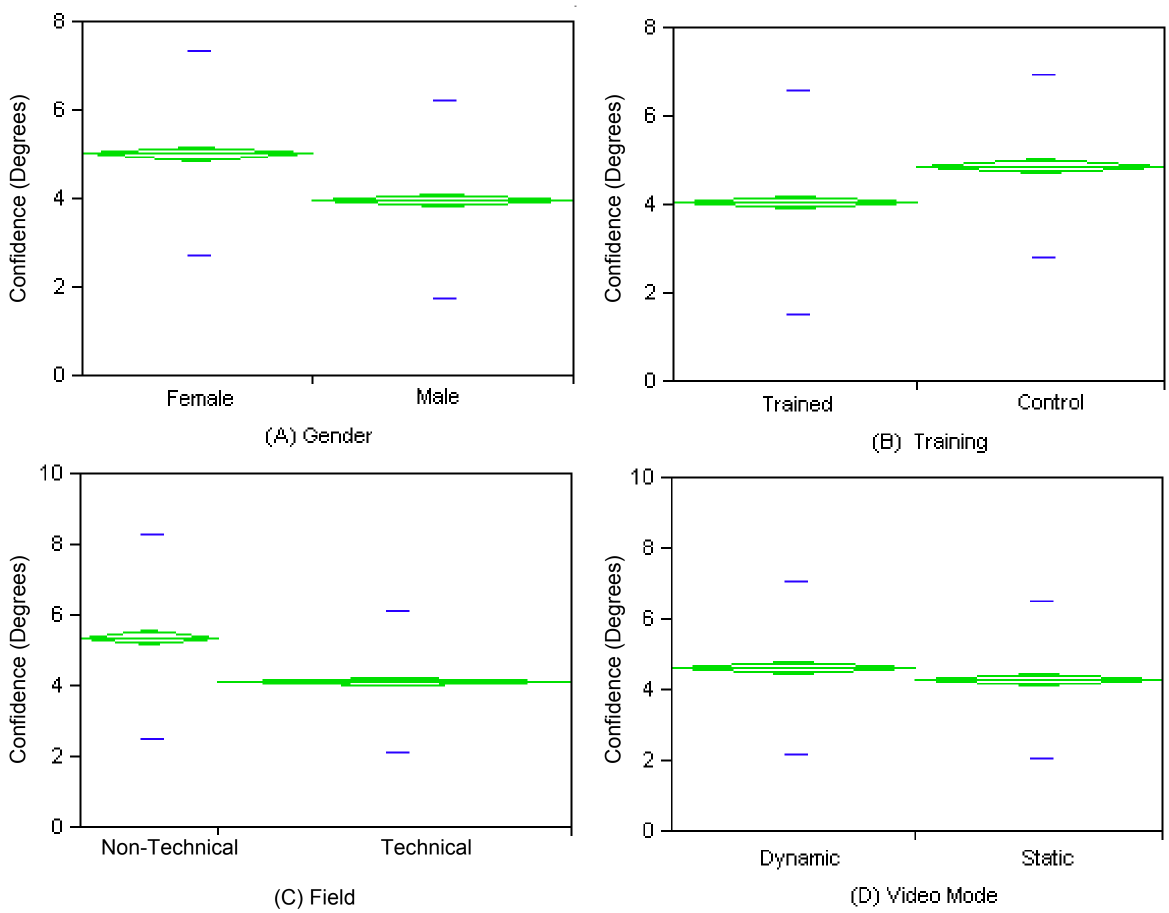


Figure 3.7: Significant main effects (A) gender (B) training (C) field and (D) video mode with confidence

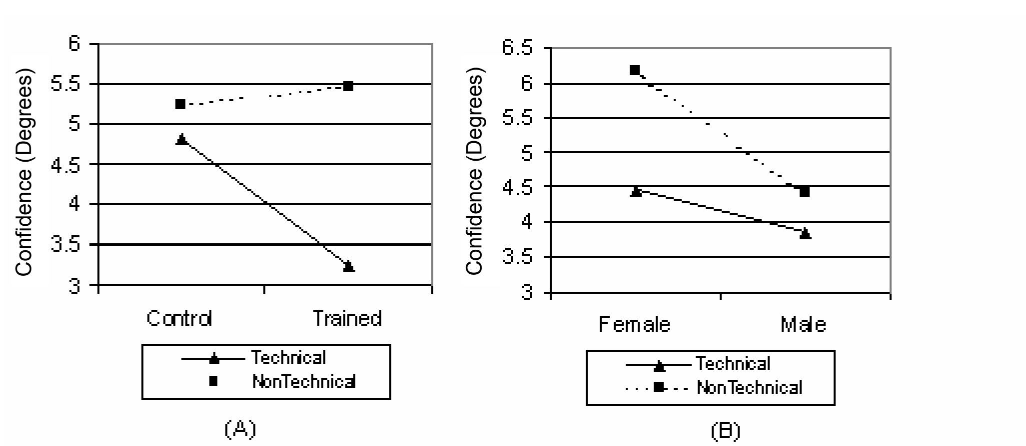


Figure 3.8: Significant two-way interactions (a) field*training and (c) field*gender with confidence

Table 3.25: Significant interaction effect field*gender with confidence

Male	Non-Technical	Technical
Non-Technical	-	+
Technical	+	-

Female	Non-Technical	Technical
Non-Technical	-	+
Technical	+	-

Table 3.26: Significant interaction effect field*training with confidence

Trained	Non-Technical	Technical
Non-Technical	-	+
Technical	+	-

Control	Non-Technical	Technical
Non-Technical	-	+
Technical	+	-

4 Discussion

The objectives of this study were to 1) assess people's ability to visually estimate postural angles and 2) determine the effect of training and other observer characteristics on an observer's ability to visually estimate the postural angles of the trunk, shoulder, and wrist. This process consisted of several steps. First, a simple methodology for measuring postural angles was developed based on the data requirements for the Strain Index, RULA, and REBA. Next, a twenty-minute training program was designed to help improve participant accuracy at postural angle estimation. Third, test materials were developed and administered. Finally, the resulting data were analyzed for trends in angle error estimation and confidence in those estimations relative to the important variables under consideration.

The first hypothesis proposed that training would reduce errors in angle estimations. However, training was found to have no significant effect on angle estimation error. In addition, comparison of variances showed that untrained participants were more precise than trained participants. The lack of positive training effects suggests one of three conclusions: training was not long enough, the training technique was inadequate, or training does not improve performance. Previous researchers of visual estimation techniques trained observers for up to three days (Douwes and Dul, 1991) or used highly experienced observers (Ericson et al., 1991). However, since these researchers did not use a control group, specific effects on performance are not known. Waller and Wright (1965) found that only thirty minutes of training with flashcards was needed to improve the accuracy of navigators. However, black and white angles drawn on flashcards are visually simpler then the task required of observers in this experiment.

The training technique and materials may not be adequate to improve performance. The adequacy of the training materials used in this experiment is difficult to determine. Researchers of visual estimation used photographs and video taped recordings of work tasks to train participants. However, no evaluations of postural estimation training were found in the literature. It is assumed that the training is appropriate, since the training required the participants to perform the same activities that they would in the experiment.

Concluding that any amount or type of training will not have an effect on performance seems unreasonable. Though the task was simple, Waller and Wright (1965) were able to improve performance with training by 1.34°. However, the task of visually estimating postural angles is complex. Therefore, the amount of training needed to improve visual estimation may be unreasonable for most practical applications.

Though field was not properly balanced in this experiment, the variable illustrates an interesting trend that should be examined in future research. When examining algebraic error, non-technical participants performed similarly regardless of training. However, trained participants from technical fields performed worse, tending to underestimate body angles, than control group participants from technical fields. Training appeared to have some effects on technical participants abilities.

The second hypothesis stated that higher errors would occur under dynamic video conditions when compared to static video conditions. No significant effects were found between static and dynamic conditions. However, an interaction existed between video mode and body part but only for the trunk. There was a 1.4° improvement in absolute error between means for static trunk video over dynamic trunk video. The mean absolute error was 8.02° (7.00°) and 6.65° (6.83°) for dynamic video and static video, respectively. No video mode effects were found for algebraic error. Ericson et al. (1991) found an increase in algebraic error of 11° with dynamic video. However, this shift only occurred for the shoulder and not for the trunk or neck. Though not significant, mean error was actually slightly lower for dynamic shoulder and wrist video.

No easy explanation exists to explain why stronger video mode effects were not observed. However, the dynamic video shown to participants was very predictable and repetitive. Faster, more varied movements may have made dynamic video estimation more difficult. Additionally, the participants focused on the maximum angle, which formed a distinct moment in the video. If the participants were asked to estimate the average angle that occurred or the mid-point angle, error may have been worse. Other methodologies attempt to capture the time dependency of posture analysis (Corelett et al., 1980). However, these methodologies have not been evaluated for associated errors.

The third hypothesis focused on the differences in angle estimation ability between men and women. As hypothesized, females performed worse than male participants. Overall, females' mean absolute error was 8.35° (7.75°) and males' mean error was 7.62° (6.96°). However, this is less than 1° difference between means. Therefore, gender difference appears to be of little practical consequence.

Though the two-way interaction between gender and body part was not significant, females tended to perform worse on wrist postures. There was a 1.5° difference between male and female average absolute error wrist estimations. Interestingly, the two-way interaction between gender*field was also not significant. Therefore, technical females did not perform significantly better than technical males as suggested in previous literature (Quaiser-Pohl and Lehmann, 2002).

The fourth hypothesis considered the effect that the specific joint under consideration would have on angle estimation error. As hypothesized, error varied depending on body part. The absolute error was 7.01° (6.05°), 7.28° (6.81°), and 9.59° (8.67°) for the shoulder, trunk, and wrist, respectively. However, only the differences between the shoulder and wrist and between the trunk and wrist were significant. The mean absolute error was 2.5° more for the wrist than the other body parts.

These findings agree with previous research that found that the mean percentage of subjects correctly categorizing wrist postures was 8 percentage points lower than shoulder postures categorization and 21 percentage points lower than lower back postures categorization (Baluyut et al., 1995). This increase in error associated with the wrist may be caused by the lack of clearly visible landmarks on the hand and wrist. Though the video clips focused on the specific body part, the wrist is still smaller and the target wrist angles come from a smaller range of motion. In addition, trunk and shoulder postures are long and more closely resemble lines drawn on paper. Since the wrist has a smaller range of motion than the other body parts observed, the size of the body part may affect the ease of visual estimation.

Trends in over and underestimation also varied based on body part. The mean algebraic error was –2.96° (8.98°), -0.54° (10.09°), and 2.72° (12.84°) for the shoulder, trunk, and wrist, respectively. The body part algebraic errors were significantly different from one another. While the shoulder and trunk were underestimated, the wrist was overestimated.

In addition to the effects that directly applied to the stated hypotheses there were also some other interesting effects that are worthy of discussion. Overall, the target angle was found to affect accuracy. For the trunk and the shoulder, angles around 0° and 90° tended to be associated with less error. These trends agree with previous research, which found less error occurred with 0° and 90° angles (Ericson et al., 1991). For the wrist, the lowest error occurs near zero and increases monotonically as the wrist flexion increased. A similar trend occurred as the wrist extension angle increased. Such trends in error suggest that participants are more comfortable with what a 0°, 45° and 90° angle should look like and therefore these angles serve as anchor points when judging other angles. However, participants' reported confidence did not vary based on the angle being observed.

Trends in over- and underestimation also appear in relation to the size of the angle being observed. The largest shoulder and trunk angles are underestimated. This finding contradicts previous research that found shoulder flexion was most often slightly overestimated (Ericson et al., 1991). These researchers also found a link between direction of movement and estimation. Underestimation tended to occur during flexion while overestimation occurred during extension for the trunk and neck.

It is interesting to note the effects of the errors found in the current study on the RULA scores. With mean error rates spanning from approximately 1° to 14° (median ~7°) for the shoulder and trunk and from 1° to 18° (median ~10°) for the wrist, it is likely that error will increase or decrease the score by one category. In fact, 8%, 12%, and 14% of shoulder, trunk, and wrist answers, respectively, shifted the RULA score by one point. Users should be aware of such limitations when using RULA or other subjective posture analysis tools.

The Strain Index was even more sensitive to angle error of the wrist with participants misclassifying 40 percent of postures. Since the wrist is split into four differently scored categories instead of three, angle ranges are narrower making it more likely for a moderate angle to become extreme or vice versa. The Strain Index scores are multiplied together to reach the final index, this error could change the index score by two-thirds.

There were a number of interesting trends in participant reported confidence. Confidence did not increase over time as might be expected. However, the main effects of training, gender, field, and video type were all significant. Trained individuals were more confident though they did not perform any better than their control group counter parts. Not surprisingly, participants from technical fields were more confident with the activity than participants in non-technical fields. However, except for in the incidence of algebraic error, field did not result in any significant differences in error between participants. This suggests that confidence was more of an indicator of perception of ability than an assessment of actual ability.

Females judged themselves 2° less confident on average than males. In this case, the participants' lack of confidence was supported by the absolute error. Females were significantly worse than males, though the difference was slight. Interestingly enough, technical females overcame this hesitancy in their abilities and were nearly as confident as males. However, technical females were still not as confident as technical males and did not perform any better in reality than non-technical females. It is assumed that individuals with more of a background in math and science are not as intimidated by math-related exercises.

Again not surprisingly, overall, participants were more confident about static video estimation than dynamic. Dynamic video added an additional element of movement to the task that forced participants to have a strong sense of how the postural angles are formed and measured. However, again confidence is more of an indicator of perception of ability than an assessment of ability, since video mode did not significantly affect accuracy.

The procedure used to assess angle estimation in this research was intentionally oversimplified when compared to a real world setting. The video model wore tight fitting clothing that allowed the muscles and bony landmarks of the limbs to be viewed clearly. Under less ideal situations, including deviated viewing positions, baggy clothing, poor lighting, faster speeds, and more variation of movement, the angle error is only expected to increase. Future research should verify the use of posture analysis tools under less idealistic circumstances.

In addition, algebraic error findings for the wrist should be examined further. Since so many participants failed to indicate wrist flexion during testing, negative signs had to be artificially inserted before processing the algebraic error data. It was assumed that participants could tell the difference between wrist flexion and extension. However this may not be the case, participants may not be able to tell the difference between flexion and extension further increasing the algebraic error associated with the wrist.

Confidence findings should also be interpreted with care. Since a constricted five or ten point scale was not used, confidence results may be biased by the confidence range selected by each individual. However, participants seemed to naturally place a cap of 10 degrees error on their answers.

The current research raises additional questions for future study. First, the research employed only one methodology for measuring postures. Other methodologies may exist that are easier to use and thus would reduce estimation error. In addition, participants may benefit from longer training times or different training styles. Future research should also examine the long-term effects of training. In Waller and Wright's (1965) navigator study, training effects lasted for at least 1 week after the initial training. Ericson et al. (1991) found no significant differences between estimations made two weeks apart. Time effects on error were not examined in the current study.

5 Conclusions

In summary, this research explored the error associated with simple posture angle estimation of the shoulder, trunk, and wrist. This research has determined that visual estimation error can be expected to range between 7 and 10 degrees. Such estimation error is high given that the experiment was conducted under idealized circumstances. Clothing distortions, viewing angle deviations, and movement variability were all minimized to capture angle estimation error under the best possible circumstances. It can be assumed that angle error will increase in real world scenarios.

Due to estimation errors, analysis tools that include posture are less accurate in predicting worker risk. Error shifted RULA scores by one point while the Strain Index scores could change as much as a two-thirds. Current posture components only increase the overall error associated with the tools. Either posture should be removed from risk analysis tools or the way in which posture is recorded should be modified. However, since posture, especially when combined with movement and force, is a primary suspect for causing MSDs in the workplace the former suggestion does not seem reasonable.

Future research should focus on ways of improving postural angle estimation by either changing the angle estimation methodology, improved training, or some other yet unidentified factor. No strong predictors of both absolute and algebraic error were found other than body part in this research. This suggests that some other independent variable exists that may explain why and when visual estimation errors will occur. In the mean time, since risk assessment tool users are unlikely to take on the expense of more objective posture estimation techniques, these tools should be modified to accommodate the error that occurs with subjective estimation of postural angles.

References

Amponsah, B. A. (2000). Comparison of sex differences in visual-spatial performance from preadolescence to adulthood in Ghana and Norway. *South African Journal of Psychology*, 30 (4): 25-31.

Armstrong, T.J., Foulke, J.A., Joseph, B.S., and Goldstein, S.A. (1982). Investigation of cumulative trauma disorders in a poultry processing plant. *American Industrial Hygiene Association Journal*, 43: 103-116.

Arthur, E.J., Hancock, P.A., AND Chrysler, S.T. (1997). The perception of spatial layout in real and virtual worlds. *Ergonomics,* 40 (1): 69-77.

Baluyut, R., Genaidy, A.M., Davis, L.S., Shell, R.L., and Simmons, R.J. (1995). Use of visual perception in estimating static postural stresses: magnitudes and sources of error. *Ergonomics*, 38(9) 1841-1850.

Barbe, M.F., Barr, A.E., Gorzelany, I., Amin, M., Gaughan, J.P., and Safadi, F.F. (2003). Chronic repetitive reaching and grasping results in decreased motor performance and widespread tissue responses in a rat model of MSD. *Journal of Orthopaedic Research*, 21 (1): 167-176.

Burt, S. and Punnett, L. (1999). Evaluation of interrater reliability for posture observations in a field study, *Applied Ergonomics*, 30, 121-135.

Corlett, E.N. & Manenica, I. (1980). The effects and measurement of working postures. *Applied Ergonomics*, 11(1), 7-16.

Dawson, W.J. (2002). Upper- extremity problems caused by playing specific instruments. *Medical Problems of Performing Artists*, 17 (3): 135-140.

Douwes, M. and Dul, J. (1991). Validity and reliability of estimating body angles by direct and indirect observations. Designing for Everyone: Proceedings of the 11th Congress of the International Ergonomics Association (pp. 885-887). London: Taylor & Francis.
.

Ericson, M., Kilbom, A., Wiktorin, C., Winkel, J., and Stockholm (1991). Validity and reliability in the estimation of trunk, arm and neck inclination by observation. In *Designing for Everyone and Everybody: Proc. 11th Congress of the Intentional Ergonomics Association* (pp. 245-247). London: Taylor & Francis.

Genaidy, A.M., Al-Shedi, A.A., and Karwowski, W. (1994). Postural stress analysis in industry. *Applied Ergonomics*, 25(2), 77-87.

Genaidy, A.M., Simmons, R.J., Guo, L., Hidalgo, J.A. (1993). Can visual-perception be used to estimate body part angles. *Ergonomics,* 36(4) 323-329.

Hignett, S. and McAtamney, L. (2000). Rapid Entire Body Assessment (REBA). *Applied Ergonomics*, 31, 201-205.

Issever, H., Aksoy, C., Sabuncu, H., Karan, A. (2003). Vibration and its effects on the body. *Medical Principles and Practice*, 12 (1): 34-38.

Juul-Kristensen B, Fallentin N, and Ekdahl C. (1997). Criteria for classification of posture in repetitive work by observation methods: A review. *International Journal of Industrial Ergonomics*, 19 (5): 397-411.

Kattel, B.P., Fredricks, T.K., Fernandez, J.E., and Lee, D.C. (1996). The effect of upper extremity posture on maximum grip strength. *International Journal of Industrial Ergonomics,* 18 (5-6): 423-429.

Ketola, R., Toivonen, R., and Viikari-Juntura, E. (2001). Interobserver repeatability and validity of an observation method to assess physical loads imposed on the upper extremities. *Ergonomics*, 44(2) 119-131.

Keyserling, W.M. (1986). A computer-aided system to evaluate postural stress in the workplace. *American Industrial Hygiene Association Journal*. 47 (10): 641-649.

Liberty Mutual. (2001a). Liberty mutual reveals first-ever ranking of the leading causes of workplace injuries. Press Release (February 26). [http://www.libertymutual.com/about/pressclub/press_workplace.html]

Liberty Mutual (2001b). Majority of U.S. Businesses report workplace safety delivers a return on investment. Press Release (August 28). [http://www.libertymutual.com/about/pressclub/safety_survey.html]

National Institute of Occupational Safety and Health. (1997). Musculoskeletal disorders and workplace factors: a critical review of epidemiologic evidence for work related musculoskeletal disorders of the neck, upper extremity, and low back. Bernard, B.P. 97-14. Cincinnati: Department of Health and Human Services. NIOSH technical report DHSS(NIOSH).

MacLean, I.E. and Stacey, B.G. (1971). Judgment of angle size: An experimental appraisal. *Perception and Psychophysics,* 9(6): 499-504.

Mallory, M. and Bradford, H. (1992). An invisible work place hazard gets harder to ignore. *Business Week*. January 30, 92-95.

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. and 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.

Moore, J.S., Rucker, N.P**.**, Knox, K. (2001). Validity of generic risk factors and the strain index for predicting nontraumatic distal upper extremity morbidity. AIHAJ, 62 (2): 229-235.

Montgomery, D.C. (1997). Design and Analysis of Experiments. New York: John Wiley & Sons.

Nordvik, H. and Amponsah, B. (1998). Gender differences in spatial abilities and spatial activity among university students in an egalitarian educational system. *Sex Roles*, 38 (11-12): 1009-1023.

Quaiser-Pohl, C. and Lehmann, W. (2002). Girls' spatial abilities: Charting the contributions of experiences and attitudes in different academic groups. *British Journal of Educational Psychology*, 72: 245-260.

Rucker, N. and Moore, J.S. (2002). Predictive Validity of the Strain Index in Manufacturing Facilities. Applied Occupational and Environmental Hygiene, 17(1): 63-73.

Safety Line Institute (1998). Vibration Effects. WorkSafe Western Australia. [http://www.safetyline.wa.gov.au/institute/level2/course19/lecture60/l60_02.asp]

Voyer, D., Nolan, C., and Voyer, S. (2000). The relation between experience and spatial performance in men and women. *Sex Roles*, 43 (11-12): 891-915.

Waller, T.G. and Wright, R.H. (1965). The Effect of Training on Accuracy of Angle Estimation. *HumRRO-Technical-Report*, 65-8(2), 24pp.

Weimer, L.H., Yin, J., Lovelace, R.E., and Gooch, C.L. (2002). Serial studies of carpal tunnel syndrome during and after pregnancy. *Muscle & Nerve*, 25 (6): 914-917.

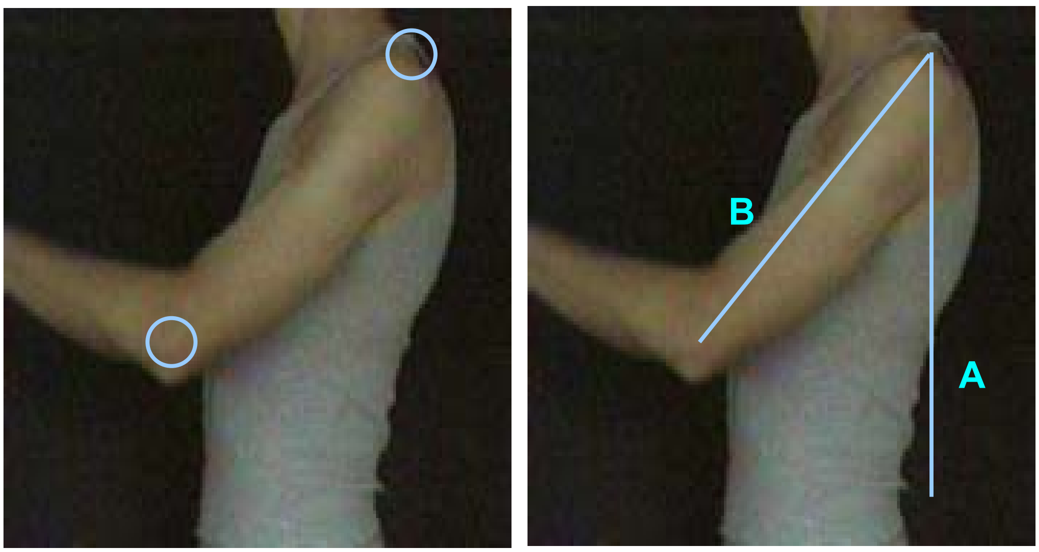
Wenderoth, P. and Johnson, M. (1984). The effects of angle-arm length on judgments of angle magnitude and orientation contrast. *Perception and Psychophysics,* 36(6): 538 544.

Appendices

Appendix A – Written Instructions

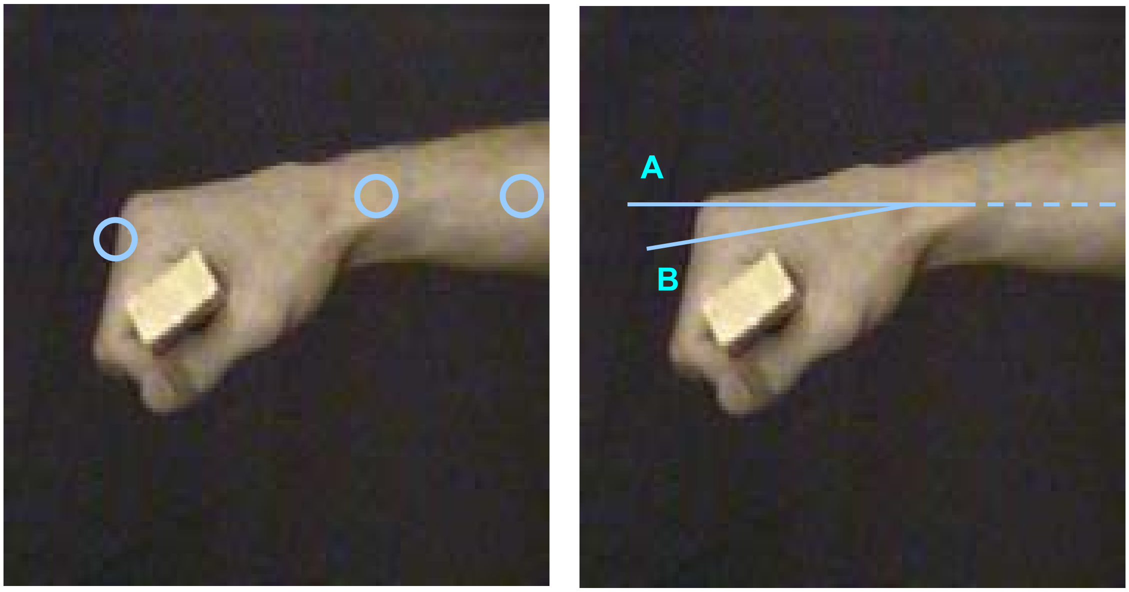
Please read the following instructions: You have been given CDs containing video and photos of work tasks. You will be asked to visually estimate postural angles of the shoulder, trunk, or wrist that occur during the video clip. Please estimate the postural angles according to the methodology described below.

Shoulder Angles: The shoulder angle is considered to be the angle between a line perpendicular to the ground and through the top of the shoulder (line A) and a line drawn from the top of the shoulder through the head of the radius at the elbow (line B).



Trunk Angle: The trunk angle is considered to be the angle between a line perpendicular to the ground and through the hip rotation point (line A) and a line through the hip rotation point and the top of the shoulder (line B).

Wrist: The wrist angle is considered to be the angle between a line drawn between the center of the forearm (line A) and a line from the center of rotation of the wrist through the center of the palm of the hand (line B). Angles below line A are negative angles. Angles above line A are positive angles.



Appendix B – Plot of Error Residuals

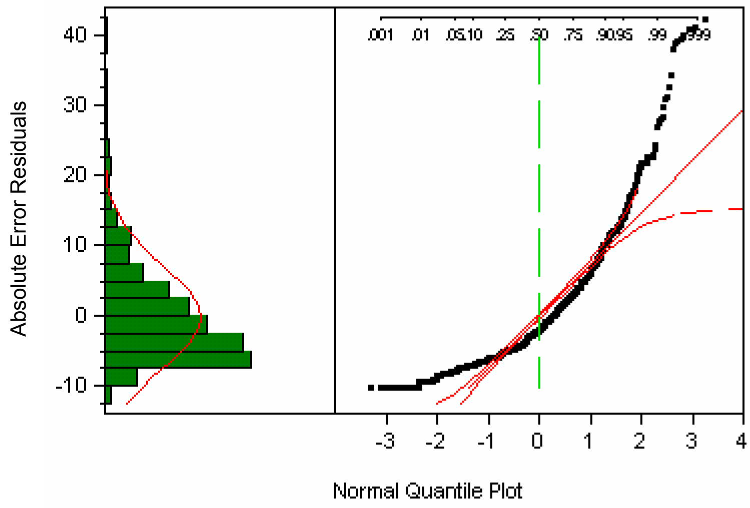


Figure B.1: Absolute error residuals distribution

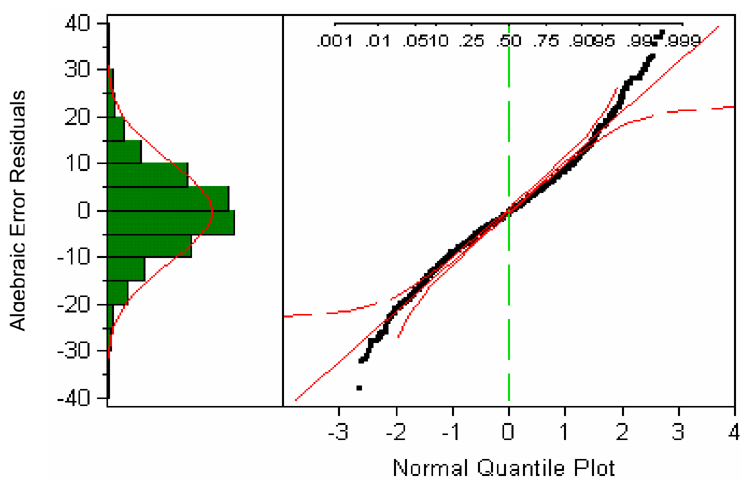


Figure B.2: Algebraic error residuals distribution

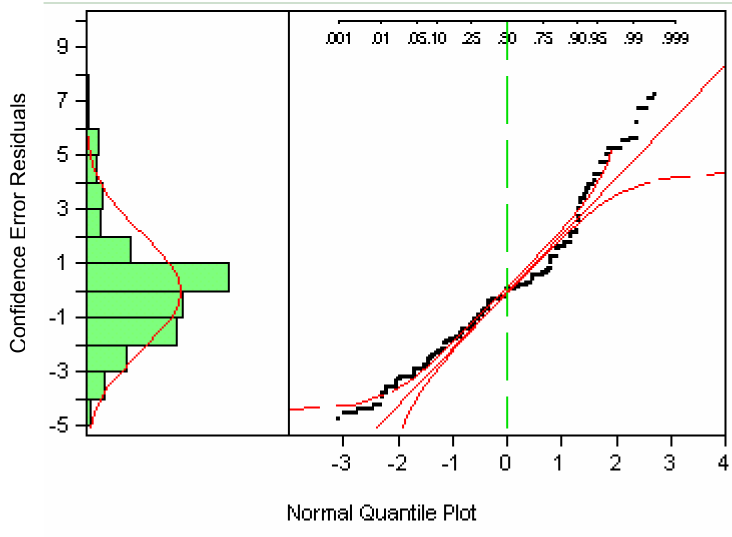


Figure B.3: Confidence residuals distribution

Appendix C – Residuals Versus Order

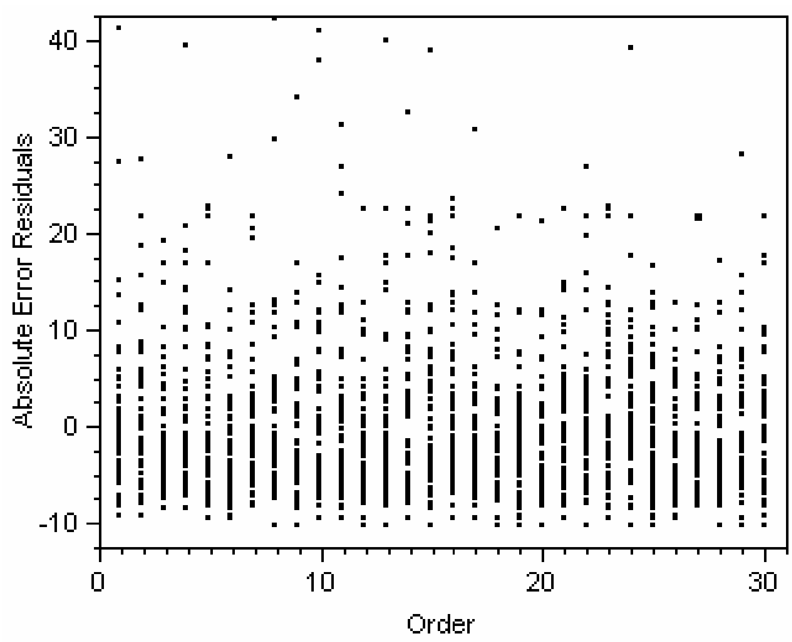


Figure C.1: Absolute error residuals versus order

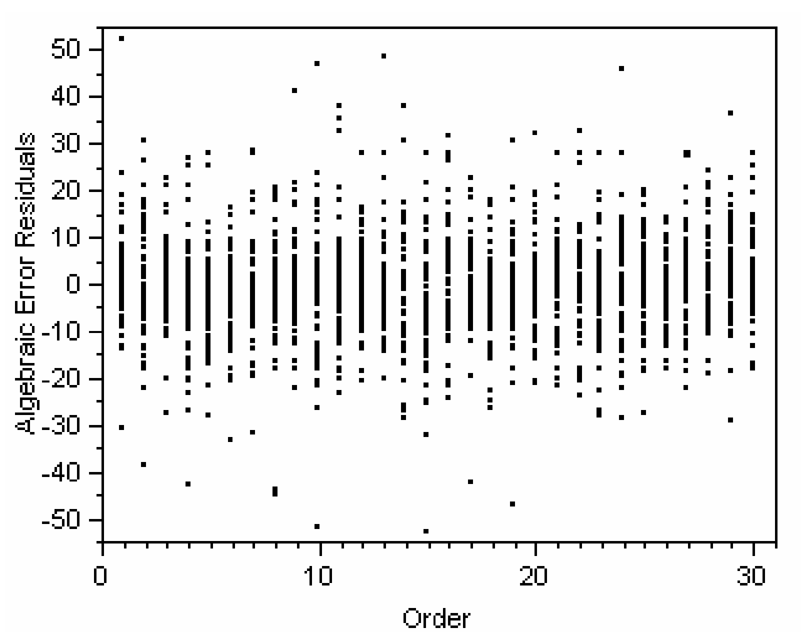


Figure C.2: Algebraic error residuals versus order

Figure C.3: Confidence residuals versus order

Appendix D - Homogeneity of Variances

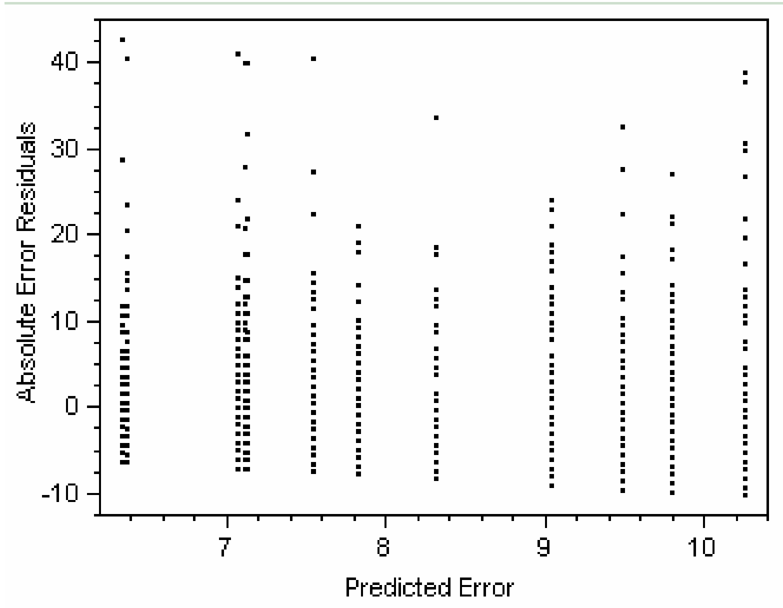


Figure D.1: Absolute error residuals versus fitted values

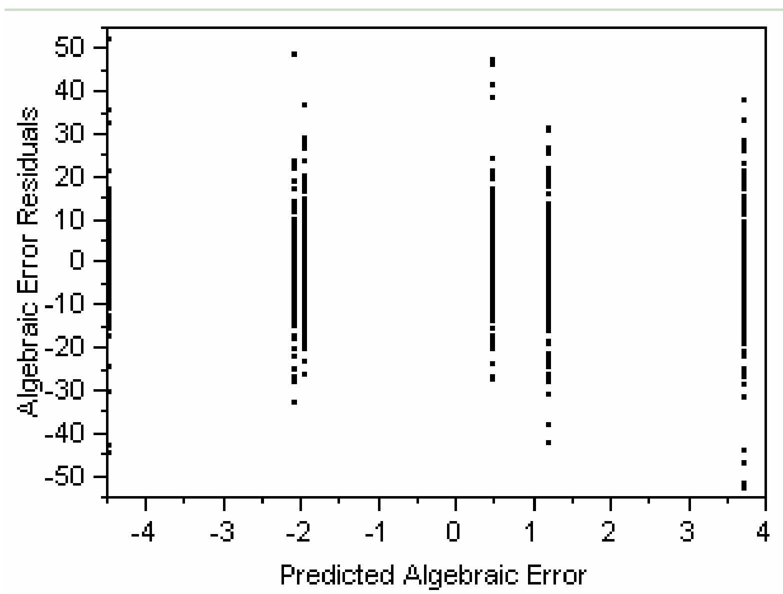


Figure D. 1: Algebraic error residuals versus fitted values

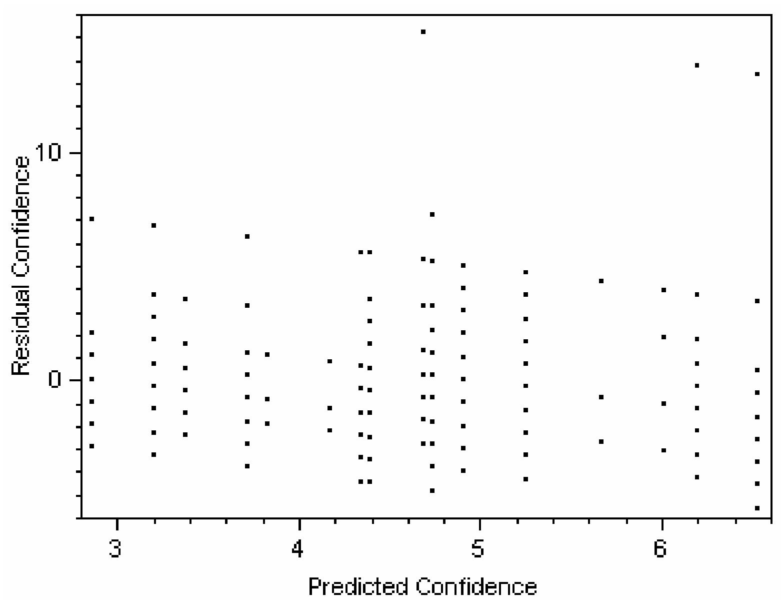


Figure D.3: Confidence residuals versus fitted values