DEVELOPMENT AND VALIDATION OF THE SELF-REPORT ERGONOMIC ASSESSMENT TOOL (SEAT)

A Dissertation

by

PAUL DANIEL RITCHEY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PUBLIC HEALTH

Chair of Committee,	S. Camille Peres
Co-Chair of Committee,	Ranjana K. Mehta
Committee Members,	Roger Zoh
	Mark Benden
Head of Department,	Mark Benden

August 2017

Major Subject: Epidemiology and Environmental Health

ABSTRACT

Despite considerable advances in the practice of office ergonomics, office workers are still suffering from musculoskeletal disorders (MSDs). These disorders, like carpal tunnel syndrome, can lead to high medical costs for employers and intense pain and discomfort for employees. The design of software office workers use could be a contributing factor to their risk of developing MSDs and a tool sensitive enough for evaluating ergonomic risks associated with the design of software is needed. Presented here are the results of a series of three studies focused on the development, improvement, and validation of a Self-report Ergonomic Assessment Tool (SEAT). The SEAT was found to comprise two important factors, stress and strain, and was found to be sufficiently consistent and sensitive to the exertions and postures related to office work. Data from two studies were used to validate stress components of the SEAT, e.g., postures, by using recorded videos and comparing participants' responses on the SEAT to those of trained raters. Results showed that participants were unable to reliably self-report stressors. Data from one study was used to validate the strain components of the SEAT by comparing participants' self-reported discomforts to muscle activity measured via surface electromyography and muscle oxygenation measured via near infrared spectroscopy. Participants' self-reported discomfort did correlate with these physiological measures, however, important exceptions revealed opportunities for future development and testing of the SEAT.

DEDICATION

This dissertation is dedicated to my mother.

And father.

But mostly my mother.

ACKNOWLEDGMENTS

I would like to acknowledge and thank the Office of Ergonomics Research Council for funding this research and the National Institute of Occupational Health and Safety (NIOSH) Trainee program for partially funding my graduate studies.

I would also like to thank my advisors, Dr. S. Camille Peres and Dr. Ranjana Mehta, for their contributions, feedback, and unwavering support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professors S. Camille Peres, chair, and Ranjana Mehta, co-chair, and Mark Benden of the Department of Environmental and Occupational Health and Professor Roger Zoh of the Department of Epidemiology and Biostatistics.

All work for the dissertation was completed by the student, in collaboration with Drs. S. Camille Peres and Ranjana Mehta of the Department of Environmental and Occupational Health.

Funding Sources

This work was made possible in part by the Office Ergonomics Research Committee (OERC) under Contracts M1401259 and M1601242.

Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the OERC.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

INTRODUCTION

Since the 1980s, ergonomic researchers have been examining associations between musculoskeletal disorders (MSDs) and computer work in the office environment (Murray et al., 1981; Knave et al., 1985). MSDs for office workers are created, in part, because of the low force, highly repetitive movements computer users perform while also maintaining non-neutral postures for long periods of time (IJmker et al., 2007; Punnett & Bergqvist, 1997). For example, more than 40% of injury claims of Washington State Employees were from office workers, costing over $12 million per year and accounting for 60% of the cost of all workers' compensation claims (OEAC, 2002). An appreciable amount of work has been done to further the understanding of the associations between computer work and MSDs (Andersen et al., 2003; Blatter & Bongers, 2002; Heinrich, Blatter, & Bongers, 2004; Newburger, 2001; Wahlström, 2005).

One potential source of ergonomic risk could be the software being used, more specifically, the software's interaction design (Peres, et al., 2015). While effectively designed keyboards, mice, chairs, *etc.* can help ameliorate the hazards of computer work, the ways in which a software's design requires its users to interact with their devices ultimately create those hazards. For example, if a software requires a high amount of mouse travel because of frequently used menus, icons, or other aspects of the software are located far apart from each other, changing the pointer devices to an upright mouse could lower the level of risk. However, redesigning the software to reduce overall mouse travel distances controls the risk at its source. Human Factors methods (e.g., User Centered Design) can inform software design by improving efficiency, satisfaction, and ease of learning. These methods can similarly be applied to the problem of ergonomic risks but only if there is a tool through which these risks can be *measured* and effectively integrated into human factors methods.

Controlling ergonomic risks and developing effective ergonomic interventions depends on the ability to detect ergonomic risk. Ergonomic researchers have made significant progress towards developing the knowledge and tools necessary to mitigate the ergonomic risks imposed on computer workers (Gerr, Monteilh, & Marcus, 2006; Brewer et al., 2006; Kennedy et al., 2010). However, no tools have been designed to be specifically sensitive to the ergonomic risks put into place by software design. The most effective ergonomic interventions are comprehensive (Kennedy et al., 2009), and the first step towards introducing software design as part of a comprehensive ergonomic intervention program is the development of a tool that can be used to measure and quantify the subtle ergonomic risks created from the software-driven interactions between humans and computers. The control of software design induced ergonomic risk factors would be most effective if applied during the development of software when making design changes is much easier compared to later in production or even post-production. As such, there is need for a quick, easily understandable, and self-administrable ergonomic assessment tool for use during the software design lifecycle.

The Self-report Ergonomic Assessment Tool (SEAT) is a new ergonomic assessment tool for evaluating upper body ergonomic risks associated with the use of modern computing devices. A series of studies were conducted in which the SEAT was developed, improved, and validated. These studies are presented here as three papers:

1. The SEAT was developed based on a review of existing ergonomic assessment tools by selecting components that address a breadth of ergonomic stressors (i.e., postures and repetition) and strain (i.e., discomfort) relevant to the risks imposed by computer work. A large study of professional geoscientists (an at-risk population) and students (Study 1: N = 159) was conducted to determine how items on the SEAT group with one another, and if items on the SEAT are sensitive to differences between tasks and input methods. The results from this analysis were used to update and revise the SEAT.
2. Video recordings of participants (N = 42) from Study 1 were used to determine how participants' self-reporting of stressors compare to those of trained raters. Results and insights from these analyses were used to inform the second iteration of the SEAT. Validity of participants' self-reporting of stressors with the second iteration of the SEAT were assessed using video recordings from a follow-up study (Study 2: N = 30).
3. The second iteration of the SEAT was administered to participants (Study 2: N = 30) after they complete an emailing and a calendaring task on four different computing devices. To test the validity of the participants' responses to the strain components of the SEAT, physiological measures of muscle activity (surface electromyography and near-infrared spectroscopy) were compared to participants' self-reported discomfort.

The goals of these three papers were to develop, improve, and validate the SEAT by showing that the users without formal ergonomic training, when using the SEAT, can accurately self-report computer-related stressors and strains.

DEVELOPMENT OF THE

SELF-REPORT ERGONOMIC ASSESSMENT TOOL*

Introduction

With the increased use of computers has been an increase in the occurrence of upper extremity musculoskeletal disorders, e.g., cumulative trauma disorders or CTDs (Andersen et al., 2003; Blatter & Bongers, 2002; Newburger, 2001; Wahlström, 2005). This has been associated with the awkward postures, low force and repetitive movements required for keyboard and mouse use, as well as looking at a visual display for extended periods of time (Gerr et al., 2002; IJmker et al., 2006; Punnett & Bergqvist, 1997; Village, Rempel, & Teschke, 2005). For many, interventions such as taking breaks and redesigning workstations have been successful at reducing ergonomic risks (Bohr, 2001; Dainoff, Maynard, Robertson, & Andersen, 2012). However, some workers using these ergonomic interventions still get injured (Bishea, Wood, & Muddimer, 2007; Stern, 2015). For instance, geoscientists—who work for extended periods of time with software that have large, complex graphical interfaces—get injured even though many of them have ergonomically adjusted offices and take regular breaks (*Land Geophysical Safety Manual*, 2012; Taylor, 2007). These users typically work with two to three large monitors with an image (typically of seismic data they are interpreting) spread across the monitors (see Figure 1). The software they use requires a constant series of iterative

* Reprinted with permission from "Assessing ergonomic risks of software: Development of the SEAT" Peres, S. C., Mehta, R. K., & Ritchey, P., 2017. *Applied Ergonomics*, 59, 377-386, Copyright [2017] by Elsevier.

actions involving clicking on icons and then clicking and dragging the mouse over the entire image (e.g., across both monitors seen in Figure 1) similar to the work of graphic artists using programs like Photoshop® (Stern, 2015). This interaction design (i.e., having users interact with the software by clicking icons and dragging the mouse) could be contributing to the risk of CTDs because 1) the users are constantly dragging the mouse across two large screens—requiring static activation of the muscles to keep the mouse in a clicked position that is a high-risk factor for the development of CTDs (Anghel, Argesanu, Talpos-Niculescu, & Lungeanu, 2007; Bernard & Putz-Anderson, 1997; Tittiranonda, Burastero, & Rempel, 1999)—and 2) the users must leverage a large number of very small icons (see Figure 1)—requiring precise honing and fine motor movements in order to successfully click the target object.

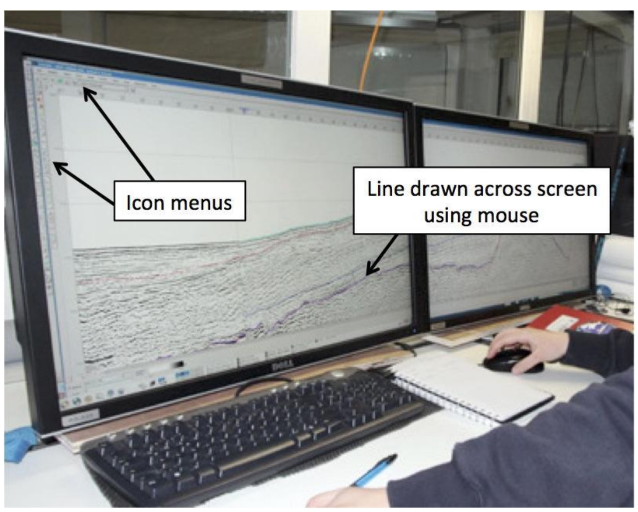


Figure 1. Example of a geoscientist's desk setup for interpreting images of seismic data. The icon menu toward the far-left side of the left screen is one that is used frequently and changes given the contexts of the specific task the person is performing. Reprinted with permission from (Peres et al., 2017).

Current office ergonomic interventions focus on changing equipment, modifying other aspects of the physical setup of the computing environment, providing training, and implementing break requirements—not on how the design of the software affects how a person must interact with their computer. This could explain why previously successful interventions are not consistently effective for workers like geoscientists (Bishea et al., 2007; Stern, 2015). Dennerlein and Johnson reported findings (2006b) that provided further support for this. They found that for two different tasks (graphing and web browsing), the levels of risks associated with non-neutral posture were similar. However, the way the mouse was used and the risk associated with muscle activity for the two tasks were significantly different. Their results suggest that it is not just the input device used (here, the mouse) or the posture that influences ergonomic risk but how that device is being used. Specifically, it could be that software interaction methods, i.e., using the mouse to drag versus click, may place different ergonomic risks on the user.

If these suggestions regarding interaction method affecting ergonomic risk are true, software developers need to consider the ergonomic implications of their interaction designs during the development process. However, current ergonomics assessment methods are not conducive to integration into software development lifecycles. Agile is one of the more commonly used development methodologies and is an extremely fast-paced iterative process where teams work to develop functional iterations of the product during "sprints" that are sometimes only two weeks long (Martin, 2003). Thus, evaluation of any part of the design (ergonomic or otherwise) must fit within the constraints of this process. Current ergonomic assessments typically require extended periods of time and special expertise (Dainoff et al., 2012) and therefore would not work within the software development paradigm.

Usability professionals have struggled with this issue and found a successful method for integrating usability assessment into the development lifecycle—the use of subjective usability measures, *e.g.*, System Usability Scale (SUS: Brooke, 1996). This is not a replacement for formal usability testing but instead provides values indicating whether the usability of the software is at acceptable levels. We submit that a similar model could be applied to assessing ergonomic risks of software design to let developers know when their software is above acceptable ergonomic risk levels and thus, they need to leverage more formal ergonomic assessment methods to identify less risky interaction methods. This type of ergonomic assessment method would need to be not only reliable and valid for software use but also easy to administer and interpret. Although there are currently several existing self-report measures regarding office ergonomics (Dane et al., 2002; Heinrich, Blatter, & Bongers, 2004; Li & Buckle, 2000; McAtamney & Corlett, 1993; Robertson et al., 2009; Sonne, Villalta, & Andrews, 2012; Speklé et al., 2009), they focus primarily on the office setup and overall work posture and movement. These measures generally give a score for how well the physical office is set up, how to change the set up to reduce ergonomic risks, or the level of risk that the person may have with regard to his or her office arrangement (Sonne et al., 2012; Speklé et al., 2009). None of the existing self-report measures include considerations or provide any information at the software level of interaction and thus would not be sufficient for use evaluating ergonomic impacts of software design.

The objective of this paper is to present the development of the Self-report Ergonomic Assessment Tool (SEAT). To develop this measure:

- We leveraged items from existing ergonomic assessment measures—both self-report and administered by an ergonomic expert. Although most were developed for industrial applications and not the office environment, *e.g.*, Strain Index (SI: Moore & Garg, 1995) and Hand Activity Level (HAL: Latko et al., 1997), these measures have been widely used, are reliable, and some are validated for industrial settings. Thus, we chose to use items from these measures as the foundation for our new measure. Given the substantial differences between the office and industrial settings, e.g., intensity of force applied, repetition, and duration of exposure (Bruno Garza et al., 2012), we selected only the specific items from the measures that were relevant for the office environment.
- We performed typical psychometric testing of the SEAT, including principal component analysis (to determine the number of underlying dimensions or factors in the measure) and Cronbach's alpha (for assessing internal consistency).
- We tested the discriminant validity of the SEAT with an experimental design that had two conditions that have been shown to have different ergonomic risks (Duffield, Peres, Amonette, & Ritchey, 2013).

Our first step toward building the SEAT measure was a conservative, proof of concept approach focused on confirming: 1) people would be able to *consistently* respond to a self-report measure regarding ergonomic risks associated with software interaction; and 2) a self-report measure would be sufficiently *sensitive* to identify tasks that have known differences in ergonomic risks.

Methods

Participants

A total of 166 participants (61 females) were recruited from a three-day professional convention in Texas for geoscientists (N = 56: age range 18 to 72 years) and through flyers and emails at Texas A&M University (N = 110: age range 18 to 80). For the convention group, all sessions were completed on site in a conference room. For the university group, sessions were conducted in a usability lab. Seven participants from the convention sample had incomplete data and were not included in certain analyses for the study. Of the 159 participants included in the final data set, eight were left-handed and were instructed to complete the tasks as they would normally use a computer. One used their left hand to manipulate the mouse and the rest used their right. The participants age ranged from 18 to 80 (M = 31.58, SD = 13.00). The convention average age was similar to the university average age, although the convention sample was, on average, older (M = 38.65, SD = 13.05) compared to the university sample (M = 28.44, SD = 11.73). The convention group also had a higher self-reported hours per week spent on a computer while at work (M = 11.04, SD = 10.38) compared to the university sample (M = 7.97, SD = 7.98). To assess whether any differences existed between these groups in terms of effect on responses to the measure, Group (2 levels: Convention, University) was included in the statistical analyses.

Input Methods

A desktop personal computer with touch capabilities (Dell Optiplex 9020 All-In-

One PC running Windows 7) was used and 1) the direct input method required the participant to use the touch interface to complete the task (Direct input) and 2) the indirect input method required the participant to use peripheral devices (Indirect input: Dell KM623 Wireless Keyboard and Mouse) to complete the task. Figure 2 illustrates the setup and postures typical when performing the experimental tasks with the two different input methods. Previous studies have found that Direct input results in higher of discomfort when it is the only input method used as a person must raise their hand to the screen and maintain that position (Shin & Zhu, 2011) and thus, is considered the higher load condition (particularly for the more postural upper extremity muscles such as the shoulders and upper arms). These two extreme input conditions were chosen because they were shown to have reliable and meaningfully different ergonomic risks for the muscles in the shoulders and upper arms (Duffield et al., 2013). For the self-report measure to be sufficiently sensitive people must be able to identify these differences using the self-report measure, i.e., there must be a main effect of Input method for any responses associated with the shoulders and upper arms.

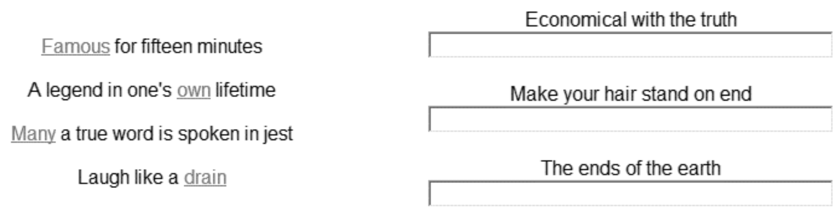
Direct Input
(High load)

Indirect Input
(Low load)

Figure 2: Postures typically adopted in the Direct input and Indirect input conditions. Reprinted with permission from (Peres et al., 2017).

Tasks

For each Input method, the participant completed two three-minute long tasks (Selection and Typing) that have interaction features similar to those used by geoscientists. The tasks were completed on an HTML page created for the study (Figure 3) and consisted of participants either clicking or touching on a hyperlinked word in a sentence (Select: see Figure 3a) or using both hands to type a phrase into a form field on a page (Type: see Figure 3b). A task time of three minutes was chosen because in previous studies similar time periods have been found sufficient to ensure exposure to stress on the body (Duffield et al., 2013; Ritchey, Peres, & Duffield, 2012) and this task time would fit within software development cycles such as Agile.



3a. Selection 3b. Typing

Figure 3: Images of the two different tasks participants completed. a) The selection task where participants selected, or clicked, on the hyperlinked word. b) The typing task where participants typed the phrase above the text box into the text box itself. Reprinted with permission from (Peres et al., 2017).

Measures

We built the SEAT by combining items from established measures of ergonomic risks and outcomes and the final self-report is provided in Appendix A. The items were drawn from modified versions of the Borg CR10 (BORG: Borg, 1990), Strain Index (SI: Moore & Garg, 1995), Body Discomfort Diagram (BDD: Ritchey, Peres, & Duffield 2012; Megasari, 2009; Cameron, 1996), the Rapid Upper Limb Assessment (RULA: McAtamney & Corlett, 1993), and Hand Activity Level (HAL: Ebersole & Armstrong, 2002) measures. We specifically included items that would address each of the ergonomic risk factors and body regions and Table 1 shows this. All items were modified to ensure that they applied to an office work environment and that people who were not trained in ergonomics could understand the terms used on the survey. To do this, we presented the items to 5 people who use graphically intensive software as part of their daily work. We had them respond to each item using a “Think a Loud” methodology while we were sitting with them. This method involves their answering the item, verbally describing their understanding of what the item was asking for, and indicating whether it applied to their work (Virzi, 1993). This resulted in 30 items, specifically, 16 items on body discomfort (which included items for both right and left (RL) sides), 2 on hand activity level (1 item: RL), 4 on wrist position (2 items: RL), 4 on shoulder position (2 items: RL), 1 on level of effort, 1 on rating of perceived exertion, 1 on speed of work, and 1 on level of precision required for the task. The subjective force and load items from the RULA and HAL were not included in the measure because similar items included in the Borg CR10 and the SI these were easier to answer. All participants completed the body discomfort items at the beginning of the study to establish a baseline to control for any preexisting discomfort. Difference scores from this baseline were used in all calculations.

Table 1. Risk factors and specific regions of SEAT. This table shows the risk factors assessed and specific regions of the body included in each of the measures used to create the SEAT. Some items of these measures were modified or not included and the entire self-report used in this study is included in Appendix A. Those indicated with an asterisk (*) were included in SEAT. Reprinted with permission from (Peres et al., 2017).

	Intended as Self Report	Risk Factors Considered				Symptoms	Specific Body Parts Addressed						
Measure		Posture	Force/ Load, Exertion	Duration	Repetition	Discomfort	Wrist	Hand	Lower Arm	Upper Arm	Head/ Neck	Trunk	Lower Body
Strain Index (SI)		X	X*	X	X		X	X					
Body Discomfort Diagram (BDD)	X					X	X*	X*	X*	X*	X*	X*	X*
TLV Hand Activity Level (HAL)			X*	X*	X*			X*					
Borg CR10 (Borg)	X		X*										
Rapid Upper Limb Assessment (RULA)		X*	X		X		X*		X	X*	X	X	X
SEAT	X	X	X	X	X	X	X	X	X	X	X	X	X

Procedures

Before the study began, all participants were provided an informed consent form—reviewed and approved by the Texas A&M University Institutional Review Board—and were given a short introduction to the nature of the study. After agreeing to participate and signing the consent form, participants were seated at the workspace and were instructed to self-adjust the chair to their own comfort. The monitor position was adjusted for the two different tasks based on the ANSI/HFES standards (2007). For the Direct task, the monitor was placed wrist length away (i.e., with the arm extended out fully, parallel to desk surface). For the Indirect task, the monitor was placed arm's length away (i.e., fingers just touching the screen, arm fully extended, parallel to the desk surface). Participants were also allowed to adjust the placement of the monitor if they preferred. Following workspace adjustments, participants provided demographic information and audio/video recordings were started. The audio and video information was collected throughout the experiment (note: the analysis of the video data is beyond the scope of this study). Both the input method and task type were counterbalanced to control for order effects with the tasks blocked so participants completed both tasks with one input method and then the completed the tasks in the same order using the other input method. Participants were randomly assigned to one of the four counterbalance conditions. After completing each task, the self-report measure was verbally administered to participants, which also allowed for a rest period of 3 to 4 minutes as researchers filled in the survey based on the participant's verbal responses. The entire session took approximately 45 minutes.

Analyses

Principal Component Analysis (PCA), with Varimax orthogonal rotation with Kaiser Normalization, was conducted on the data obtained from all participants (N = 166) to determine the underlying structure of the items on the survey, and pairwise exclusion was used to adjust for those with missing data. All other analyses were run on the 159 participants with complete data. Cronbach's alpha was then conducted on the overall measure and each component to assess the internal consistency of the measures. An a priori decision was made that analyses would be done separately for each component found in the PCA and that the alpha level for the analyses would be adjusted for the number of components found in the PCA. Factorial Repeated Measure Analyses of Variance (RMANOVAs) were performed on the parametric components and Wald Chi-Square analyses were performed for the non-parametric components that included only dichotomous responses. For the discriminant validity analysis of the SEAT, there were two independent variables—Input method (Direct and Indirect) and Task type (Type and Select). The original design was a fully within repeated measure design but the between-subject factor Group (Convention and University) was added to determine if any meaningful differences existed between these two samples. All effect sizes are listed in Table 3 and thus will not be included in the text of the Results section for parsimony and clarity.

Results

Measurement Consistency

The PCA resulted in the 30 items of the SEAT loading onto eight components. A one-item component was excluded (Discomfort of the right ring and pinky finger) and another eliminated from a component that was not conceptually related to other items in that component (Floyd & Widaman, 1995). We retained four components with two items, as they are important for understanding specific ergonomic stressors. This left 28 items and is the current version of the SEAT. Table 2 shows the rotated component matrix, bolded component coefficients denote component loadings. See Table 3 for eigenvalues of the final components with the percent variance explained by each and the final seven-component solution (64.6%).

Table 2. Rotated Component Matrix from PCA using Varimax Rotation. *Indicates items not included in final components. Reprinted with permission from (Peres et al., 2017).

Items	Rotated Component Coefficients							
	Com. 1	Com.2	Com.3	Com.4	Com.5	Com.6	Com.7	Com.8
Discomfort Left Lower Arm	**.844**	.147	.043	.172	.067	.008	.031	.015
Discomfort Left Wrist	**.817**	.160	.009	.253	.096	-.005	.024	.010
Discomfort Left Shoulder	**.786**	.362	.089	-.126	-.027	-.054	-.006	.105
Discomfort Left Upper Arm	**.784**	.250	.128	-.017	.045	-.080	.097	.053
Discomfort Left TIM	**.771**	.215	.046	.264	.058	-.048	-.001	-.029
Discomfort Left RP	**.727**	.112	-.096	.422	.004	-.018	.018	.040
Discomfort Left Trapezius	**.701**	.378	.010	-.138	-.036	-.156	.051	.093
*HAL Left	**.651**	-.286	.098	-.247	.318	.160	-.060	.081
Discomfort Right Shoulder	.089	**.797**	.249	.085	-.030	.001	.102	.078
Discomfort Neck	.245	**.792**	.001	-.053	.062	.048	-.021	-.007
Discomfort Right Trapezius	.154	**.788**	.139	.046	-.069	-.078	.055	.082
Discomfort Eye	.186	**.745**	-.116	-.070	-.017	.125	-.047	-.058
Discomfort Right Upper Arm	.051	**.742**	.284	.162	-.055	-.019	.086	.084
Discomfort Right Lower Arm	.179	**.697**	.167	.392	.024	.073	.012	.011
Discomfort Right TIM	.139	**.668**	.124	.492	.014	.027	-.001	-.029
Discomfort Right Wrist	.178	**.663**	.115	.471	.061	.092	-.026	-.006
SI Effort Level	.070	.342	**.691**	.261	-.118	.066	.200	.108
Borg Rating of Perceived Exertion	.145	.419	**.629**	.238	-.109	.087	.229	.086
Wrist Position Right	-.074	.062	**.569**	-.046	.433	-.121	-.131	-.052
Precision	.086	.230	**.524**	-.202	-.074	.229	.084	.153
Wrist Position Left	.424	-.240	**.426**	-.205	.374	-.066	-.123	.053
*Discomfort Right RP	.286	.325	.013	**.674**	-.020	.111	-.076	.123
Wrist Deviation Right	-.019	.104	-.094	.102	**.819**	.021	.159	.043
Wrist Deviation Left	.371	-.129	.047	-.078	**.741**	-.071	.096	.058
HAL Right	-.017	.073	.157	.063	.043	**.816**	.021	.029
Speed of Work	-.132	.025	-.050	.030	-.096	**.815**	-.019	-.057
Shoulder Abduction Right	-.132	.122	.062	-.024	.037	.043	**.836**	.021
Shoulder Abduction Left	.247	-.058	.092	-.025	.155	-.047	**.766**	.045
Shoulder Position Right	-.067	.178	.164	.141	-.034	-.005	.085	**.873**
Shoulder Position Left	.457	-.121	.015	-.089	.204	-.037	-.021	**.734**

Table 3. Listing of the different components, the % variance explained, number of items for each component and the effect sizes (for the significant effects). Wrist deviation and Shoulder abduction were broken into analysis for the right and left side due for non-parametric analyses. Reprinted with permission from (Peres et al., 2017).

Component	Eigen Values	% Var	Cronbach's α	Items	Input	Task	Task X Input
					Effect sizes (partial eta squared)		
Left side discomfort	5.61	18.71	0.93	7	0.22*	0.43*	0.21*
Right side discomfort	5.49	18.31	0.92	8	0.41*	0.37*	
Work demand	2.01	6.70	0.59	5	0.56*		
Task activity	1.54	5.12	0.45	2		0.09*	0.09*
Shoulder flexion	1.43	4.77	0.65	2	0.44*	0.38*	
					Effect sizes (phi)		
Wrist deviation (RL)	1.80	5.99	0.67	2			
Wrist deviation (R)						0.19*	0.11*
Wrist deviation (L)						0.41*	
Shoulder abduction (RL)	1.51	5.03	0.57	2			
Shoulder abduction (R)					0.13*	0.11*	0.14*
Shoulder abduction (L)						0.13*	
Total (Overall)		64.6	0.90	28			

* Indicates that the effect associated with the effect size was significant at the 0.007 level.
Partial eta square effect size conventions are small = 0.01, medium = 0.06 and, large = 0.14.
Phi effect size conventions are small = 0.1, medium = 0.3 and, large = 0.5 (Cohen, 1988).

Principle component 1 (PC1) was categorized as the Left Side Discomfort (LSD) as all left (L) side discomfort questions loaded onto it, with the HAL of the L hand. For these analyses, the HAL for the L side was excluded from the LSD to increase the interpretability of the results. Similarly, PC2 was categorized as the Right Side Discomfort (RSD) and consisted of right (R) side discomfort questions and neck and eye discomfort. PC3 was comprised of questions related to Work Demand (WoD) and PC4 was related to Wrist Deviation (WrD), both R and L. PC5 included two questions that addressed Task Activity (TA), the HAL of the R hand and the reported speed of work required for the given task. PC6 and PC7 were related to Shoulder Abduction (SA) and Shoulder Flexion (SF), respectively. For clarity of interpretation, all component-based scores were created by summing the values for each item in the component instead of using the coefficients from the PCA. Table 4 provides a summary of the components.

Table 4. Listing and description of the components. Reprinted with permission from (Peres et al., 2017).

Component	Contents of the component
Strain	
Left side discomfort (LSD)	Items in this component all describe participants' discomfort they experience in the left arm (upper and lower), trapezius, shoulder, wrist, and fingers.
Right side discomfort (RSD)	Items in this component all describe participants' discomfort they experience in the left arm (upper and lower), trapezius, shoulder, wrist, and fingers.
Stress	
Work demand (WoD)	Items in this component were all associated with the demands or effort level in the work itself and describe the participants' effort level, perceived exertion, precision required, and the position of the left and right wrist.
Task Activity (TA)	Items in this component described the activity level of the right hand and the speed of the work itself.
Shoulder flexion (SF)	Items in this component describe the degree to which the right and left shoulders were flexed during the task.
Wrist deviation (WrD)	These items indicated whether or not the right and left wrists were deviated.
Shoulder abduction (SA)	These items indicated whether or not the right and left shoulders were abducted.

Table 3 shows the Cronbach's alpha scores for the overall measure and the seven component-based scores individually. The Left and Right Discomfort components had

Cronbach's alphas of 0.97 and 0.92, respectively and the full measure had an alpha of 0.90. The other five components all had alphas below 0.70.

Discriminant Validity

WrD and SA were comprised of dichotomous responses for both sides (e.g., participant reported their wrist was deviated or not) and so were analyzed separately for the R and L side using chi-square tests. The remaining scores were analyzed using separate 2 (Input method: Direct, Indirect) X 2 (Task: Typing, Selection) X 2 (Group: Convention, University) RMANOVAs with an alpha of 0.0071 (i.e., alpha of 0.05 adjusted for seven statistical tests, one for each principal component using Bonferroni correction of 0.05/7). All analyses were completed using IBM SPSS V.20.

Differences between Samples and Order Effects. For Right Side Discomfort, the Convention group reported less discomfort (M = 11.71, SD = 25.5) than the University group (M = 18.55, SD = 17.0) and this effect was significant, $F(1,157) = 7.88$, $p = 0.006$, $\eta_p^2 = 0.05$. There was also an interaction between Group and Input for Task Activity, $F(1,157) = 9.23$, $p = 0.003$, $\eta_p^2 = 0.06$. No other effects of group or interactions with group were significant (p's > 0.04).

There were no main effects of the order of presentation for Task or Input (p's > 0.3) and there were interactions for Task X Task Order (F (1,155) = 9.26, $p = 0.003$, $\eta_p^2 = 0.06$) and Input X Input Order (F (1,155) = 13.56, $p < 0.001$, $\eta_p^2 = 0.08$). For the interaction of Task X Task Order, participants had lower overall SEAT scores (less strain/stress) for the first task they performed. For the interaction related to Input, participants had lower overall scores on the first condition they did for the Direct condition but not the Indirect condition.

Left Side Discomfort (LSD). LSD consisted of seven discomfort items with a maximum possible score of 70. As seen in Figure 4, participants reported more discomfort when using Direct input (M = 6.43, SD = 8.7) compared to using Indirect input (M = 2.81, SD = 7.1) and this effect was significant, $F(1,157) = 45.44$, $p < 0.001$. Figure 4 also shows that for the LSD, discomfort was highest after completing the Typing task (M = 9.14, SD = 11.43) compared to the Selection task (M = 0.11, SD = 5.29) and this effect was also significant, Task $F(1,157) = 116.58$, $p < 0.001$. An interaction between Task and Input was found, $F(1,157) = 41.69$, $p < 0.001$, with the difference between the Direct and Indirect being larger for the Typing task (D = 6.9) than for the Selection Task (D = 0.35).

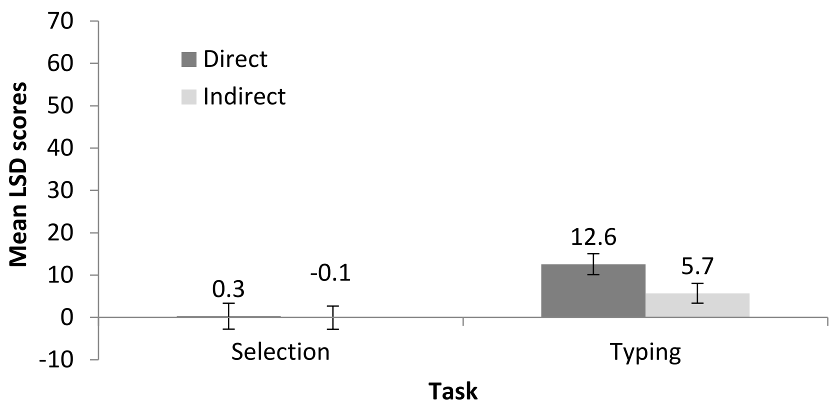


Figure 4. Mean scores for the Left Side Discomfort (LSD) for Task and Input method. LSD consisted of seven discomfort items with a combined maximum possible score of 70. There were main effects for Input Method ($F(1,157) = 45.44$, $p < .001$, $\eta_p^2 = .22$) and Task ($F(1,157) = 116.58$, $p < .001$, $\eta_p^2 = .43$). An interaction between Task and Input was found, $F(1,157) = 41.69$, $p < .001$, $\eta_p^2 = .21$, as well. Error bars represent the 95% confidence interval. Reprinted with permission from (Peres et al., 2017).

Right Side Discomfort (RSD). RSD consisted of eight discomfort items with a maximum possible score of 80. Figure 5 shows the main effects for the RSD of Input and Task. Participants reported higher discomfort after using Direct input to complete the tasks (M = 19.96, SD = 17.9) compared to Indirect method (M = 10.29, SD = 14.8) and this effect of Input was significant, $F(1,157) = 109.72$, $p < 0.001$. Also shown in Figure 5, participants reported more discomfort for Selection (M = 17.74, SD = 16.6) compared to Typing (M = 12.52, SD = 14.9) and this effect of Task was also significant, $F(1,157) = 90.15$, $p < 0.001$. There was no interaction between Task and Input method ($p = 0.05$) for RSD.

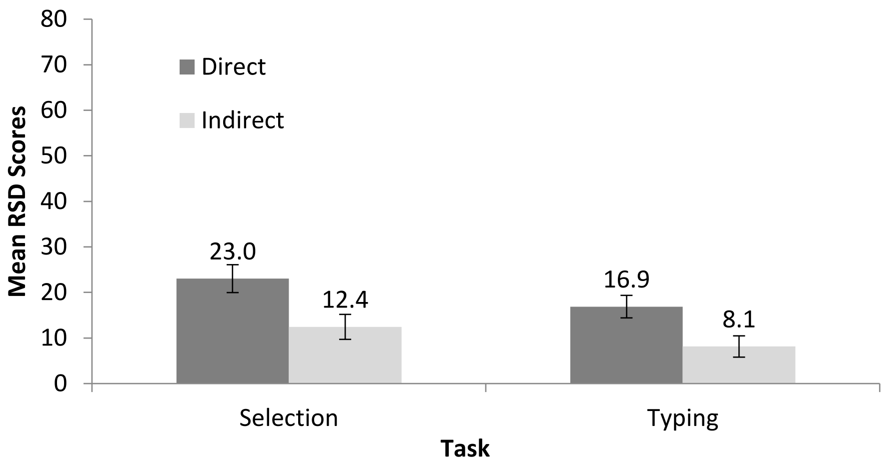


Figure 5. Mean scores of the Right Side Discomfort (RSD) for Task and Input method. RSD has a maximum score of 80. There were main effects for Input method ($F(1,157) = 109.72$, $p < .001$, $\eta_p^2 = .411$) and Task ($F(1,157) = 90.15$, $p < .001$, $\eta_p^2 = .365$). Error bars represent the 95% confidence interval. Reprinted with permission from (Peres et al., 2017).

Work Demand (WoD). This measure consisted of five items with a maximum possible score of 31. As seen in Figure 6, WD was reported to be the highest for Direct input (M = 17.06, SD = 4.2) compared to Indirect input (M = 12.67, SD = 3.8) and this difference was significant, $F(1,157) = 201.41$, $p < 0.001$. No effect of Task, or an interaction between Task and Input were found for the WD component ($p = 0.01$ and $p = 0.06$ respectively).

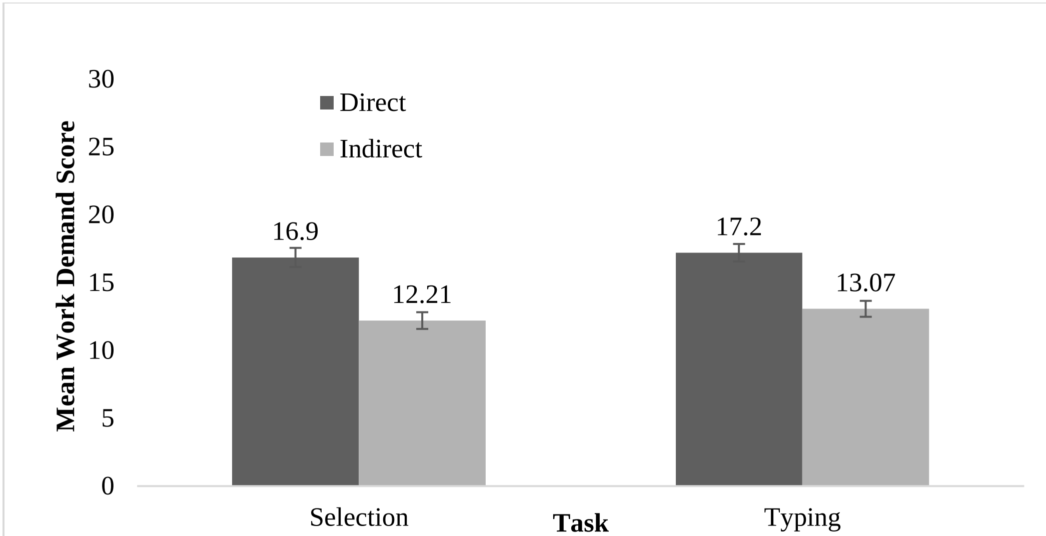


Figure 6. Mean scores for the Work Demand (WoD) component, by Task and Input method. WoD consisted of five items with a combined maximum possible score of 31. There is a main effect of Input Method, $F(1,157) = 201.41$, $p < .001$, $\eta_p^2 = .56$ and no effect of Task, nor an interaction between Task and Input ($p > 0.05$). Error bars represent the 95% confidence interval. Reprinted with permission from (Peres et al., 2017).

Task Activity (TA). There were two items in this component and the maximum score was 15. As seen in Figure 7, TA was not significantly different for Input method ($p = 0.46$) but there was a main effect of Task with participants reporting less task activity for Typing (M = 9.15, SD = 1.7) compared to Selection (M = 9.76, SD = 2.1), $F(1,157) = 14.71$, $p < 0.001$. An interaction between Task and Input method was found, $F(1,157) = 15.42$, $p < 0.001$, with Direct having a higher value than Indirect for the Selection task (D = 0.43) while Direct had a lower value than Indirect for the Typing task (D = -0.67).

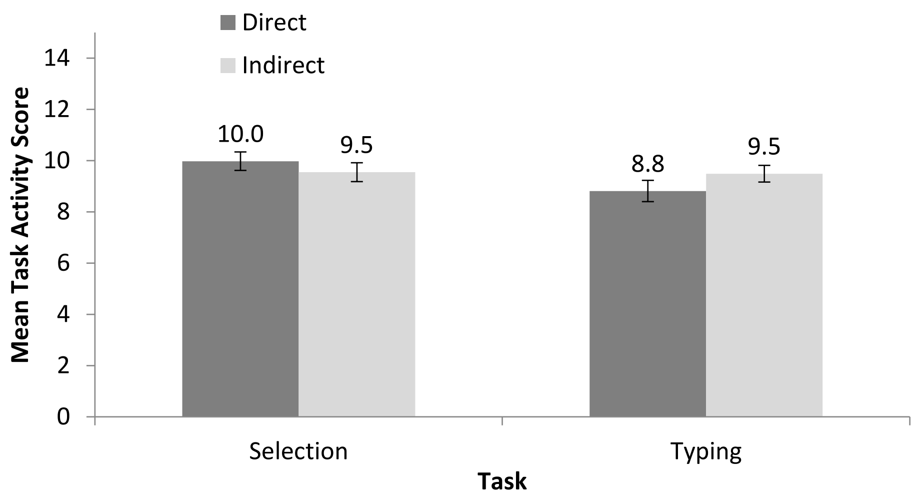


Figure 7. Mean scores for Task Activity (TA) by Task and Input Method. There was no effect for Input method (p = .46) but there was a main effect of Task, F(1,157) = 14.71, p < .001, η_p^2 = .09, and an interaction between Task and Input method, F(1,157) = 15.42, p < .001, η_p^2 = .09. Error bars represent the 95% confidence interval. Reprinted with permission from (Peres et al., 2017).

Shoulder Flexion (SF). SF had two items in it and the maximum possible score was 10. Participants reported a greater degree of SF when using Direct input (M = 5.39, SD = 1.3) compared to Indirect input (M = 4.30, SD = 1.4) irrespective of Task and this effect was significant, F(1,157) = 122.25, p < 0.001 (See Figure 8). Participants also reported a greater degree of shoulder flexion for Typing (M = 5.21, SD = 1.4) compared to Selection (M = 4.48, SD = 1.2) and this effect was also significant, F(1,157) = 96.59, p < 0.001. There was no interaction between Input and Task for SF (p = 0.21).

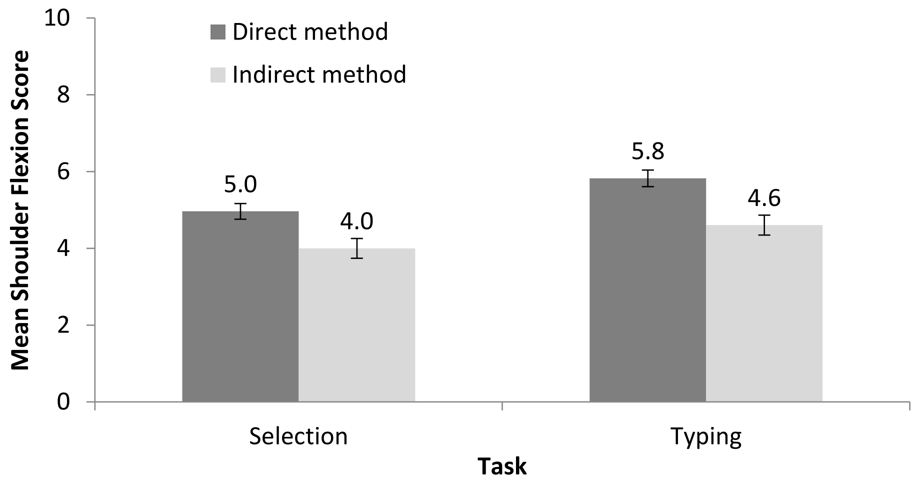


Figure 8. Mean scores for the Shoulder Flexion (SF) component by Input method and Task. SF consisted of two items with a combined maximum possible score of 10. There is a main effect of Input method, $F(1,157) = 122.25$, $p < .001$, $\eta_p^2 = .44$ and for Task $F(1,157) = 96.59$, $p < .001$, $\eta_p^2 = .38$. However, there was no interaction between Input and Task ($p = .21$). Error bars represent the 95% confidence interval. Reprinted with permission from (Peres et al., 2017).

Discriminant Validity – Non-parametric measures. The percent "yes" responses for Wrist Deviation (WrD) and Shoulder Abduction (SA) by Input method and Task are shown in Figures 9a and 9b. Wrist Deviation (WrD). For WrD (Figure 9a), a greater percentage of participants reported deviating their right wrist during the Typing (70%) task compared to the Selection (53%) task, $\chi^2(1, N = 636) = 24.20$, $p < 0.001$. There was also an interaction between Task and Input method, $\chi^2(1, N = 636) = 8.15$, $p = 0.004$. Specifically, the percentage of right WrD for Typing was greater when using the Indirect method (73.6%) compared to the Direct method (67.3%), however, the opposite trend was found for Selection, where more right wrist deviation was reported for the Direct method (59.1%) compared to the Indirect method (47.8%). For the left wrist, reported deviations occurred more often during Typing (61%) than Selection (8%), $\chi^2(1, N = 636) = 105.92$, $p < 0.001$. For both the right and left wrist, there was no effect of Input method on wrist deviation.

Shoulder Abduction (SA). As seen in Figure 9b, for left SA, there was a significantly greater percentage of left SA during Typing (15%) compared to Selection (5%), $\chi^2(1, N = 636) = 10.34$, p = 0.001. For the right shoulder, there was a greater percentage right SA when using the Direct method (36%) compared to the Indirect method (19%; $\chi^2(1, N = 636) = 10.64$, $p = 0.001$). Further, an effect of Task was found ($\chi^2(1, N = 636) = 8.24$, $p = 0.004$), with a greater number of participants reported right SA during Selection (32%) compared to Typing (22%). An interaction between Task and Input method was also found for right SA, $\chi^2(1, N = 636) = 11.59$, $p < 0.001$, where the difference between the Direct and Indirect input methods was significantly larger for Selection (D = 27%) than Typing (D = 6%).

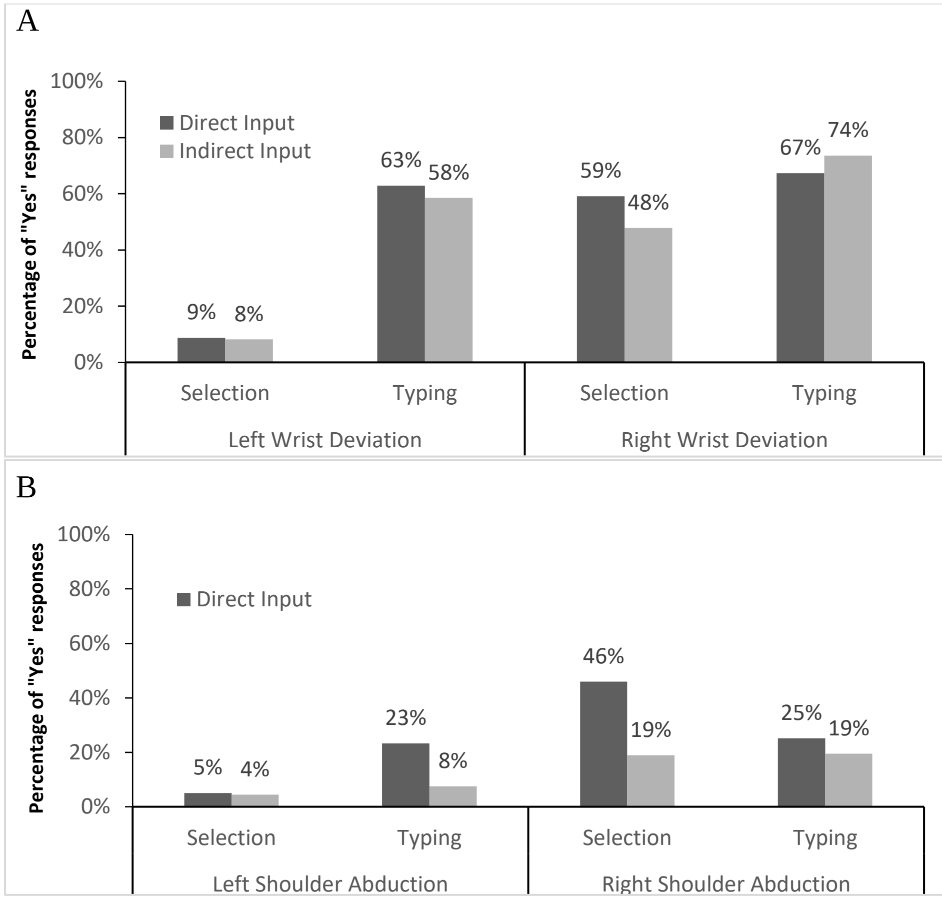


Figure 9. Percentage of “Yes” responses to a) Wrist Deviation and b) Shoulder Abduction. For right WrD (a), there was a main effect of Task ($\chi^2(1, N = 636) = 24.19$, $p < .001$, $\varphi = .19$) and an interaction between Task and Input method, $\chi^2(1, N = 636) = 8.15$, $p = .004$, $\varphi = .11$. For left WrD, there was a main effect of Task ($\chi^2(1, N = 636) = 105.92$, $p < .001$, $\varphi = .41$). For both the right and left WrD, there was no effect of Input method. For left SA (b), there was a main effect of Task, $\chi^2(1, N = 636) = 10.34$, $p = .001$, $\varphi = .13$. For right SA, there was a main effect of Input method ($\chi^2(1, N = 636) = 10.64$, $p = .001$, $\varphi = .13$), a main effect of Task ($\chi^2(1, N = 636) = 8.24$, $p = .004$, $\varphi = .11$), and an interaction between Task and Input method, $\chi^2(1, N = 636) = 11.59$, $p < .001$, $\varphi =.14$. Reprinted with permission from (Peres et al., 2017).

Discussion

Our effort in this study was to develop a self-report measure that met two measurement criteria, consistency and sensitivity (i.e., discriminant validity). The results of the study clearly indicate that SEAT meets these criteria. Although there were some differences between the two samples (convention and university), these differences were not evident in the effects on the measures of interest in this study. Further, the SEAT was relatively easy to answer for participants, as they were able to answer questions quickly and needed clarifications on only few questions (i.e., only those related to wrist deviation and shoulder flexion). This feedback will be incorporated in improving items in the SEAT. By the end of the experimental session, people took approximately three to four minutes to complete the SEAT, which supports the notion that this tool would be feasible in the iterative software development cycle.

Measurement Consistency

Internal consistency. With an overall Cronbach's alpha of 0.89 (Table 3), the SEAT meets established consistency criteria of an alpha greater than 0.70 (Cronbach, 1951; Cronbach & Shavelson, 2004). The Left and Right side components also meet this criterion but the other components do not (alphas < 0.66). This may be due to the small number of items in each of these components as five items or fewer in a scale often does not provide sufficient variability for calculating a true Cronbach's alpha (Cronbach, 1951; Cronbach & Shavelson, 2004). Future iteration of this survey will increase the number of items in these scales with the goal of improving their internal consistency.

Loading of the components. The PCA analysis revealed two interesting findings that have important implications for development and effectiveness of ergonomics self-report measures. One possible expectation of the groupings would have been that users' responses would group by body regions, i.e., both stressors (e.g., right wrist deviation) and the resulting strain (e.g., discomfort in the right wrist) from those stressors—

indicating that the measure was addressing aspects associated with a region of the body. Instead, the users' responses consistently grouped separately for stress and strain.

As reported in Table 3, of the seven factors that explained 64.64% variability in the self-reported data, the Left and Right Discomfort factors measure strain experienced on the left and right upper extremity musculature, whereas the other five factors (Work Demand, Task Activity, Wrist Deviation, Shoulder Abduction, and Shoulder Flexion) provide information on the stresses placed on the user. One explanation of this is that users provided their perception of discomfort (strain) immediately after the 3-minute task, thus they reported their current state. Whereas their perception of stress was retrospective in nature and required them to provide a cumulative assessment within the 3-minute task, specifically, they were using their memory to report a past state, albeit an immediately past state.

Another surprising grouping in the PCA was that participants' responses for strain consistently grouped based on the side of the body (Right vs. Left) but did not for stressors, particularly given that participants completed both unilateral tasks (selection tasks) and bilateral tasks (typing tasks). This lack of consistency between stressors of the left and right side could be an effect of the experimental design used or the design and presentation of the items on the SEAT, and thus needs further investigation.

Future work must be conducted that explores whether these groupings are found: to be valid (based on external observations); useful in the software development lifecycle; and generalize to different populations, devices, tasks, and durations.

Discriminant Validity

Previous findings (Duffield et al., 2013; Shin & Zhu, 2011) have found a difference between Input methods with regard to the strain experienced by the body. Given this and the components identified in the PCA, we would expect to find an effect of Input for components that describe discomfort (Left and Right side discomfort) or demand (Work Demand), and—since the arm had to be raised for the direct input condition and lowered for the indirect condition—the position of the shoulders (Shoulder flexion and Right shoulder abduction). As seen in Table 3, the expected effect of Input was found for those five components, thus indicating that the measure is sensitive to this effect and indeed the effect sizes are large (large effect: $\eta_p^2 > 0.14$) by Cohen's standards (Cohen, 1988). Interestingly, there was also an effect of Input method for Shoulder abduction with the Direct method having more than the Indirect method—particularly for the Selection task. This may be due to participants needing to abduct their shoulders to position their hand so that it does not obstruct their view of the display. If observational analyses validate this result, it would be an example of how SEAT could be used to discover strain and stressor relationships that are not expected.

Given the experiences of workers like geoscientists and graphic artists, it was possible that there would be a difference for Task with regard to strain on the body but there was no previous research on this. Further, we were not sure whether participants could self-report any difference between the tasks that might exist. Given the effects and effect sizes of Task for the Discomfort components, it appears there is an effect of Task design on discomfort. The effect of Task was also present for several of the stressor variables, i.e., Shoulder flexion, Wrist deviation (RL) and Shoulder abduction (RL)[1]. This indicates that participants may be able to report the stressors that are associated with their perceived strain. Further, the interactions of Task by Input method for both Wrist Deviation and Shoulder Abduction on the right side indicate that the stressors associated with the two tasks differed based on the Input method. This is the type of information software designers could incorporate into their design process—e.g., avoid incorporating numerous selection type tasks if the input method will be primarily touch (Direct). The next steps will be to validate these findings against the video observations of postures. If validated, this would indicate that the measure is sensitive to not only subtle strain components such as discomfort, but also stress components (such as posture). It is important to note that, for Left side discomfort, the effect of Task and the interaction of Task and Input are likely because one task (Selection) required much less use of the left side than the other task (Typing).

Conclusion

This paper presents an innovative method for translating ergonomic research to the domain of software interaction design—using a short, self-report measure that can be reliably completed by non-trained individuals. This is an important addition to the existing assessment methods for two reasons:

First, it allows for specific investigations on a potential source of ergonomic risk that currently has not been well explored. Previous studies and interventions have

[1] An effect of Task was found for Task activity but the effect sizes are much smaller than other effects, making them difficult to interpret. These effects may be a result of the power of the study or more about the pace someone chooses or needs to work than the design of the interface.

focused primarily on the physical aspects of the working environment, e.g., keyboard and mouse design; desk, monitor, and chair positioning; and resting practices (Blatter & Bongers, 2002; Ciccarelli, Straker, Mathiassen, & Pollock, 2011; Dennerlein & Johnson, 2006a; Dennerlein & Johnson, 2006b; Lee, Fleisher, 1 An effect of Task was found for Task activity but the effect sizes are much smaller than other effects, making them difficult to interpret. These effects may be a result of the power of the study or more about the pace someone chooses or needs to work than the design of the interface. McLoone, Kotani, & Dennerlein, 2007; Robertson et al., 2009; Shaw & Hedge, 1977; Sonne et al., 2012; Wu, Liu, & Chen, 2010). The results of this study suggest that the interaction design of software may indeed be a source of ergonomic risk and presents a potential measure for investigating these risks.

Second, it expands the methods available for assessing ergonomic risk to those untrained individuals and into the face-paced world of software development. Most current assessment methods are designed for trained ergonomists and, while ideal, this is not always practical. For example, when GUI designers are developing their software, it is simply not feasible for ergonomist to participate in every iteration of a software design. However, having designers test a prototype for a short period and then complete the SEAT fits well within the constraints of development cycles like Agile. When risks are identified, designers can then work with trained ergonomists to identify mitigation methods. This is the method currently used to identify designs that result in more efficient and effective interactions and with the SEAT it could be possible to design interactions that have less ergonomic risks.

The next stage in the development of the SEAT will be an additional validation effort including biomechanical and physiological correlates of the measure when users interact with different tasks, devices, and interaction methods. How the results of the SEAT are best communicated to the different user groups also needs to be established. This study does have a few limitations that warrant discussion. First, while participants were given adequate rest between conditions, development of muscle fatigue was not assessed. Also, the internal reliability (Cronbach's Alpha) for some of the scales is below typical standards, and there were differences (albeit small) between the two samples used. Although order effects were found, they had extremely small effect sizes and the use of a counterbalance likely controlled for these effects. Regardless of these limitations, the results presented here indicate that the SEAT can reasonably translate ergonomic assessment research to software interaction design.

VALIDATION OF THE STRESS COMPONENT OF THE SEAT

Introduction

Computer work involves low force, highly repetitive movements that, at times, users complete while maintaining awkward postures for long durations. These forces and movements are some of the primary risk factors responsible for computer work related upper extremity musculoskeletal disorders (MSDs; IJmker et al., 2007; Punnett & Bergqvist, 1997). Computer work has been long associated with an increase in the prevalence of upper extremity MSDs (Andersen et al., 2003; Blatter & Bongers, 2002; Heinrich et al., 2004; Newburger, 2001; Wahlström, 2005). While methodological limitations muddy the epidemiological evidence connecting computer use to the development of MSDs, associations between adverse upper extremity musculoskeletal outcomes and factors such as hours spent keying and poor placement of keyboards have provided guidance for the development of ergonomic interventions (Gerr, Monteilh, & Marcus, 2006). However, systematic reviews have found little evidence to support the effectiveness of singularly focused interventions like workstation adjustments, rest breaks, exercise, and job stress training (Brewer et al., 2006). There is evidence for a positive effect of comprehensive interventions that combine ergonomic training, workspace adjustments, and rest breaks (Dainoff, Maynard, & Robertson, 2012; Kennedy et al., 2010; Benden & Pickens, 2009). Comprehensive ergonomic interventions target several risk factors and it is likely that this multi-factor approach is why they are more effective than singularly focused interventions.

One often ignored and understudied potential source of ergonomic risks is that of software interaction design (Peres et al., 2016). The design of computer software applications dictates the ways in which users are physically required to provide input to those applications—therefore, interaction design is one of the fundamental sources of ergonomic risk when using computers. For example, if an application design requires repetitively clicking the mouse for long periods of time, an alternative mouse may be an effective tool for mitigating an ergonomic risk of low force repetitive movements. However, the risk would never have been introduced had the interaction design of the application precluded the requirement of rapid clicking.

Before risks like those borne from software interaction design can be controlled, they must be measured. The Self-report Ergonomic Assessment Tool (SEAT), developed by Peres, Mehta, and Ritchey (2017), is an ergonomic assessment tool intended to be used during the software development lifecycle as a tool to identify ergonomically risky interaction designs, or design elements, providing the ability for developers to design out a risk before the product ever reaches consumers. The SEAT is a 30-item self-report tool that was created through leveraging and modifying existing ergonomic assessment tools. All items included on the SEAT were re/written in the form of self-report items. The SEAT was specifically designed to be a self-report tool that is usable without considerable training or ergonomic expertise as reliance on trained personnel to use the tool would significantly complicate design processes by being costly in terms of both time and money. Software development is a quick and iterative process that could be disrupted by lengthy or otherwise costly ergonomic assessments at each iteration

(Martin, 2003). Designing the SEAT as a self-report tool with a focus on a quick completion time ensures that it can be an effective tool for use during software development.

In the domain of computer work, current self-report tools have been developed that focus on postures, movements, and office setup (Dane et al., 2002; Heinrich, Blatter, & Bongers, 2004; Li & Buckle, 2000; McAtamney & Corlett, 1993; Robertson et al., 2009; Sonne, Villalta, & Andrews, 2012; Speklé et al., 2009) whereas the SEAT is being developed to ensure enough sensitivity to detect ergonomic risks differences between interaction designs or individual design elements. The first use of the SEAT in a study demonstrated that participants' responses on the SEAT differed between a clicking task and a typing task, as well as between the use of touch as an input method and the use of a standard keyboard and mouse (Peres et al., 2017). This finding of sensitivity is promising, however, SEAT requires further validations before the measure is deployable.

One of the first steps in the continued development of the SEAT is to assess the reliability of participants' self-reporting of ergonomic stressors and associated strains. The SEAT comprises several items from existing ergonomic assessment tools that measure postures, speed, force, activity level, i.e., the stressors. The SEAT also measures discomfort, i.e., strain, across different body regions. The present paper focuses on the reliability and validity of the stressor items of the SEAT. For the SEAT to be a viable self-report tool that non-ergonomists can use, it is important to compare participants' self-reported stressors (i.e., was their wrist as flexed as they reported it to be) to that of external, formally trained raters.

General Methods and Analyses

Overview

The methods and analyses presented here are from two studies that assessed the reliability of participants' self-reported responses to the SEAT as compared to that of trained raters. In Study 1, trained raters completed the SEAT using video recordings of a subset of participants from Peres et al. (2017). These ratings were compared with participants' own self-reports to determine if participants can reliably report each stressor item within the SEAT. Results from these comparisons and from Peres et al. (2017) were used to inform the creation of a second iteration of the SEAT. In Study 2, task videos from a follow-up study in which the second iteration of the SEAT was administered were reviewed by a single trained rater who used them to complete the SEAT. The rater's responses were then compared to participant's self-reported responses.

Procedure

For each of these studies, a version of the SEAT was administered after participants completed tasks on a computer In Study 1, participants completed a total of four tasks and in Study 2 participants completed a total of eight tasks. The protocols as well as the methods used to process and analyze the videos from both studies are presented here.

Data Analyses

Computing inter-rater reliability statistics investigated the relationship between participants' self-reported responses of the stress items on the two versions of the SEAT and raters. Specifically, linearly weighted kappas were computed for Likert-type response items and non-weighted kappa for dichotomous response items. In addition to a comparison of absolute agreement (i.e., kappa statistic), Spearman's Rho correlations were computed (and similarly, Pearson's Phi was computed for dichotomous response items) to investigate the rank correlations between participants' self-reported responses on the SEAT and the raters' responses. Differences between participant-rater agreements on SEAT 1.0 (Study 1) were used to inform design changes and the creation of a second version, SEAT 2.0, which was used in Study 2.

Study 1

Participants

Video data of participants completing a series of tasks with the computer were collected during experimental sessions conducted at two sites: a geosciences convention in Houston, Texas, USA during April of 2014 (N = 56: age range 18 to 72 years) and at Texas A&M University in a usability lab (N = 110; age range 18 to 80) during the summer of 2014. The participants recruited at the convention site were, on average, 10 years older and spent three hours more per week working on a computer. There were no significant differences between the two groups otherwise (Peres et al., 2017). This study was approved by the Texas A&M Institutional Review Board.

In this study, the videos of 42 participants were randomly sampled using computer generated random numbers from the final data set of 159 participants. Participants were sampled equally from both sites and participant demographic data for the sampled participants by data collection site are shown in Table 5.

Table 5. Participant demographics of the 42 participants sampled, by data collection setting. Mean values are reported with standard deviation cited parenthetically. Gender is given as the percentage of the samples who identified as female.

	N	Age (years)	Hours/Week Spent Working on Computer	Years of Computer Use	Gender (% female)	Weight (kg)	Height (cm)
Field Setting	21	36.3 (11.6)	11.9 (10.7)	13.4 (6.9)	38.1%	79.6 (17.0)	176.1 (11.7)
Lab Setting	21	26.2 (08.4)	9.9 (08.3)	10.3 (3.9)	33.3%	70.2 (13.8)	170.1 (10.7)
Total Sample	42	31.3 (11.2)	10.9 (09.5)	11.9 (5.7)	35.5%	74.8 (16.1)	173.1 (11.4)

Workstation and Input Methods

Participants were seated at a computer workstation and instructed to self-adjust the chair to their comfort. Participants completed all tasks on a desktop personal computer with touch capabilities (Dell Optiplex 9020 All-In-One PC running Windows 7), and interacted with the computer via the touchscreen or a mouse and keyboard (Dell KM623 Wireless Keyboard and Mouse) depending on the task.

Tasks

Each participant completed a clicking task and a typing task—once using only the touchscreen to interact with the computer and once using only the keyboard and mouse to interact with the computer. The 2x2 repeated measures factorial design (2 Tasks X 2 Interaction Methods) created a total of four tasks that participants completed. Each task lasted for three minutes, and the order in which tasks were presented was constant across interaction method, with that order and the order of interaction method being counterbalanced. For the typing task participants transcribed short phrases into text fields located directly under each phase. Detailed task descriptions are presented in Peres et al., (2017). The clicking task required participants to click on a hyperlinked word with a short phrase and then click “Ok” on a subsequent pop-up dialogue box. Participants completed the tasks at their own pace and continued to transcribe or click/touch for the full three minutes.

Measures

During all tasks, participants were video recorded as they complete each task with a digital camera placed on their right side and positioned to capture a side shot (i.e., parasagittal view), at about chest height. After completing each task, participants completed the SEAT 1.0. The SEAT 1.0 comprised 30 items created through the modification of existing assessment measures (Table 6). The Borg CR10 was used as a measure of whole body exertion (BORG: Borg, 1990) as was the exertion item from the Strain Index (SI: Moore & Garg, 1995). A modified version of the Body Discomfort Diagram (BDD) was used to measure participants' discomfort for seven body parts bilaterally (wrists, lower arms, upper arms, shoulders, trapeziuses, and the left ulnar and radial sides of the hands), as well as the eyes and neck (Cameron, 1996). Posture questions taken from the Rapid Upper Limb Assessment (RULA) included wrist extension/flexion, wrist abduction, shoulder extension/flexion, and shoulder abduction (McAtamney & Corlett, 1993). The Hand Activity Level (HAL) measure was included to capture speed and repetitiveness of motions for both the right and left hands (Ebersole & Armstrong, 2002). Additionally, the SI's item for speed of work was included. One additional item, developed specifically for the SEAT, was added to determine self-reported task precision requirements. The full SEAT 1.0 used for Study 1 is included in Appendix A.

Table 6. SEAT 1.0's items grouped by question type. *Indicates that the item is answered for the right and left side of the body, e.g., Hand Activity Level is asked for the right hand and then the left hand. For body part discomfort, 2 items (neck and eyes) are not bilateral.

Question Type	Existing Measure	Number of Items
Body part discomfort	BDD	16*
Whole body exertion	BORG C10 & SI	2
Hand Activity Level	TLV for HAL	2*
Wrist Flexion/Extension	RULA	2*
Wrist Abduction/Adduction	RULA	2*
Shoulder Flexion/Extension	RULA	2*
Shoulder Abduction	RULA	2*
Speed of Work	SI	1
Task Precision Demand	n/a	1

Video Clip Sampling Procedure

For each of the three minute experimental tasks, a 20 s segment of video was sampled from each minute. The 20 s video clip was taken from the same location for each of the three minutes for any given participant’s task, creating three video samples per task. Figure 10 shows the three sampling strategies used to generate the 20 s segments: A) the 20 s clip was taken at the start of each minute of the task (i.e., 0:00 to 0:20; 1:00-1:20; 2:00-2:20), B) the 20 second clip was taken during the middle of each minute, and C) the 20 s clip was taken at the end of each minute. These three sampling strategies were chosen so that, with respect to raters’ responses, difference across each minute of the tasks could be investigated to determine if participant's postures, effort, or task speed differed between the minutes of the task. The sampling strategy used for each participant's video data was randomly assigned such that an equal number of each sampling strategy were assigned to participants.

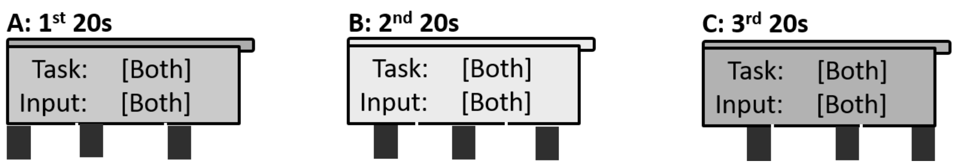


Figure 10. The three sampling strategies used: A, B, and C. The three blocks at the bottom of each strategy represent the points during the tasks a 20 s video clip was made, either the first 20s, second 20s, or third 20s of each minute of the task.

Video Clip Rating Procedure

Two raters were trained on how to code the video clips by answering each stressor item of the SEAT 1.0 (i.e., all items except the body discomfort items) based on their assessment of the video clips. Both raters were graduate research assistants in an occupational safety and health program at Texas A&M University who had successfully completed structured practicum industrial experiences (minimum of 200 hours) in the field of human factors and ergonomics as well as taken master-level courses in ergonomics and occupational safety. Rater training consisted of collectively operationalizing and agreeing on how each question should be responded to, informed by how those questions were to be answered in the measures from which they originated (e.g., BORG CR10, Strain Index). The two raters independently coded trials until the differences between their responses were within the following predefined threshold: for questions with scales that had a range of greater than five, the rating threshold was defined as within one rating point (i.e., a rating of a 5 and 6 would be considered acceptable agreement); for question with 5 or fewer response options, the rating threshold was defined as an exact match. After raters completed 13 trials, created from participants' video that were not included in the experimental sample, the specified threshold was met and they completed their reviews of the video clips independently.

Analyses

Sampling Strategy Analysis. Repeated measures analyses of variance (RMANOVAs) were conducted on both raters' responses to test for effects of sampling strategy and task minute (see Figure 10). For each item on the SEAT 1.0, a mixed RMANOVA was conducted with sampling strategy (3 levels: A, B, C) as a between-subject variable and three within-subject variables: task minute (3 levels: 1^{st}, 2^{nd}, 3^{rd}), task (2 levels: clicking, typing), and input (2 levels: Direct method, Indirect method).

Reliability and Validity Analysis. Reliability of the SEAT 1.0 was assessed by examining the level of agreement between the two raters (inter-rater agreement) by computing weighted kappas and Spearman's rhos for each of the ordinal stress items (non-weighted kappas and Pearson's phi were computed for dichotomous items). Validity of the SEAT 1.0 was similarly assessed by examining the level of agreement between the raters' responses and participant's response. The categories outlined in Table 7 were used to aid in the interpretation of the kappa values (Viera & Garret, 2005).

Table 7. Qualitative description of agreement levels. Based on the values of weighted kappas taken from Viera & Garret (2005).

Descriptor	Weighted Kappa Value Thresholds
Very Weak	0.00 to 0.19
Weak	0.20 to 0.39
Moderate	0.40 to 0.59
Strong	0.60 to 0.79
Very Strong	0.80 to 1.00

Results

Sampling Strategy Results. For both raters, the only effect of sampling strategy was found for the Right HAL item. For rater 1 ($F(2,155) = 4.66$, $p = .011$, $\eta_p2 = .057$), the second (M = 3.51 , SD = 0.84) and third (M = 3.46, SD = 0.84) 20 s interval samples were rated higher than the first 20 s interval (M = 3.07, SD = 0.84). Similarly, for rater 2 ($F(2,149) = 6.31$, $p = .002$, $\eta_p2 = .078$), the second (M = 4.73, SD = 0.94) and third (M = 4.51, SD = 0.92) 20 s interval samples were rated higher than the first 20 s interval (M = 4.10, SD = 0.96). A similar trend was found for rater 2 with respect to an effect of task minute on Right HAL ($F(2,149) = 3.47$, $p = .032$, $\eta_p2 = .023$) with lower scores being given for the first minute (M = 4.36, SD = 0.62) compared to the second (M = 4.52, SD = 0.60) and third minutes (M = 4.45, SD = 0.59).

For rater 2, an effect of task minute was found for Left Shoulder Flexion ($F(2,150) = 7.797$, $p < .001$, $\eta_p2 = .049$), Right Shoulder Flexion ($F(2,150) = 9.77$, $p < .001$, $\eta_p2 = .061$), and Speed of Work ($F(2,149) = 5.87$, $p = .003$, $\eta_p2 = .038$). For each of these effects, the same trend was found in which scores increased as task time increased, with the highest level of stressors reported in the third minute. Given the small sizes of the effect found ($\eta_p 2$ range = .023 – .078), the reliability and validity analyses are presented, and subsequently interpreted, collapsed across sampling strategy and task minute.

Reliability Results. Weighted kappa's and Spearman's rhos for each ordinal stress item and non-weighted kappa and Pearson's Phi for the four dichotomous variables are reported in Table 8. The Left Wrist Deviation ratings provided by rater 1 lacked sufficient variability to be used. Agreements between raters (as measured by kappas) were moderate to strong for all items except the Strain Index (SI; $\kappa_w = 0.14$), Speed of Work ($\kappa_w = 0.04$), Right Hand Activity Level ($\kappa_w = 0.18$), Left ($\kappa_w = 0.24$) and Right Shoulder Abduction ($\kappa_w = 0.33$). Similarly, correlations (a measure of more ordinal versus exact agreement) between rater responses was also high, only three correlations were considered weak (SI, rho = 0.28; Left Shoulder Abduction, phi = 0.32; Right Shoulder Abduction, phi = 0.38).

Validity Results. Given their similarity, agreement scores (kappa and rho) for each pair (i.e., rater 1 vs. participant and rater 2 vs. participant) were averaged together and presented in Table 8. All agreements were considered none or very weak ($\kappa_w < .20$) except for a moderate agreement for both Left Hand Activity Level (κ_w 0.50) and Left Wrist Extension/Flexion (κ_w 0.40).

A strong correlation was found between raters' responses and participants for Left Hand activity Level ($rho_{rater1} = 0.80$; $rho_{rater2} = 0.79$). Moderate correlations were found for the following items on the SEAT 1.0: Borg (rho = 0.44), Left Wrist

Extension/Flexion (rho = 0.48), Left Shoulder Flexion (rho = 0.46), and Right Shoulder Flexion (rho = 0.46).

Table 8. Kappa Values for SEAT 1.0. Weight kappa for ordinal variables, non-weighted for binary and Spearman Rho for ordinal variables, Pearson Phi for binary variables.

	Inter-Rater Reliability		Raters† vs. Participant	
	Kappa	Rho	Kappa	Rho
Borg (exertion)	0.62	0.95	0.10	0.44
Left HAL	0.66	0.94	0.50	0.79
Right HAL	0.18	0.60	0.06	0.16
Left Shoulder Flexion	0.66	0.72	0.18	0.38
Right Shoulder Flexion	0.68	0.78	0.19	0.46
Left Wrist Extension/Flexion	0.62	0.76	0.40	0.48
Right Wrist Extension/Flexion	0.51	0.59	0.13	0.14
Left Wrist Deviation*	x	x	-0.01	-0.06
Right Wrist Deviation*	0.57	0.58	-0.01	-0.04
Speed of Work	0.04	0.51	0.11	0.29
Left Shoulder Abduction*	0.24	0.32	0.03	0.09
Right Shoulder Abduction*	0.33	0.38	0.05	0.08
SI (exertion)	0.14	0.28	0.01	0.04

*non-weighted kappas and Pearson's phi were computed because these are dichotomous variables.

†Kappas and rhos for rater 1 vs. participant and rater 2 vs. participant are presented here as an average.

Discussion

Agreement between the two independent raters was, on average, moderate (ranging from 0.41 to 0.60; Viera & Garrett, 2005). For right and left shoulder abduction agreement was fair (0.21-0.40). However, for some items, agreement was poor or even non-existent (≤ 0.20): Right HAL, Speed of Work, SI (Exertion). These results show that for some items the trained raters had a reasonable amount of agreement. However, no items were found to have above a moderate amount of agreement. This could indicate an issue with how some items on the SEAT were designed or reflect on the complicated nature of subjective evaluations of ergonomic risks in the office. Ordinal agreement between raters was very high, as reflected by the Spearman Rho values for some of the items, indicating that while raters may not absolutely agree on the right answer for a stressor item, they were in agreement regarding the relative differences between the stressors present during the tasks. Similarly lack of agreement between ergonomic raters has been found in application of ergonomic assessments of office workers (Liebregts, Sonne, & Potvin, 2016; Pereira, Straker, Comans, & Veneria, 2016) as well as for non-office workers (Park et al., 2009; Eliasson et al., 2017). The use of the same rater for any given set of SEAT responses may be sufficient for assessing the reliability of participants' responses.

Agreement between raters and participants showed that for most items there was no to poor specific agreement between participant and raters. There was moderate agreement for the Left HAL. This agreement for the Left HAL, but not for the Right HAL, is likely because for some of the tasks there was no use of the left hand, i.e., the left hand remained idle. Evaluations of an idle hand would understandably be easier to agree on than, for example, determining if a hand is should be rated as a 5 or 6 on a 10 point scale.

Creating SEAT 2.0. Changes to the SEAT 1.0 were made based on results of Peres et al. (2017) and further informed by the results presented here. The wrist ulnar/radial deviation and shoulder abduction items were changed from dichotomous response items (i.e., yes or no) to 5-point Likert-type response items. This change was made to address the complete lack of agreement found for these items (κ ranged from -0.01 to 0.05) likely caused by confusion around how much deviation or abduction is needed for an affirmative response to be given. Adding more response options could create concrete, intermediary threshold of abduction/deviation that participants can use for their response and could improve participants' ability to correctly respond to this item. Additionally, changing these items to include Likert-type response scales could increase the consistency between all of the items on the SEAT, as the radial/ulnar wrist deviation and shoulder abduction questions were the only two dichotomous response items. The SI exertion item was also removed due to a complete lack of agreement between participants and raters ($\kappa_w = 0.01$). Furthermore, the SI and Borg items are redundant as both address the same construct of exertion and do so similarly. The difference between these two exertion items is that the Borg offers a larger scale for participants to use which could be the main factor in the much higher kappa and rho values for the Borg compared to the SI.

While the strain components of the SEAT (i.e., body discomfort) are beyond the scope of this study, as self-reported discomfort ratings cannot meaningfully be compared to those of an external rater, some changes were identified based on previous analyses and experiment observations. For the body discomfort diagram, the eye and finger items were removed. Some participants found the juxtaposition of eye discomfort/strain and body part discomfort to be jarring. The finger items that differentiated between 'ring and pinky' and 'thumb, index, and middle finger' were reduced to one discomfort item that represented the entire hand. Based on the principal component analysis of the SEAT conducted by Peres et al., (2017) it is likely that the differentiation of the hand in to two, finger-specific sections is unnecessary and problematic as, for example, right ring and pinky finger discomfort did not load onto the same factor with the other right side discomfort items. The SEAT 2.0, used in study 2 presented below, is included in Appendix B.

Based on the findings of the effects of sampling strategy and task minute, we also hold that the method of using two raters and multiple samples per task and per minute is not necessary. Overall, the effects found for sampling strategy and minute of task show that speed of work was greatest and postures were their worse during the last third minute of the tasks. Video samples be taken toward the end of the task and given the small effect sizes found ($\eta_p 2$ range = .023 – .078) only one sampling point is needed when observing repetitive tasks like the ones used here.

Limitations. A lack of more diverse and representative devices, as well as tasks, limits the generalizability of this study. The modern office environment is no longer comprised uniformly with desktop computers. Notebooks, tablets, and smart phones are important device categories to consider when developing an assessment tool like the SEAT. The tasks used in this study were contrived and not related to the common types of tasks completed using electronic devices. Assessing the SEAT using more representative office computer tasks is needed.

Study 2

Participants

A total of 32 participants were recruited from respondents to a recruitment email sent to a university-wide student email distribution list using a random selection process. Only those participants who reported that they have never been diagnosed with a musculoskeletal disorder of the upper extremities, neck, or shoulder were included in the study. Two participants' data were excluded because of data collection issues, creating a final sample of 30 participants. Participant demographics are shown in Table 9. The Texas A&M Institutional Review Board approved this study, and participants provided their consent before the start of the experiment.

Table 9. Participant demographics of the 30 participants in Study 2. Mean values are reported with standard deviation cited parenthetically. Gender is given as the percentage of the samples who identified as female.

	Mean (SD)
Age (years)	20.5 (2.7)
Hours/Week Spent on Desktop or Notebook Computer	38.0 (22.2)
Years of Daily Computer Use	8.0 (4.3)
Gender (% female)	73.3%
Weight (kg)	68.3 (21.6)
Height (cm)	165.9 (9.6)

Workstation Setup

Participants were seated at a workstation with an ergonomic, adjustable task chair, and a height adjustable desk for all devices and tasks. Prior to beginning their first task, the chair and desk were adjusted such that the work station was complaint with HFES/ANSI 100-2007 standards. After the workstation was adjusted, participants were allowed to further adjust to their own comfort.

Devices

For the purpose of this study we used four computer devices, presented in a counterbalanced order, that represent the most common devices used in both occupational and personal settings: iPad Air 2 Tablet, iPhone 5s Smart Phone, MacBook Pro Notebook, and a traditional Desktop setup (Docked MacBook Pro with external Monitor, keyboard, and mouse). Apple® devices were chosen to eliminate as much difference as possible between devices with respect to the application used and operating system. All tasks were completed using the native apps Mail and Calendar, which were connected to a Google Account created for the study. The tablet and the smart phone were placed in protective cases, as this reflects how these devices are commonly used (NPD, 2013).

Tasks

Two tasks were completed by participants on each of the four devices—an emailing task and a calendaring task (Figure 11). Both tasks were three-minutes long. These tasks were designed to be representative of commonly performed tasks on the four devices that differ with respect to interaction method. During the emailing task participants spent about 90% of their time typing and during the calendaring task participants spent about 90% of their time clicking/touching. The order in which the tasks were performed was counterbalanced between participants but held constant across devices for each participant. For the emailing task, participants responded to emails and were given instructions to do so "as fast and as accurately as possible" for the entire duration of the three-minute task. Participants responded to emails by transcribing the subject line of the email into the body of their response. For the calendar task, participants created calendar events based on event details provided for them on a printed sheet of paper which was placed in a document holder on the desk. Between each task there was a five-minute rest period, during which the experimenter verbally administers the SEAT 2.0 and participants were required to respond verbally to each item.

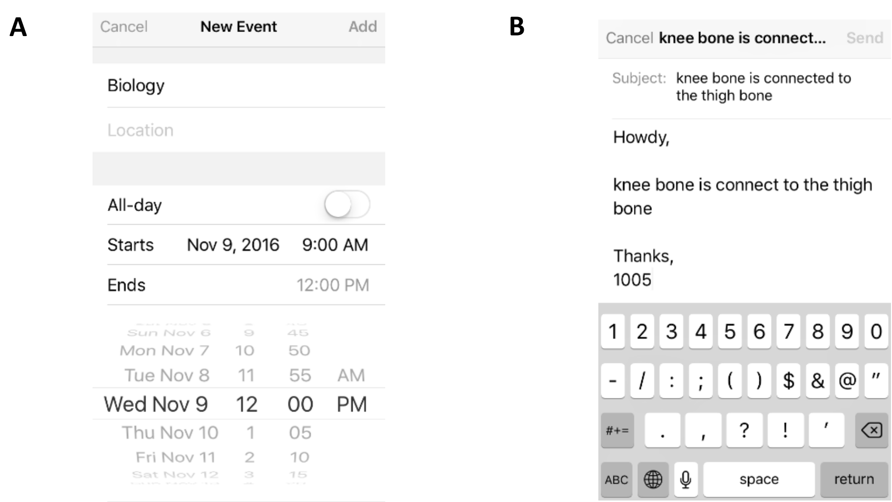


Figure 11. Screen shots of the two tasks. a) Screenshot of how participants created calendar events using the smartphone during the calendaring task. b) Screenshot of how participants replied to the emails using the smartphone during the emailing task.

Measures

Video recordings of each participant were taken during the experimental sessions using two digital video cameras. One camera was placed perpendicular to the participant's sagittal plane, on the right side of their body and at elbow height, to capture a side-shot as they completed the tasks. A second camera was placed on their left side, angled downward and placed at shoulder height which provided a view that aided in the rating of wrist abduction angles and hand activity levels. The placement of these two cameras was to ensure that all the participants' relevant body parts were visible by raters who would later be watching the videos.

Physiological data were also collected continuously during the experimental tasks. Surface Electromyography was used to collect muscle activity data and Near Infrared Spectroscopy was used to collect muscle oxygenation data. The methods for the collection of these physiological data and their analyses are beyond the scope of this study.

Video Clip Sampling Procedure

The video data from all 30 participants' sessions were included in the analyses. One 20 s clip from the start of the third minute of each task was taken from each task for each participant. Results from Study 1 indicated that the highest levels of stressors were found in the last minute of the tasks and that and multiple sampling points with a single minute were unnecessary.

Video Clip Rating Procedure

Based on results from Study 1, one occupational health and safety graduate student rated all of the video clips instead of two because of the similar level of agreement between participant responses and either of the two raters found in Study 1. The rater reviewed the training material used for the first study as well as operationalized guidance for the rating of new or altered items of the SEAT 2.0. These new guidelines were reviewed and agreed upon by two faculty researchers. The rater completed the SEAT 2.0 using each of the video clips pairs (i.e., side-shot and shoulder-shot for a given task).

Analyses

Validity of the SEAT was assessed through the level of agreement between participant's self-reporting of stress items on the SEAT 2.0 and the trained rater. Agreement was determined by computing weight kappas for each stress item. Spearman's Rho correlations were also computed to assess ordinal relationships between participant and rater responses. For the purpose of interpreting weight kappas, the categorization thresholds shown in Table 13 were used (Viera & Garrett, 2005).

Results

As in Study 1, the level of agreement between participants' self-reports and the trained raters scoring was assessed by computing weight kappa's and Spearman's rhos for each stress item on the SEAT 2.0. Weighted kappas and Spearman's rhos for each stress item are reported in table 10. The Borg item of the SEAT 2.0 lacked sufficient variability with respect to the trained raters' responses (a rating of 1 was given for all but 8 of the samples) and as such kappas and spearman's rhos were unable to be computed.

Only Neck Extension/Flexion item had a kappa that indicated a fair amount of agreement between participant's self-report and the trained rater, $\kappa_w = 0.25$ (95% CI, 0.16 to 0.34), $p < .001$. All other weighted kappas indicate that the trained rater's scoring of the SEAT 2.0 had very little or no agreement with participants self-reported stressors.

Weak correlational relationships were found between participants' self-reporting and the raters evaluation for Neck Flexion/Extension, Rho = 0.37 (95% CI, 0.25 to 0.48), $p < 0.001$, Left HAL, Rho = 0.32 (95% CI, 0.10 to 0.35), $p < 0.001$), Right HAL, Rho = 0.23 (95% CI, 0.20 to 0.43), $p = 0.001$), and Right Shoulder Flexion/Extension, Rho = 0.20 (95% CI, 0.07 to 0.32), $p = 0.003$).

Table 10. Weighted kappas and Spearman's rhos for each item on the SEAT 2.0. Agreement statistics could not be computed for the Borg item because the rater's responses lacked sufficient variability.

	Rater vs. Participant	
	Weighted Kappa	Spearman's Rho
Borg (exertion)	[no variance in Rater responses]	
Neck Extension/Flexion	0.25	0.37
Left HAL	0.12	0.32
Right HAL	0.10	0.23
Left Shoulder Flexion	0.09	0.19
Right Shoulder Flexion	0.08	0.20
Speed of Work	0.09	0.18
Left Shoulder Abduction	0.09	0.11
Right Shoulder Abduction	0.15	0.16
Left Wrist Deviation	0.02	0.05
Right Wrist Deviation	0.04	0.07
Left Wrist Extension/Flexion	-0.09	-0.16
Right Wrist Extension/Flexion	0	-0.02

Discussion

Overall, agreement between participants' self-reporting of stressors and the evaluation of an external rater was lacking, with only Neck Flexion having an agreement greater than none-to-poor. These results indicate that participants may not be able to accurately self-report their own postures and hand activity levels. However, correlations between participant and rater responses were markedly better. While participants are unable to accurately report the precise levels of stressors, they are able to self-report ordinal changes in their postures and activity across different tasks and devices.

Changes made to the SEAT based on the results from Study 1 did result in small improvements in agreements for those items that were updated, however, these improvements were not substantial. Left and Right Wrist Deviation, which were changed from dichotomous response items to Likert-type response items, showed only a small increase in rater-participant agreement for SEAT 1.0 (Right and Left = -0.01) compared to SEAT 2.0 (Right = 0.04; Left = 0.02). Right and Left Shoulder Abduction, similarly changed from dichotomous to Likert-type response items, exhibited a far greater increase in agreement when comparing SEAT 1.0 (Right = 0.05; Left = 0.03) to SEAT 2.0 (Right = 0.15; Left = 0.09). While these changes in agreement do not elevate these items out of the category of "very weak" agreement, they do provide support for the changes made.

Validity of Self-reporting Stressors with the SEAT. Participants' inability to self-report the office ergonomic stressors while interacting with different tasks/devices using the SEAT, as evidenced by the lack of agreement between participants and raters, was unexpected but have important implications. One very important factor to note is that Study 2, which resulted in the worst agreement (Mean Agreement for SEAT 1.0 = 0.13; Mean Agreement for SEAT 2.0 = 0.08), included tasks and devices that were more representative to the office environment than the tasks and devices used in Study 1. The highly contrived and intentionally different interaction methods used in Study 1 (using the touch screen vs. keyboard and mouse) created differences that were large and easily detectable and reportable, whereas in Study 2, the differences between tasks and devices with respect to stressors was subtle and more representative of real world conditions. It is possible that agreement between participants' self-reports and that of trained raters is difficult to achieve when the levels of stressors present are more subtle compared to pronounced differences like in Study 1. It may be difficult for users to self-report stressors in the office environment due the lack of range those stressors exhibit.

Physical stressors, like wrist and shoulder flexion/extension, are important factors to assess during office ergonomic interventions. However, assessing stressors in an office environment using self-report techniques has been shown to result in low levels of reliability. For example, Heinrich, Blatters, & Bongers (2004) found that computer workers' self-reported work postures were not reliable. In a study where officer workers' self-reported scores on the Rapid Office Strain Assessment (ROSA) were compared to observers, Sonne & Andrews (2012) found significant differences between the two sets of scores. While self-reported risk factors for MSDs are the least reliable when compared to observational and direct measurements, self-reported data may be made more reliable if adjusted to account for self-reporting biases (Spielholz et al., 2001). For tasks outside of the context computer work, like manual material handling, that involve postures and movements with greater ranges of motion, self-report methods can be useful (Wiktorin et al., 1993). Self-reported stressors are only well reported when postures are very general, whereas when workers are asked to self-report subtle and specific body angles, these stressors are not well reported (Stock et al., 2005). The results from Study 1 and 2 lead to a similar conclusion. The subtle and small differences between postures adopted during computer use make self-reporting stressors related to computer work very difficult, as shown by the low agreement between the participants and raters with the lowest occurring during the more subtle tasks and device differences in Study 2. However, participant and raters responses did correlate, which indicates the possibility of using score adjustments to render self-report stressor data more useful (Spielholz et al., 2001).

*Limitations***.** It is possible that the raters using the SEAT were a significant source of error and any interpretation of the validity of self-reported stressors when compared against the raters becomes problematic. However, the inter-rater reliability of the SEAT is comparable to that of other assessment methods, including those assessment tools which were used to develop the SEAT. A study conducted using participants from manufacturing and health care industries found similar inter-rater reliability for HAL (kappa = 0.44, Spearman rho = 0.65) as we found for SEAT 1.0 (with the exception of the kappa value for the HAL item on the SEAT, kappa = 0.18) and a lower correlation between raters for the BORG CR-10 (Spearman rho = 0.28; Spielholz et al., 2008). In other studies of industrial tasks, similar inter-rater reliabilities have been found for HAL

(Paulsen et al., 2014). However, high levels of agreement for industrial tasks have also been reported for both hand activity (Ebersole & Armstrong, 2002) and exertion (Stevens, Vos, Stephens, & Moore, 2004).

Other tools with which trained raters assessed ergonomic stressors via observations have shown similar inter-rater reliability as well, such as the PATH method and ROSA (Park et al., 2009; Liebregts, Sonne, & Potvin, 2016). Inter-rater agreement with the SEAT was superior compared to studies where observations were made in real time (Burt & Punnett, 1999) or without a systematic approach (Eliasson et al., 2017). However, some studies have reported higher levels of inter-rater agreement, like the original validation study for the ROSA (Sonne, Villalta, & Andrews, 2012) and for the Assessment of Repetitive Task method (ART; Roodbandi, Choobineh, & Feyzi, 2015). One source of rater error Liebregts et al. (2016) noted during their work with the ROSA is the parallax effect. When the viewing angle of an observer, or video recording, is not perfectly orthogonal, the ability to determine the angle of a joint, i.e., how much a wrist is extended or flexed, is negatively impaired. In Study 1 and Study 2, cameras were not positioned in such a way to completely control for this parallax effect. The parallax effect could account for some of the lack of agreement between the raters in Study 1 and Study 2, however, it is unlikely this effect was great enough to seriously complicate interpretations of rater-participant agreements.

Conclusion

Based on the results of Study 1 and 2, participants are unable to reliably self-report stressors like posture and hand activity when using the SEAT. Further iterations of the SEAT could address the lack of participant-rater agreement either through instruction or item redesign or through post-completion adjustments to responses. However, for the purpose of the SEAT (i.e., for use during software development), it is important to note that strain, e.g., discomfort, is also a valuable factor to assess and may be more easily, quickly, and reliably assessed. Strains are the most proximal causes of MSDs in the office environment as strain results from a person's exposure to stressors. Given the difficulties and lack of validity of self-reported stressors when using the SEAT, it may be that evaluating stressors via self-report is less useful compared to assessing strain. Assessing participant's ability to self-report strain is an important next step in the development of the SEAT.

VALIDATION OF THE STRAIN COMPONENT OF THE SEAT

Introduction

The development of the Self-report Ergonomic Assessment Tool (SEAT; Peres et al., 2017) was a step towards controlling the risks put into play due to software interaction design. The SEAT is currently in its second iteration, version 2.0, and it comprises items modified and adapted from existing measures. The SEAT is unique compared to other ergonomic assessment tools in that it is intended to be useable without formal training in ergonomics—the SEAT asks users to self-report relevant ergonomic stressors (e.g., repetitions, postures, speeds) and strains (e.g., discomfort).

For the SEAT to be reliable and valid, evidence is needed to show that participants using the SEAT are able to meaningfully report the stressors and strains they experience when using computer devices. With respect to the reliability of self-reported stressors, participants' responses to the SEAT items on repetitions, postures, exertions, etc. were largely incongruent to responses given by trained raters who reviewed video recordings of the participants (Ritchey, Mehta, & Peres, 2017). It is possible that participants are not able to accurately self-report their exposure to the stressors, or the observation-based ergonomic stressor assessment method is not appropriate for determining ergonomic risks associated with computing tasks, or both. Thus, for determining the ergonomic risk associated with computing device usage, self-reported stressors may not be a reliable source of information (in preparation: Ritchey, Mehta, & Peres, 2017). In this paper, we explored the validity of the strain items of the SEAT.

These items consist of self-reported discomfort responses items for 13 different body parts and serve as a proxy for strain.

The postures, forces, and repetitive movements (i.e., s*tressors*) associated with computer work interact and result in physical loads placed on musculoskeletal tissues and elicit physiological responses (i.e., *strain*), musculoskeletal pain or discomfort, and ultimately increase the risk of MSDs (Tittiranonda, Burastero, & Rempel, 1999; Wahlstrom, 2005). Chronic inflammatory responses triggered by muscular overexertion, repetitive contractions, musculotendon overstretching, and repetitive movements can result in pain and loss of function because of fibrosis and the breakdown of tissues (Barbe & Barr, 2005). While computer work involves much lower levels of force compared to industrial tasks, the overloading of type I fibers through sustained low-intensity activity can cause muscle tissue disorders and chronic pain (Wahlstrom, 2005; Visser & van Dieen, 2006). Repetitive, low-intensity activity can also cause tissue damage and pain through a proposed fatigue-induced inhibition of the intercellular release of calcium ions necessary for muscle contractions (Westerlad, Bruton, Allen, & Lännergren, 2000). Non-neutral postures and repetitive movements can increase intermuscular pressure causing a decrease in intermuscular blood flow, reducing the availability of oxygenated hemoglobin in muscle tissues (Visser & van Dieen, 2006) as well as placing pressure on nerves resulting in pain and an increased risk MSDs (Werner, Armstrong, Bir, & Aylard, 1997; Keir, Bach, & Rempel, 1999). Armstrong et al., (1993) proposed a model of MSDs that allows for the examination of multiple exposure factors, pathways, and complex interactions through cascading dose-response relationships.

Surface electromyography (sEMG) is a commonly used method for measuring muscle activity and has frequently been used to measure muscular effort within the context of modern office job exposures (Kim et al., 2014; Lin, Young, & Dennerlein, 2015; Rietveld et al., 2007; Szeto, Straker, & O'Sullivan 2005). Muscle activity measured using sEMG is a common technique used to assess muscular load (Wahlström, 2005) and has been shown to have strong and positive correlations with both subjective measures of exertion (e.g., Borg's CR-10) and objective measures of exertion like grip force (Grant, Habes, & Putz-Anderson, 1994), isometric contraction forces (Lawrence & DeLuca, 1983), and keyboard typing forces (Martin et al., 1996). Hermans & Spaepen (1995) showed that during visual display terminal work, sEMG activity and perceived discomfort for the upper right and left trapezius increased and decreased together. While similar relationships between muscle activations and self-reported discomfort have been found (Vasseljen & Westgaard, 1995; Wells et al., 1997), proposed causal relationships between sEMG activity and discomfort during computer work are not well established nor reliable (Knardahl, 2002; Harvey & Peper, 1997). The Rapid Upper Limb Assessment (RULA) survey was shown to be able to differentiate high risk postures though the use of self-reported discomfort scores while sEMG measurements were not significantly different between the tested postures (Fountain, 2003). While the relationship between objectively measured muscular activity and self-reported discomfort are not perfect nor completely understood, reductions of either have the same likely effect of reducing the risk of MSDs (Dul, Douwes, & Smitt, 1994; Miedema, Douwes, & Dul, 1997; Wahlstrom, 2005).

Another method for measure muscle activity is near infrared spectroscopy (NIRS) and it uses the differential absorption of near infrared light within tissues, like muscles, as a means of measuring the amount of oxygenated hemoglobin in tissue. NIRS is a valid method, with respect to sensitivity and specificity, to assess large muscles as well as the small forearm muscle used during computer work due to its good spatial resolution (Cole et al., 2012; Perry, Thedon, & Bringard, 2010). NIRS provides unique insight into MSD as low levels of oxygenated hemoglobin in tissues can result in localized muscle fatigue, tissue ischemia, and ultimately tissue death (Boushel et al., 2001; Perry, Thedon, & Bringard, 2010; Rolfe, 2000; Visser & van Dieen, 2006). NIRS has been used to investigate the effects of posture, repetition, and force on oxygenated hemoglobin levels across a wide range of contexts: an automobile assemble task (Ferguson et al., 2013), lumbar massage systems for prolonged driving (Durkin, Harvey, Hughson, & Callaghan, 2006), downhill walking (Ahmadi, Sinclair, & Davis, 2008), computer mousing tasks (Aasa et al., 2011; Crenshaw, Djupsjöbacka, & Svedmark, 2006; Nielsen et al., 2000). Jones & Cooper (2014) even investigated muscle oxygenation during swimming by modifying a NIRS system to be waterproof. Although NIRS is not always related discomfort, (e.g., a study of prolonged sitting or standing did not find a relationship between NIRS measures of oxygenated hemoglobin and low back discomfort, Callaghan, Gregory, & Durkin, 2010), discomfort has been connected to low levels of muscle oxygenation both theoretically (Visser & van Dieen, 2006) and empirically (Le et al., 2014).

Based on the previous literature presented here, the validity of the SEAT's strain component requires the finding of two results: 1) positive correlations between self-reported discomfort and muscle activity (sEMG), and 2) negative correlations between self-reported discomfort oxygenated hemoglobin (NIRS).

The aim of this study is to validate the SEAT's strain component across four different devices and two tasks by assessing the relationships between self-reported discomfort and objective measures of muscular activity (sEMG) and oxygenation (NIRS) to validate the SEAT's strain component.

Methods

Participants

Participants were selected for enrollment in the study from a pool of respondents to a university-wide call for participants. To be eligible for participation, respondents must have been 18-years of age or older and self-reported that they have never been diagnosed with a musculoskeletal disorder of the upper extremities, neck, or shoulder. Through a random selection process, 32 participants were recruited from the pool of all eligible respondents. Two participants' data were excluded because of data collection errors. The mean age of participants was 20.5 (range: 18 – 28) and demographic information for the final sample of 30 participants are presented in Table 11. This study was approved by the Texas A&M Institutional Review Board.

Table 11. Selected demographic information for the 30 participants included in analyses. Mean values are reported with standard deviation cited parenthetically. Gender is given as the percentage of the samples who identified as female.

	Mean (SD)
Age (years)	20.5 (02.7)
Hours/Week Spent on Desktop or Notebook Computer	38.0 (22.2)
Years of Daily Computer Use	8.0 (04.3)
Gender (% female)	73.3%
Weight (kg)	68.3 (21.6)
Height (cm)	165.9 (09.6)

Workstation Setup

Participants were seated at a workstation with an ergonomic, adjustable task chair and a height adjustable desk for all devices and tasks. Prior to beginning their first task, the chair and desk were adjusted such that the work station was HFES/ANSI 100-2007 compliant. After the workstation was adjusted, participants could adjust to their own comfort.

Devices

The devices selected for this study were representative of one of the four types of computer devices commonly used in, and outside of, the office environment: tablets, smart phones, notebooks, and desktops. All Apple® devices were used to minimize possible confounds due to inter-device and operating system differences. Placements for devices without applicable HFES/ANSI standards were informed by preferences data reported in observational and lab studies. Example placements are shown in Figure 12.

Desktop. The same MacBook Pro that was used for the Notebook condition was connected to an external monitor (24” Dell), a wired external keyboard (Apple Keyboard with Numeric Keypad), and a wireless mouse (Apple Magic 2). Use of a standard posture and position of a desktop computer was employed, specifically the HFES/ANSI 100-2007 standards (Kim et al., 2014).

Notebook. An Apple MacBook Pro was used for the notebook condition. Participants used both the native keyboard and touchpad to complete the tasks. Participants could reposition the notebook after researchers placed it, however, the notebook had to be flat and completely on the desk. Participants could adjust the screen angle to their liking. Participants were seated based on HFES/ANSI 100-2007 standards and the notebook was placed such that the home row keys are 12 cm away from the edge of the desk. Participants could adjust location of the notebook, their chair, and the desk, but the notebook had to stay on the table.

Tablet. An Apple iPad Air 2 was used for the tablet condition. The iPad Air 2 was placed in an Apple Smart Cover (for the iPad Air). Stawarz & Benedyk (2013) conducted a small human factors/ethnographic study among office workers and found that few people reported using accessories like external keyboards (15%), docking stations (27%), and styluses. However, 85% of participants reported using a tablet cover. This cover allows for the device to be propped up at an angle of 10 degrees, what Apple describes as “Keyboard stand” mode, to angle the screen towards the participant while still having the device at a comfortable angle for using the on-screen keyboard for typing. Participants were instructed that they could change the position of the tablet on the desk, however, the tablet must stay completely on the desk.

Smartphone. An Apple iPhone 5S was used for the smart phone condition. The iPhone 5S was placed in a Spigen Tough Armor iPhone 5S case. Participants were instructed to hold the smart phone in a portrait orientation, with both hands holding the phone and resting on the desk. This phone posture was chosen because several observational studies have shown people predominantly use their phone in the portrait orientation (Liang & Hwang, 2016; Shirazi et al., 2013), holding the phone with both hands and using their thumbs to interact with the touchscreen (Gold et al., 2012). The decision to have participants rest their hands on the desk, rather than fully support the phone without support, was to standardize postures across participants and to eliminate differences in strain that would occur from between-participant upper limb weight differences.

Tasks

Two, three-minute tasks were created that are 1) representative of typical computer tasks, 2) commonly completed using the devices in this study, and 3) are different in terms of how participants need to interact with the device to complete the task. The emailing task (Email) is predominantly completed by typing, whereas the calendaring task (Calendar) is predominantly completed by interacting with the graphical user interface using a mouse, touchpad, or touchscreen, depending on the device being used to complete the task.

Email task. The Email task was a typing dominant task designed such that participants spent the majority of the three-minute task using a keyboard for word production. Participants completed the task by replying to a series of emails using the Mac iOS native app Mail. Each email had a short phrase in the subject line and participants were instructed to reply to each email "as quickly and as accurately as possible".

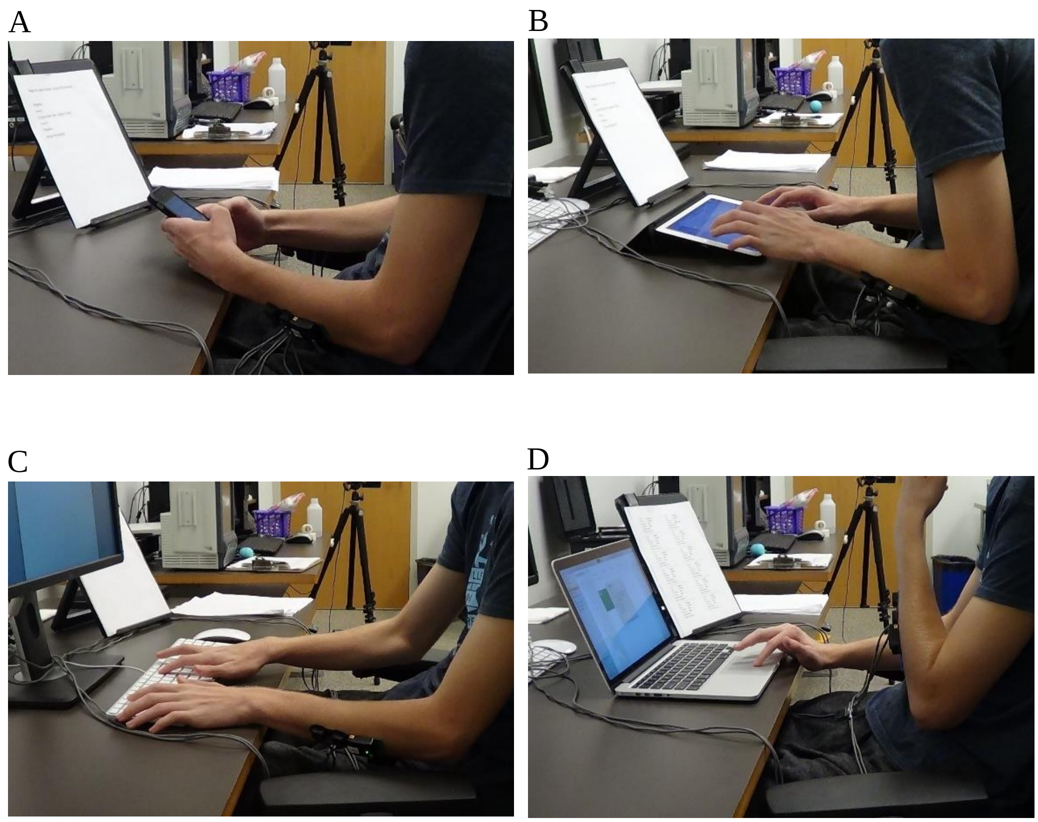


Figure 12. The four devices used for this study: a) smartphone, b) tablet, c) desktop, and d) notebook.

The phrases used to generate the email subjects were sampled from a 500-phrase set created by Mackenzie and Soukoreff (2003). Average phrase length was 28.61 letters with an average word length of 4.46 letters. The subject line of the email being replied to was visible on the display while participants typed their reply. Participants were told to reply to as many emails as they could in the three-minutes of the task. Presentation order of emails was the same for all participants and independent of task or device.

Calendar task. The Calendar task was an interaction dominant task designed so that participants spent the majority of the time interacting (e.g., clicking, touching) with the device's interface, i.e., participants spent the majority of the task not typing. Participants were given a list of calendar events to create using the Mac iOS native app Calendar. Calendar events to be created were printed on a sheet of paper. The printed sheet of paper was placed in a document holder, and participants were instructed to adjust the placement of the document holder to their liking.

Measures

Self-reported strain. The SEAT 2.0 was administered verbally immediately after participants completed each 3-minute task. Participants were shown a paper copy of the SEAT 2.0 while the experimenter read through and recorded participants' responses. This method of administering the SEAT 2.0 allows for participants to rest their muscles while the SEAT 2.0 is being administered. The body discomfort items of the SEAT 2.0 are the primary dependent variables for the analyses presented here as they represent the amount of strain reported by the participant. Participants completed a discomfort question for each of the 13 body parts included on a body discomfort diagram. The discomfort question asked participants to indicate "the amount of discomfort felt on a scale from 0 to 10, with 0 being no discomfort and 10 being very, very intense discomfort".

Muscle activity. Muscle activity for four upper extremity muscles was measured bilaterally using sEMG. Principle component analysis of the SEAT conducted by Peres, Mehta & Ritchey (2017) showed that strain items (i.e., discomfort of specific body parts) that ranked highest in terms of component loadings within the right and left side strain components were the lower arm and wrist for the left-side strain component and shoulder and neck for the right-side component. This analysis provided the rationale to select locations that recorded the forearm, shoulder, and neck muscle activity in the present study. Two muscles in the forearm were selected—the Extensor Carpi Ulnaris (ECU) and Flexor Carpi Ulnaris (FCU)—and then one muscle related to shoulder/arm movements, the Anterior Deltoid (AD), and one related to the neck, the Upper Trapezius (UT). These specific muscles were selected for their involvement in the use of computer devices.

The ECU was selected because it has the highest activity compared to other muscles across multiple studies that include a wide range of devices, postures, and tasks that are relevant to the office environment (Dennerlein & Johnson 2006a; Szeto & Lin, 2011; Werth & Babski-Reeves, 2014; Won et al. 2003; Young et al., 2013; Lin, Young & Dennerlein, 2015). Similarly, the FCU was selected to compliment the inclusion of the ECU (i.e., a flexor and extensor pair) and because it has been found to have higher levels of activity compared to the flexor carpi radialis during computing tasks (Lin, Young & Dennerlein, 2015; Young et al., 2013; Simoneau, Marklin, & Bergman, 2003; Dennerlein et al., 2002; Won et al., 2003). The AD was selected for its prominence in the anterior flexion and rotation of the shoulder that would be expected when using computer devices in different postures in the sagittal plane (e.g., Figure 12b vs 12c). The

UT is commonly measured in studies that look at muscle activity during computer work (Blangsted, Hansen, & Jensen, 2003; Chiu et al., 2015; Shin & Zhu, 2011; Szeto, Straker, O'Sullivan, 2005; Werth & Babski-Reeves, 2014; Young, et al., 2013; Xie et al., 2016).

Placement of the Trigno Wireless sEMG electrodes (Delsys, Inc., MA, USA) was determined through boney landmarks and palpation and confirmed by visual inspection of the sEMG signal when participants activated those muscles (Dennerlein & Johnson, 2006b; Perotto, 1994). The ECU probe was located just above the shaft of the ulna and one thirds distance of the forearm distal to the elbow. The FCU probe was located approximately two fingerbreadths volar to the ulna and in-line with the ECU probe (Werth & Babski-Reeves, 2013). The AD probe was placed on the ventral side of the shoulder approximately 50 mm from the acromion (Lin, Young & Dennerlein, 2015; Young et al., 2013). The UT probe was placed approximately 50 mm above the midpoint between the neck and acromion (Young et al., 2013).

The sEMG data were recorded at 2000 Hz and post-collection filtering was applied. The sEMG signals were band pass filtered at 10 Hz – 500 Hz (4th order Butterworth filter) and full wave rectified. The filtered sEMG data were used to compute root mean square (RMS) values with static windows of 30 ms with no overlap. For each muscle, an average RMS (aRMS) value was calculated for each maximum voluntary contraction (MVC) exercise, described later, as well as each task. Task aRMS values were normalized using the greatest aRMS of the three MVC exercises, creating percentage MVC (%MVC) values. Muscle activity in units of percent of a participant's maximum exertion allows for comparisons across devices and tasks and %MVC is a commonly used metric for assessing muscle activity and has been shown to correlate with subjective measures of exertion (e.g., Borg Ratings of Perceived Exertion; Grant, Habes, & Putz-Anderson, 1993) as well as objective measures of muscle force (Lawrence and DeLuca, 1983).

Muscle oxygenation. Muscle oxygenation was measured using NIRS (NIRO 200 NX, Hamamatsu Photonics, Japan). NIRS probes were used to measure oxygenation of the ECU muscle (bilaterally) during tasks. The ECU was selected because of its high level of activity during computer tasks. One study gives guidance and precedence for the co-location of a NIRS and EMG probe on a forearm extensor (ECR; Elcadi & Forsman, 2011). However, in that study, primary placement considerations were given to the NIRS probe and the sEMG probe was located laterally to the placement of the NIRS probe. Given the physical size of both probes used in this study and the average width of the ECU, the NIRS probe was placed immediately distal of the sEMG probe. The co-location of NIRS and EMG probes over the same muscle has been done in at least one other study as well (Crenshaw, Djupsjobacka, & Svedmark, 2006). Relative changes in oxygenated hemoglobin levels, HbO, were averaged over each task and used as a measure of muscle oxygenation.

Procedures

After probe sites were located, the skin above the belly of the muscles was prepared by shaving any hair if present and then cleaning the area with an alcohol swab. Then, sEMG and NIRS probes were placed on those sites. Participants then completed a series of three very short (5 sec) MVC trials for each muscle, with a 1-minute rest break in between each MVC trial. The maximum of the three MVC trials per muscle were used to normalize participants' muscle activity obtained during the tasks. For the ECU, while seated and in a neutral posture with elbow at 90 degrees and shoulders relaxed, participants extended their wrist while the experimenter resists the movement. Similarly, for the FCU, participants flexed at the wrist while the experimenter resists the movement. For the AD, while keeping the upper arm near the neutral posture (i.e., at rest and vertically aligned with the torso) the experimenter resisted shoulder flexion (Dennerlein & Johnson, 2006b). For the UT, participants were told to attempt to lift/shrug their shoulders as hard as they can while the researcher resisted the movement and applied force downward at the right and left acromion (Dennerlein & Johnson, 2006b).

Participants rested for 5 minutes after completing the last MVC trial before beginning tasks. The order in which participants used each device was counterbalanced with the presentation of tasks being the same for a given participant across all devices. Once participants completed a task they rested for 5 minutes while the researcher verbally administered the SEAT to which participants verbally respond. Once participants completed the emailing and calendaring task on all four devices they completed a demographic form and were debriefed.

Analyses

To assess the ability of participants to self-report musculoskeletal strain via the strain items of the SEAT (i.e., body discomfort items), we compared mean %MVC values, as well as mean HbO values, to participants' self-reported discomforts. Spearman Rho correlations were conducted based on an *a priori* decision to compare physiological measurements with body discomfort scores that are biomechanically the most closely related to the muscles measured in this study. The pairings are listed in Table 12. Correlations were conducted with the data collapsed across all devices and tasks, as well as for each individual device. Given that the total number of correlations computed was high (160), to control for an increase in Type 1 errors, the Benjamini-Hochberg procedure (B-H; Benjamini & Hochberg, 1995) was used with a false discovery rate set at 10%, which resulted in a significance criteria of $\square \leq .034$.

Table 12. Correlation pairs determined a priori. CU = Carpi Ulnaris

	Physiological Strain	
Self-Report Strain	sEMG	NIRS
Unilateral Body Parts		
Neck	Right Upper Trap Left Upper Trap	N/A
Upper Back	Right Upper Trap Left Upper Trap	N/A
Lower Back	Right Upper Trap Left Upper Trap	N/A
Left Side Body Parts		
Left Shoulder	Left Anterior Deltoid Left Upper Trapezius	N/A
Left Upper Arm	Left Anterior Deltoid Left Upper Trapezius	N/A
Left Lower Arm	Left Extensor CU Left Flexor CU	Left Extensor CU
Left Wrist	Left Extensor CU Left Flexor CU	Left Extensor CU
Left Hand	Left Extensor CU Left Flexor CU	Left Extensor CU
Right Side Body Parts		
Right Shoulder	Right Anterior Deltoid Right Upper Trapezius	N/A
Right Upper Arm	Right Anterior Deltoid Right Upper Trapezius	N/A
Right Lower Arm	Right Extensor CU Right Flexor CU	Right Extensor CU
Right Wrist	Right Extensor CU Right Flexor CU	Right Extensor CU
Right Hand	Right Extensor CU Right Flexor CU	Right Extensor CU

Results

Rho values for all correlations identified as significant via the B-H procedure (α = .034) for mean %MVC sEMG and mean HbO NIRS are shown in Table 14 and 15, respectively, and device specific correlations are identified parenthetically (D = Desktop, N = Notebook, P = Phone, T = Tablet). See Table 13 for qualitative descriptions of relationship strengths.

Table 13. Qualitative description of relationship based on absolute value of Spearman's Rho. Qualitative descriptors were adapted from Viera & Garret, (2005).

Descriptor	Rho
Very Weak	0.00 to 0.19
Weak	0.20 to 0.39
Moderate	0.40 to 0.59
Strong	0.60 to 0.79
Very Strong	0.80 to 1.00

Muscle activity (sEMG). All significant correlation coefficients for the Neck, Lower Back, and Upper Back were negative and weak, ranging from -0.216 to -0.320, indicating that as the Left and Right Upper Trapezius muscle activity increased, participants reported decreasing levels of discomfort in those body parts. No significant correlations were found between mean %MVC of the Right Upper Trapezius and either Lower Back or Upper Back discomfort.

For Left side proximal body parts (Left Shoulder and Left Upper Arm), all significant correlation coefficients between discomfort in these body parts and the Left

Upper Trapezius were negative and weak, ranging from -0.201 to -0.363, again, indicating that as Left Upper Trapezius mean %MVC increased, reported discomfort in those body parts decreased. However, for the Left Anterior Deltoid, as mean %MVC increased, reported discomfort increased. All significant correlations between Left Shoulder and Left Upper Arm discomfort and mean %MVC of the Left Anterior deltoid were positive and weak, ranging from 0.186 to 0.324.

For the Left side distal body parts, moderate, positive relationships were found—only for the smartphone—between Left Lower Arm, Left Wrist, and Left Hand Discomfort and mean %MVC values for the Left Extensor Capri Ulnaris (range: 0.417 to 0.489). Aside from a moderate, positive relationship between Left Hand discomfort and mean %MVC values for the Left Flexor Carpi Ulnaris when using the smartphone, all relationships between left side distal discomfort and the Left Flexor CU were weak.

Table 14. Significant Spearman Rho values for %MVC. Correlations were computed between discomfort and Mean %MVC scores for 8 muscles measured in the study. Significant correlations were determined using the B-H procedure.

Self-reported Strain	Physiological Strain (%MVC sEMG)	
	Left Upper Trapezius	Right Upper Trapezius
Neck	-0.216	-0.274 (N)
	-0.320 (T)	
Lower Back	-0.216	
	-0.297 (P)	
Upper Back	-0.254	
	-0.315 (N)	
	-0.311 (T)	
Left Proximal	Left Upper Trapezius	Left Anterior Deltoid
Left Shoulder	-0.201	0.186
	-0.335 (D)	0.305 (N)
Left Upper Arm	-0.232	0.218
	-0.363 (D)	0.324 (T)
		0.303 (N)
Left Distal	Left Extensor CU	Left Flexor CU
Left Lower Arm	0.167	0.172
	0.417 (P)	
Left Wrist	0.489 (P)	0.170
Left Hand	0.431 (P)	0.216
		0.376 (P)
Right Proximal	Right Upper Trapezius	Right Anterior Deltoid
Right Shoulder		
Right Upper Arm		
Right Distal	Right Extensor CU	Right Flexor CU
Right Lower Arm	0.382	0.316
	0.514 (N)	0.461 (P)
	0.427 (D)	0.421 (T)
	0.400 (T)	
	0.283 (P)	
Right Wrist	0.365	0.249
	0.502 (D)	0.358 (P)
	0.434 (T)	0.303 (T)
	0.411 (N)	
Right Hand	0.374	0.256
	0.434 (D)	0.309 (P)
	0.407 (T)	0.309 (P)
	0.399 (N)	
	0.362 (P)	

No significant relationships between mean %MVC and discomfort were found for the Right side proximal body parts, *i.e.*, neither the Right Upper Trapezius nor Right Anterior Deltoid were significantly correlated with Right Shoulder or Right Upper arm discomfort. For the Right Side distal body parts, moderate and positive relationships were found between both Right Extensor and Flexor Carpi Ulnaris mean %MVC and each of the three right side distal body parts, Right Lower Arm, Right Wrist, and Right Hand discomfort (range: 0.249 to 0.514).

Muscle oxygenation (NIRS). Significant negative correlations were found for mean HbO of the Right Extensor CU and all right side distal body parts (range: -0.158 to -0.318). Weak negative correlations were found for the tablet specific relationships between mean HbO of the Right Extensor CU and Right Lower Arm (rho = -.306), Right Wrist (rho = -0.318) and Right Hand (rho = -0.295). A notebook-specific, weak negative correlation was also found for the Right Wrist (rho = -0.317). No significant correlations were found between left distal body parts and the Left Extensors CU.

Table 15. Significant Spearman Rho values for HbO. Correlations were computed between discomfort the Left and Right Extensor Carpi Ulnaris mean HbO. Significant correlations were determined using the B-H procedure.

	Physiological Strain (Mean HbO NIRS)
Left Distal Body Parts	Left Extensor CU
Left Lower Arm	
Left Wrist	
Left Hand	
Right Distal Body Parts	Right Extensor CU
Right Lower Arm	-0.233
	-0.306 (T)
Right Wrist	-0.248
	-0.318 (T)
	-0.317 (N)
Right Hand	-0.158
	-0.295 (T)

Discussion

The aim of this study was to validate the SEAT's strain component by satisfying two requirements: demonstrating 1) positive correlations between self-reported discomfort and muscle activity (sEMG) and 2) negative correlations between self-reported discomfort oxygenated hemoglobin (NIRS). These findings confirm that the expected correlations with self-reported discomfort were found (with exceptions) for both sEMG and NIRS measures. Participants using the SEAT can self-report strain via the proxy of self-reported discomfort.

sEMG Correlations. The greatest number of significant correlations between sEMG and discomfort measures was found for the right distal body parts. Especially for the Right Extensor Carpi Ularnis, all device specific correlations were significant (except for the smartphone and the right wrist) for each of the following body parts: right lower arm, right wrist, and right hand. Furthermore, the magnitudes of these correlations were the highest. These device specific correlations support the cross-device validity of the SEAT, a necessary attribute given the diversity of devices used in the office environment. Differences in strength between device-specific correlations does, however indicate that muscular activity and self-reported discomfort are more closely related for some devices compared to others. The pattern of significant correlations for the Right Flexor CU can offer some insight here, with only the smartphone and tablet exhibiting significant device-specific correlations. Similarly, for the left distal body parts, the only device specific correlations are for the smartphone. It is likely that device size or input method play a role in the relationship between perceived discomfort and muscular activity. Compared to the notebook and desktop, the smartphone and tablet are both more similar in form factor and require touchscreen interaction.

Significant correlations between self-reported discomfort and both sEMG of the Left Extensor and Flexor Carpi Ulnaris (CU) were found for all the Left Distal body parts, which includes the Left Lower Arm, Left Wrist, and Left Hand. However, those relationships that were not device-specific were weak (Rhos < 0.22) and not present for all pairings. Smartphone only correlations were much stronger (range: 0.376 to 0.489) and especially prominent for the left extensor. The discrepancy between the completeness of right-sided, device-specific correlation pairs and the sparse left-sided pairs could be a function of hand dominance. All participants in this study self-reported as right hand dominant, which could create a difference in the levels of activation between the left and right forearm muscles. This activation bias, coupled with the possibility that dominant side discomfort scores influenced the self-reported discomfort of the non-dominant side could explain these findings.

NIRS correlations. For the NIRS measure of oxygenated hemoglobin (HbO), we were expecting to find negative correlations between discomfort and HbO levels. All significant correlations found did indicate a negative relationship between HbO levels and discomfort. As the blood within the muscle tissue became less and less oxygenated, participants reported higher and higher levels of discomfort. However, we only found significant correlations for the right Extensor CU, and the significant correlations were predominantly tablet-specific (three out of four device-specific correlations). At this moment it is unclear as to why the tablet-specific effects would be more pronounced, nor why the left side correlations were not similarly significant. One limitation for the NIRS data is that primary placement of the sensor was given to the sEMG probe, and the NIRS probe was placed more distal to the ideal location (i.e., belly of the muscle), and this compromised placement could have affected the NIRS measurements.

Limitations. Left Upper Trapezius muscle activity was significantly correlated with Neck, Upper Back, and Lower Back. However, all significant relationships with the Upper Trapezius (both right and left) were negative, indicating a decrease in reported discomfort as muscle activity increased. As muscle activity increases, strain increases, and thus it was expected that the amount of discomfort reported would increase as well. The observed significant negative relationship could be attributed to several factors. First, the trapezius is a strong postural musculature when compared to smaller and more distal muscles of the forearm. Because of the trapezius's larger size (i.e., resistant to fatigue) and role as a postural stabilizer, had the tasks been longer than three minutes, there might have been a significant relationship between mean %MVC and discomfort for the Left and Right Upper Trapezius. Also, we only measured the EMG activity of the upper section of the trapezius. Distribution of musculoskeletal load from the Upper Trapezius to other parts of the muscle or other nearby muscles could have resulted in a conflation of discomfort sources. Second, and not limited to the trapezius, weak, non-existent, or unexpected correlations could also be a function of poor 1:1 mapping between body discomfort regions and the placement of the sEMG probes. For example, the Upper Trapezius is not the only, and for some body parts, not even the primary muscle responsible for the movements of the neck, upper back, lower back, left shoulder, and upper arms. Further validation studies with either modified body discomfort diagrams or different sensor placements may be needed if participants' feelings of discomfort in these body parts did not align well enough with our intended mappings of probe placements to body part discomfort questions.

No significant relationships were found between right proximal body part discomforts and the muscle activity of either the right Upper Trapezius or Anterior Deltoid. The potential mismatching of which muscles were measured and which body part discomforts were self-reported could also explain the lack of significant correlations between muscle activity and discomfort for the Right Shoulder and Right Upper Arm. Other studies have reported findings of no significant relationships between discomfort and EMG (Harvey & Peper, 1997; Corlett, Manenica, and Goillau, 1983), and Knardahl

(2002) has challenged the fundamental associations between EMG and muscle pains like that of neck and shoulder pain. However, the RULA was still useful in differentiating between different high-risk postures even though sEMG measurements were not able to do the same (Fountain, 2003). It is possible that, for the SEAT, the conceptualization of self-reported discomfort as a strict correlate of sEMG muscle activity is problematic unless other known MSD factors are accounted for, such as psychosocial, work organizational, and individual factors like sex (Wahlstrom, 2005). The effect of sex/gender differences on neck and shoulder MSDs (Côté, 2012) could be particularly important to account for in this specific study's results as our sample was predominantly female (70%).

Conclusion

While the relationships between physiological measurements and self-reported discomfort were inconsistent between body parts and devices, these results overall support the idea that participants' self-reported discomfort is a useful metric for assessing strain as it occurs in an office environment, across four very typical electronic devices. However, further work needs to be done to fully understand the device-specific differences found here. Similarly, while we have presented explanations for the negative relationships between sEMG and self-reported discomfort for the right and left trapezius, additional analyses or follow-up studies are needed to confirm the causes of these unexpected results.

These results, contrasted with findings that participants are unable to reliably self-report stressors using the SEAT (in preparation: Ritchey, Mehta, Peres, 2017), have implications for further development of the SEAT. This suggests that it may be more appropriate to collect information on stressors through more objective techniques when assessing ergonomic risks associated with computing tasks. In many modern office environments stressor information is already being collected using applications that record the number of keystrokes, mouse clicks, and many other interactions between employees and their computers. The SEAT's strain component can be used to supplement these types of objective stressor data, and data from wearable sensors, to provide a rich measurement of ergonomic risks in the office or for specific office computer tasks. The findings reported here imply that self-reported discomforts, obtained via the SEAT, have moderate correlations with physiological strain associated with computing tasks and thus may serve as early indicators of MSD risk – even for short duration tasks or observation periods. The ability to quickly and reliably assess the ergonomic strain created from software interactions can enable software designers to not only limit the amount of risk in a product before it is deployed, but to also continually update existing software to remove ergonomic risks throughout the lifetime of the software.

CONCLUSION

There are numerous assessment tools and checklists and surveys available to ergonomists. However, most all focus intently on assessing the physical environment around the office worker. For example, by providing detailed checklists to determine if the chair the office worker has is sufficiently adjustable. The three papers presented here represent the groundwork being laid for a different approach to measuring, and eventually controlling, ergonomic risks in the office environment. Most work-related physical exertions and movements in the office environment are dependent on how office workers must interact with their computer devices, and more specifically the software on those devices. Ergonomic risks are first created by the ways in which software designs dictate how workers must interact with their computer devices.

Controlling the ergonomic risks created from the design of software could be a crucial, and currently missing component from ergonomic interventions. The SEAT was developed to be an ergonomic assessment tool sensitive to the risks created from software designs. The results from the three papers presented here show that the SEAT has potential to be an effective self-report assessment tool. The self-report feature of the SEAT is vital to maximize the potential impact of the tool. Not all companies have or can afford to employ trained ergonomists and a strength of the SEAT is that it is an assessment tool that, from its inception, has been developed as a self-report measure.

Results from the studies presented here have further illustrated how difficult accurate assessment of stressors like posture and hand activity can be in the office environment. However, the correlations found between discomfort and physiological measures show promise that self-reported discomfort could be an easy construct through which to measure ergonomic risks of software design. With the varied and transient postures and movements in an office environment, obtaining accurate assessments of stressors during computer work can be difficult. Alternatively, a better approach might be to use a simple survey like the SEAT to collect discomfort data which can be used to quickly identify ergonomic issues. To fully utilize self-reported discomfort, it is likely that other known MSD risk factors should be account for, like sex and psychosocial stress.

More work needs to be done before the SEAT is a functioning assessment tool. The SEAT needs to be applied to real tasks and software use that are not experimentally constructed and constrained in a lab setting. Once the SEAT has been sufficiently refined, it can be used to generate design heuristics that identify specific software design elements and quantify the amount of ergonomic risk they create. The SEAT has the potential to be an easy to use, quick, and informative tool for the assessment of computer work related ergonomic risks.

Public Health Impact

The goal of the SEAT is for it to be used as a tool during the development of software as a method to eliminate ergonomic risks that exist within the design of the software. The ability to eliminate software design related ergonomic risks could result in significant reductions in the incidence of MSDs for computers workers and for the general public given most people interact with computer software on a daily basis, regardless of their occupation. Just as developers update software to fix bugs and add features, the SEAT could be used to develop updates focused on decreasing the ergonomic risks of software. The SEAT could also be used to inform both personal and business software purchasing decisions by allowing for the comparison of software options based on level of ergonomic risk.

In 2003, 77 million U.S. workers reported that they used a computer at work (BLS, 2005). In 2002, the Office Ergonomics Advisory Committee reported that for the state of Washington, medical costs and costs due to time away from work amounted to a total cost of around $12 million per year. In 2012, 91% of children in the United States (ages 12 to 15) reported daily use of a computer (Herrick et al., 2014). The ubiquitous use of computers, and the software installed on them, within and outside of the workplace creates an opportunity for design changes based on information gained using the SEAT to have wide reaching public health benefits for millions of people.

REFERENCES

Aasa, U., Jensen, B. R., Sandfeld, J., Richter, H., Lyskov, E., & Crenshaw, A. G. (2011). The impact of object size and precision demands on fatigue during computer mouse use. Advances in Physiotherapy, 13(3), 118-127.

Ahmadi, S., Sinclair, P. J., & Davis, G. M. (2008). Muscle oxygenation after downhill walking-induced muscle damage. Clinical physiology and functional imaging, 28(1), 55-63.

Andersen, J. H., Thomsen, J. F., Overgaard, E., Lassen, C. F., Brandt, L. P. A., Vilstrup, I., Kryger, A. I., & Mikkelsen, S. (2003). Computer use and carpal tunnel syndrome: a 1- year follow-up study. Jama, 289(22), 2963-2969.

Anghel, M., Argesanu, V., Talpos-Niculescu, C., & Lungeanu, D. (2007). Musculoskeletal disorders (MSDs)-consequences of prolonged static postures. Journal of Experimental Medical & Surgical Research, 4, 167-172.

ANSI/HFES 100-2007 (2007). Human Factors Engineering of Computer Workstations. Human Factors and Ergonomics Society, Santa Monica, CA, USA.

Armstrong, T. J., Buckle, P., Fine, L. J., Hagberg, M., Jonsson, B., Kilbom, A., ... & Viikari-Juntura, E. R. (1993). A conceptual model for work-related neck and upper-limb musculoskeletal disorders. Scandinavian journal of work, environment & health, 73-84.

Bao, S., Howard, N., Spielholz, P., & Silverstein, B. (2006) Quantifying repetitive hand activity for epidemiological research on musculoskeletal disorders – Part II:

comparison of different methods of measuring force level and repetitiveness, Ergonomics, 49:4, 381-392.

Barbe, M. F., & Barr, A. E. (2006). Inflammation and the pathophysiology of work-related musculoskeletal disorders. Brain, behavior, and immunity, 20(5), 423-429.

Benden, M., & Pickens, A. (2009). Evaluation day: measuring office ergonomics training can show impact. Industrial Engineer, 41(9), 46-50.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), 289-300.

Bernard, B. P., & Putz-Anderson, V. (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: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health Cincinnati, OH, USA.

Bishea, D. M., Wood, S. B., & Muddimer, A. (2007). Ergonomic computing in geophysical interpretation. Paper presented at the Society for Exploration Geologists Technical Program Expanded Abstracts, San Antonio. 26(1), 2673-2677.

Blangsted, A. K., Hansen, K., & Jensen, C. (2003). Muscle activity during computer-based office work in relation to self-reported job demands and gender. European journal of applied physiology, 89(3-4), 352-358.

Blatter, B., & Bongers, P. (2002). Duration of computer use and mouse use in relation to musculoskeletal disorders of neck or upper limb. International Journal of Industrial Ergonomics, 30(4), 295-306.

(BLS) U.S. Department of Labor, Bureau of Labor Statistics. (2005). Computer and internet use at work in 2003. Retrieved from https://www.bls.gov/news.release/pdf/ciuaw.pdf

Bohr, P. C. (2001). Office ergonomics education: a comparison of traditional and participatory methods. Work (Reading, Mass.), 19(2), 185-191.

Borg, G. (1990). Psychophysical scaling with applications in physical work and the perception of exertion. Scandinavian Journal of Work, Environment & Health, 16(Suppl. 1), 55–58.

Boushel, R., Langberg, H., Olesen, J., Gonzalas-Alonzo, J., Bulow, J., & Kjaer, M. (2001). Monitoring tissue oxygen availability with near infrared spectrocopy (NIRS) in health and disease. Scandinavian Journal of Medicine & Science in Sports, 11, 213–222.

Brewer, S., Van Eerd, D., Amick III, B.C., Irvin, E., Daum, K.M., Gerr, F., Moore, J.S., Cullen, K. and Rempel, D., (2006). Workplace interventions to prevent musculoskeletal and visual symptoms and disorders among computer users: a systematic review. Journal of occupational rehabilitation, 16(3), 317-350.

Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

Bruno Garza, J. L., Eijckelhof, B. H. W., Johnson, P. W., Raina, S. M., Rynell, P. W., Huysmans, M. A., van Dieën, J. H., van der Beek, A. J., Blatter, B. M., & Dennerlein, J. T. (2012). Observed differences in upper extremity forces, muscle efforts, postures, velocities and accelerations across computer activities in a field study of office workers. Ergonomics, 55(6), 670-681.

Burt, S., & Punnett, L. (1999). Evaluation of interrater reliability for posture observations in a field study. Applied ergonomics, 30(2), 121-135.

Callaghan, J. P., Gregory, D. E., & Durkin, J. L. (2010). Do NIRS measures relate to subjective low back discomfort during sedentary tasks?. International Journal of Industrial Ergonomics, 40(2), 165-170

Cameron, J. A. (1996). Assessing work-related body-part discomfort: Current strategies and a behaviorally oriented assessment tool. International Journal of Industrial Ergonomics, 18, 389-398.

Chiu, H. P., Tu, C. N., Wu, S. K., & Chien-Hsiou, L. (2015). Muscle activity and comfort perception on neck, shoulder, and forearm while using a tablet computer at various tilt angles. International Journal of Human-Computer Interaction, 31(11), 769-776.

Cole, A. L., Herman, R. A., Heimlich, J. B., Ahsan, S., Freedman, B. A., & Shuler, M. S. (2012). Ability of near infrared spectroscopy to measure oxygenation in isolated upper extremity muscle compartments. The Journal of hand surgery, 37(2), 297-302.

Corlett, E. N., Manenica, I., & Goillau, P. J. (1983). The relationship between EMG activity of the sacrospinalis and reported back discomfort. European journal of applied physiology and occupational physiology, 50(2), 213-222.

Côté, J. N. (2012). A critical review on physical factors and functional characteristics that may explain a sex/gender difference in work-related neck/shoulder disorders. Ergonomics, 55(2), 173-182

Crenshaw, A. G., Djupsjöbacka, M., & Svedmark, Å. (2006). Oxygenation, EMG and position sense during computer mouse work. Impact of active versus passive pauses. European Journal of Applied Physiology, 97(1), 59-67.

Dainoff, M., Maynard, W., Robertson, M., & Andersen, J. H. (2012). Office Ergonomics. Handbook of Human Factors and Ergonomics, Fourth Edition, 1550-1573.

Dane, D., Feuerstein, M., Huang, G. D., Dimberg, L., Ali, D., & Lincoln, A. (2002). Measurement properties of a self-report index of ergonomic exposures for use in an office work environment. Journal of Occupational and Environmental Medicine, 44(1), 73-81.

Dennerlein, J. T., & Johnson, P. W. (2006a). Changes in upper extremity biomechanics across different mouse positions in a computer workstation. Ergonomics, 49(14), 1456-1469.

Dennerlein, J.T., Johnson, P.W. (2006b). Different computer tasks affect the exposure of the upper extremity to biomechanical risk factors. Ergonomics 49 (1), 45-61.

Dennerlein, J., DiMarino, M. H., Becker, T., & Johnson, P. (2002). Wrist and shoulder muscle activity changes across computer tasks. In Proceedings of the Human

Factors and Ergonomics Society Annual Meeting. SAGE Publications, 46(13), 1129-1132.

Duffield, T. J., Peres, S. C., Amonette, W., & Ritchey, P. (2013). Standard Deviation of sEMG: Measuring the dynamicity of muscle activity. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications 57(1), 1400-1404.

Dul, J., Douwes, M., & Smitt, P. (1994). Ergonomic guidelines for the prevention of discomfort of static postures based on endurance data. Ergonomics, 37(5), 807-815.

Durkin, J. L., Harvey, A., Hughson, R. L., & Callaghan, J. P. (2006). The effects of lumbar massage on muscle fatigue, muscle oxygenation, low back discomfort, and driver performance during prolonged driving. Ergonomics, 49(1), 28-44.

Ebersole, M. L., & Armstrong, T. J. (2002). Inter-rater reliability for hand activity level (HAL) and force metrics. In the Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications 46(13), 1037-1040.

Elcadi, G. H., Forsman, M., & Crenshaw, A. G. (2011). The relationship between oxygenation and myoelectric activity in the forearm and shoulder muscles of males and females. European journal of applied physiology, 111(4), 647-658.

Eliasson, K., Palm, P., Nyman, T., & Forsman, M. (2017). Inter-and intra-observer reliability of risk assessment of repetitive work without an explicit method. Applied Ergonomics, 62, 1-8.

Ferguson, S. A., Allread, W. G., Le, P., Rose, J., & Marras, W. S. (2013). Shoulder muscle fatigue during repetitive tasks as measured by electromyography and near-infrared spectroscopy. Human factors, 55(6), 1077-1087.

Fountain, L. J. (2003). Examining RULA’s postural scoring system with selected physiological and psychophysiological measures. International Journal of Occupational Safety and Ergonomics, 9(4), 383-392.

Gerard, M. J., Armstrong, T. J., Martin, B. J., & Rempel, D. A. (2002). The effects of work pace on within-participant and between-participant keying force, electromyography, and fatigue. Human factors, 44(1), 51-61.

Gerr, F., Marcus, M., Ensor, C., Kleinbaum, D., Cohen, S., Edwards, A., Gentry, E., Ortiz, D.J., Monteilh, C., 2002. A prospective study of computer users: I. Study design and incidence of musculoskeletal symptoms and disorders. Am. J. Ind. Med. 41 (4), 221-235.

Gerr, F., Monteilh, C. P., & Marcus, M. (2006). Keyboard use and musculoskeletal outcomes among computer users. Journal of Occupational Rehabilitation, 16(3), 259.

Gold, J. E., Driban, J. B., Thomas, N., Chakravarty, T., Channell, V., & Komaroff, E. G. (2012). Postures, typing strategies, and gender differences in mobile device usage: An observational study. Applied ergonomics, 43(2), 408-412.

Grant, K. A., Habes, D. J., & Putz-Anderson, V. (1994). Psychophysical and EMG correlates of force exertion in manual work. International Journal of Industrial Ergonomics, 13(1), 31-39.

Harvey, R., & Peper, E. (1997). Surface electromyography and mouse use position.

Ergonomics, 40(8), 781-789.

Heinrich, J., Blatter, B., & Bongers, P. (2004). A comparison of methods for the assessment of postural load and duration of computer use. Occupational and environmental medicine, 61(12), 1027-1031.

Hermans, V., & Spaepen, A. (1995). Perceived discomfort and electromyographic activity of the upper trapezius while working at a VDT station. International Journal of Occupational Safety and Ergonomics, 1(3), 208-214.

Herrick, K. A., Fakhouri, T. H., Carlson, S. A., & Fulton, J. E. (2014). TV Watching and Computer Use in US Youth Aged 15-16, 2012. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Hughes, L. E., Babski-Reeves, K., & Smith-Jackson, T. (2007). Effects of psychosocial and individual factors on physiological risk factors for upper extremity musculoskeletal disorders while typing. Ergonomics, 50(2), 261-274.

IJmker, S., Huysmans, M., Blatter, B. M., van der Beek, A. J., van Mechelen, W., & Bongers, P. M. (2006). Should office workers spend fewer hours at their computer? A systematic review of the literature. Occupational and environmental medicine.

Jones, B., Dat, M., & Cooper, C. E. (2014). Underwater near-infrared spectroscopy measurements of muscle oxygenation: laboratory validation and preliminary observations in swimmers and triathletes. Journal of biomedical optics, 19(12), 127002-127002.

Keir, P. J., Bach, J. M., & Rempel, D. (1999). Effects of computer mouse design and task on carpal tunnel pressure. Ergonomics, 42(10), 1350-1360.

Kennedy, C.A., Amick III, B.C., Dennerlein, J.T., Brewer, S., Catli, S., Williams, R., Serra, C., Gerr, F., Irvin, E., Mahood, Q., Franzblau, A., Van Eerd, D., Evanoff, B. & Rempel, D. (2010). Systematic review of the role of occupational health and safety interventions in the prevention of upper extremity musculoskeletal symptoms, signs, disorders, injuries, claims and lost time. Journal of Occupational Rehabilitation 20, 127-162.

Kim, J. H., Aulck, L., Bartha, M. C., Harper, C. A., & Johnson, P. W. (2014). Differences in typing forces, muscle activity, comfort, and typing performance among virtual, notebook, and desktop keyboards. Applied ergonomics, 45(6), 1406-1413.

Knardahl, S. (2002). Psychophysiological mechanisms of pain in computer work: the blood vessel-nociceptor interaction hypothesis. Work & Stress, 16(2), 179-189.

Knave, B. G., Wibom, R. I., Voss, M., Hedström, L. D., & Bergqvist, U. O. (1985). Work with video display terminals among office employees: I. Subjective symptoms and discomfort. Scandinavian journal of work, environment & health, 457-466.

Land Geophysical Safety Manual. (2012). (10 Ed.). Houston, TX USA: International Association of Geophysical Contractors.

Latko, W. A., Armstrong, T. J., Foulke, J. A., Herrin, G. D., Rabourn, R. A., & Ulin, S. S. (1997).Development and Evaluation of an Observational Method for

Assessing Repetition in Hand Tasks. American Industrial Hygiene Association Journal, 58(4), 278-285.

Lawrence, J.H. & DeLuca, C.J., 1983. Myoelectric signal versus force relationship in different human muscles. Journal of Applied Physiology, 54(6): 1653-1659.

Le, P., Rose, J., Knapik, G., & Marras, W. S. (2014). Objective classification of vehicle seat discomfort. Ergonomics, 57(4), 536-544.

Li, G., & Buckle, P. (2000). Evaluating Change in Exposure to Risk for Musculoskeletal Disorders—A Practical Tool. In the Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, 44(30), 407-408.

Liang, H. W., & Hwang, Y. H. (2016). Mobile phone use behaviors and postures on public transportation systems. PloS one, 11(2), e0148419.

Liebregts, J., Sonne, M., & Potvin, J. R. (2016). Photograph-based ergonomic evaluations using the Rapid Office Strain Assessment (ROSA). Applied Ergonomics, 52, 317-324.

Lin, M. Y., Young, J. G., & Dennerlein, J. T. (2015). Evaluating the effect of four different pointing device designs on upper extremity posture and muscle activity during mousing tasks. Applied ergonomics, 47, 259-264.

MacKenzie, I. S., & Soukoreff, R. W. (2003). Phrase sets for evaluating text entry techniques. In CHI'03 extended abstracts on Human factors in computing systems, ACM, 754-755.

Martin, B. J., Armstrong, T. J., Foulke, J. A., Natarajan, S., Klinenberg, E., Serina, E., & Rempel, D. (1996). Keyboard reaction force and finger flexor electromyograms during computer keyboard work. Human Factors: The Journal of the Human Factors and Ergonomics Society, 38(4), 654-664.

Martin, R. C. (2003). Agile software development: principles, patterns, and practices: Prentice Hall PTR.

Megasari, A. (2009). Study of the correlation between subjective and objective measurement among different computer tasks. University of Houston-Clear Lake.

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.

Miedema, M.C., Douwes, M. & Dul, J. (1997). Recommended maximum holding times for prevention of discomfort of static standing postures. Int. J. Ind. Ergon. 19, 9-18.

Moore, S. J., & Garg, A. (1995). The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. American Industrial Hygiene Association, 56(5), 443- 458.

Murray, W. E., Moss, C. E., Parr, W. H., & Cox, C. (1981). Potential health hazards of video display terminals. National Institute for Occupational Health. Cincinnati, OH. 81-129.

Newburger, E. C. (2001). Home computers and Internet use in the United States, August 2000: US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Nielsen, P. K., Laursen, B., Murthy, G., & Jensen, B. R. (2000). Tissue Oxygenation and

EMG during Computer Mouse Work and Submaximal Isometric Wrist Extensions. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, 44(6), 644-647.

(NPD) The NPD Group (2013). One-in- four Smartphone Owners Don’t Use a Case to Protect Their Phone According to NPD, December 2013.

(OEAC) Office Ergonomics Advisory Committee (2002). Office Ergonomics, practical solutions for a safer workplace. Seattle, WA: State of Washington Department of Labor and Industries.

Park, J. K., Boyer, J., Tessler, J., Casey, J., Schemm, L., Gore, R., ... & Healthy, P. (2009). Inter-rater reliability of PATH observations for assessment of ergonomic risk factors in hospital work. Ergonomics, 52(7), 820-829.

Paulsen, R., Schwatka, N., Gober, J., Gilkey, D., Anton, D., Gerr, F., & Rosecrance, J. (2014). Inter-rater reliability of cyclic and non-cyclic task assessment using the hand activity level in appliance manufacturing. International journal of industrial ergonomics, 44(1), 32-38.

Pereira, M. J., Straker, L. M., Comans, T. A., & Johnston, V. (2016). Inter-rater reliability of an observation-based ergonomics assessment checklist for office workers. Ergonomics, 59(12), 1606-1612.

Peres, S. C., Kortum, P. T., Akladios, M., & Muddimer, A. (2016). Developing and validating a self-report assessment tool for software biomechanics. Work, 53(1), 193-204.

Peres, S. C., Mehta, R. K., & Ritchey, P. (2017). Assessing ergonomic risks of software:

Development of the SEAT. Applied Ergonomics, 59, 377-386.

Perotto, A., 1994, Anatomical Guide for the Electromyographer: The Limbs and Trunk. Springfield, IL: Charles C. Thomas.

Perrey, S., Thedon, T., & Bringard, A. (2010). Application of near-infrared spectroscopy in preventing work-related musculoskeletal disorders: Brief review. International Journal of Industrial Ergonomics, 40(2), 180-184

Punnett, L., & Bergqvist, U. (1997). Visual display unit work and upper extremity musculoskeletal disorders. Stockholm: National Institute for Working Life, 997.

Rietveld, S., Van Beest, I., & Kamphuis, J. H. (2007). Stress-induced muscle effort as a cause of repetitive strain injury?. Ergonomics, 50(12), 2049-2058.

Ritchey, P., Mehta, R., & Peres, S.C., (2017) Validation of the Stressor Components of the Self-report Ergonomic Assessment Tool. Manuscript in preparation.

Ritchey, P., Peres, S.C., Duffield, T.J., (2012). Differences in muscle activity for 4 touch devices. In the Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, 56 (1), 1594-1598.

Robertson, M., Amick, B. C., DeRango, K., Rooney, T., Bazzani, L., Harrist, R., & Moore, A. (2009). The effects of an office ergonomics training and chair intervention on worker knowledge, behavior and musculoskeletal risk. Applied Ergonomics, 40(1), 124-135.

Rolfe, P. (2000). In vivo near-infrared spectroscopy. Annual Review of Biomedical Engineering, 2, 715–754.

Roodbandi, A. J., Choobineh, A., & Feyzi, V. (2015). The Investigation of Intra-rater and Inter-rater Agreement in Assessment of Repetitive Task (ART) as an Ergonomic Method. Occup Med Health Aff, 3(217), 2.

Shin, G., & Zhu, X. (2011). User discomfort, work posture and muscle activity while using a touchscreen in a desktop PC setting. Ergonomics, 54(8), 733-744.

Shirazi, S. A., Henze, N., Dingler, T., Kunze, K., & Schmidt, A. (2013). Upright or sideways?: analysis of smartphone postures in the wild. In Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services. ACM, 362-371.

Simoneau, G. G., Marklin, R. W., & Berman, J. E. (2003). Effect of computer keyboard slope on wrist position and forearm electromyography of typists without musculoskeletal disorders. Physical Therapy, 83(9), 816.

Sonne, M., Villalta, D. L., & Andrews, D. M. (2012). Development and evaluation of an office ergonomic risk checklist: ROSA–Rapid office strain assessment. Applied Ergonomics, 43(1), 98-108.

Speklé, E., Hoozemans, M., Van der Beek, A., Blatter, B., Bongers, P., & Van Dieën, J. (2009). Internal consistency, test–retest reliability and concurrent validity of a questionnaire on work-related exposure related to arm, shoulder and neck symptoms in computer workers. Ergonomics, 52(9), 1087-1103.

Spielholz, P., Bao, S., Howard, N., Silverstein, B., Fan, J., Smith, C., & Salazar, C. (2008). Reliability and validity assessment of the hand activity level threshold limit value and strain index using expert ratings of mono-task jobs. Journal of occupational and environmental hygiene, 5(4), 250-257.

Spielholz, P., Silverstein, B., Morgan, M., Checkoway, H., & Kaufman, J. (2001). Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors. Ergonomics, 44(6), 588-613.

Stawarz, K., & Benedyk, R. (2013). Bent necks and twisted wrists: Exploring the impact of touch-screen tablets on the posture of office workers. In Proceedings of the 27th international BCS human computer interaction conference. British Computer Society, 41.

Stern, A., (2015). What Drives Ergonomics Problems and Success at Pixar?. COEH Summer Institute.

Stevens, E. M., Vos, G. A., Stephens, J. P., & Moore, J. S. (2004). Inter-rater reliability of the strain index. Journal of occupational and environmental hygiene, 1(11), 745-751.

Stock, S. R., Fernandes, R., Delisle, A., & Vézina, N. (2005). Reproducibility and validity of workers' self-reports of physical work demands. Scandinavian journal of work, environment & health, 409-437.

Szeto, G. P., & Lin, J. K. (2011). A study of forearm muscle activity and wrist kinematics in symptomatic office workers performing mouse-clicking tasks with different precision and speed demands. Journal of Electromyography and Kinesiology, 21(1), 59-66.

Szeto, G. P., Straker, L. M., & O'Sullivan, P. B. (2005). The effects of speed and force of keyboard operation on neck–shoulder muscle activities in symptomatic and asymptomatic office workers. International Journal of Industrial Ergonomics, 35(5), 429-444.

Taylor, K. (2007). Oil industry dangers lurk in the office environment. Oil & Gas Network Magazine.

Tittiranonda, P., Burastero, S., & Rempel, D. (1999). Risk factors for musculoskeletal disorders among computer users. Occupational Medicine-Philadelphia, 14, 17-38.

Toy, K. J., Peres, S. C., David, T. Y., Nery, A., & Phillips Jr, R. G. (2012). Examining user preferences in interacting with touchscreen devices. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, 56(1), 1862-1866).

Vasseljen, O., & Westgaard, R. H. (1996). Can stress-related shoulder and neck pain develop independently of muscle activity?. Pain, 64(2), 221-230.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. Fam Med, 37(5), 360-363.

Village, J., Rempel, D., Teschke, K. (2005). Musculoskeletal disorders of the upper extremity associated with computer work: a systematic review. Occup. Ergon. 5 (4), 205-218.

Virzi, R. A., Sorce, J. F., & Herbert, L. B. (1993). A comparison of three usability evaluation methods: heuristic, think-aloud, and performance testing. Paper presented at the Human Factors and Ergonomics Society 37th Annual Meeting. 309-313.

Visser, B., & van Dieën, J. H. (2006). Pathophysiology of upper extremity muscle disorders. Journal of Electromyography and Kinesiology, 16(1), 1-16.

Wahlström, J. (2005). Ergonomics, musculoskeletal disorders and computer work. Occupational Medicine, 55(3), 168-176.

Wells, R., Norman, R., Neumann, P., Andrews, D., Frank, J., Shannon, H., & Kerr, M. (1997). Assessment of physical work load in epidemiologic studies: common measurement metrics for exposure assessment. Ergonomics, 40(1), 51-61.

Werner, R., Armstrong, T. J., Bir, C., & Aylard, M. K. (1997). Intracarpal canal pressures: the role of finger, hand, wrist and forearm position. Clinical Biomechanics, 12(1), 44-51.

Werth, A. J., & Babski-Reeves, K. (2013). Electromyography & Portable Computing Devices What Forearm Muscles Should be Measured?. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, 57(1) 892-896.

Werth, A. J., & Babski-Reeves, K. (2014). Effects of portable computing devices on posture, muscle activation levels and efficiency. Applied ergonomics, 45(6), 1603-1609.

Westerblad, H., Bruton, J. D., Allen, D. G., & Lännergren, J. (2000). Functional significance of Ca 2+ in long-lasting fatigue of skeletal muscle. European journal of applied physiology, 83(2), 166-174.

Wiktorin, C., Karlqvist, L., & Winkel, J. (1993). Validity of self-reported exposures to work postures and manual materials handling. Scandinavian journal of work, environment & health, 208-214.

Won, E., Johnson, P., Punnett, L., Becker, T., & Dennerlein, J. (2003). Gender differences in exposure to physical risk factors during standardized computer tasks. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications 47(10) 1155-1158.

Xie, Y., Szeto, G. P., Dai, J., & Madeleine, P. (2016). A comparison of muscle activity in using touchscreen smartphone among young people with and without chronic neck–shoulder pain. Ergonomics, 59(1), 61-72.

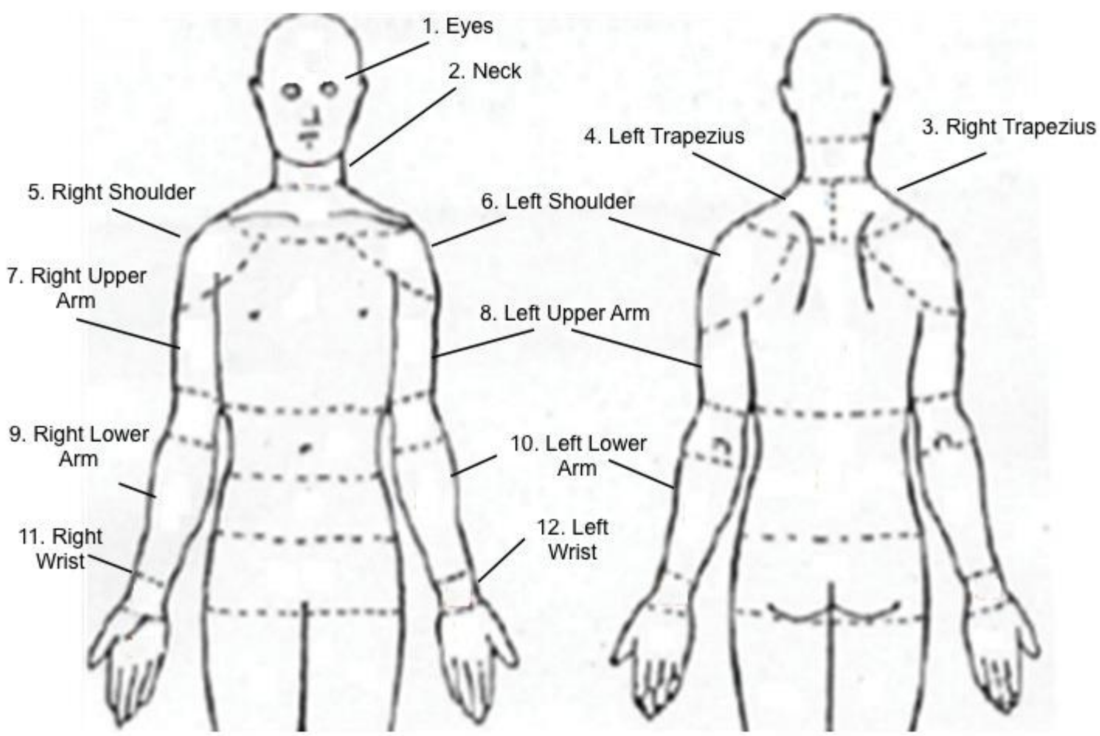
Young, J. G., Trudeau, M. B., Odell, D., Marinelli, K., & Dennerlein, J. T. (2013). Wrist and shoulder posture and muscle activity during touch-screen tablet use: Effects of usage configuration, tablet type, and interacting hand. Work, 45(1), 59-71.

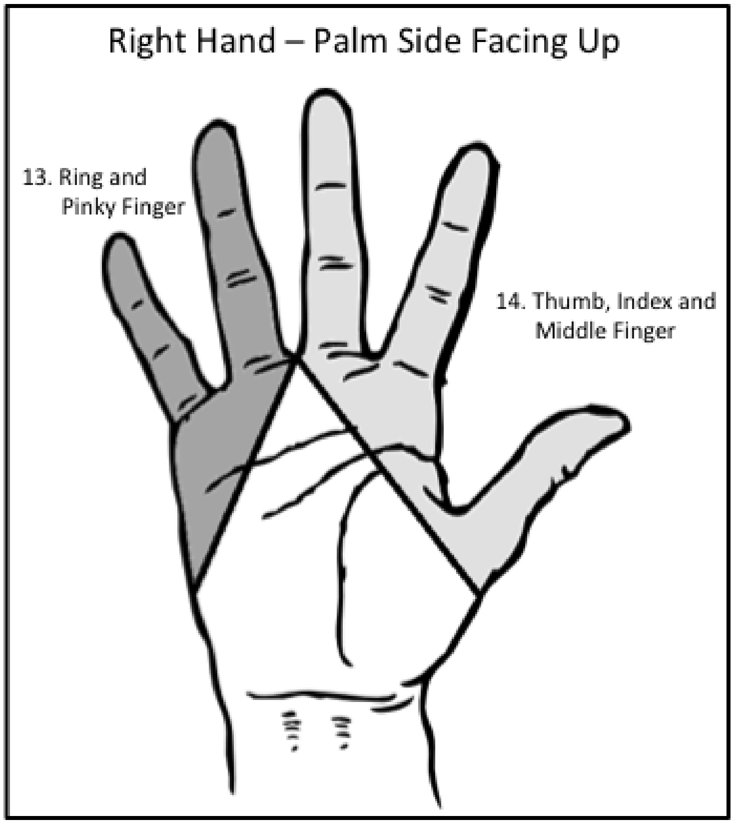
APPENDIX A

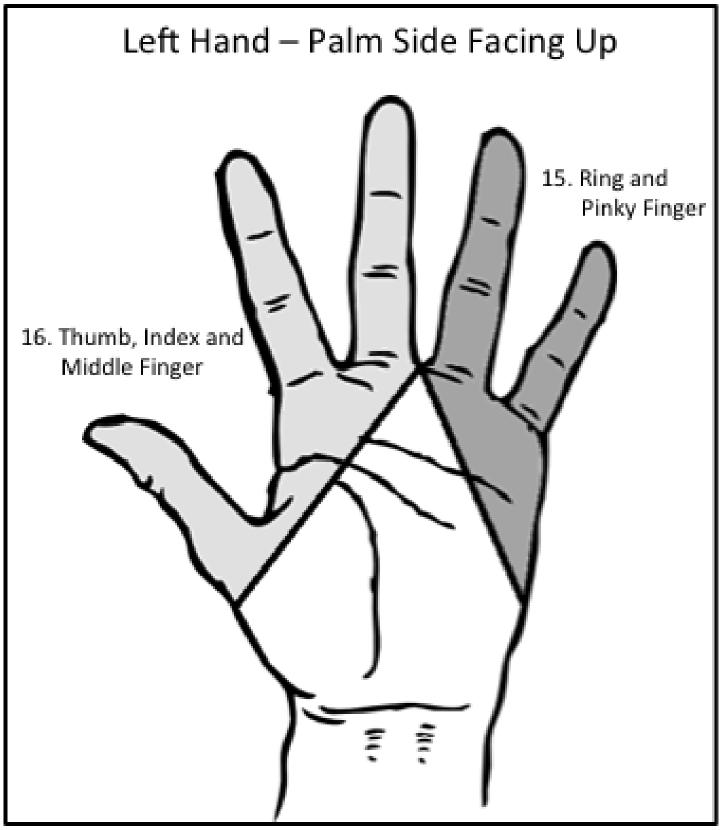
Used in Study 1, the first version of the Self-report Ergonomic Assessment Tool (SEAT 1.0).

For each body part during the task you just completed, indicate the amount of discomfort perceived on a scale from 0 to 10, with 0 indicating that there was no perceived discomfort and 10 indicating very, very intense discomfort. For the eyes and neck, indicate the amount of strain you experienced using the same 0 to 10 scale. Also, for each body part indicate the percentage of the time you feel that body part was used during the task you just completed.

0	1	2	3	4	5	6	7	8	9	10
Nothing at all (no discomfort)					**Moderate**					**Very Intense discomfort**




Right Hand – Palm Side Facing Up
13. Ring and Pinky Finger
14. Thumb, Index and Middle Finger


Left Hand – Palm Side Facing Up
15. Ring and Pinky Finger
16. Thumb, Index and Middle Finger

1. Please indicate the rating of physical effort that you feel best describes your amount of physical effort, in other words how hard your muscles were working, during the task you just completed using any number from 0 to 10, 0 being no effort at all and 10 being extremely strong, almost maximum effort. You can use decimals, such as 1.5 or 2.5:

Rating	Verbal Anchor
0	Nothing at all (At Rest)
1	
2	
3	
4	
5	Strong
6	
7	
8	
9	
10	Extremely Strong (Almost max, lifting 350 lbs. of weights)

2. For the task you just completed, on a scale from 0 to 10 please select the level of activity that best describes your hands. You will give a score for each hand. Consider a pause to be when your hand was idle or at rest.

Rating	Verbal Anchor
0	Hand is idle most of the time, no regular physical efforts
1	
2	Consistent and noticeable long pauses or very slow movements
3	
4	Slow and steady movements with frequent but brief pauses
5	
6	Steady movements and physical effort with few pauses
7	
8	Fast and repetitive movements with no pauses
9	
10	Fast and repetitive movements that cannot be maintained.

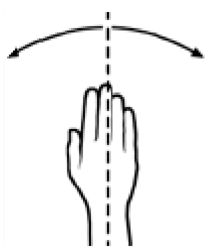
3. Select a level of effort that best describes your physical effort, how hard you felt your muscles worked, during the task you just completed:

1	Light	(Barely noticeable or relaxed effort)
2	Somewhat Hard	(Noticeable or definite effort)
3	Hard	(Obvious effort)
4	Very Hard	(Substantial effort)
5	Near Maximal	(Uses shoulder or trunk for force)

4. Select the image that best reflects the position of your hand and wrist during the task:

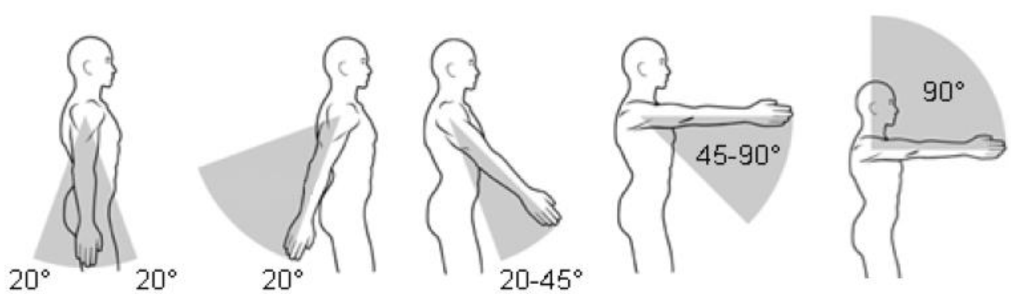
1. Wrist Straight **2. Wrist Slightly Bent** **3. Wrist Bent**

5. Were your hands deviated from the midline ever?



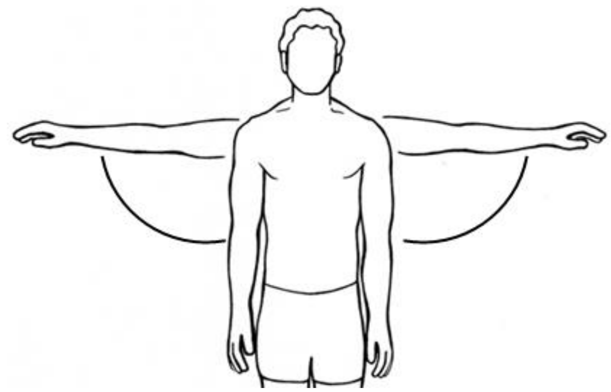
Yes or No

6. Select the image the best reflects the position of your arms during the task:



1 **2** **3** **4** **5**

7. Were your shoulders ever raised out to the side of your body?



Yes or No

8. Please select the option that best describes your speed of work (how fast you were typing, clicking, or using the touch screen) during the task you just completed.

1	Very Slow	(Extremely relaxed pace)
2	Slow	(Taking one's own time)
3	Fair	(Normal speed of motion)
4	Fast	(Rushed, but able to keep up)
5	Very Fast	(Rushed and barely/unable to keep up)

9. On a scale from 0 to 10, 0 being no precision required to 10 being a great amount of precision required, how much precision would you say the task required?

Rating	Verbal Anchor
0	No precision required
1	
2	
3	
4	
5	Moderate amount of precision required
6	
7	
8	
9	
10	Great amount of precision required

APPENDIX B

The second version of the Self-report Ergonomic Assessment Tool (SEAT 2.0).

For each body part during the task you just completed, indicate the amount of discomfort you felt a scale from 0 to 10, with 0 being no discomfort and 10 being very, very intense discomfort.

0	1	2	3	4	5	6	7	8	9	10
Nothing at all (no discomfort)					**Moderate**					**Very Intense discomfort**

For each body part, indicate how hard you felt the muscles in that body part were working during the task on a scale from 1 not at all, if ever, to 5 very hard.

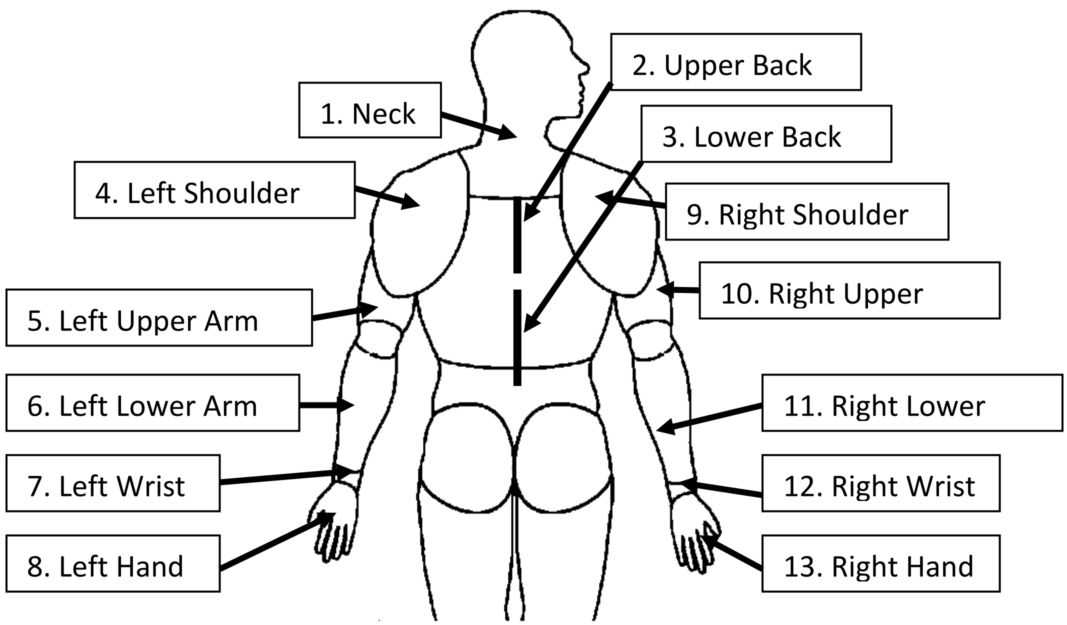
1	2	3	4	5
Not at all		**A moderate amount**		**Very Hard**

For each body part, indicate how often the muscles in that body part were being used, regardless of the level of that use.

1	2	3	4	5
Not used at all		**Used about half the time**		**Used all, or all most all of the time**

For each body part, indicate how repetitive you felt the use of the body part was.

1	2	3	4	5
Not repetitive at all		**Moderately Repetitive**		**Highly repetitive**


1. Neck
2. Upper Back
3. Lower Back
4. Left Shoulder
5. Left Upper Arm
6. Left Lower Arm
7. Left Wrist
8. Left Hand
9. Right Shoulder
10. Right Upper
11. Right Lower
12. Right Wrist
13. Right Hand

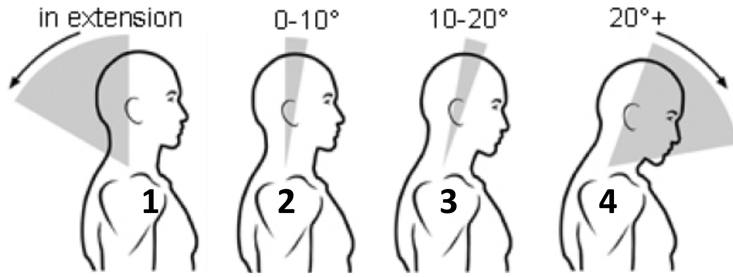
14. How hard were your muscles working (in other words, how much effort did you have to put forth) during the task you just completed?
Select a number from 0 to 10, with 0 being no effort at all and 10 being extremely strong, almost maximum effort.

Rating	Description
0	Nothing at all (At Rest)
1	
2	
3	
4	
5	Strong
6	
7	
8	
9	
10	Extremely Strong (Almost max, lifting 350 lbs. of weights)

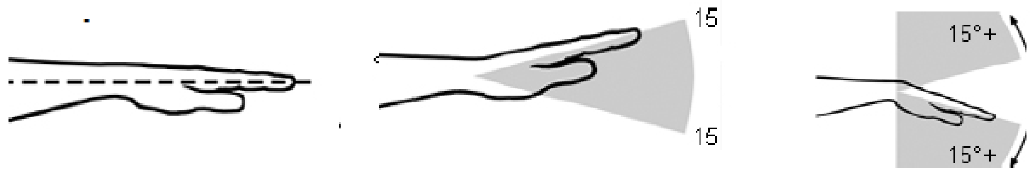
15. For the task you just completed, on a scale from 0 to 10 please select the level of activity that best describes your hands. You will give a score for each hand. Consider a pause to be when your hand was idle or at rest.

Rating	Verbal Anchor
0	Hand is idle most of the time, no regular physical efforts
1	
2	Consistent and noticeable long pauses or very slow movements
3	
4	Slow and steady movements with frequent but brief pauses
5	
6	Steady movements and physical effort with few pauses
7	
8	Fast and repetitive movements with no pauses
9	
10	Fast and repetitive movements that cannot be maintained.

16. Select the image the best reflects the position of your neck during the task:

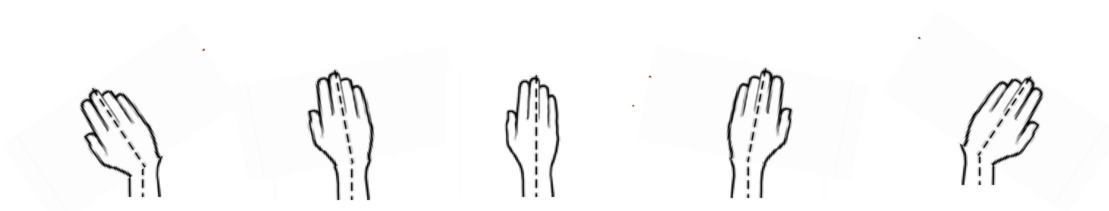


17. Select the image that best reflects the most common bend of your wrist during the task: Pick a number for your right and then also left wrist.



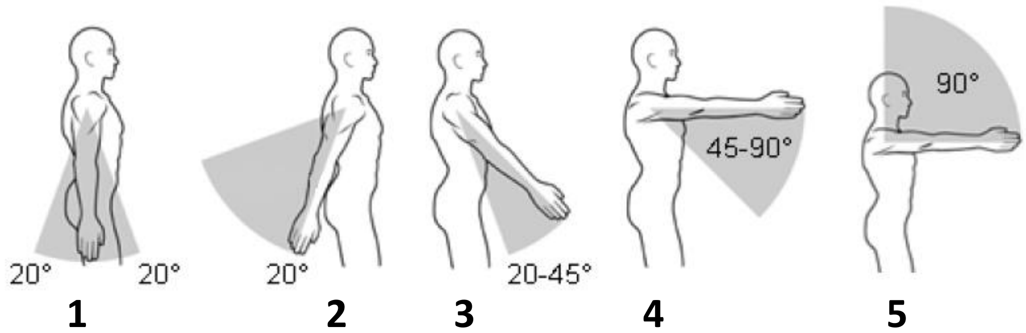
1. Wrist Straight **2. Wrist Slightly Bent** **3. Wrist Bent**

18. Select the image that best reflects the most position of your hand during the task. Pick a number for your right and then also left hand.

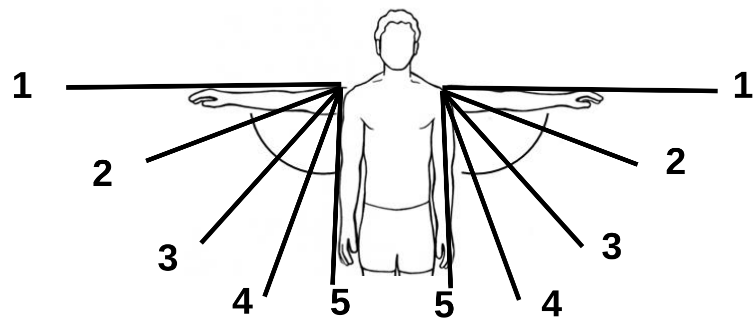


1 **2** **3** **4** **5**

19. Select the image the best reflects the position of just your **upper arm.** Pick a number for your right and then also left arm.



20. Were your arms ever raised out to the side of your body? Pick a number for your right and then also left arm.



21. Please select the option that best describes your speed of work (how fast you were typing, clicking, or using the touch screen) during the task you just completed.

1	Very Slow	(Extremely relaxed pace)
2	Slow	(Taking one's own time)
3	Fair	(Normal speed of motion)
4	Fast	(Rushed, but able to keep up)
5	Very Fast	(Rushed and barely/unable to keep up)

22. On a scale from 0 to 10, 0 being no precision required to 10 being a great amount of precision required, how much precision would you say the task required?

Rating	Verbal Anchor
0	No precision required
1	
2	
3	
4	
5	Moderate amount of precision required
6	
7	
8	
9	
10	Great amount of precision required