Assessing the Relationship between Occupational Injury Risk and Performance: the Efficacy of Adding Adjustability and Using Exoskeletons in the Context of a Simulated Drilling Task

Saad Alabdulkarim

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In
Industrial and Systems Engineering

Maury A. Nussbaum, Chair
Deborah E. Dickerson
Sunwook Kim
Divya Srinivasan

November 1, 2017
Blacksburg, Virginia

Keywords: exoskeleton, adjustability, intervention effectiveness, performance, injury risk
Assessing the Relationship between Occupational Injury Risk and Performance: the Efficacy of Adding Adjustability and Using Exoskeletons in the Context of a Simulated Drilling Task

Saad Alabdulkarim

Abstract

Work-related musculoskeletal disorders (WMSDs) continue to occur despite an increasing understanding of the risk factors that initiate these disorders. Ergonomics is commonly seen as a health and safety approach that has no influence on performance, a perspective potentially hindering intervention proposals in practice. Highlighting potential performance benefits can facilitate intervention cost-justification, along with the traditional focus on reducing exposure to injury risk.

The main objective of this research was to examine the dual influences (i.e., on performance and injury risk) of two distinct types of interventions: adding adjustability, as a commonly advocated approach when considering ergonomics early in the (re)design phase to change task demands; and using exoskeletons to enhance worker capacity. A simulated drilling task was used, which was considered informative as it entailed diverse demands (precision, strength, and speed) and permitted quantifying two dimensions of task performance (productivity and quality).

The dual influences of three levels of workstation adjustability were examined first; increasing adjustability improved performance, with this benefit occurring only when a given level of adjustability also succeeded in reducing ergonomic risk. Across examined conditions, several significant linear associations were found between risk (e.g., Strain Index score) and performance metrics (e.g., completion time), further supporting an inverse relationship between these two outcomes. The dual influences of three distinct passive exoskeletal designs were investigated/compared subsequently, in a simulated overhead drilling task and considering the potential moderating effects of tool mass and precision requirements. Specific designs were: full-body (Full) and upper-body (Arm) exoskeletons with attached mechanical arms; and an upper-body (Shl) exoskeleton providing primarily shoulder support. Both designs with mechanical arms increased static and median total muscle activity while deteriorating quality. The Shl design reduced shoulder loading while increasing dominant upper arm loading and deteriorating quality in the highest precision requirements. Influences of both increasing precision and tool mass were fairly consistent across the examined designs. As such, no single design was obviously superior in both physical demands and performance. Although future work is needed under more diverse/realistic scenarios, these results may be helpful to (re)design interventions that achieve dual benefits on performance and injury risks.
Assessing the Relationship between Occupational Injury Risk and Performance: the Efficacy of Adding Adjustability and Using Exoskeletons in the Context of a Simulated Drilling Task

Saad Alabdulkarim

General Audience Abstract

Occupational injuries continue to occur despite an increasing understanding of what can cause these injuries. A potential reason for this continuing problem is that Ergonomics is commonly seen as a health and safety approach that has no influence on performance. This perspective can hinder approving/funding interventions that aim to reduce injury risk in workplaces. Emphasizing that these interventions can also improve performance (and lead to more profits) can better convince decision makers to approve them, and subsequently reduce injury risk. The primary objective of this dissertation was to examine the influences of two types of ergonomic interventions on both performance and injury risk. Three experiments were completed in the context of a simulated drilling task. This specific task was considered informative, as it entailed diverse demands (precision, strength, and speed) and permitted quantifying two dimensions of task performance (productivity and quality). Examined interventions were: 1) adding adjustability, as a commonly advocated approach when considering Ergonomics early in the (re)design phase to change task demands; and, 2) using exoskeletons, to potentially enhance worker capacity. For the former, increasing adjustability improved performance. However, this benefit occurred only when a given level of adjustability also succeeded in reducing injury risk. This suggests that reducing injury risk is associated with increasing performance. For exoskeletons, the dual influences of three distinct exoskeletal designs were investigated/compared in a simulated overhead drilling task. The potential effects of tool mass and precision requirements were also considered. Specific designs were: full-body and upper-body exoskeletons with attached mechanical arms; and an upper-body exoskeleton providing primarily shoulder support. Both designs with mechanical arms increased overall muscle activity while deteriorating quality. The shoulder-focused design reduced shoulder demands while increasing demands on the dominant upper arm and deteriorating quality in the highest precision requirements. Effects of both increasing precision requirements and tool mass were fairly consistent across the examined designs. As such, no single design was obviously superior in both physical demands and performance. These results may be helpful to (re)design interventions that achieve dual benefits on performance and injury risks. However, future work is needed under more diverse/realistic scenarios.
Acknowledgements

First, I would like to sincerely thank my advisor, Dr. Maury Nussbaum, for his exceptional guidance, continuous support, and great patience. Working under your supervision was beneficial on multiple dimensions. I am greatly thankful for all the knowledge and wisdom you shared with me not only to successfully complete my dissertation, but also helping me to succeed in my academic career, which you were a great role model for. I also would like to thank my committee members, Dr. Deborah Dickerson, Dr. Sunwook Kim, and Dr. Divya Srinivasan for all their advice and feedback. Special thanks to Dr. Kim, who, despite his busy schedule, was always open to helping me in brainstorming or solving any issues raised during the experiments.

Additionally, I would like to thank the current and graduated members of the Occupational Ergonomics and Biomechanics Lab, in particular my officemate Dr. Alireza Sedighi, Dr. Ehsan Rashedi, Mohammad Esfahani, and Jang-Ho Park. I am also grateful to Randy Waldron and Will Vest, from the ISE department staff, for their timely and technical help in setting up the experiments.

Special thanks and gratitude goes to my parents in Saudi Arabia, who have continuously motivated me to work harder. For my mother, thank you for supporting me and my family nearly every day, while I have been studying abroad. Additionally, I am very thankful for my beloved wife, Bassmh Alabdulkarim, for greatly easing me through the many tough times I encountered during this journey. Last but not least, for my dear son, Abdulmohsen, having you made me work substantially harder; looking at your happy smile after a stressful workday was a tremendous relief, thank you.

Many thanks go to The Boeing Company and Ford Motor Company for loaning equipment needed in my studies and communicating aspects deemed important from an industrial perspective. Lastly, I would like to thank my sponsor, King Saud University, in Riyadh, Saudi Arabia, for giving me the opportunity to pursue my graduate studies. I am also grateful to the Saudi Arabian Cultural Mission in the United States for all their efforts for me and my family while being abroad.
# Table of Contents

**Chapter 1: Introduction** ................................................................. 1  
  Accounting for Individual Variability in Workstation Design .................. 4  
  Wearable Exoskeletons ........................................................................ 6  
References .............................................................................................. 9  

**Chapter 2: Impact of Task Design on Task Performance and Injury Risk: Case Study of a Simulated Drilling Task** ................................................................. 13  
Abstract ............................................................................................... 13  
Introduction ............................................................................................ 14  
Methods .................................................................................................. 17  
  Participants ............................................................................................ 17  
  Task Description .................................................................................... 17  
  Independent Variable ........................................................................... 20  
  Procedures and Data Collection ............................................................ 21  
  Dependent Measures ............................................................................ 23  
  Statistical Analyses .............................................................................. 28  
Results ..................................................................................................... 29  
  Productivity-related Metrics ................................................................. 30  
  Quality-related Metrics ....................................................................... 30  
  Risk Metrics ......................................................................................... 31  
  Perceived Exertion and Discomfort ....................................................... 33  
  Relationships between Risk and Performance ...................................... 34  
  Perceived versus Measured Performance ............................................ 37  
Discussion ............................................................................................... 39  
References ............................................................................................... 46  

**Chapter 3: Influences of Different Exoskeleton Designs and Tool Mass on Physical Demands and Performance in a Simulated Overhead Drilling Task** .................................................. 51  
Abstract ............................................................................................... 51  
Introduction ............................................................................................ 52  
Methods .................................................................................................. 55  
  Participants ............................................................................................ 55  
  Task Description .................................................................................... 55  
  Independent Variables ........................................................................... 57  
  Procedures and Data Collection ............................................................ 58  
  Dependent Measures ............................................................................ 63  
  Statistical Analyses .............................................................................. 64  
Results ..................................................................................................... 65  
  Maximum Acceptable Frequency and Quality Metrics .......................... 65  
  Ratings of Perceived Discomfort Metrics ............................................ 67  
  Peak Loading Metrics (90th percentile EMG) ........................................ 69  
  Median Loading Metrics (50th percentile EMG) ...................................... 71
Chapter 5: Conclusion and Recommendations

Introduction .................................................................................................................. 120
Effects of Workstation Adjustability ........................................................................ 121
Effects of Exoskeleton Design ..................................................................................... 121
Future Directions ......................................................................................................... 124

Appendices .................................................................................................................. 125
Appendix A: Participants Instructions (modified from Snook et al., 1995) .......... 125
Appendix B: Usability and Discomfort Questionnaire ............................................. 126
List of Figures

Figure 1. A view of the mock fuselage used in the study. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 18

Figure 2. The commercial drill (a) used with the load cell and the drill bit attached to it as illustrated in figure (b). Note that the drill was wrapped with an elastic band to cover all wires during the experiment. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 19

Figure 3. Examples of adopted postures for the Some and None conditions. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 22

Figure 4. Productivity-related performance results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 30

Figure 5. Quality-related performance results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 31

Figure 6. RULA risk results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 32

Figure 7. Strain Index risk results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 33

Figure 8. Ratings of perceived exertion (RPE) and discomfort (RPD). Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 34

Figure 9. Relationships between performance measures and RULA scores. Note, several transformations were used to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 35

Figure 10. Relationships between performance measures and SI scores. Note, a square root transformation on FTE data was needed to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 36

Figure 11. Mean (95% confidence intervals) values of performance measures and RULA scores. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 36

Figure 12. Mean (95% confidence interval) values of performance measures and SI scores. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 37

Figure 13. Relationships between perceived and measured performance assessed using simple linear regression models. FTE data were transformed to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017). ................................................................. 38
Figure 14. Illustrations of participants using: (a) the FORTIS™ Exoskeleton (*Full*); (b) the SuitXTM ShoulderXTM (Shl); and (c) the Fawcett Exsovest™ with a zeroG2 mechanical arm (*Arm*)

Figure 15. Maximum Acceptable Frequency (MAF) for the four Exo levels (a). The symbol * indicates a significant difference between genders within Exo levels. Letters specify groupings obtained from pairwise comparisons between Exo conditions (among females in figure a). Quality results, in terms of the number of errors made (b). Letters specify groupings obtained from pairwise comparisons between conditions. In both graphs, error bars indicate 95% confidence intervals.

Figure 16. Ratings of perceived discomfort (RPD) for different body regions. Letters specify groupings obtained from pairwise comparisons between Exo conditions (within Gender for figure d), and error bars indicate 95% confidence intervals.

Figure 17. Peak loading EMG metrics (as a percentage of maximum levels). The symbol * indicates a significant difference between genders (figure c) and between Tmass levels within Exo levels (figure e). Letters specify groupings obtained from pairwise comparisons between Exo conditions (within Gender for figure c and Tmass for figure e), and error bars indicate 95% confidence intervals.

Figure 18. Median loading EMG metrics (as a percentage of maximum levels). The symbol * indicates a significant difference between genders within Exo levels. Letters specify groupings obtained from pairwise comparisons between Exo conditions (within Gender for figures a-c and h), and error bars indicate 95% confidence intervals.

Figure 19. Static loading EMG metrics (as a percentage of maximum levels). The symbol * indicates a significant difference between genders within Exo levels. Letters specify groupings obtained from pairwise comparisons between Exo conditions (within Gender for figures a, b, g, and h), and error bars indicate 95% confidence intervals.

Figure 20. Usability questions responses on a Likert type scale (1= strongly disagree, 5= strongly agree). Q1: “It meets my needs for completing the task,” Q2: “It makes the things I want to accomplish easier to get done,” Q3: “It is easy to use,” Q4: “I am satisfied with it,” Q5 “It works the way I want it to work, and Q6: “Overall, I prefer working with it more than without it.” Letters specify groupings obtained from pairwise comparisons between Exo conditions, and error bars indicate 95% confidence intervals.

Figure 21. A view of the height adjustable fixture (rung) that was used in the study (a). The commercial drill (d) that was used, with the load cell and the drill bit attached. Note that the drill was wrapped to cover all wires during the experiment. Photos of participants using: (b) the Fawcett Exsovest™ with a zeroG2 mechanical arm (*Arm*); (c) the EksoWorks™ Vest (Shl); and (c) the FORTIS™ Exoskeleton (*Full*).

Figure 22. Quality measure (total number of errors made) for the four Exo conditions. Letters indicate groupings obtained from pairwise comparisons between Exo levels. Error bars show 95% confidence intervals.

Figure 23. Ratings of Perceived Exertion (RPE) for several body parts. Letters indicate groupings obtained from pairwise comparisons between Exo levels (among males in figure f). Error bars show 95% confidence intervals. The symbol * specifies a significant difference between genders within Exo conditions in figure f.
Figure 24. Total (across all monitored muscles) peak, median, and static EMG loading measures, each as percentages of MVC. Letters indicate groupings obtained from pairwise comparisons between Exo levels. Error bars show 95% confidence intervals. The symbol * specifies a significant difference between genders within Exo conditions for static values.

Figure 25. Peak loading measures as percentages of MVC. Letters indicate groupings obtained from pairwise comparisons between Exo levels. Error bars show 95% confidence intervals.

Figure 26. Median loading measures as percentages of MVC. Letters indicate groupings obtained from pairwise comparisons between Exo levels (within Gender for figures b, e, and f). Error bars show 95% confidence intervals.

Figure 27. Static loading measures as percentages of MVC. Letters indicate groupings obtained from pairwise comparisons between Exo levels (within Gender for figures a, b, and e). Error bars show 95% confidence intervals. The symbol * specifies a significant difference between genders within Exo conditions for the AD-D.

List of Tables

Table 1. Statistical results for the main and interactive effects of adjustability condition and force. Specific transformations of the dependent variables used are indicated (to meet model assumptions). P values are presented along with effect sizes (partial $\eta^2$) in parentheses for each effect, with bold font highlighting significant effects ($p<0.05$). Reprinted with permission from Alabdulkarim et al., (2017).

Table 2. P values, $r^2$, and slopes of fitted lines for the simple linear regression models of risk-performance relationships. Bold font indicates significant slopes ($p<0.05$). Note, data from all three conditions. Reprinted with permission from Alabdulkarim et al., (2017).

Table 3. Results from linear regression modeling of the relationship between perceived and measured performance Note that FTE data were square-root transformed to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017).

Table 4. Statistical results for the main and interactive effects of Exo, $T_{mass}$, and Gender on maximum acceptable frequency (MAF), Errors, and ratings of perceived discomfort (RPD). Any transformations applied to the dependent measures are indicated. P values are reported along with effect sizes ($\eta^2_p$) in parentheses for each effect, with bold font highlighting significant effects ($p<0.05$).

Table 5. Statistical results for the main and interactive effects of Exo, $T_{mass}$, and Gender on peak loading EMG metrics. Any transformations applied to the dependent measures are indicated. P values are reported along with effect sizes ($\eta^2_p$) in parentheses for each effect, with bold font highlighting significant effects ($p<0.05$).

Table 6. Statistical results for the main and interactive effects of Exo, $T_{mass}$, and Gender on median loading EMG metrics. Any transformations applied to the dependent measures are indicated. P values are reported along with effect sizes ($\eta^2_p$) in
parentheses for each effect, with bold font highlighting significant effects ($p<0.05$).

Table 7. Statistical results for the main and interactive effects of Exo, $T_{mass}$, and gender on static loading EMG metrics. Any transformations applied to the dependent measures are indicated. $P$ values are reported along with effect sizes ($\eta^2_p$) in parentheses for each effect, with bold font highlighting significant effects ($p<0.05$).

Table 8. Results from linear regression modeling of the relationships between MAF and other physical demand metrics. $P$ values are reported along with coefficient of determination ($r^2$) in parentheses, with bold font indicating significant relationships ($p<0.05$).

Table 9. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on errors and ratings of perceived exertion (RPE). Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) are reported for each effect in parentheses, with bold font indicating significant effects ($p<0.05$).

Table 10. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the total peak, median, and static loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$).

Table 11. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the peak loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$).

Table 12. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the median loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$).

Table 13. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the static loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$).
Chapter 1: Introduction

The number of work-related musculoskeletal disorders (WMSDs) reached approximately 365,000 cases in the United States in 2014, and the manufacturing sector, in contrast to most other sectors, had a 5% increase in WMSD cases compared to 2013 (BLS, 2015). These disorders involve a substantial personal and societal burden (NRC, 2001). WMSDs continue to occur despite an increasing understanding of the risk factors that initiate these disorders, and the reasons why such understanding has not consistently translated into a reduction in WMSDs are likely diverse, and are as yet not completely known. One likely reason is that operations management research is traditionally separated from that in physical ergonomics (Neumann & Dul, 2010), along with increasing management disinterest in issues with ethical implications and a more exclusive focus on profitability (Walsh et al., 2003). An important consequence of the former separation is the (mis)conception that ergonomics is strictly a health and safety tool that has no influence on the return of investment (Hägg, 2003; Zare et al., 2015). Such a belief makes it more difficult to justify ergonomic interventions from a direct cost-savings standpoint.

WMSDs can have both tangible and intangible costs. While tangible costs, such those related to worker’s compensation costs and replacing (and training) injured workers, are relatively easy to track, it is more challenging to quantify intangible costs, such as reduced performance due to WMSD symptoms. Although the former may be easier to quantify, it can be misleading to measure the effect of (poor) ergonomic conditions based solely on these (e.g., the number of lost workdays). Musculoskeletal complaints also reduce productive time (Ricci et al., 2006), contribute to lower work ability (Neupane et al., 2011), and cause early retirement (van den Berg et al., 2010). Such intangible costs can be amplified by poor ergonomic conditions. In a strategy developed for the
future of human factors/ergonomics discipline by the International Ergonomics Association (IEA), Dul et al. (2012) indicated that common applications that do not highlight the dual influence of ergonomics on performance and human well-being, and in a broader sense do not use a holistic approach in design, play a role in under-exploiting the potential benefits of ergonomics.

Ergonomic interventions, particularly those done proactively, typically require economic justification, yet there are factors that make this challenging. For example, a common focus is on injury-related outcomes, yet these tend to be lagging (not leading) indicators. Further, workers are variable in both their exposures and responses to given tasks. Focusing solely on potential injury reduction may also miss potential positive impacts of ergonomic interventions on worker performance (i.e., productivity and quality).

In a review of 250 case studies, Goggins et al. (2008) showed that positive effects, including increased productivity and quality and reduced turnover and absenteeism, subsequent to ergonomic interventions across a variety of industrial sectors. In a recent review of the influence of ergonomics on quality, 12 studies were summarized as demonstrating a strong relationship between physical ergonomics risk and quality (Zare et al., 2015). However, it is important to note that many of these studies (e.g., Eklund, 1995; Falck et al., 2010; Falck et al., 2014; Fritzsche et al., 2014) were done to quantify the association between predefined risk groups (e.g., high, moderate and low risk) and quality metrics, for several tasks in targeted manufacturing plants. Given that task design variables, and the subsequent risk levels, were not “manipulated” in the same task, such studies provide only limited support for causal relationships between task (re)design and both performance improvement and risk reduction.
Other studies (e.g., González et al., 2003; Yeow & Sen, 2006; Das et al., 2007) involved “interventions”, in which risk level was manipulated through some modifications to a task. These studies, however, typically included only two levels of ergonomic risk and associated quality outcomes (i.e., before and after an intervention). While providing stronger support for the noted causality, these studies also provide incomplete support for ergonomic interventions, for two reasons. First, two-level studies do not necessarily capture the shape (or, functional form) of the relationship between task design factors (e.g., posture) or ergonomic risk and performance. Understanding such relationships would be valuable for purposes of intervention justification, particularly for tasks involving low or moderate levels of risks (for which existing evidence is more limited in the context of worker performance). Second, in many of the noted studies the intervention substantially changed the nature of the task, often such that human involvement was minimized (e.g., adding a conveyor instead of a worker manually handling boxes). In these cases, performance improvement could not clearly be tied to reduced task-related ergonomic risk.

Despite these limitations, however, existing evidence does support at least an association between an ergonomic intervention and a subsequent positive effect on both ergonomic risk and performance (i.e., productivity and quality). From a proactive perspective, such relationships are of value, as they can facilitate cost/benefit analyses of an intervention prior to any capital investments. When (re)designing interventions (or tasks), ergonomists typically aim to improve the balance between task demands and worker capacity to reduce or minimize the risk of developing WMSDs. The main objective of this dissertation was to investigate two factors that can influence such a balance: 1) designing with adjustability, as a factor amendable to design; and, 2) use of wearable exoskeletons, as an intervention that can enhance worker capacity, while
considering two potential moderating factors relevant for exoskeletons occupational application, precision requirements, as a task design variable, and tool mass. Considering these moderating factors is important as they can modify the noted balance. The overall hypothesis was that the examined ergonomic interventions (i.e., adding adjustability and using exoskeletons) would both improve task performance and reduce injury risk. In the context of a simulated drilling task, three studies were completed to examine this hypothesis. The following sections elaborate on each of the examined factors.

**Accounting for Individual Variability in Workstation Design**

Workers have highly variable functional characteristics such as strength, endurance, and anthropometry. Such variability can influence how they interact with a design (Garneau & Parkinson, 2011). Generally, there is a tradeoff between body parts in tasks (re)design (Marras, 2006); as an example, it may not be feasible to reduce postural extremes for all body parts in a given (re)design, particularly for complex tasks. Such design challenges can be magnified by the noted variability between workers.

There are three fundamental approaches to deal with anthropometric variability: design for extreme, design for the average, and providing adjustability (Bhattacharya & McGlothlin, 1996). In the design for extreme approach, dimensions of the largest or smallest individuals (e.g., 5th or 95th percentiles) are used, such as to determine clearance dimensions (e.g., height of an overhead conveyer) or reach capabilities (e.g., maximum height of shelves). Designing for the average may lead to the greatest proportion of users being insufficiently accommodated, and may lead these users to adopt more non-neutral postures, potentially causing discomfort, compromised
performance, and/or increased risk of developing WMSDs. Providing adjustability is done to accommodate the maximum proportion of users. However, adding adjustability features typically requires more expensive components, may increase maintenance costs and make breakdown times longer, and may decelerate production due to the additional time needed to (re)adjust the task.

There are two additional important factors related to designing adjustability, which are the extent (or range) and location of adjustability. For the former, workstation height is typically designed to accommodate from a 5th percentile female through 95th percentile male, such as using elbow height data (Bhattacharya & McGlothlin, 1996). For the latter, adjustability should be “optimally” allocated (Nadadur & Parkinson, 2013). Factors defining such optimization problems include proportion of population targeted for accommodation, budget constraints, location constrains and expected performance gains. When defining the population, it should be recognized that designs should aim to accommodate potential users rather than existing ones, for several reasons including the possibility that existing ones can be “survivors” from a larger group of workers who left the job because of WMSDs caused by the design (Chengalur et al., 2004). Also, there are “sub-groups” of workers that should be recognized. de Vries and Parkinson (2014) discussed that designs should ensure a minimum “rate of disaccommodation” for specific sub-groups of works (e.g., females in a male-dominated workplace).

Because of the different costs associated with each design approach, it is important to examine the efficacy of these approaches on both accommodating the “representative” variability (and subsequently reducing injury risk) and performance. Performance results can be used to cost-justify such (re)designs. It should be noted that accommodation is not a binary variable. Rather,
there are many indicators to detect the extent of accommodation, such as regional ratings of perceived discomfort and postural extremes.

**Wearable Exoskeletons**

In the United States, overexertion was the leading exposure resulting in occupational injuries or illnesses, accounting for 33% of total non-fatal occupational cases involving days away from work, with a median of 13 days away from work in 2014 (BLS, 2015) and costing one insurer ~$15 billion in worker compensation in 2012 (Liberty Mutual Research Institute for Safety, 2014). Overexertion injuries are typically considered to result when task demands exceed worker capacity (Chaffin et al., 2006). When tasks are physically demanding, it is common to use (or attempt) automation as an engineering control in industry. However, this approach sometimes is infeasible and/or unjustifiably expensive. Common assistive devices include mechanical manipulators and wearable exoskeletons. The focus in this dissertation is on the latter, as they can be used in a wider range of tasks for a given job, might requires less changes to the workplace and, thus, can be less expensive.

Exoskeletons are designed to enhance user capacity (de Looze et al., 2015), and can improve the balance between task demands and worker capacity. They can broadly be classified as passive or active devices (de Looze et al., 2015). While active devices can provide more support, they are usually heavier and more expensive. Active devices (e.g., Kobayashi et al., 2007) use one or more actuators to augment human power with different technology, such as pneumatic muscles, hydraulic or electric motors (Gopura & Kiguchi, 2009). Power augmentation in completely passive devices (e.g., Abdoli-E et al., 2006), however, comes from springs, dampers, or materials
that can store the energy generated by user motion and discharge it as designed (de Looze et al., 2015). Exoskeletons can also be categorized by the body part(s) they are design to support (Lee et al., 2012; de Looze et al., 2015), such as upper body (e.g., Kobayashi et al., 2007), lower body (e.g., Kim et al., 2009), and full body (e.g., Toyama & Yonetake, 2007) devices. Finally, exoskeleton designs can fundamentally differ in their approach to transferring external loads (e.g., from the shoulder to the waist/hip or to the ground).

Wearable exoskeletons have been investigated to date mainly for rehabilitation/medical purposes, such as to support weak, injured, or disabled patients (Viteckova et al., 2013) or for military applications (Lee et al., 2012). For occupational use, a recent review highlighted a potential for these devices and showed that they can reduce physical demands considerably (de Looze et al., 2015). However, while they can reduce demands on specific part(s) of the wearer’s body, they can have unintended consequences such as increasing loading and/or discomfort on “other” regions of the wearer’s body. In a simulated assembly task with the trunk in a forward-bending posture, a passive exoskeleton reduced loading on the back and legs but increased discomfort in the chest (Bosch et al., 2016). Discomfort normally occurs in areas where an exoskeleton contacts the wearer’s body. Even if a biomechanical advantage was confirmed, eliminating/minimizing such discomfort is fundamental, particularly because discomfort can prevent workers from using the devices (de Looze et al., 2015). In a simulation of overhead work, using an upper limb passive exoskeleton reduced loading on the upper arm and shoulder but increased it on the low back (Rashedi et al., 2014), further highlighting potential inconsistent influences of exoskeletons.
In addition, wearing exoskeletal devices can modify working techniques, which may (partly) explain observed changes in the physical loading. For example, when lifting with the personal lift-assist device (PLAD), participants changed their lifting technique more towards squat-like method by reducing lumbar and thoracic flexion and increasing hip and ankle flexion (Sadler et al., 2011). Such changes may explain specifically the increased legs muscle activities found when using passive exoskeletons (de Looze et al., 2015).

Precision demands are present to differing degrees in most occupational tasks, and can have important physiological and biomechanical effects on workers as well influencing task performance. In a seated posture, for example, increasing precision caused participants to adopt “poorer” postures for a manual task (Li & Haslegrave, 1999). Increasing precision can also increase muscle activation; in a task simulating dentistry work, Milerad and Ericson (1994) found that increased precision demands led to higher activity levels in muscles with a stabilizing function (mainly in neck, shoulder and wrist). In addition to increasing co-contraction, precision can increase loading by modifying movement kinematics; with precision, movement time increases, a phenomena commonly described as the speed-accuracy tradeoff, and widely explained by Fitts’ Law (Fitts, 1954; Fitts & Peterson, 1964). More specifically in goal-oriented movements, as target size decreases, there is a longer decelerative phase during which more corrective sub-movements occur (e.g., MacKenzie et al., 1987; Thompson et al., 2007; Temprado et al., 2013). In repetitive lifting tasks, precision requirements also led to changes in posture and/or muscle activation (Collier et al., 2014; Joseph et al., 2014; Mehta et al., 2015), further supporting the occupational significance of precision. Repetitive movements requiring precision are also recognized as a risk factor for neck and shoulder injury (Ekberg et al., 1994). Along with the noted effects of precision on physical demands, increasing precision requirements can reduce task performance, as was
found in a computer-simulated crane operation (Huysmans et al., 2006) and an aiming task using a computer mouse (Visser et al., 2004). Higher precision demands also increased completion time and reduced quality in a simulated assembly task common in the automotive industry (Wartenberg et al., 2004). Wearing exoskeletons can restrict movement and modify working strategies; such changes may interact with precision effects and lead to influences on physical demands and precision task performance, two dimensions of effectiveness that are important for eventual occupational application.

As argued earlier, a better understanding of the association between ergonomic risk and performance can be useful to both cost-justify ergonomic interventions and to highlight the potential performance decrements due to the sub-optimal ergonomic conditions. This dissertation sought to provide such knowledge in three laboratory studies, with each study presented here in a chapter. The first study investigated the impact of workstation adjustability on performance and injury risk; the second and third compared three different passive exoskeletal designs in terms of performance and injury risk in a simulated overhead work considering the potential moderating effects of tool mass (Study 2) and precision requirements (Study 3). The last chapter (Chapter 5) provides overall conclusions, practical implications, and directions for future research.

References


Chapter 2: Impact of Task Design on Task Performance and Injury Risk: Case Study of a Simulated Drilling Task

Abstract

Existing evidence is limited regarding the influence of task design on performance and ergonomic risk, or the association between these two outcomes. In a controlled experiment, we constructed a mock fuselage to simulate a drilling task common in aircraft manufacturing, and examined the effect of three levels of workstation adjustability on performance as measured by productivity (e.g., fuselage completion time) and quality (e.g., fuselage defective holes), and ergonomic risk as quantified using two common methods (RULA and the Strain Index). The primary finding was that both productivity and quality significantly improved with increased adjustability, yet this occurred only when that adjustability succeeded in reducing ergonomic risk. Supporting the inverse association between ergonomic risk and performance, the condition with highest adjustability created the lowest ergonomic risk and the best performance; while there was not a substantial difference in ergonomic risk between the other two conditions, in which performance was also comparable.

Keywords: Ergonomic risk, performance, quality, adjustability, task design
Introduction

Work-related musculoskeletal disorders (WMSDs) reached approximately 365,000 cases in the United States in 2014, and the manufacturing sector, in contrast to most other sectors, had a 5% increase in WMSD cases compared to 2013 (BLS, 2015). These disorders involve a substantial personal and societal burden (e.g., NRC, 2001). WMSDs continue to occur despite an increasing understanding of the risk factors that initiate these disorders, and the reasons why such understanding has not consistently translated into a reduction in WMSDs are likely diverse and as yet not completely known. One likely reason is that operations management research is traditionally separated from that in physical ergonomics (Neumann & Dul, 2010), along with increasing management disinterest in issues with ethical implications and a more exclusive focus on profitability (Walsh et al., 2003). An important consequence of the former separation is the (mis)conception that ergonomics is strictly a health and safety tool that has no influence on the return of investment (Hägg, 2003; Zare et al., 2015). Such a belief makes it more difficult to justify ergonomic interventions from a direct cost-savings standpoint.

WMSDs can have both tangible and intangible costs. While tangible costs, such those related to worker’s compensation costs and replacing (and training) injured workers, are relatively easy to track, it is more challenging to quantify intangible costs, such as reduced performance due to WMSD symptoms. Although the former may be easier to quantify, it can be misleading to measure the effect of (poor) ergonomic conditions based solely on these (e.g., the number of lost workdays). Musculoskeletal complaints also reduce productive time (Ricci et al., 2006), contribute to lower work ability (Neupane et al., 2011), and cause early retirement (van den Berg et al., 2010). Such intangible costs can be amplified by poor ergonomic conditions. In a strategy developed for the
future of human factors/ergonomics discipline, Dul et al. (2012) indicated that common applications that do not highlight the dual influence of ergonomics on performance and human well-being, and in a broader sense do not use a holistic approach in design, play a role in under-exploiting the potential benefits of ergonomics.

Ergonomic interventions, particularly those done proactively, typically require economic justification, yet there are factors that make this challenging. For example, a common focus is on injury-related outcomes, yet these tend to be lagging (not leading) indicators. Further, workers are variable in both their exposures and responses to given tasks. Focusing solely on potential injury reduction may also miss potential positive impacts of ergonomic interventions on worker performance (i.e., productivity and quality).

In a review of 250 case studies, Goggins et al. (2008) showed that positive effects, including increased productivity and quality and reduced turnover and absenteeism, subsequent to ergonomic interventions across a variety of industrial sectors. In a recent review of the influence of ergonomics on quality, 12 studies were summarized as demonstrating a strong relationship between physical ergonomics risk and quality (Zare et al., 2015). However, it is important to note that many of these studies (e.g., Eklund, 1995; Falck et al., 2010; Falck et al., 2014; Fritzsche et al., 2014) were done to quantify the association between predefined risk groups (e.g., high, moderate and low risk) and quality metrics, for several tasks in targeted manufacturing plants. Given that task design variables, and the subsequent risk levels, were not “manipulated” in the same task, such studies provide only limited support for causal relationships between task (re)design and both performance improvement and risk reduction.
Other studies (e.g., González et al., 2003; Yeow & Sen, 2006; Das et al., 2007) involved “interventions”, in which risk level was manipulated through some modifications to a task. These studies, however, typically included only two levels of ergonomic risk and associated quality outcomes (i.e., before and after an intervention). While providing stronger support for the noted causality, these studies also provide incomplete support for ergonomic interventions, for two reasons. First, two-level studies do not necessarily capture the shape (or, functional form) of the relationship between task design factors (e.g., posture) or ergonomic risk and performance. Understanding such relationships would be valuable for purposes of intervention justification, particularly for tasks involving low or moderate levels of risks (for which existing evidence is more limited in the context of worker performance). Second, in many of the noted studies the intervention substantially changed the nature of the task, often such that human involvement was minimized (e.g., adding a conveyor instead of a worker manually handling boxes). In these cases, performance improvement could not clearly be tied to reduced task-related ergonomic risk.

Despite these limitations, however, existing evidence does support at least an association between an ergonomic intervention and a subsequent positive effect on both ergonomic risk and performance (i.e., productivity and quality). From a proactive perspective, such relationships are of value, as they can facilitate cost/benefit analyses of an intervention prior to any capital investments. The purposes of the current study were twofold, and were addressed within the context of a specific repetitive drilling task that is common in aircraft manufacturing and was simulated in a laboratory environment. First, to evaluate if causal relationships existed between workstation adjustability, as a task variable amendable to (re)design, and both performance and ergonomic risk. Second, to assess whether ergonomic risk was associated with performance. We
hypothesized that the addition of workstation adjustability will improve performance and reduce ergonomic risk and that performance and risk will be inversely associated.

**Methods**

**Participants**

A total of 34 individuals (18 males, 16 females) were recruited from the local community and University to participate in the study. All completed an informed consent procedure approved by the Virginia Tech Institutional Review Board. Of these, 16 (1 male, 15 females) were removed during an initial screening session, as they were unable to consistently perform all of the experimental tasks. The remaining 18 participants (17 males, 1 female) completed the study. All but one male participant reported being right handed, and all reported no recent (past year) or current musculoskeletal problems and having normal or corrected-to-normal vision. Their mean (SD) age, stature, and body mass were 23.4 (2.6) years, 176.5 (7.8) cm, and 78.2 (17.5) kg, respectively.

**Task Description**

A repetitive drilling task was simulated in the laboratory, and was chosen because it approximates a realistic manufacturing task, is repetitive, involves both precision and strength demands, and it was feasible to implement different levels of adjustability. A mock cylindrical fuselage was constructed for use in the study (Figure 1), the configuration of which was considered representative of small aircraft. An 80 cm diameter disc was at either end of the cylinder, and these were connected by six rungs to simulate longitudinal stringers (5.08 x 5.08 cm hollow rectangular stock, 121.9 cm long) spaced evenly around the circumference of the discs. During the experiment,
the initial fuselage configuration was such that there were two rungs each in the “upper”, “middle”, and “lower” positions, and two rails at the height of the central axis of the cylinder. Each rung had six evenly-space holes that were 1.3 cm in diameter, for a total of 36 holes. The structure was covered with heavy canvas fabric, to simulate the fuselage skin, with cut-outs at each hole. The fuselage was attached to vertical structures that allowed for changes in height and rotation about the central axis.

![Figure 1. A view of the mock fuselage used in the study. Reprinted with permission from Alabdulkarim et al., (2017).](image)

A commercial pneumatic drill (Figure 2) was employed (mass = 1.10 kg; model #A2801447, Atlas Copco, Hungary). A simulated drill bit was used, with four components. At the base, a hexagonal portion was “chucked” in the drill, and connected to a uniaxial load cell (Interface, SML-100, Scottsdale, AZ). The load cell, in turn, was connected to a “probe”, consisting of sequential portions of steel and nylon with respective lengths of 3.3 and 2.5 cm. The diameter of the probe was 1 cm, and was used to mimic the quality requirements from an aircraft manufacturer’s
specifications for drilling fuselage fastener holes. Specifically, drilling for subsequent riveting requires that holes be within an angularity tolerance of ±2° of normal to the surface. Holes drilled outside this tolerance range are considered errors or defects as excessive hole angularity adversely impacts rivet quality. Given the diameters of the rung holes, the probe diameter, and the rung dimensions, any deviation beyond 2° of normal led to contact between the probe and a rung, and was captured as an “error” as described below.

Figure 2. The commercial drill (a) used with the load cell and the drill bit attached to it as illustrated in figure (b). Note that the drill was wrapped with an elastic band to cover all wires during the experiment. Reprinted with permission from Alabdulkarim et al., (2017).

The experimental task involved completing a sequence of discrete simulated drilling actions. Each action involved three steps. The probe was first inserted into a hole and centered. Participants then needed to generate sufficient force for a set duration to “complete” a hole. Exertions were isometric because the probe did not make any displacement through the rung. Finally, the probe was removed and the participant moved to the next hole. Two target forces used were 66.7 N and 111.2 N, which
represents forces required in practice to complete “pilot” and “full size” drilling in a typical aluminum fuselage skin structure – determined consulting an industry subject matter expert. Pilot drilling was completed first, for all holes in the fuselage, followed by final drilling after a rest break (~10 min). The same sequence of holes was completed throughout the experiment, and by all participants, with all holes numbered to facilitate this.

The accumulated duration for a complete hole was set to 2.5 sec. Forces being generated needed to exceed the relevant target for this duration, and were monitored using the load cell (at 1000 Hz). When forces were below the target, the duration of this did not count toward the 2.5 sec required total for a hole. No maximal force level was specified. For each hole, participants were provided with three auditory feedback types indicating when the generated force was above the target (tone 1) and when the 2.5 sec duration was completed (tone 2). Feedback was also generated when an “error” was made (tone 3), consisting of contact of the probe with a rung. Both the probe and the fuselage were wired together; contact closed the circuit and this was used within the data collection system to generate the error tone when the noted contact was made. A custom Labview™ program (National Instruments, Austin, TX, USA) was developed to manage the experiment.

**Independent Variable**

A single independent variable was manipulated in the experiment, which was the extent of “workstation adjustability”. This variable and the levels used were selected partly because this variable is feasibly modifiable in actual settings. Three adjustability conditions were created:

- No adjustability (None): the midline of the fuselage was set at mean elbow height, or 108.2 cm (mixed gender sample: Marras & Kim, 1993). Participants were allowed to choose any
posture in this (and any) condition; they often squatted, or sat on the floor, for the lower rungs in this condition (Figure 3d and 3e).

- **Some adjustability (Some):** the midline of fuselage was set at mean shoulder height, or 144.0 cm (Marras & Kim, 1993). This condition made it easier to access the lower rungs, and a 2-step stool was available on each side of the fuselage to aid in reaching the higher rungs. When completing the lower rungs, many participants sat on the stool (Figure 3a-c).

- **High adjustability (High):** the fuselage height was set to participant preference, following initial practice. The mean (SD) of chosen heights was 69.4 (5.5) % of individual stature. In this condition, the fuselage could also be rotated by the participant about the long axis. The fuselage was locked to prevent rotation during drilling.

**Procedures and Data Collection**

Participants completed both training and testing sessions, separated by at least two days to minimize any effect from residual muscle fatigue. In the training session (~2.5 hours), participants were asked to practice drilling the entire fuselage at least 10 times. Practice was performed for the three conditions and at the two force levels. More practice was provided as needed, until participants reported being comfortable and proficient with the procedures. Participants were also instructed during practice (and again during the testing session) to work as quickly but as accurately as possible. Participants were allowed to use one or both hands during the tasks, though not allowed to grab/hold the probe during drilling, and freely chose their postures. When practicing the Some condition, participants were instructed to make sure they know when and how to move the stools.
In the testing session (~2 hours), participants were reminded about the procedures and completed additional practice (~15 min). The order of exposure to the three adjustability conditions was completely counterbalanced. Rest breaks of at least 10 minutes were provided after completing each condition. Generated forces were collected (at 1000 Hz) throughout the testing sessions, and participants’ postures were monitored (at 60 Hz) using a wearable inertial motion capture system (MVN, Xsens technologies B.V., Enschede, the Netherlands).

![Figure 3](image)

Figure 3. Examples of adopted postures for the Some and None conditions. Reprinted with permission from Alabdulkarim et al., (2017).
**Dependent Measures**

To evaluate productivity and quality, several measures were obtained. Productivity was quantified by two metrics, both based on the time series of recorded probe forces. First, Fuselage Completion Time (FCT), defined as the time between starting the first hole to ending the last hole in the fuselage. Second, the Time Between Rungs (TBR), or the time spent transitioning between rungs, and calculated by subtracting rung completion times from a FCT, for a given fuselage. Quality was assessed using two metrics. First, the number of defective holes for the fuselage (FDH) was obtained. A hole was considered defective if at least one error was made. Second, the total number of errors for the fuselage (FTE) was obtained.

Ergonomic risk can be assessed using diverse approaches that can be broadly classified into three categories. They are, in order of increasing precision of the collected data and invasiveness, self-reports, observational methods, and direct measurements (David, 2005). Despite the relative ease of use, self-reports (e.g., questionnaires and interviews) can have low validity (Wiktorin et al., 1993) and reliability (Burdorf & Laan, 1991). In observational methods, workers are observed and a systematic approach used to classify pre-specified risk factors (e.g., posture or force). These approaches are different in the factors they evaluate, and different in the validity and intra-observer and inter-observer repeatability (for recent reviews on observational methods, see NIOSH, 2014 and Takala et al., 2010). Direct measurements require more expertise and appropriate technology, yet generally yield more accurate and valid measures.

The purpose of an investigation and the nature of the evaluated work determine the appropriate assessment method(s). Here, two observational methods were used, the Rapid Upper Limb
Assessment (RULA: McAtamney & Corlett, 1993) and the Strain Index (SI: Moore & Garg, 1995). Direct measures were used, where possible (i.e., direct measures of posture using the inertial motion capture system). Because the two methods differ in their approach and emphasis (Drinkaus et al., 2002), both were used to provide a broader assessment. Two cross-sectional field studies have shown an association between higher RULA scores and self-reported musculoskeletal disorders (Shuval & Donchin, 2005; Breen et al., 2007). Also, two laboratory studies found an association between RULA scores and discomfort (McAtamney & Corlett, 1993; Fountain, 2003). The inter-observer repeatability of RULA was found to be good (McAtamney & Corlett, 1993), though there is insufficient evidence concerning its intra-observer repeatability (Takala et al., 2010). Several retrospective studies have clearly demonstrated an association between SI scores and upper limb disorders (Moore & Garg, 1995; Moore et al., 2001; Rucker & Moore, 2002; Spielholz et al., 2008). The inter-observer and intra-observer repeatability of SI have been reported to be moderate to good (Stevens et al., 2004; Stephens et al., 2006).

RULA and SI risk scores at both the rung level and fuselage levels were obtained as estimates of ergonomic risk (note, fuselage-level risk was the sum of risk scores at the rung level). The following discussion is a brief summary of the two methods, and the reader is referred to the original papers cited above for more complete details.

RULA was developed to evaluate musculoskeletal loads resulting from posture, forces exerted, and muscle use. It yields a “grand score” from combining estimated risk in two body regions, specifically neck/trunk/legs (score D, abbreviated as NTLS) and upper arm/lower arm/wrist (score C, abbreviated as ULS standing for upper limb score). For each body part, there are different risk
classifications. For example, a risk value of 4 is added to the upper arm score when the upper arm is flexed more than 90° from neutral. And, this risk is increased by one when the upper arm is abducted. ULS and NTLS are achieved using tables that combine the risk values for the specified body parts and by adding the muscle use and force scores (discussed below). Because this method evaluates each side of the body separately, we analyzed the side that was used in the task.

In the original RULA, there are risk factors described but often with little specification. For example, the method requires increasing the risk on the upper arm if it was abducted, without specifying a threshold for the magnitude (e.g., degrees) of abduction. We attempted to increase the specificity of the method by adopting the following missing thresholds:

1. Wrist ulnar and radial deviations were considered “risky” when they exceed 14.5° and 21.8°, respectively, from neutral. An earlier study indicated that these angles were set to protect 75% of some healthy participants from reaching a critical pressure (30 mmHg) for nerve injury in the carpal tunnel (Keir et al., 2007).

2. Wrist pronation and supination were considered “risky” when the deviation exceeded 45° in either direction from neutral. This was also used because carpal tunnel pressure exceed the noted critical pressure beyond these thresholds (Rempel et al., 1998).

3. In a prospective cohort study of electronics assembly workers, it was found that the percentage of a work cycle in which the shoulder is abducted more than 30° was an indicator for symptoms of cervicobrachial disorders (Kilbom & Persson, 1987). Therefore, we considered abduction “risky” when this threshold was exceeded.
Concerning the muscle use and force scores, a score of 1 needs to be added to both ULS and/or NTLS when the action was repeated more than 4 times per minute to account for risk from repetition (McAtamney & Corlett, 1993). This was the case for both ULS and NTLS in the simulated task, because both body regions repeated actions more than the specified threshold. For force scores, respective scores of 1 and 3 were added to the ULS for the 66.7 and 111.2 N conditions, to include risk from forceful actions, and these particular scores were consistent with descriptions of the RULA method.

The SI is a semiquantitative method that was developed from principles of biomechanics, physiology, and epidemiology, and focuses on the risk of distal upper extremity (elbow, forearm, wrist, and hand) disorders. Demands of an analyzed task are evaluated by six task variables: intensity of exertion, duration of exertion, efforts per minute, hand/wrist posture, speed of work, and duration of task per day. Each of the six variables is assigned a multiplier based on the associated risk. The SI “grand risk” score results from multiplying the six multipliers. Intensity of exertion is the most influential variable.

Participants rated physical effort in the hand/wrist area at each of the three rung levels after each trial using the Borg 10-point scale (Borg, 1990). The intensity multiplier was derived from these ratings, as described in the original SI paper (Moore & Garg, 1995). For the duration of exertion, the multiplier was derived based on the mean duty cycle in each rung (duty cycle = exertion duration divided by cycle time). It should be noted that participants varied here in their response speed to the completion tone. Exertion duration extends as long as the target force level was exceeded. The efforts per minute (or frequency) multiplier was assigned based on the mean cycle
time in each rung. The hand/wrist posture (of the dominant side that was used in the task) multiplier was assigned according to ranges specified in the method for wrist extension/flexion angle and ulnar deviation angle. These angles were obtained from the motion capture system. Based on the original SI method description, the task speed of work was not considered to be a risk-amplifying factor. For the duration per day variable, we assumed the task was performed for a consistent 4 hours per day.

When evaluating tasks with a variable physical exposure, the authors of the two methods proposed assessing “peak” stresses, as they are potentially responsible for exceeding injury thresholds. Based on this, we evaluated the 90th percentile angles within each rung for RULA variables and the hand/wrist pasture variable in SI; this approach avoided peak values which were likely influenced by noise.

Perceptual measures were obtained using a questionnaire administered after each condition was completed. Global ratings of perceived exertion (RPE) were obtained using the 20-point Borg scale (Borg, 1982), and participants were asked to use this scale to provide an overall evaluation of the condition just completed. Ratings of perceived discomfort (RPD) experienced in the hand/wrist, upper arm, and shoulder were also obtained, using Borg’s 10-point scale (Borg, 1990). To help conceptualize the latter scale, as well as to normalize ratings over the full range, participants performed a static endurance task in the training session, and provided RPDs every 5 seconds until they reach maximum discomfort (Sood et al., 2007; Rashedi et al., 2014). Using the same scale, participants rated the level of physical effort required in hand/wrist area for each of the three rung levels (high, middle, low).
After completing a fuselage, participants were asked to rate the extent to which they believed that the condition they just completed (i.e., the adjustability and force levels) supported their efficiency (ratings of perceived efficiency, RPEF), and accuracy (ratings of perceived accuracy, RPA). These ratings were done using two 10-cm visual analogue scales (VASs). RPEF and RPA were obtained as the location (in cm) of the lines drawn between the left (minimal extent) and right (labeled: maximal extent) verbal anchors. These ratings were used to explore the relationship between perceived and actual performance in the examined conditions.

**Statistical Analyses**

Separate $2 \times 3$ repeated measures analyses of variance (ANOVA) were used for each dependent variable to assess the effects of drilling force and adjustability. Presentation order for the latter was included as a blocking variable. All statistical tests were considered significant when $p<0.05$. Post-hoc comparisons were performed using Tukey’s HSD test. Partial $\eta^2$ was used to determine effect sizes. To meet parametric model assumptions, appropriate transformations of the dependent variables were used. The association between ergonomic risk and performance was assessed both across and within conditions. For the former, simple linear regression models were developed to assess the correspondence between each performance metric (e.g., FCT) and each risk metric (e.g., RULA score). Coefficients of determination ($r^2$) were used to evaluate the level of correspondence. Within conditions, a qualitative comparison was done between mean performance and mean ergonomic risk in each condition (across the two force levels).
Results

Statistical results are summarized in Table 1, and subsections below expand on these.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Transformation</th>
<th>Condition</th>
<th>Force</th>
<th>Force × Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCT</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.774)</td>
<td>(0.323)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>TBR</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>0.483</td>
<td>0.803</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.526)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>FDH</td>
<td>--</td>
<td>0.0003</td>
<td>0.955</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.175)</td>
<td>(0.0004)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>FTE</td>
<td>Square root</td>
<td>0.0004</td>
<td>0.577</td>
<td>0.524</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.170)</td>
<td>(0.003)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>RULA</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.706)</td>
<td>(0.683)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>SI</td>
<td>Square root</td>
<td>0.004</td>
<td>&lt;0.0001</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.124)</td>
<td>(0.355)</td>
<td>(0.0315)</td>
</tr>
<tr>
<td>RPE</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.917</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.470)</td>
<td>(0.578)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>RPD shoulder</td>
<td>--</td>
<td>0.0002</td>
<td>&lt;0.0001</td>
<td>0.976</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.185)</td>
<td>(0.205)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>RPD upper arm</td>
<td>--</td>
<td>0.0002</td>
<td>&lt;0.0001</td>
<td>0.442</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.186)</td>
<td>(0.289)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>RPD hand/wrist</td>
<td>--</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
<td>0.801</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.199)</td>
<td>(0.338)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>RPEF</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.636</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.639)</td>
<td>(0.214)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.636</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.605)</td>
<td>(0.177)</td>
<td>(0.011)</td>
</tr>
</tbody>
</table>
**Productivity-related Metrics**

Fuselage completion time (FCT) was affected by Condition and Force, with no significant interactive effects (Figure 4a). High adjustability led to the shortest FCTs, with the longest times required in the Some condition. The higher Force condition led to longer FCTs, and this effect was fairly consistent between conditions (i.e., no interaction effect). A similar pattern of results was found for time between rungs (TBR), but without a significant effect of Force (Figure 4b).

Figure 4. Productivity-related performance results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017).

**Quality-related Metrics**

Both the total number of defective holes (FDH) and errors (FTE) significantly differed between Conditions, but were not affected by Force or the interaction (Figure 5). Both FDH and FTE were lowest in the High condition.
**Risk Metrics**

For RULA scores, there was a significant Condition x Force interaction effect. RULA scores were lowest in the High condition and with lower force, with the influence of force being most substantial in the High vs. the other two conditions (Figure 6).

Figure 5. Quality-related performance results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017).
Figure 6. RULA risk results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017).

SI scores differed between Conditions and Force levels (Figure 7). SI scores were lowest in the High condition and lower at the 66.7 N level.
Figure 7. Strain Index risk results. Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017).

**Perceived Exertion and Discomfort**

All measures of perceived exertion and discomfort were significantly affected only by the main effects of Conditions and Force, with no significant interactive effects. Each was smallest in the High condition and smaller at the lower force (Figure 8).
Figure 8. Ratings of perceived exertion (RPE) and discomfort (RPD). Light and dark bars indicate the 66.7 and 111.2 N force levels, respectively. Capital letters indicate groupings obtained from pairwise comparisons between adjustability conditions, and error bars indicate 95% confidence intervals. Reprinted with permission from Alabdulkarim et al., (2017).

**Relationships between Risk and Performance**

Results from fitting simple regression models to the risk and performance metrics are summarized below (Table 2). Figures 9 and 10 show graphical representations of the significant models for RULA and SI risk metrics, respectively.
Table 2. $P$ values, $r^2$, and slopes of fitted lines for the simple linear regression models of risk-performance relationships. Bold font indicates significant slopes ($p<0.05$). Note, data from all three conditions. Reprinted with permission from Alabdulkarim et al., (2017).

<table>
<thead>
<tr>
<th></th>
<th>RULA</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P$ value ($r^2$)</td>
<td>Slope (95% CI)</td>
</tr>
<tr>
<td>FCT</td>
<td>$&lt;0.0001$</td>
<td>0.0236</td>
</tr>
<tr>
<td></td>
<td>(0.251)</td>
<td>(0.0157:0.0314)</td>
</tr>
<tr>
<td>TBR</td>
<td>0.0004</td>
<td>0.0196</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.0090:0.0301)</td>
</tr>
<tr>
<td>FDH</td>
<td>0.0095</td>
<td>0.3739</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.0934:0.6543)</td>
</tr>
<tr>
<td>FTE</td>
<td>0.0211</td>
<td>0.1696</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.0259:0.3132)</td>
</tr>
</tbody>
</table>

Figure 9. Relationships between performance measures and RULA scores. Note, several transformations were used to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017).
Figure 10. Relationships between performance measures and SI scores. Note, a square root transformation on FTE data was needed to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017).

Graphical representations of the relationships between mean performance and RULA and SI scores (across the two force levels) are provided in Figure 11 and Figure 12, respectively.

Figure 11. Mean (95% confidence intervals) values of performance measures and RULA scores. Reprinted with permission from Alabdulkarim et al., (2017).
Figure 12. Mean (95% confidence interval) values of performance measures and SI scores.

Reprinted with permission from Alabdulkarim et al., (2017).

Perceived versus Measured Performance

Both ratings of perceived accuracy (RPA) and efficiency (RPEF) were largest in the High condition, followed by None and Some, and both were significantly higher when exerting 66.7 N vs. 111.2 N. Simple linear regression models of the relationship between measured and perceived performance are summarized in Table 3 and illustrated in Figure 13. While there were significant relationships between perceived and measured performance, the level of correspondence was relatively low.
Table 3. Results from linear regression modeling of the relationship between perceived and measured performance. Note that FTE data were square-root transformed to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017).

<table>
<thead>
<tr>
<th></th>
<th>RPA</th>
<th></th>
<th>RPEF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P$ value</td>
<td>Slope (95% CI)</td>
<td>$P$ value</td>
<td>Slope (95% CI)</td>
</tr>
<tr>
<td>FDH</td>
<td>0.0221</td>
<td>-0.7163 (-1.3277 to -0.1048)</td>
<td>0.0005</td>
<td>-7.633 (-11.813 to -3.452)</td>
</tr>
<tr>
<td>FTE</td>
<td>0.0305</td>
<td>-0.34502 (-0.65697 to -0.03307)</td>
<td>0.0468</td>
<td>-0.9058 (-1.798 to -0.0132)</td>
</tr>
</tbody>
</table>

Figure 13. Relationships between perceived and measured performance assessed using simple linear regression models. FTE data were transformed to meet model assumptions. Reprinted with permission from Alabdulkarim et al., (2017).
**Discussion**

The purposes of this study were to 1) evaluate if causal relationships exist between task design, specifically workstation adjustability here, and both performance and ergonomic risk in the simulated drilling task; and 2) to investigate the association between these outcomes (i.e., ergonomic risk and performance). As discussed below, the primary finding from this study was that both productivity and quality significantly improved when increasing adjustability, but only when that adjustability succeeded in reducing ergonomic risk.

Regarding the relationship between task design and performance, it was hypothesized that increasing workstation adjustability will improve both productivity and quality. However, the results suggest that the influence of increasing workstation adjustability depends on the level of that added adjustability. The use of High adjustability significantly and substantially improved both quality and productivity. Yet, despite the expectation that Some adjustability would reduce postural extremes, the Some condition reduced productivity (both metrics) and did not have a significant influence on either of the quality metrics (although there was evidence to support a possible deterioration in FDH between the Some and None conditions). There are several sources that may have contributed to the reduced productivity with the Some adjustability. In particular, participants were required to manually move the stools whenever and wherever they decided to use them. That movement time included time spent optimizing the stool’s location (i.e., horizontal and vertical distance in relation to the fuselage). Also, participants sometimes needed to re-locate the stool after attempting to drill a hole. While we attempted to minimize such effects, through practice and reminders to focus on stool placement, it might have been more effective to more precisely guide stool movement through, for example, lines on the workstation floor. In addition,
several participants felt less stable on the stool when completing the upper rungs (Figure 3c) and that this instability was a hindrance to performance in that condition.

Quality appeared to be independent of Force, though completion time increased with Force. Supporting the former, Finneran and O'Sullivan (2013) found that force did not affect precision performance, unlike posture, in a precision task (Fitts-type tapping task). In a repetitive gripping task, increasing force was found to reduce self-selected duty cycle (Finneran & O'Sullivan, 2010; Finneran & O'Sullivan, 2014), consistent with the noted main effect of Force on FCT found here. Also, Potvin (2011) showed a negative exponential relationship between maximum acceptable effort and duty cycle.

Regarding the relationship between task design and ergonomic risk, it was hypothesized that increasing workstation adjustability will reduce ergonomic risk. Similar to the earlier findings for performance, the effect of increasing workstation adjustability depended on the level of that added adjustability. For both RULA and SI risk assessment methods, the High condition resulted in the lowest ergonomic risk while non-significant differences were found between Some and None. As for performance, we expected that the Some condition would reduce ergonomic risk as a result of providing “better”, or less extreme, reaches to the holes. Although the validity of the risk assessment methods used here have been supported in earlier work (e.g., Moore et al., 2001; Rucker & Moore, 2002; Fountain, 2003; Shuval & Donchin, 2005; Spielholz et al., 2008), it is possible that these methods were not sufficiently sensitive to detect the (small) magnitude of differences in ergonomic risk between the Some and None conditions. Differences in exposure in the High condition were likely detected because the impacts were more substantial.
Despite the noted lack of differences in ergonomic risk between the Some and None conditions, there is evidence that adjustability in the Some condition improved certain risk factors pertaining to distal upper extremity, as quantified by the SI. In particular, the Some condition reduced the specific risk (SI multiplier) related to hand/wrist posture, repetition, and duration of exertion by 6.2, 8.6 and 2.4%, respectively. In contrast, risk related to the intensity of exertion in the hand/wrist increased by ~25% compared to the None condition. The former improvements were offset by the latter increase in risk, since the SI method is most influenced by the intensity of exertions, as discussed earlier. Given that this critical variable was measured subjectively, though, it is possible that participants overestimated their perceived exertion because of, for example, the repeated need to move the stool and/or the feeling of instability when using the stools. Note that the intensity of exertion captures the proportion of maximum capacity (strength) required to perform a task (Moore & Garg, 1995). It was quantified subjectively here, because use of objective methods, such as directly measuring muscles activation levels through electromyography, was considered infeasible in the current study. Bao et al. (2006), however, found a poor correlation between self-reports, ergonomist-estimates, and direct measurements (using instrumentation) of exertion intensity. While we believe that the calibration task used here improved the validity of our measures (using Borg’s 10-point scale), the possibility of overestimation remains present. Supporting this possibility, Spielholz et al. (2001) found that self-reported hand force, using a visual analogue scale, consistently overestimated force levels compared to both observational and direct measurement methods. For estimating gripping forces, McGorry et al. (2010) indicated that the correlation between the noted three methods is best for simple tasks. DiDomenico and Nussbaum (2008) found Borg CR10 ratings to be sensitive to changes in physical workload for lifting tasks, and this sensitivity was not altered by the addition of mental workload. This also suggests that the
mental workload due to precision requirements in the current simulated task likely did not substantially affect perceived exertion.

A subsequent analysis on postural extremes (90\textsuperscript{th} percentile angles) found inconsistent effects of the added adjustability in the Some condition, compared to None. For example, neck and trunk postures generally improved, but upper and lower arm postures had larger extremes. This inconsistency can be considered a manifestation of the potential tradeoff between body parts in tasks (re)design (Marras, 2006). Generally, it may not be feasible to reduce postural extremes for all body parts in a given (re)design, particularly for a relatively complex task.

The association between ergonomic risk and performance was assessed at two levels, first across conditions and second between the three conditions. For the former, the majority of the derived regression models (Table 2) showed a significant, indirect relationship between ergonomic risk and performance (Figures 9 and 10). Because participants responded differently to the different adjustability conditions (cf. Figure 3a and b), these models have the advantage of looking at each data point, irrespective of the adjustability condition. While RULA scores were inversely associated with both productivity and quality, SI scores were inversely associated only with quality metrics (FTE and FDH). These findings suggest that reducing risk in the distal upper extremity, as quantified by SI in the simulated task, should be associated particularly with quality improvement; reducing risk in the upper limb overall, as quantified by RULA, should be associated with productivity and quality improvements. It should be noted, though, that there was relatively large variability that was not explained by the risk metrics, highlighting the need to investigate other variables that influence performance, ergonomic risk, and their inter-relationships.
Between the three conditions, there was some support for the hypothesized inverse relationship between ergonomic risk and performance (Figures 11 and 12). The lowest ergonomic risk in the High condition was associated with the highest productivity and quality, though the latter was only evident for FTE, and not for FDH. Further, the lack of difference in ergonomic risk between the None and Some conditions was associated with a lack of difference in quality (both metrics), though contrary to our hypothesis there was a deterioration in productivity (both metrics). These findings indicate a need to test the efficacy (in terms of risk reduction and performance improvement) of any (re)design prior to actual implementation.

Our findings are generally consistent with results reported in the literature. In a field study, Loo and Yeow (2015) tested two physical ergonomics interventions for a brazing task. An inverse association was evident between the three RULA scores (pre-intervention, intervention 1, and intervention 2) and both productivity and quality. However, it should be noted that RULA assessments were done on only one operator in that study. In a metal factory, a physical ergonomic intervention was applied based on RULA risk evaluations and this reduced the rejected and reprocessed parts by 45 and 22%, respectively (González et al., 2003). Subsequent to a redesign based on principles of work design and ergonomics, productivity and quality in a simulated drill press workstation was improved by 22 and 50%, respectively (Das et al., 2007). These studies support our findings that task (re)design can improve performance when it reduces ergonomic risk, as well as the inverse relationships between risk and performance. In assembly tasks, Ivarsson and Eek (2015) provided some support for the inverse relationship between ergonomic risk and quality. However, it should be noted that a company-specific assessment tool quantified ergonomic risk, and that its validity and reliability were not tested. Two studies in the automotive industry also
showed that quality was worse for tasks that had higher ergonomic risk, as estimated by ergonomists (Falck et al., 2010; Falck et al., 2014). However, these later three studies provided less support for the noted inverse association because comparisons were made across different tasks.

In addition to performance and injury risk, different levels of adjustability also affected perceptual responses (RPE and RPD). Such perceptual measures have been suggested as indicators of muscle fatigue (Hummel et al., 2005; Sood et al., 2007), and thus the current results suggest that fatigue was minimized when the added adjustability reduced ergonomic risk (i.e., the High condition). This finding was consistent for the majority of RPD variables. Minimizing discomfort may be an important goal, since its presence may indicate the risk of future development of a WMSD (Madeleine, 2010). Across the three conditions, a significant positive linear association was observed between RPE and each of FCT, TBR, and FDH, and between RPDs (all three RPDs) and each of FCT and TBR (results not reported). A similar regression model was reported between discomfort and self-selected duty cycle for a repetitive gripping task (Finneran & O'Sullivan, 2010). This is generally consistent with the concept that more rest is required for more fatiguing tasks (Rhomert, 1973). Although studying the effect of fatigue on performance was beyond the scope of this study, these findings suggest an inverse association between fatigue and performance. However, mixed evidence has been reported for the effect of fatigue on performance (Mehta & Agnew, 2010).

Although this was not a fundamental research question examined here, we found a significant association between perceived and measured performance. This association suggests that
individual perception could be helpful to quantify performance, particularly when objective measures may not be feasible or practical. However, firm conclusions are not warranted here regarding this approach, especially given the large residual variability present (cf. Figure 13), until more evidence is available regarding validity and reliability (for more on self-reported productivity, see Prasad et al., 2004).

There are a few limitations in this study that should be noted. While the use of a controlled experimental design ensured good internal validity, the level of external validity is unknown. For example, vibration exposure is a known WMSD risk factor (Armstrong, 1986) and present in actual drilling, yet it was not simulated here. Additionally, actual drilling operations produce torques that intensify as as cutters break through the exit surface of a drilled material. These forces are reacted to by the operator in manual drilling tasks, but our simulation did not account for this effect. The current results may thus not necessarily generalize to actual drilling, nor it is clear if the results regarding adjustability will generalize to other tasks. The lack of differences found between the None and Some conditions, along with the unforeseen limitations inherent in the latter condition noted above, limited our ability to address the “shape” of the relationships between adjustability and performance/risk. Because recruiting experienced drilling workers was considered infeasible in the experiment, we trained novices on the task and used a repeated measures design. However, it is unknown if experienced workers would give the same pattern of major results.

We limited our definition of ergonomic risk to “physical” exposure, which is an influential determinant of ergonomic risk (e.g., Winkel & Mathiassen, 1994) and perhaps the most easily
modifiable in practice. However, it should be noted that WMSDs are multi-factorial in nature (van der Beek & Frings-Dresen, 1998). In addition to physical aspects, earlier research has identified important individual (Armstrong et al., 1993) and psychosocial/organizational (Devereux et al., 2002) occupational risk factors for WMSDs development. Consideration of the latter factors, and their possible interactive effect (e.g., Widanarko et al., 2014, 2015), however, was beyond the scope of this study.

In summary, increasing the adjustability in a simulated drilling task improved performance (quality and productivity), but only when that adjustability succeeded in also reducing ergonomic risk. Along with performance improvements, the “successful” adjustability also reduced discomfort in several body parts and perceived exertion. Across the examined adjustability conditions, several significant inverse relationships were found between ergonomic risk and performance. These results provide support that ergonomic “improvements” can have dual benefits in some cases, by enhancing performance in parallel with reducing injury risk. Future work is needed, however, involving a wider range of tasks, task conditions, and participants, to assess the generality and applicability of these results.

References


Devereux, J., Vlachonikolis, I., & Buckle, P. (2002). Epidemiological study to investigate potential interaction between physical and psychosocial factors at work that may increase the risk of symptoms of musculoskeletal disorder of the neck and upper limb. *Occupational and Environmental Medicine, 59*(4), 269-277.


Chapter 3: Influences of Different Exoskeleton Designs and Tool Mass on Physical Demands and Performance in a Simulated Overhead Drilling Task

Abstract

We compared different passive exoskeletal designs in terms of physical demands (maximum acceptable frequency = MAF, perceived discomfort, and muscular loading) and quality in a simulated overhead drilling task, and the moderating influence of tool mass (~2 and ~5 kg). Three distinct designs were used: full-body and upper-body exoskeletons with attached mechanical arms; and an upper-body exoskeleton providing primarily shoulder support. Participants \( n = 16 \), gender-balanced) simulated drilling for 15 minutes to determine their MAF, then maintained this pace for three additional minutes during which the remaining outcome measures were obtained. The full-body/upper-body devices led to the highest/lowest MAF for females and the lowest quality. The shoulder support design reduced peak shoulder muscle loading but did not influence quality. Differences between exoskeleton designs were largely consistent across the two tool masses. These results may be helpful to (re)design exoskeletons that reduce injury risk and improve performance.

Keywords: Exoskeleton, wearable assistive device, quality, performance, psychophysics
Introduction

Overexertion is a leading exposure or event resulting in occupational injuries or illnesses, which remain both prevalent (BLS, 2015) and costly (Liberty Mutual Research Institute for Safety, 2014). Overexertion injuries are typically considered to result when task demands exceed worker capacity (Chaffin et al., 2006). When occupational tasks are physically demanding, one intervention approach is to use (or attempt) automation as an engineering control. However, this approach can be infeasible and/or unjustifiably expensive. An alternative is assistive devices, which include mechanical manipulators and, especially recently, wearable exoskeletons. The focus here is on the latter, as they may be applicable for a wide range of tasks and might require only limited changes, if any, to the workplace (and thus can be less expensive). Supporting the potential of exoskeletons in manufacturing, de Looze et al., (2017) found that exoskeletons are preferred to maintain human flexibility and creativity, as automation can be challenging especially when frequent changes in production are expected (e.g., to product types or task locations).

Exoskeletons are designed to enhance user capacity (de Looze et al., 2015), and thus can improve the balance between task demands and worker capacity. They can broadly be classified as passive or active devices (de Looze et al., 2015). Active devices (e.g., Kobayashi et al., 2007) use one or more actuators to augment human strength with different technologies, such as pneumatic muscles, hydraulics, or electric motors (Gopura & Kiguchi, 2009). Such augmentation in completely passive devices (e.g., Abdoli-E et al., 2006), however, comes from springs, dampers, or materials that can store energy generated by user motion and discharge it effectively (de Looze et al., 2015). While active devices can provide a higher degree of augmentation, they are usually heavier and more expensive. The focus of this study, therefore, was on passive exoskeletal designs,
particularly given their broad potential for industrial application, the fact that several new technologies are entering the market, and that their design can be challenging from an ergonomics perspective. Exoskeletons can also be categorized by the body part(s) they are design to support (Lee et al., 2012; de Looze et al., 2015), such as the upper body (e.g., Kobayashi et al., 2007), lower body (e.g., Kim et al., 2009), or full body (e.g., Toyama & Yonetake, 2007) devices. Finally, exoskeleton designs can fundamentally differ in their approach to transferring external loads (e.g., from the shoulder to the waist/hip or to the ground).

Wearable exoskeletons have been investigated thus far mainly for rehabilitation/medical purposes, such as to support weak, injured, or disabled patients (Viteckova et al., 2013), or for military applications (Lee et al., 2012). For industrial use, a recent review highlighted the potential for these devices and showed that they can reduce physical demands considerably (de Looze et al., 2015). However, while they can reduce demands on specific body parts, they can also have unintended consequences such as increasing loading and/or discomfort on “other” regions of the wearer’s body. In a simulated assembly task with the trunk in a forward-bending posture, for example, a passive exoskeleton reduced loading on the back and legs but increased discomfort in the chest (Bosch et al., 2016). Discomfort normally occurs in areas where an exoskeleton contacts the wearer’s body. Even if a biomechanical advantage was confirmed, eliminating/minimizing such discomfort is fundamental, particularly because discomfort can prevent workers from using these devices (de Looze et al., 2015). In a simulation of overhead work, using an upper limb passive exoskeleton reduced loading on the upper arm and shoulder but increased it slightly on the low back (Rashedi et al., 2014), further highlighting the possibility of inconsistent influences of exoskeletons.
Furthermore, wearing exoskeletal devices can modify working techniques, which may (partly) explain reported changes in physical loading. For example, when lifting with the personal lift-assist device (PLAD), participants changed their lifting technique more towards a squat-like method, by reducing lumbar and thoracic flexion and increasing hip and ankle flexion (Sadler et al., 2011). Such behavioral changes may explain the increased leg muscle activity found when using passive exoskeletons (de Looze et al., 2015).

Different passive exoskeletal design approaches were compared here using a psychophysical approach. Ideally, these design approaches should be tested in prolonged sessions to reflect anticipated use in industry (i.e., enhance external validity). However, such extended sessions can be expensive for experimenters and tiring for participants. Although relying on the assumption that subjective limits are below injury thresholds, psychophysical methods offer a realistic/holistic evaluation approach that can require less cost and time (Snook, 1999). Specifically, the holistic consideration when setting psychophysical judgments is of great interest here, since exoskeletal designs can have different unintended consequences (e.g., regional discomforts). Detecting such consequences is considered critical, as they can hinder acceptance in the workplace. From comparing different exoskeletal design approaches, generic exoskeletons design recommendations can be drawn to facilitate achieving the potential dual benefits intended from using the devices (i.e., reduce injury risk and improve performance). Because of the growing trend to use exoskeletons in industry, where economic benefits are often a key input to decision making, and because they have mainly been assessed in terms of reducing physical demands, it is critical to examine their efficacy in terms of both improving performance and reducing injury risk.
To the best of our knowledge, this is the first study investigating multiple exoskeleton design approaches while simultaneously considering the noted two dimensions of effectiveness. This study was designed to address the following hypotheses, in the context of a simulated overhead occupational task (drilling): 1) different exoskeleton designs will lead to different psychophysical limits (assessed here by the maximum acceptable frequency, or MAF) and differences in task performance (assessed here by quality); and 2) tool mass will moderate the influence of different exoskeleton designs on MAF and quality.

Methods

Participants
A total of 16 individuals (gender-balanced) completed the study, and were recruited from the University and local community. Prior to any data collection, all participants completed an informed consent procedure that was approved by the Virginia Tech Institutional Review Board. They all reported having no current or recent (past year) musculoskeletal problems, having normal or corrected-to-normal vision, and being right handed. Participant mean (SD) age, stature, and body mass were 23.0 (2.1) years, 175.8 (5.3) cm, and 78.4 (12.8) kg, respectively, for males; and 23.0 (1.9) years, 168.7 (2.5) cm, and 66.6 (12.0) kg, respectively, for females.

Task Description
A repetitive, overhead drilling task was simulated in a laboratory. This task was intended to be representative of common demands in manufacturing, specifically involving both strength and precision. This task was also a simplified version of the one presented by Alabdulkarim et al., (2017). Representing a longitudinal stringer, a single rung was constructed (5.08 x 5.08 cm hollow
rectangular stock, 121.9 cm long). The rung had six, evenly-spaced holes that were 1.6 cm in diameter, and it was attached to vertical structures that allowed for changes in height. Working height in the experiment was set relative to each individual’s anthropometry, at a “low” overhead level, and this was specified when the shoulder and elbow were flexed at 90° as described in Sood et al. (2007). This particular height was chosen to represent tasks requiring arm elevation, a situation known to increase shoulder injury risk (e.g., Punnett et al., 2000; Björkstén et al., 2001; Van Rijn et al., 2010). And, all three of the exoskeletal designs examined (see below) are relevant to this type of task.

A commercial drilling tool (mass= 2.26 kg; Model # R71111, Ridgid Company) with two handles was used (Figure 14), and a load cell (Interface, SML-100, Scottsdale, AZ) and a probe, simulating a drill bit, were attached to the tool. Drilling force (as measured by the load cell) had to be at least 66.7 N and needed to be sustained for an accumulated 2.5 seconds to complete a single hole. During a single simulated drilling effort, any deviations more than 2° of normal resulted in a contact between the probe and the rung, and was captured as an “error.” Participants received three auditory feedback types: when the required force level was reached, when an error was made, and when the accumulated 2.5 seconds was completed. These simulation parameters were considered representative of a drilling task common in aircraft manufacturing (for more details on the task simulation, see Alabdulkarim et al., 2017). Participants repeated drilling the rung by moving back and forth along the rung, repeatedly completing the set of six holes in opposite order for each set. The time at which each hole was started, and thus the task pacing, was specified using a metronome (a fourth auditory feedback type). When finishing a given hole, participants were
free to lower their arms to facilitate recovery and to put the tool down on a table in their proximity (if applicable).

**Independent Variables**

Two within-subject independent variables were manipulated in the experiment: Exoskeleton (Exo) and Tool Mass ($T_{mass}$). For the former, three commercially-available passive exoskeletons were used, and were selected to represent designs with fundamentally different approaches to reducing physical demands. A control condition was also included (i.e., without assistive device). Specific exoskeletons were: 1) the FORTIS™ (Lockheed-Martin), a full-body exoskeleton that transfers tool loads to the ground (Figure 14a); 2) the SuitX™ ShoulderX™, an upper body vest that supports the arms (Figure 14b); and, 3) an exoskeletal vest (Fawcett Exovest™; The Tiffen Company, Hauppauge, NY, USA) with a mechanical arm (zeroG²; Equipois Inc., Manchester, NH, USA) attached to it (Figure 14c), which supports a tool and transfers loads to the hip/waist. Subsequently, these conditions will be abbreviated as “Full”, “Shl”, and “Arm”, respectively, and “No” for the control condition. To attach the tool to the mechanical arms used in the Full and Arm conditions, a gimbal and mounting system were used (Saturn-2, medium, 135° configuration; Equipois Inc., Manchester, NH, USA). Including the mechanical arm and mounting system weights (if used), the Shl, Arm, and Full weights were ~ 6, 10.3, and 19-28 kg, respectively. The Full weight is reported as a range, as this depended on the number of counterbalancing weights used.

Because the efficacy of these devices was expected to be moderated by the mass of the tool used, two tool masses (Lighter: ~2 kg and Heavier: ~5 kg) were examined. Two objectives were considered for selecting these specific masses. First, it was desired that the lighter tool also be
heavy enough to justify using an assistive device, especially under more physically-demanding requirements (e.g., high repetition rate and/or non-neutral postures). Second, we wanted the heavier tool mass to be such that most volunteer participants could complete the experimental tasks under the control (unassisted) condition.

Figure 14. Illustrations of participants using: (a) the FORTIS™ Exoskeleton (Full); (b) the SuitX™ ShoulderX™ (Shl); and (c) the Fawcett Exsovest™ with a zeroG² mechanical arm (Arm).

Procedures and Data Collection

A repeated-measures design was used, in which participants completed the simulated drilling task in all eight experimental conditions (combinations of Exo and T\textsubscript{mass}). The study was completed in two testing sessions, one for each tool mass, and each session required ~3.5 hours. The order of exposure to tool mass was counterbalanced (i.e., half of the participants started with each tool mass), and 4×4 Balanced Latin Squares were used to counterbalance the order of exoskeleton (or control) conditions. Sessions were separated by at least two days to minimize any effects from residual muscle fatigue.
In the first session, participant’s anthropometric “fit” to the three exoskeletons was determined initially. For each exoskeleton, the respective manufacturer’s manuals were followed to determine and adjust this fit. All 16 participants had acceptable fit to the three exoskeletons. Then, Borg’s 10-point scale (Borg, 1990) was explained to participants, which was used later in the experiment to report ratings of perceived discomfort (RPDs) experienced in several body parts. To help them normalize these ratings over the whole range and better conceptualize this scale, participants performed a static endurance task at the beginning of the session, and provided RPDs every 5 seconds until they reached maximum discomfort (Sood et al., 2007; Rashedi et al., 2014; Alabdulkarim et al., 2017).

Subsequently, surface electromyography (EMG) electrodes were placed over the following four muscles bilaterally (note: R and L will be used hereinafter to denote the right and left side, respectively): the anterior and middle deltoids (AD and MD), triceps brachii (TB), and iliocostalis lumborum pars lumborum (ILL). These specific muscles were selected because they were accessible across all conditions (i.e., some muscles were not accessible due to the structure or straps of one or more exoskeletons), and were considered major contributors during overhead work based on earlier studies (Nussbaum et al., 2001; Chopp et al., 2010; Rashedi et al., 2014). Before placing the electrodes, the skin was shaved, gently abraded, and cleaned using 70% alcohol. Following this preparation, pairs of bipolar Ag/AgCl electrodes (AccuSensor, Lynn Medical, MI, USA) with a 2.5-cm inter-electrode distance were placed on the skin surface following procedures described by Hermens et al. (1999). Using a telemetered system (TeleMyo 900, Noraxon, AZ, USA), raw EMG signals were monitored at 1500 Hz.
Participants were then asked to carefully review written instructions that described the approach to selecting MAF (see Appendix A, modified from Snook et al., 1995). These instructions also noted that productivity and quality should be considered equally important. Initial practice on the task (without assistance) was completed to ensure familiarization with the task requirements. Participants then performed three trials of a task-specific maximum voluntary contraction (MVC). In these MVCs, participants pushed upward on a load cell (AMTI MC3A-6) using a handle that was constructed to simulate the posture used when using the drilling tool without an assistive device. Participants were instructed to use mainly the upper extremity muscles. An additional three MVC trials were completed for the back extensors, using a 45° Roman chair. When positioned on the chair, the participant’s back was connected to the ground using a modified construction harness while keeping the lumbar spine flexed at approximately 45°. Participants were instructed to gradually extend their back as hard as possible, attempting to pull the fixture from the ground. Nonthreatening verbal encouragement was provided during all MVC trials and a minimum of one-minute rest separated each. Rest was provided after completing all MVCs (~10 min), and rest continued until participants reported that RPDs for all body parts had returned to baseline. After rest, participants performed the testing conditions according to the noted counterbalancing approach.

Before beginning each condition, practice was provided to (re)adjust exoskeleton fit (if used) and to remind participants of the task requirements. When practicing with each of the exoskeletons, participants were exposed to different adjustment levels and asked to select their preferred levels. For the Full and Arm conditions, this involved different stiffness levels of the mechanical arm. For the Full, participants also explored the use of different numbers of counterweights (if any)
needed. Finally, participants were exposed to both near maximum and minimum support that can be provided by the Shl (the device has knobs to control the level of support for each shoulder). Once an individual selected a level for these adjustments, it was sustained throughout the remainder of practice and during subsequent testing. Participants were encouraged more generally, when practicing, to explore different working strategies, to best benefit from the support provided by the exoskeletons and to improve their performance (productivity and quality). They were also encouraged to complete as much practice as needed.

After this practice, participants were asked to complete the repetitive drilling task for a total of 18 minutes. In the first 15 minutes, they were asked to converge to a MAF that they believed they could sustain for a two-hour shift (a common duration for job rotation schedules), without leading to numbness, pain, or excessive discomfort in any body part, and while keeping errors at a minimum. Participants started randomly at either a relatively fast (16 holes/min) or slow (2 holes/min) pace. To allow for both coarser and finer levels of adjustments, four changes to the pacing could be made by the participants, by saying “a little faster”, “faster”, or “a lot faster”; these were then used to increase the pace by 10, 20, or 30%, respectively. Slowing the pace was done in a similar manner. To allow for returning to the same pace after a specific change, a fourth adjustment could be made by saying “go back.” During the 15-minute adjustment period, an audio recording was played to brief the instructions to the participant every three minutes. Specifically, this indicated: “remember, you are to find your maximum pace that you think you can do for a two-hour shift without getting excessive discomfort and while keeping your errors at a minimum.” After the adjustment period, an additional three minutes of repetitive drilling was completed,
during which quality was assessed as described below; the pace during these final three minutes was maintained at the MAF identified at the end of the 15-minute adjustment period.

To determine the minimum adjustment duration for selecting the MAF, a pilot study was completed with eight participants (5 males and 3 females). In that pilot, participants performed the task for one hour (repeatedly drilling one rung back and forth), starting randomly from a fast or slow pace. This was done in two separate sessions, with each session involving either of two conditions assumed to be the most (No exoskeleton, with the heavier tool) or the least (Full exoskeleton, with the lighter tool) demanding among those described above. This approach, involving testing on a subset of the more extreme conditions, is similar to that used by Sood et al. (2017). Based on the observed convergence of MAF values over time, it was concluded that 15 minutes was a reasonable adjustment duration.

Rest breaks of at least 10 minutes were provided between conditions, and more rest was encouraged as needed. After completing each condition, participants immediately responded to a custom questionnaire with items regarding usability and discomfort (Appendix B); responses regarding usability were not solicited after the No condition. The second session was completed as described for the first one, except that there was no need to determine exoskeleton fit or to perform the static endurance task for the RPD scale.
Dependent Measures

Physical demands were quantified using both subjective and objective metrics. Metrics for the former were the MAF and RPDs for the examined body parts (see Appendix B for the list of body parts). MAF was the pace to which the participant converged by the end of the adjustment duration (15 minutes) and which was sustained for the additional three minutes. It was assumed here that the MAF selected in a given experimental condition reflected the physical demands imposed by that condition. Or, more specifically, that there is an inverse relationship between MAF and physical demands. Surface EMGs served as an objective assessment. Raw EMG signals were first band-bass filtered (20-450 Hz). Then, EMG root mean square (RMS, time constant 100 ms) values were calculated during the three additional minutes (at a constant MAF) and these were subsequently normalized (nEMG) to maximal RMS values obtained during the MVC trials. From the nEMG data, the 10th, 50th, and 90th percentile values were determined, and were considered representative of the static, median, and peak loadings in a given condition, respectively (cf. Jonsson, 1982).

The quality dimension of task performance was quantified based on the total number of “errors” generated during the additional three-minute periods after adjusting MAF (again, during this duration the pace was no longer adjustable). From the noted questionnaire (Appendix B), six usability measures were obtained (e.g., ease of use) in the first question, using Likert-type scales (1= strongly disagree, 5= strongly agree). Finally, after completing each session participants ranked the four Exo conditions from least to most physically demanding, and any expressed concerns/suggestions regarding the usability and design of the three exoskeletons were documented.
Statistical Analyses

Separate 2×4 repeated-measures analyses of variance (ANOVAs) were used for each dependent variable to assess the effects of Exo and T_{mass}, and with Gender as a blocking variable. Starting pace and the presentation order of Exo conditions were explored as additional potential blocking variables, and stature and strength (as measured in the task-specific MVC) were tested as possible covariates. Starting pace was not found to significantly influence any of the dependent measures. However, within-session presentation order (of Exo) significantly influenced three measures (TB-R 50%, AD-L 10%, and the lower leg and foot RPD). Thus, Order was maintained as a blocking variable for only these specific measures. Stature and strength were both significant covariates for MAF, and the latter was significant for Errors, MD-R 90%, and ILL-R 10, 50, and 90% variables. Particularly because these covariates, albeit significant, did not change the pattern of statistical results for the noted measures, they were dropped from the final models to ease interpretation.

Parametric model assumptions were assessed, and data transformations were used to meet model assumptions when needed. Post-hoc comparisons were performed using Tukey’s HSD and simple-effects tests, and partial eta-squared ($\eta_p^2$) was used to quantify effect sizes. To quantify the association between MAF and the other metrics of physical demands (RPDs and EMG peak, median, and static values), simple linear regression models were developed, and coefficients of determination ($r^2$) were used to evaluate the levels of correspondence.

A mixed-factor, multivariate analysis of variance (MANOVA) was used to evaluate the main and interactive influences of Exo, T_{mass}, and Gender on the six usability metrics (i.e., question #1 in the questionnaire). Only the main effects of Exo and Gender were found to be significant (with
Wilk’s Lamda of 0.568 and 0.845, and $P$ values = $<0.0001$ and 0.034, respectively); subsequently, separate repeated-measures ANOVAs were used to assess the main effects of Exo and Gender on each usability measure. All ANOVA analyses were completed in JMP Pro 12 (SAS, Cary, NC). RStudio (version 1.0.136) statistical software (RStudio, Inc., Boston, MA) was used to: 1) perform the MANOVA, using the `manova` function from the `stats` package (R Core Team, 2017); and 2) to analyze significant interactive effects, using the `lsmeans` function from the `lmerTest` package (Kuznetsova et al., 2016), along with the `lmer` function from the `lme4` package (Bates et al., 2014). All statistical tests were considered significant when $p<0.05$.

**Results**

**Maximum Acceptable Frequency and Quality Metrics**

MAF was significantly influenced by Exo and $T_{\text{mass}}$, and the Exo x Gender interaction (Table 4). For the latter interaction, MAF significantly differed between Exo conditions only for female participants (Figure 15a). Specifically, for female participants the *Full* condition led to a significantly lower MAF than all other conditions, with the *Arm* leading to a higher MAF (though not significantly) compared to both the *Shl* ($p=0.1118$) and *No* conditions ($p=0.1295$). For the main effect of $T_{\text{mass}}$, the heavier tool reduced MAF by $\sim25\%$. Errors were significantly influenced by Exo and the Gender x $T_{\text{mass}}$ interaction. Regarding the former, the *Shl* condition led to significantly lower Errors than both the *Full* and *Arm*, and the *No* condition led to significantly lower Errors than the *Full* (Figure 15b). Regarding the $T_{\text{mass}} \times$ Gender interactive effect, the lighter tool led to more errors among males, while females made more errors using the heavier tool.
Table 4. Statistical results for the main and interactive effects of Exo, $T_{mass}$, and Gender on maximum acceptable frequency (MAF), Errors, and ratings of perceived discomfort (RPD). Any transformations applied to the dependent measures are indicated. $P$ values are reported along with effect sizes ($\eta^2_p$) in parentheses for each effect, with bold font highlighting significant effects ($p<0.05$).

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>$T_{mass}$ (T)</th>
<th>E x T</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>T x G</th>
<th>E x T x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAF</td>
<td>Log</td>
<td>0.0001</td>
<td>&lt;0.0001</td>
<td>0.372</td>
<td>0.248</td>
<td>0.019</td>
<td>0.351</td>
<td>0.644</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.188)</td>
<td>(0.403)</td>
<td>(0.031)</td>
<td>(0.112)</td>
<td>(0.096)</td>
<td>(0.009)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Errors</td>
<td>Square root</td>
<td>&lt;0.0001</td>
<td>0.945</td>
<td>0.131</td>
<td>0.469</td>
<td>0.183</td>
<td>0.022</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.206)</td>
<td>(0.00005)</td>
<td>(0.056)</td>
<td>(0.057)</td>
<td>(0.048)</td>
<td>(0.052)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>RPD hand/wrist</td>
<td>--</td>
<td>0.004</td>
<td>0.001</td>
<td>0.277</td>
<td>0.954</td>
<td>0.308</td>
<td>0.589</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.124)</td>
<td>(0.107)</td>
<td>(0.038)</td>
<td>(0.0005)</td>
<td>(0.036)</td>
<td>(0.003)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>RPD upper arm</td>
<td>--</td>
<td>0.001</td>
<td>0.006</td>
<td>0.900</td>
<td>0.910</td>
<td>0.941</td>
<td>0.238</td>
<td>0.918</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.153)</td>
<td>(0.075)</td>
<td>(0.006)</td>
<td>(0.002)</td>
<td>(0.004)</td>
<td>(0.014)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>RPD shoulder</td>
<td>--</td>
<td>0.069</td>
<td>0.037</td>
<td>0.593</td>
<td>0.204</td>
<td>0.178</td>
<td>0.889</td>
<td>0.853</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.070)</td>
<td>(0.044)</td>
<td>(0.019)</td>
<td>(0.147)</td>
<td>(0.049)</td>
<td>(0.0002)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>RPD neck</td>
<td>--</td>
<td>0.196</td>
<td>0.307</td>
<td>0.720</td>
<td>0.417</td>
<td>0.827</td>
<td>0.117</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.046)</td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.089)</td>
<td>(0.009)</td>
<td>(0.025)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>RPD low back</td>
<td>Log</td>
<td>0.003</td>
<td>0.317</td>
<td>0.817</td>
<td>0.484</td>
<td>0.018</td>
<td>0.854</td>
<td>0.626</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.134)</td>
<td>(0.010)</td>
<td>(0.009)</td>
<td>(0.078)</td>
<td>(0.097)</td>
<td>(0.0003)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>RPD thigh</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>0.010</td>
<td>0.621</td>
<td>0.949</td>
<td>0.107</td>
<td>0.367</td>
<td>0.727</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.260)</td>
<td>(0.066)</td>
<td>(0.018)</td>
<td>(0.001)</td>
<td>(0.060)</td>
<td>(0.008)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>RPD lower leg and foot</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>0.022</td>
<td>0.938</td>
<td>0.871</td>
<td>0.123</td>
<td>0.636</td>
<td>0.470</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.539)</td>
<td>(0.054)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.059)</td>
<td>(0.002)</td>
<td>(0.026)</td>
</tr>
</tbody>
</table>
Figure 15. Maximum Acceptable Frequency (MAF) for the four Exo levels (a). The symbol * indicates a significant difference between genders within Exo levels. Letters specify groupings obtained from pairwise comparisons between Exo conditions (among females in figure a). Quality results, in terms of the number of errors made (b). Letters specify groupings obtained from pairwise comparisons between conditions. In both graphs, error bars indicate 95% confidence intervals.

**Ratings of Perceived Discomfort Metrics**

Mean RPDs overall ranged from 1.9 in the thigh to 3.8 in the shoulder, corresponding to “weak” and more than “moderate” levels of discomfort. Most RPDs were influenced by Exo and $T_{mass}$ (Table 4). Compared to the *No* condition, all three exoskeletal designs led to lower levels of hand/wrist and upper arm RPDs (Figure 16a and b), while the *Shl* condition yielded the highest
shoulder RPDs (Figure 16c). For low back RPDs, there was a significant Exo x Gender interactive effect (Table 4). Among females, designs with mechanical arms (Full and Arm) led to higher levels of low back RPDs (Figure 16d). For males, the Shl condition significantly reduced low back RPDs compared to the No. The Full device caused significantly higher RPDs in the thigh and lower leg/foot than all other conditions (Figure 16e and f). $T_{mass}$ had significant effects on RPDs at the hand/wrist, upper arm, shoulder, thigh, and lower leg/foot, with the heavier tool leading to RPDs that were higher by 25, 15, 13, 15, and 14%, respectively.

Figure 16. Ratings of perceived discomfort (RPD) for different body regions. Letters specify groupings obtained from pairwise comparisons between Exo conditions (within Gender for figure d), and error bars indicate 95% confidence intervals.
**Peak Loading Metrics (90th percentile EMG)**

The mean level of peak loading across all conditions and muscles was 21.6% of MVC. Most peak loading metrics were influenced by Exo and $T_{mass}$ (Table 5). For the AD-R and MD-R, the Shl condition led to lower levels compared to all other conditions (Figure 17a and b). Similar results were found for the AD-L and MD-L, but with the Full condition leading to higher levels than the No (Figure 17e and f). There was also a significant $T_{mass} \times$ Exo interactive effect on AD-L, in which the influence of increasing $T_{mass}$ was greater when using designs with mechanical arms (Full and Arm; Figure 17e). For the TB-R, there was a significant Exo x Gender interactive effect; among females, the Full condition caused significantly lower peak loading than all other conditions (Figure 17c). For the ILL-R, the Full yielded significantly higher levels than all other conditions (Figure 17d), while for the ILL-L designs with mechanical arms (Full and Arm) led to significantly higher levels (Figure 17h).

There were significant $T_{mass} \times$ Gender interactive effects on peak loading in all tested muscles except the ILL-R, TB-L, and MD-L (Table 5). Specifically, the heavier tool led to significantly higher peak loading, and this effect was more prominent among female participants (excepting the ILL-L, where it was more prominent for males). When the noted interaction was not significant (i.e., in the ILL-R, TB-L, and MD-L), there was a significant main effect of $T_{mass}$, and the heavier tool led to a significantly higher peak loading.
Table 5. Statistical results for the main and interactive effects of Exo, $T_{mass}$, and Gender on peak loading EMG metrics. Any transformations applied to the dependent measures are indicated. $P$ values are reported along with effect sizes ($\eta_p^2$) in parentheses for each effect, with bold font highlighting significant effects ($p<0.05$).

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>$T_{mass}$ (T)</th>
<th>E x T</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>T x G</th>
<th>E x T x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD-R</td>
<td>--</td>
<td><strong>0.003</strong></td>
<td>&lt;0.0001</td>
<td>0.415</td>
<td>0.328</td>
<td>0.197</td>
<td><strong>0.018</strong></td>
<td>0.455</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.131)</td>
<td>(0.480)</td>
<td>(0.029)</td>
<td>(0.076)</td>
<td>(0.046)</td>
<td>(0.056)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>MD-R</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.359</td>
<td>0.732</td>
<td>0.053</td>
<td><strong>0.026</strong></td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.208)</td>
<td>(0.219)</td>
<td>(0.032)</td>
<td>(0.011)</td>
<td>(0.075)</td>
<td>(0.050)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>TB-R</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.052</td>
<td><strong>0.047</strong></td>
<td><strong>0.006</strong></td>
<td><strong>0.001</strong></td>
<td>0.344</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.196)</td>
<td>(0.384)</td>
<td>(0.075)</td>
<td>(0.215)</td>
<td>(0.120)</td>
<td>(0.109)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>ILL-R</td>
<td>Square root</td>
<td>&lt;0.0001</td>
<td><strong>0.032</strong></td>
<td>0.287</td>
<td>0.705</td>
<td>0.689</td>
<td>0.831</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.328)</td>
<td>(0.046)</td>
<td>(0.038)</td>
<td>(0.020)</td>
<td>(0.015)</td>
<td>(0.0005)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>TB-L</td>
<td>--</td>
<td>0.053</td>
<td>&lt;0.0001</td>
<td>0.728</td>
<td>0.090</td>
<td>0.184</td>
<td>0.448</td>
<td>0.621</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.075)</td>
<td>(0.366)</td>
<td>(0.013)</td>
<td>(0.174)</td>
<td>(0.048)</td>
<td>(0.006)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>MD-L</td>
<td>--</td>
<td>&lt;0.0001</td>
<td><strong>0.033</strong></td>
<td>0.674</td>
<td>0.817</td>
<td>0.358</td>
<td>0.118</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.248)</td>
<td>(0.045)</td>
<td>(0.015)</td>
<td>(0.004)</td>
<td>(0.032)</td>
<td>(0.025)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>AD-L</td>
<td>--</td>
<td>&lt;0.0001</td>
<td><strong>0.0005</strong></td>
<td><strong>0.003</strong></td>
<td><strong>0.039</strong></td>
<td>0.235</td>
<td><strong>0.008</strong></td>
<td>0.549</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.374)</td>
<td>(0.118)</td>
<td>(0.130)</td>
<td>(0.263)</td>
<td>(0.042)</td>
<td>(0.070)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>ILL-L</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>0.530</td>
<td>0.747</td>
<td>0.540</td>
<td><strong>0.008</strong></td>
<td>0.821</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.206)</td>
<td>(0.275)</td>
<td>(0.022)</td>
<td>(0.016)</td>
<td>(0.022)</td>
<td>(0.071)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>
Median Loading Metrics (50th percentile EMG)

Median loading across all conditions and muscles was 5.0% of MVC. Most median loading metrics were influenced by Exo, while the influence of T\textsubscript{mass} was less evident than was the case for peak loading metrics (Table 6). For both AD-R and MD-R, there was an Exo x Gender interactive effect (Table 6). Among females, the Full condition led to significantly higher loading than all other conditions (Figure 18a and b). For both MD-L and AD-L, the Shl and No conditions led to significantly lower median loading than either the Full or Arm conditions (Figure 18e and f). For TB-R, there was an Exo x Gender interactive effect that approached significance; for females, the Shl led to significantly higher median loading than all other conditions (Figure 18c).
Designs with mechanical arms (*Full* and *Arm*) appeared to increase median metrics of the bilateral ILL (Figure 18d and h), though this difference was significant for the ILL-L only among females.

There was a significant $T_{mass} \times$ Gender interactive effect on the median loading of ILL-L; the heavier tool led to a significantly higher loading only for male participants ($p = 0.0011$). There was also a significant main effect of $T_{mass}$ on the median loadings of ILL-R and AD-L, with the heavier tool causing significantly higher loadings.

**Table 6.** Statistical results for the main and interactive effects of Exo, $T_{mass}$, and Gender on median loading EMG metrics. Any transformations applied to the dependent measures are indicated. *P* values are reported along with effect sizes ($\eta_p^2$) in parentheses for each effect, with bold font highlighting significant effects ($p < 0.05$).

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>$T_{mass}$ (T)</th>
<th>E x T</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>T x G</th>
<th>E x T x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD-R</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.471</td>
<td>0.119</td>
<td>0.001</td>
<td>&lt;0.0001</td>
<td>0.306</td>
<td>0.795</td>
</tr>
<tr>
<td>MD-R</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.782</td>
<td>0.306</td>
<td>0.009</td>
<td>&lt;0.0001</td>
<td>0.174</td>
<td>0.331</td>
</tr>
<tr>
<td>TB-R</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.814</td>
<td>0.942</td>
<td>0.141</td>
<td>0.057</td>
<td>0.251</td>
<td>0.562</td>
</tr>
<tr>
<td>ILL-R</td>
<td>Square root</td>
<td>&lt;0.0001</td>
<td>0.029</td>
<td>0.798</td>
<td>0.566</td>
<td>0.733</td>
<td>0.707</td>
<td>0.214</td>
</tr>
<tr>
<td>TB-L</td>
<td>--</td>
<td>0.641</td>
<td>0.069</td>
<td>0.693</td>
<td>0.099</td>
<td>0.461</td>
<td>0.169</td>
<td>0.669</td>
</tr>
<tr>
<td>MD-L</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>0.103</td>
<td>0.976</td>
<td>0.707</td>
<td>0.555</td>
<td>0.899</td>
<td>0.929</td>
</tr>
<tr>
<td>AD-L</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.002</td>
<td>0.183</td>
<td>0.403</td>
<td>0.940</td>
<td>0.799</td>
<td>0.540</td>
</tr>
<tr>
<td>ILL-L</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.008</td>
<td>0.664</td>
<td>0.940</td>
<td>0.004</td>
<td>0.006</td>
<td>0.440</td>
</tr>
</tbody>
</table>
Across all conditions and muscles, the mean level of static loading was 1.1% of MVC. Most static loading metrics were influenced by Exo, but with only one muscle influenced by $T_{\text{mass}}$ (Table 7). A majority of results obtained for the static loading metrics were similar to those described above for median loading. A few differences, though, were evident. For TB-R, static loading in the Shl and Full conditions was higher than in the Arm and No conditions (Figure 19c). For the TB-L, there was a significant Exo $\times$ Gender interactive effect; for females, the Full condition led to significantly higher static loading than all other conditions (Figure 19g). While ILL-R loading did

Figure 18. Median loading EMG metrics (as a percentage of maximum levels). The symbol * indicates a significant difference between genders within Exo levels. Letters specify groupings obtained from pairwise comparisons between Exo conditions (within Gender for figures a-c and h), and error bars indicate 95% confidence intervals.

**Static Loading Metrics (10th percentile EMG)**

Across all conditions and muscles, the mean level of static loading was 1.1% of MVC. Most static loading metrics were influenced by Exo, but with only one muscle influenced by $T_{\text{mass}}$ (Table 7). A majority of results obtained for the static loading metrics were similar to those described above for median loading. A few differences, though, were evident. For TB-R, static loading in the Shl and Full conditions was higher than in the Arm and No conditions (Figure 19c). For the TB-L, there was a significant Exo $\times$ Gender interactive effect; for females, the Full condition led to significantly higher static loading than all other conditions (Figure 19g). While ILL-R loading did

73
not significantly differ between Exo conditions overall (Figure 19d), the Arm condition led to significantly higher ILL-L loading than all other conditions among females (Figure 19h).

Table 7. Statistical results for the main and interactive effects of Exo, \(T_{\text{mass}}\), and gender on static loading EMG metrics. Any transformations applied to the dependent measures are indicated. \(P\) values are reported along with effect sizes (\(\eta^2_p\)) in parentheses for each effect, with bold font highlighting significant effects (\(p<0.05\)).

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>(T_{\text{mass}}) (T)</th>
<th>E x T</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>T x G</th>
<th>E x T x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD-R</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>(0.403)</td>
<td>0.089</td>
<td>0.010</td>
<td>0.004</td>
<td>0.256</td>
<td>0.600</td>
</tr>
<tr>
<td>MD-R</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>(0.378)</td>
<td>0.047</td>
<td>0.299</td>
<td>0.001</td>
<td>0.278</td>
<td>0.825</td>
</tr>
<tr>
<td>TB-R</td>
<td>Log</td>
<td>0.0001</td>
<td>(0.193)</td>
<td>0.645</td>
<td>0.790</td>
<td>0.604</td>
<td>0.466</td>
<td>0.223</td>
</tr>
<tr>
<td>ILL-R</td>
<td>Log</td>
<td>0.074</td>
<td>(0.068)</td>
<td>0.459</td>
<td>0.924</td>
<td>0.678</td>
<td>0.067</td>
<td>0.119</td>
</tr>
<tr>
<td>TB-L</td>
<td>--</td>
<td>0.035</td>
<td>(0.084)</td>
<td>0.218</td>
<td>0.068</td>
<td>0.010</td>
<td>0.253</td>
<td>0.536</td>
</tr>
<tr>
<td>MD-L</td>
<td>--</td>
<td>0.003</td>
<td>(0.130)</td>
<td>0.219</td>
<td>0.460</td>
<td>0.814</td>
<td>0.895</td>
<td>0.532</td>
</tr>
<tr>
<td>AD-L</td>
<td>Square root</td>
<td>0.0002</td>
<td>(0.186)</td>
<td>0.636</td>
<td>0.655</td>
<td>0.066</td>
<td>0.800</td>
<td>0.754</td>
</tr>
<tr>
<td>ILL-L</td>
<td>--</td>
<td>0.009</td>
<td>(0.111)</td>
<td>0.172</td>
<td>0.561</td>
<td>0.733</td>
<td>0.009</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Relationships between MAF and Other Physical Demand Metrics

Results from fitting simple regression models to the MAF and other physical demand metrics are summarized in Table 8. The highest level of correspondence ($r^2=0.254$) was found between MAF and median loading of the AD-L. Based on the number of significant relationships, MAF appeared more associated with peak EMG loading metrics than with either median or static EMG metrics. Although detailed results are not given here, MAF was also significantly associated with all static EMG metrics except TB-R, MD-L, and ILL-L metrics, and with all RPD metrics except those at the low back and neck; while significant, the $r^2$ values for these relationships were relatively low.
(0.030 – 0.076). All significant associations with MAF were inverse, except for those with static and median measures of AD-L and the static measure of AD-L.

Table 8. Results from linear regression modeling of the relationships between MAF and other physical demand metrics. \( P \) values are reported along with coefficient of determination \( (r^2) \) in parentheses, with bold font indicating significant relationships \( (p<0.05) \).

<table>
<thead>
<tr>
<th>Peak metrics</th>
<th>P value ( (r^2) )</th>
<th>Slope (95% CI)</th>
<th>Median metrics</th>
<th>P value ( (r^2) )</th>
<th>Slope (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD-R</td>
<td>0.003</td>
<td>-0.049</td>
<td>AD-R</td>
<td>(0.069)</td>
<td>(-0.080:-0.017)</td>
</tr>
<tr>
<td>MD-R</td>
<td>&lt;0.0001</td>
<td>-0.059</td>
<td>MD-R</td>
<td>(0.151)</td>
<td>(-0.083:-0.034)</td>
</tr>
<tr>
<td>TB-R</td>
<td>0.0005</td>
<td>-0.062</td>
<td>TB-R</td>
<td>(0.108)</td>
<td>(-0.094:-0.030)</td>
</tr>
<tr>
<td>ILL-R</td>
<td>0.0005</td>
<td>-0.095</td>
<td>ILL-R</td>
<td>(0.092)</td>
<td>(-0.148:-0.042)</td>
</tr>
<tr>
<td>TB-L</td>
<td>0.0002</td>
<td>-0.050</td>
<td>TB-L</td>
<td>(0.108)</td>
<td>(-0.076:-0.025)</td>
</tr>
<tr>
<td>MD-L</td>
<td>0.450</td>
<td>-0.008</td>
<td>MD-L</td>
<td>(0.004)</td>
<td>(-0.029:0.013)</td>
</tr>
<tr>
<td>AD-L</td>
<td>0.0168</td>
<td>-0.059</td>
<td>AD-L</td>
<td>(0.001)</td>
<td>(-0.042:0.028)</td>
</tr>
<tr>
<td>ILL-L</td>
<td>0.0044</td>
<td>(-0.108:-0.011)</td>
<td>ILL-L</td>
<td>(0.024)</td>
<td>(-0.206:0.011)</td>
</tr>
</tbody>
</table>

**Usability and Rankings**

For questions 1 (“It meets my needs for completing the task”) and 4 (“I am satisfied with it”), Exo had a significant influence \( (p =0.001 \text{ and } <0.0001, \text{ respectively}) \). For both questions, the Arm condition led to a significantly higher agreement than the Full (Figure 20a and d). For questions 2 (“It makes the things I want to accomplish easier to get done”), 3 (“It is easy to use”), 5 (“It works the way I want it to work”), and 6 (“Overall, I prefer working with it more than without it”), Exo also had a significant effect \( (p =0.0001, <0.0001, <0.0001, \text{ and } 0.0004, \text{ respectively}) \). For
these questions, the *Full* condition led to a significantly lower agreement than either the *Shl* or *Arm* conditions (Figure 20). For questions 2 and 6, responses also significantly differed between genders ($p = 0.030$ and $0.0316$, respectively), with males reporting higher agreement for both. Rankings were significantly influenced only by Exo ($p = 0.0002$). Both the *Full* and *No* conditions were ranked higher than either the *Shl* or *Arm*, with respective mean rankings of 3.0, 2.8, 2.1, and 2.0.

![Figure 20](image.png)

Figure 20. Usability questions responses on a Likert type scale (1= strongly disagree, 5= strongly agree). Q1: “It meets my needs for completing the task,” Q2: “It makes the things I want to accomplish easier to get done,” Q3: “It is easy to use,” Q4: “I am satisfied with it,” Q5 “It works the way I want it to work, and Q6: “Overall, I prefer working with it more than without it.” Letters specify groupings obtained from pairwise comparisons between Exo conditions, and error bars indicate 95% confidence intervals.

**Discussion**

The main objective of this study was to compare three different exoskeletal designs, each intended to support the upper extremity, in terms of physical demands (as measured by subjective and
objective metrics) and quality (as measured by deviations from a required angular tolerance). This comparison was done in the context of a simulated overhead drilling task, while considering the potential moderating influence of tool mass. As discussed below, the primary finding was that the three exoskeletal designs tested indeed had differing effects on physical demands and quality. While these effects were influenced to some extent by tool mass, gender was an important moderating factor for more of the outcome measures obtained.

**Physical Demands and Quality**

When setting MAF in the current study, and as is generally assumed in a psychophysical approach, participants are presumed to have taken the entire task into account and integrated a range of physiological and biomechanical factors (Snook, 1999). Supporting this to some extent in our study, several significant inverse associations were found between MAF and both objective (EMG peak, median, and static loading) and subjective (RPD metrics) measures of physical demands (Table 8). Additionally, MAF can be treated as indicating potential productivity (Dempsey, 1998; Potvin, 2012), an important dimension of task performance when studying the economic effectiveness of such interventions in industry. The Arm device appeared to impose the least physical demands, though perhaps only among females, as suggested by MAF results (Figure 15a). Supporting this, participants reported the highest level of satisfaction with this particular design (usability question 4; Figure 20d). MAF differences between Exo conditions may have been more evident for females because the task demands were relatively higher for them as a group. Gender differences in were apparent in muscular demands, for example, in the TB-R (90th%; Figure 17c) and AD-R (50th%; Figure 18a). Individual strength, as measured, was not found to be a significant
covariate for most dependent measures, though potential differences related to distal upper extremity strength and related muscular demands were not assessed here.

When using the Arm device, increasing MAF was likely limited by demands on the low back, as suggested by RPDs at the low back and all measures of ILL-L. Increased demands on the low back using this design were somewhat expected, since the weight of the mechanical arm and the tool are primarily resisted by the low back extensors. Also at the low back, demands on the left side (ILL-L) were higher, likely a result of an asymmetric external moment created by the tool and mechanical arm weights (recall that all participants were right-handed, and note that they were observed to prefer positioning the tool closer to their right side). Supporting the current results related to design of the Arm device, an earlier study using a similar device also found that this design, compared to a control condition, reduced upper arm RPD and slightly increasing low back RPD in a simulated overhead task (Rashedi et al., 2014).

Although MAF selected when using the Shl device was not significantly different from the No device condition, using the former did reduce peak loading on the shoulders (as evident by the bilateral measures of AD and MD; Figure 17). Another study that used a similar exoskeleton design (a prototype EksoVest™ by EksoBionics) in a simulated overhead task (Kim et al., 2017) also found it reduced peak shoulder loading. The increased median loading on the TB-R observed here appeared to be a potential limiting factor to increasing MAF when using the Shl design (Figure 18c). This limitation was likely because the design only supports the upper arm, whereas the Full and Arm devices support the tool (thereby unloading the entire upper extremity).
While using the *Full* exoskeleton decreased peak loading on the TB-R, it increased median and static loadings of the AD-R and MD-R (for females) and notably increased RPDs for the lower extremity. Consistent with the latter increases, participants reported the least satisfaction with this device (usability question 4; Figure 20d). Regarding the increased lower extremity discomfort, this design required fixing the exoskeleton structure to the wearer’s feet to facilitate load transfer and mobility. This device was also designed to be used with large boots (the foot straps, when adjusted to the minimum size, were originally larger than what was needed for sport shoes). Because it was hard to offer boots with multiple sizes, we reconfigured the straps by adding more holes to the original straps to allow for more adjustment and standardized the shoes by asking participants to use one kind of sports shoes that were available in our lab. While we consider it unlikely, this modification may have contributed to the general increase in physical demands and the more specific increase in lower-extremity discomfort. The lower MAF and increased Errors using this device, in contrast, were likely influenced by the increased median and static muscle loadings and increased discomfort.

Potentially due to the more complex articulations involved and the associated inertial properties, designs with mechanical arms (Full and Arm) appeared to deteriorate task quality, as reflected in the number of errors made (Figure 15b). While the Shl design did not lead to higher quality compared to the No condition, the Levitate Airframe™, another conceptually similar exoskeleton, was reported to improve performance in a simulated overhead precision task (Spada et al., 2017a; Spada et al., 2017b) and in a small-scale study of a simulated painting and welding tasks (Butler, 2016). Such inconsistent results might be attributed to the different physical demands imposed by
the different tasks examined. More generally, the influence of exoskeletal design on quality might be influenced by other aspects of a task (e.g., required precision levels).

The influence of increasing tool mass ($T_{mass}$) on physical loading appeared fairly consistent across Exo conditions, with the heavier tool causing higher muscle activity (especially peak levels) and lower MAF. Peak loading on the AD-L was the only significant $T_{mass} \times$ Exo interaction found (Table 5), wherein the effect of increasing $T_{mass}$ was greater when using designs with mechanical arms (Full and Arm; Figure 17e). This suggests that demands on the left arm (non-dominant arm for all participants) were higher in the noted conditions using the heavier tool, perhaps as a result of the continuous attachment of the tool to the wearer’s body. When using the Shl device participants were able to put the tool down (between holes), whereas when using the Full and Arm designs the non-dominant arm was loaded by a need of the user to control and/or balance the mechanical arm. The dominant arm seemed equally affected by increasing $T_{mass}$ across Exo conditions.

**User Feedback**

When using the Full device, the most frequent concern expressed was that the design reduced postural stability and made it harder to control balance. In this condition, postural balance was commonly perceived to be influenced primarily at the lower leg and foot, where the exoskeletal structure contacts the wearer’s body. Such a concern was also consistent with the increased lower leg and foot RPDs reported in this condition, and may have contributed to the increased Errors in this condition. Reduced postural balance also seemed to be a major contributor in limiting MAF for this design. Additionally, it was reported that balancing the weight of the tool and articulated
arm was mentally and physically demanding. More specifically, they indicated that they needed to return the arm to the “normal” position to regain balance after completing a hole. However, most participants indicated that the counterbalancing weights in this design were helpful for reducing physical demands, especially when using the heavier tool.

Most common concerns for the Shl device were that: 1) frequent movements of the upper arm straps caused discomfort, although this effect was not reflected in the upper arm RPD metric; and 2) some participants kept their shoulders slightly abducted to avoid resisting the assistive force generated by the device when returning their upper arms to neutral. While not evident in the hand/wrist RPDs, another concern was that the design did not support the hand/wrist.

The most frequent concern noted when using the Arm device was low back discomfort (discussed above), and this was consistent with the observed levels of static and median ILL-L and the low back RPDs. Many participants highlighted the need for a counterbalancing procedure (perhaps similar to that implemented in the Full design). A second common concern was that the wearer’s shoulders (particularly the left one as it was raised higher) rubbed against the shoulder pads when raising the arms, and that this was a source of shoulder discomfort. However, this concern was not evident in the shoulder RPD metric.

Comments on Experimental Design

For this study, our initial interest was to examine the influences of exoskeleton design and tool mass on fatigue. For this purpose, two experimental designs were considered originally. The first would involve testing each condition in a separate session, and separating sessions in time to minimize any effects from residual muscle fatigue. The task used here (i.e., simulated overhead
drilling at a *fixed* pace) would then be continued until participants reached substantial levels of fatigue in each session. Because each participant would be required to attend eight sessions (4 device conditions × 2 tool masses), this design was considered infeasible, inefficient, and rather taxing to participants. A second was also design considered, in which the eight conditions would be completed in two sessions. Each condition would be completed in a shorter time, using some type of stopping criteria, such as was used by Sood et al. (2007). For example, in each condition the task would be continued until one of the following criteria is reached: when RPDs for any body part are ≥ 8 on two consecutive samples or when 15 minutes had elapsed. This design was also removed from consideration, mainly because of two reasons. First, these sessions would have required a relatively long time to allow for recovery from muscular fatigue between conditions. Second, extensive pilot testing would be required to decide on dependent measures for performance and fatigue, because the different trials were expected to have unequal durations. Specifically, we would had needed to confirm that rates of change (vs. means) could serve as dependent measures (e.g., slope of RPDs or the quality metric). Mean value would not be appropriate, since the length of the trials would likely be unequal as indicated earlier. While we still consider fatigue to be a useful basis for future work, it was decided to modify the research question for the current study to examine the influences of the two noted factors on a subjectively-determined outcome (the MAF). As discussed earlier, psychophysical methods provide a realistic/holistic assessment approach that can require less of both cost and time (Snook, 1999) and which has been used previously to examine differences between task conditions (e.g., Ciriello, 2007; Li et al., 2007). Such differences are assumed to reflect potential longer-term influences.
Limitations

There are a few limitations in this study that should be noted. Potential concerns related to external validity, specific tolerance and force levels, and use of novices, which were described in Alabdulkarim et al., (2017) for a very similar simulated task, apply to this study as well. Additionally, this study examined/compared different exoskeleton designs during only a single simulated overhead task. It is unknown if the current results will generalize, such as to other working heights, force directions, tool types/masses, or task durations. Specifically, the pattern of results may change if a more demanding task was used. As an example, the Full design is relatively robust compared to the others tested here, and may be particularly beneficial using tools heavier than those we examined. Further, only a single device was used to represent three distinct exoskeleton design approaches, and it is unknown regarding how well the selected devices represent the examined design approaches (e.g., several new devices are entering the market now, and others are likely in the near future). Because tool mass included only two levels, any nonlinearity in the relationships between tool mass and the outcome measures was not captured.

There are also three potential limitations pertaining to the adjustment duration used here. First, only two conditions (of the eight) were tested in pilot work, which were presumed to be the most and least physically demanding conditions. We did not confirm experimentally that the adjustment duration used here was sufficient for other conditions. Second, the repeatability of the current procedures for selecting MAF in a given condition was not tested. Lastly, it is unclear regarding the extent to which results from the 15-minute adjustment duration will reflect actual longer-term outcomes.
Conclusions

In summary, exoskeleton designs that included mechanical arms appeared to increase loading on the low back (i.e., the Full and Arm devices), though this effect was partially alleviated when the design allowed counterbalancing the load and transferring it to the ground (i.e., the Full device). When the tool was connected to a mechanical arm, lower quality was observed. The exoskeleton design that mainly supported the shoulder (i.e., the Shl device) reduced shoulder peak loading, but it also increased median and static loading of the dominant upper arm (i.e., TB-R muscle) and did not appear to impact quality.

To facilitate the successful implementation of exoskeletons in occupational settings, this study highlights the need to consider at least three dimensions of potential outcomes – specifically physical demands, task performance, and usability – all of which may be influential in determining the potential effectiveness of exoskeletons in the workplace. Three distinct designs were tested here, and which led to varied outcomes in these dimensions, with no one design found obviously superior across all dimensions. Given the evident potential of such technology, however, future research is needed to address some of the challenges that were identified, and to compare exoskeleton design approaches under more diverse and realistic conditions.

References


Chapter 4: Effects of Exoskeleton Design and Precision Requirements on Physical Demands and Performance in a Simulated Overhead Drilling Task

Abstract

We compared three distinct passive exoskeletal designs in a mock drilling task under three levels of precision requirements (defined by required hole sizes), in terms of physical loading (perceived exertion and muscular loading) and quality. The investigated designs were: 1) an upper-body exoskeleton mainly supporting the shoulder, and both 2) full-body and 3) upper-body exoskeletons, each with connected mechanical arms. At a fixed drilling pace, participants \( n = 12 \) repeated “drilling” two same-sized holes for two minutes. The influence of increasing precision was fairly consistent across the examined designs, and was evident as increased physical loading in some muscles and deteriorated performance. Designs with mechanical arms led to the largest reductions in quality and increased physical loading overall (mainly in the low back). The shoulder-focused device reduced shoulder demands but deteriorated quality in the condition with highest precision requirements. Although future work is needed under more diverse/realistic scenarios, these results might be useful to (re)design occupational exoskeletons.

Keywords: Exoskeleton, performance, precision
Introduction

Recent evidence suggests that the most frequent occupational injuries are in the upper extremities, with the shoulder having the highest median of days away from work among all body parts (BLS, 2016). Among diverse risk factors, working with arm elevation is a well-documented contributor to shoulder injuries (e.g., Svendsen et al., 2004; Van Rijn et al., 2010; van der Molen et al., 2017). Common control strategies for such work include administrative (e.g., job rotation) and engineering solutions (e.g., using automation). The latter, however, is often infeasible and/or unjustifiably expensive. Wearable exoskeletons are rapidly emerging for occupational application, and may be an effective alternative or aid to existing intervention approaches. Using exoskeletons could, in some cases, require limited modification to an existing workplace (and therefore, can be less expensive) and might be useful for a broad range of tasks. Because implementing automation can be difficult particularly when repeated changes are anticipated (e.g., to product characteristics or production layout), de Looze et al. (2017) noted that exoskeletons are recommended to retain the benefits of human creativity and flexibility, further supporting their potential in diverse occupations.

Nonetheless, exoskeletons mainly have been evaluated for military uses (Lee et al., 2012), or for rehabilitation and medical purposes as an aid for disabled, injured, or weak patients (Viteckova et al., 2013). For occupational application, a recent review emphasized the potential for these technologies, and highlighted that they can substantially reduce physical demands (de Looze et al., 2015). Yet, exoskeletons can also impose “unintended consequences”, such as increasing physical demands on “other” body parts while reducing demands on target body regions. For example, a
passive exoskeleton attenuated loading on the back and legs but increased discomfort in the chest in a simulated assembly task with the trunk in a forward-bended posture (Bosch et al., 2016).

Exoskeletons can be categorized according to the body regions they aim to support (Lee et al., 2012; de Looze et al., 2015), such as full body (e.g., Toyama & Yonetake, 2007), upper body (e.g., Kobayashi et al., 2007), or lower body (e.g., Kim et al., 2009) devices. They also can be classified based on the support generation mechanism(s) employed, as active and passive devices (de Looze et al., 2015). In active devices, human strength is enhanced by at least one actuator that employs diverse technologies such as electric motors, hydraulics, or pneumatic muscles (Gopura & Kiguchi, 2009). In entirely passive devices, such support results from dampers, springs, or materials that can store energy harvested by wearer movement and produce it later as an augmentation energy (de Looze et al., 2015). Although they can offer more support, active devices are often more expensive and heavier. Therefore, the focus here was on passive exoskeletal designs, especially given that multiple new exoskeletons are entering the market. Lastly, exoskeletal designs might be distinguished based on their approach to transmitting external loads (e.g., from the shoulder to the floor or the waist/hip).

Precision demands are present to differing degrees in most occupational tasks, and can have important physiological and biomechanical effects on workers as well as influencing task performance. In a seated posture, for example, increasing precision caused participants to adopt “poorer” postures for a manual task (Li & Haslegrave, 1999). Increasing precision can also increase muscle activation; in a task simulating dentistry work, Milerad and Ericson (1994) found that increased precision demands led to higher activity levels in muscles with a stabilizing function
(mainly in neck, shoulder and wrist). In addition to increasing co-contraction, precision can increase loading by modifying movement kinematics; with precision, movement time increases, a phenomena commonly described as the speed-accuracy tradeoff, and widely explained by Fitts’ Law (Fitts, 1954; Fitts & Peterson, 1964). More specifically in goal-oriented movements, as target size decreases, there is a longer decelerative phase during which more corrective sub-movements occur (e.g., MacKenzie et al., 1987; Thompson et al., 2007; Temprado et al., 2013). In repetitive lifting tasks, precision requirements also led to changes in posture and/or muscle activation (Collier et al., 2014; Joseph et al., 2014; Mehta et al., 2015), further supporting the occupational significance of precision. Repetitive movements requiring precision are also recognized as a risk factor for neck and shoulder injury (Ekberg et al., 1994). Along with the noted effects of precision on physical demands, increasing precision requirements can reduce task performance, as was found in a computer-simulated crane operation (Huysmans et al., 2006) and an aiming task using a computer mouse (Visser et al., 2004). Higher precision demands also increased completion time and reduced quality in a simulated assembly task common in the automotive industry (Wartenberg et al., 2004).

Wearing exoskeletons can restrict movement and modify working strategies; such changes may interact with precision effects and lead to influences on physical demands and precision task performance, two dimensions of effectiveness that are important for eventual occupational application. While knowledge on the former is key to verifying their effectiveness in reducing occupational injury risk, evidence regarding the later might be useful to cost-justify such interventions, where financial benefits are commonly fundamental to decision making. To our knowledge, this is the first study comparing different exoskeletal design approaches under varying
levels of task precision requirements. The study was designed to examine the following hypothesis, in the context of a simulated overhead occupational task (drilling): the influence of exoskeleton design on physical loading and performance will depend on precision requirements.

Methods

Participants

A convenience sample of 12 participants completed the study, who were recruited from the University and local community. All participants completed an initial informed consent procedure that was approved by the Virginia Tech Institutional Review Board, and reported no current or recent (previous year) musculoskeletal problems, having normal or corrected-to-normal vision, and being right-handed (except one male who reported left-handedness). Participant mean (SD) age, body mass, and stature were 20.0 (1.1) years, 63.9 (8.7) kg, and 168.9 (6.1) cm, respectively, for females (n=5); and 22 (6.4) years, 71.4 (7.8) kg, and 174.9 (7.9) cm, respectively, for males (n=7).

Task Description

An overhead, repetitive drilling task was simulated in the laboratory. This particular task imposed physical demands (strength and precision) common in occupational environments, and was a simplified version of the task described by Alabdalarim et al. (2017). Simulating a longitudinal stringer, one rung was constructed (5.08 x 5.08 cm hollow rectangular stock, 121.9 cm long) with six identical sets of three holes; each set represented three distinct levels of precision requirements, with hole diameters of 1.3, 1.6, and 1.9 cm, and pairs of holes were separated by ~16 cm). The rung was attached a structure allowing it to be height adjustable (Figure 21a). Working height was set relative to each participant’s anthropometrics; specifically, holes were set at individual eye
height. This particular height required a moderate degree of arm elevation, a factor recognized to increase injury risk in the shoulder (e.g., Punnett et al., 2000; Björkstén et al., 2001; Van Rijn et al., 2010), and was feasible for the exoskeletons used in the study (see below).

A commercial drilling tool (mass= 2.26 kg; Model # R71111, Ridgid Company) with an adjustable handle was used (Figure 21d). To simulate a drill bit, a probe and a load cell (Interface, SML-100, Scottsdale, AZ) were connected to the tool. Completing the simulated drilling of a given hole required participants to generate a drilling force (as monitored by the load cell) of at least 66.7 N, and which had to be sustained for an accumulated duration of 2.5 seconds. These simulation variables were viewed as representative of a drilling task common in aircraft manufacturing (for more information on the task simulation, see Alabdulkarim et al., 2017). Participants received different auditory feedbacks when the required force was exceeded and when the 2.5-second duration was completed. When the probe contacted the hole edges, “error” beeps were generated. Given the three hole diameters, rung dimensions, and probe diameter (1 cm), the three holes provided an increasing angular tolerance of ±2, 3.5, and 5° from normal to the surface. The simulated task involved participants repeatedly drilling two of the same-sized holes in alternate order, at a fixed pace of 7 holes/min (controlled using an acoustic metronome). Under each of the conditions described below, this repeated drilling was completed for two minutes. The two holes were set relatively close to each other to minimize mobility demands, a factor that was expected to be influenced by specific exoskeleton technologies but was not a focus of the current study. Given the tradeoff between speed and accuracy, such as explained by Fitts’ Law (Fitts, 1954; Fitts & Peterson, 1964), a fixed pace was used to emphasize the latter. Based on pilot testing, the chosen pace was used to help distinguish between the examined exoskeleton technologies but without
adding substantial fatigue. Between holes, participants were able to put the tool down on a nearby table (if feasible) and lower their arms to facilitate recovery.

Figure 21. A view of the height adjustable fixture (rung) that was used in the study (a). The commercial drill (d) that was used, with the load cell and the drill bit attached. Note that the drill was wrapped to cover all wires during the experiment. Photos of participants using: (b) the Fawcett Exosvest™ with a zeroG² mechanical arm (Arm); (e) the EksoWorks™ Vest (Shl); and (c) the FORTIS™ Exoskeleton (Full).

**Independent Variables**

Two within-subject factors were examined in the study: Exoskeleton (Exo) and Precision requirements (Precision). In addition to a control condition (i.e., without an exoskeleton), three commercial passive exoskeletons were utilized, and were chosen to exemplify designs with distinct approaches to reducing physical demands. The specific exoskeletons were: 1) an exoskeletal vest (Fawcett Exosvest™; The Tiffen Company, Hauppauge, NY, USA) with a mechanical arm (zeroG²;
Equipois Inc., Manchester, NH, USA) connected to it (Figure 21b), which supports a tool and transmits loads to the hip/waist; 2) the EksoWorks™ (v.3, Ekso Bionics Company, Richmond, CA, USA), an upper body vest that supports the arms (Figure 21e); and, 3) the FORTIS™ (Lockheed-Martin), a full-body exoskeleton that transmits tool loads to the floor (Figure 21c). Subsequently, these conditions will be referred to as “Arm”, “Shl”, and “Full”, respectively, and “No” for the control condition. To connect the tool to the mechanical arms utilized in the Full and Arm conditions, a gimbal and mounting system were used (Saturn-2, medium, 135° configuration; Equipois Inc., Manchester, NH, USA). As noted earlier, Precision has three levels of different precision demands. Subsequently, these levels will be abbreviated as Low (±5°), Middle (±3.5°), and High (±2°), respectively. Including the mechanical arm and mounting system weights (if used), the Shl, Arm, and Full weights were ~ 4.2, 10.3, and 19-28 kg, respectively. The Full weight is reported as a range, as this depended on the number of counterbalancing weights used.

**Procedures and Data Collection**

The experiment was completed in two sessions, training and testing, which were separated by at least two days to reduce any influences from residual muscle fatigue.

**Training Session (~2 hrs)**

Following the exoskeleton manufacturers’ instruction manuals, individual anthropometric fit to the devices was determined first. All recruited participants had satisfactory fit to the three devices. Next, participants were asked to practice the task under each of the 12 conditions, with each lasting at least five minutes. Participants were encouraged to ask for more practice as needed. Rest breaks of at least 5 minutes were provided between conditions, and more rest was encouraged as needed.
Participants were instructed to work as accurately as they can, following the controlled pace (7 beeps/min). Following the pace and accurate drilling were noted as equally important. When practicing any condition, participants were asked to try to find best working procedure that reduced their physical demands and improved their performance. When practicing conditions with mechanical arms (Full and Arm), participants decided how to adjust the drill handle and the stiffness of the mechanical arm. When practicing with the Full device, participants experienced using the counterbalancing weights provided by the exoskeleton and decided if/how they wanted to use them. When practicing with the Shl device, participants were exposed to near maximum and minimum levels of arm support that were provided by the exoskeleton. Subsequently, they selected the support level that they considered appropriate for the simulated task demands. Participants were asked to practice more using the parameters they chose, to confirm suitability for the task or re-adjust as needed.

Testing Session (~3 hrs)
Before starting the first testing condition, pairs of surface electromyography (EMG) electrodes were placed over three muscles bilaterally: the rectus abdominis (RA), iliocostalis lumborum pars lumborum (ILL), and anterior deltoids (AD). These particular muscles were chosen as they were accessible across all Exo conditions (i.e., certain muscles were not reachable due to components of at least one device), and were considered as important contributors during overhead work according to earlier studies (Nussbaum et al., 2001; Chopp et al., 2010; Rashedi et al., 2014). Prior to electrode placement, the skin was shaved, lightly rubbed, and wiped by 70% alcohol. Subsequently, pairs of bipolar Ag/AgCl electrodes (AccuSensor, Lynn Medical, MI, USA) with a 2.5-cm inter-electrode distance were positioned on the skin as explained by Hermens et al. (1999).
Raw EMG signals were collected at 1500 Hz using a telemetered system (TeleMyo 900, Noraxon, AZ, USA).

Participants then completed three trials of maximum voluntary contractions (MVC) for each of the monitored muscles. For the bilateral AD, participants sat in a Biodex™ dynamometer (System 3 Pro, Shirley, New York) with the device arm positioned such that the tested arm of the participant was extended and the shoulder was flexed at ~90° with the thump pointing upward. To complete a maximum exertion trial, participants were instructed to pull the system arm upward as hard as possible. The Biodex system was also used to complete MVC trials for the RA, during which participants sat on the chair with the trunk flexed forward at 45°. To complete a trial, participants were instructed to flex their trunk as hard as possible. To obtain ILL MVCs, participants were positioned on a 45° Roman chair with the back connected to the floor using a customized construction harness while the lumbar spine flexed at ~45°. To complete a trial, participants pulled on the fixture as hard as possible. Lastly, dominant hand grip strength was measured, in three trials, using a hand dynamometer (Lafayette Instrument Co., IN, USA). At least one minute of rest was provided between MVC trials, and nonthreatening verbal encouragement was given during each. Rest of at least 10 minutes was provided following the completion of all MVCs.

To minimize order-related effects, two 6x3 Balanced Latin Squares (BLSs) were used to counterbalance the presentation order of Precision and three 4x4 BLSs were used for Exo levels. At a given level of Exo, participants completed testing with all three levels of Precision before continuing to the next level of Exo. This approach was used for efficiency, as substantial time would otherwise be needed to don/doff the different exoskeletons. Prior to starting any condition...
involving an exoskeleton, brief practice (~2 min) was given using that exoskeleton (configured using the parameters selected during the training session). At least 10 minutes of rest were provided between conditions, and more was encouraged as needed. Immediately after completing a given condition, ratings of perceived exertion (RPE) were collected using the Borg’s 10-point scale (Borg, 1990) for several body parts and participant feedback on the exoskeleton (if used) were documented.

**Dependent Measures**

Physical demands were assessed using subjective and objective measures. For the former, RPEs for the hand/wrist, upper arm, shoulder, neck, low back, thigh, and lower legs and foot were collected. For bilateral body parts (e.g., shoulder), participants were asked to report the highest level of perceived exertion. Surface EMGs served as an objective evaluation. Raw EMG signals were initially band-bass filtered (20-450 Hz). Subsequently, EMG root mean square (RMS, time constant 100 ms) values were determined during the testing duration (two minutes) and these were then normalized (nEMG) to maximal RMS values found in the MVC trials. From the nEMG data, 10th, 50th, and 90th percentile values were calculated, and were viewed as illustrative of the static, median, and peak loadings in a given condition, respectively (cf. Jonsson, 1982). Because one participant was left-handed, dominant (D) and non-dominant (ND) sides were analyzed for the examined muscles. As more comprehensive muscle activation metrics, static, median, and peak loading measures were calculated for each Exo condition from the summation of nEMG across all monitored muscles. It should be noted that the summation over the dominant and non-dominant sides was found to yield results similar to those obtained from the overall totals (dominant + non-dominant). Thus, only the overall totals are reported below. The quality aspect of task
performance was evaluated using the total number of “errors” made in a given condition. Errors were counted only when the targeted drilling force was exceeded.

**Statistical Analyses**

Separate 3×4 repeated-measures analyses of variance (ANOVAs) were used for each dependent measure to evaluate the influences of Exo and Precision; Gender and presentation order of conditions were included as blocking variables. Grip strength and participant height were explored as possible covariates. The latter was significant only for peak leading of AD-D, static and median loading of RA-D, and Errors, while both covariates were significant for Hand/wrist RPE and the static and median loading of RA-ND. Because these covariates did not change the pattern of statistical results for the indicated metrics, however, they were removed from final models to simplify interpretation. Parametric model assumptions were evaluated, and data transformations were performed to meet model assumptions as required. Post-hoc comparisons were performed using Tukey’s HSD and simple-effects tests. To quantify effect sizes, partial eta-squared ($\eta_p^2$) was computed. RStudio (version 1.0.136) statistical software (RStudio, Inc., Boston, MA) was used to analyze significant interactive effects using the lsmeans and lmer functions respectively from the lmerTest (Kuznetsova et al., 2016) and lme4 (Bates et al., 2014) packages. Other analyses were performed in JMP Pro 12 (SAS, Cary, NC). All statistical tests were considered significant when $p<0.05$. 
Results

Quality Measure

Errors were significantly affected by Exo and Precision (Table 9). Specifically, designs with mechanical arms (Full and Arm) increased the number of errors made (Figure 22). Overall, increasing precision requirements led to higher errors, regardless of the Exo condition; compared to the Low condition, errors increased by ~70 and 500% in the Middle and High conditions, respectively. Although the Exo x Precision interaction was not significant (Table 9), the Shl design appeared to increase errors especially in the High precision task compared to the No condition.

Table 9. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on errors and ratings of perceived exertion (RPE). Any transformations used on the dependent measures are specified. P values and effect sizes ($\eta^2_p$) are reported for each effect in parentheses, with bold font indicating significant effects ($p<0.05$).

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>Precision (P)</th>
<th>E x P</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>P x G</th>
<th>E x P x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>Log</td>
<td>0.006</td>
<td>&lt;0.0001</td>
<td>0.280</td>
<td>0.298</td>
<td>0.188</td>
<td>0.890</td>
<td>0.226</td>
</tr>
<tr>
<td>RPE hand/wrist</td>
<td>--</td>
<td>0.162</td>
<td>0.723</td>
<td>0.891</td>
<td>0.206</td>
<td>0.361</td>
<td>0.748</td>
<td>0.708</td>
</tr>
<tr>
<td>RPE upper arm</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.320</td>
<td>0.785</td>
<td>0.227</td>
<td>0.319</td>
<td>0.246</td>
<td>0.629</td>
</tr>
<tr>
<td>RPE shoulder</td>
<td>--</td>
<td>0.217</td>
<td>0.023</td>
<td>0.031</td>
<td>0.260</td>
<td>0.035</td>
<td>0.028</td>
<td>0.042</td>
</tr>
<tr>
<td>RPE neck</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.415</td>
<td>0.869</td>
<td>0.491</td>
<td>0.097</td>
<td>0.606</td>
<td>0.747</td>
</tr>
<tr>
<td>RPE low back</td>
<td>--</td>
<td>0.338</td>
<td>0.018</td>
<td>0.024</td>
<td>0.127</td>
<td>0.061</td>
<td>0.010</td>
<td>0.034</td>
</tr>
<tr>
<td>RPE thigh</td>
<td>--</td>
<td>0.031</td>
<td>0.409</td>
<td>0.893</td>
<td>0.066</td>
<td>0.369</td>
<td>0.939</td>
<td>0.989</td>
</tr>
<tr>
<td>RPE lower leg and foot</td>
<td>--</td>
<td>&lt;0.0001</td>
<td>0.339</td>
<td>0.999</td>
<td>0.183</td>
<td>0.095</td>
<td>0.450</td>
<td>0.986</td>
</tr>
</tbody>
</table>
Figure 22. Quality measure (total number of errors made) for the four Exo conditions. Letters indicate groupings obtained from pairwise comparisons between Exo levels. Error bars show 95% confidence intervals.

**Ratings of Perceived Exertion Measures**

Mean RPEs overall ranged from 1.8 in the neck to 4.0 in the shoulder, respectively equivalent to “weak” and more than “moderate” levels of perceived exertion. Most RPE metrics were significantly influenced by Exo, while none was affected by Precision (Table 9). Compared to the No condition, the Shl reduced RPEs in the shoulder, upper arm, and low back (Figure 23). While not changing the RPEs in the upper arm and shoulder, designs with mechanical arms (Full and Arm) increased RPEs in the low back and thigh compared to the No condition (Figures 23d and e). Among males, the Full design led to a significantly higher RPEs in the legs and feet (Figure 23f).
Across all conditions, mean values of the peak, median, and static EMG loading were 66.1, 20.1, and 6.5%, respectively. While all summation metrics were affected by Exo, only the total median was influenced by Precision (Table 10). Compared to the No condition, the Shl design significantly reduced the total peak and total median loadings (Figures 24a and b). Devices with mechanical arms (Full and Arm) increased the total median and total static loadings (Figures 24b and c). Regarding the main effect of Precision on the total median loading, it increased respectively by ~5 and 11% in the Middle and High conditions, compared to the Low condition. For the main effect of Gender on the total peak, median, and static loadings, females experienced ~58, 77, and 125% higher loading than males, respectively.

**Total EMG Loading Measures**

Figure 23. Ratings of Perceived Exertion (RPE) for several body parts. Letters indicate groupings obtained from pairwise comparisons between Exo levels (among males in figure f). Error bars show 95% confidence intervals. The symbol * specifies a significant difference between genders within Exo conditions in figure f.
Table 10. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the total peak, median, and static loading EMG measures. Any transformations used on the dependent measures are specified. \( P \) values and effect sizes (\( \eta^2_p \)) in parentheses are reported for each effect, with bold font indicating significant effects (\( p<0.05 \)).

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>Precision (P)</th>
<th>E x P</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>P x G</th>
<th>E x P x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak</td>
<td></td>
<td>0.002</td>
<td>0.553</td>
<td>0.965</td>
<td>0.001</td>
<td>0.345</td>
<td>0.748</td>
<td>0.883</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.135)</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.809)</td>
<td>(0.033)</td>
<td>(0.006)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>&lt;.0001</td>
<td>0.041</td>
<td>0.534</td>
<td>0.004</td>
<td>0.308</td>
<td>0.802</td>
<td>0.952</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.475)</td>
<td>(0.062)</td>
<td>(0.049)</td>
<td>(0.756)</td>
<td>(0.036)</td>
<td>(0.004)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Static</td>
<td></td>
<td>&lt;.0001</td>
<td>0.222</td>
<td>0.654</td>
<td>0.012</td>
<td>0.006</td>
<td>0.455</td>
<td>0.729</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.417)</td>
<td>(0.030)</td>
<td>(0.040)</td>
<td>(0.671)</td>
<td>(0.118)</td>
<td>(0.016)</td>
<td>(0.035)</td>
</tr>
</tbody>
</table>

Figure 24. Total (across all monitored muscles) peak, median, and static EMG loading measures, each as percentages of MVC. Letters indicate groupings obtained from pairwise comparisons between Exo levels. Error bars show 95% confidence intervals. The symbol * specifies a significant difference between genders within Exo conditions for static values.

**Peak EMG Loading Measures**

Across all conditions, mean levels of peak loading ranged from 5.6% in the ILL-ND to 20.3% in the AD-D. While none of the peak loading measures was significantly affected by Precision, most
were influenced by Exo (Table 11). Compared to the No condition, the Shl design significantly reduced peak loading of the AD-ND while not changing bilateral measures of ILL and RA (Figure 25). The Arm design increased loading on the ILL-ND and RA-ND, while reducing the loading on ILL-D compared to the No condition (Figure 25). The Full condition led to a significantly lower peak loading on RA-D but increased loading on the bilateral measures of ILL. Regarding the main influence of Gender on AD-D, AD-ND, and RA-D (Table 11), females experienced ~130, 70, and 110% higher peak loading than males, respectively. The influence of Precision on the ILL-ND peak loading approached significance ($p = 0.07$); compared to the Low condition, it increased by ~2.2 and 15% in the Middle and High conditions, respectively.

Table 11. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the peak loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$).
Figure 25. Peak loading measures as percentages of MVC. Letters indicate groupings obtained from pairwise comparisons between Exo levels. Error bars show 95% confidence intervals.

**Median EMG Loading Measures**

Mean median loading overall varied from 1.6% in the RA-D to 6.0% in the AD-ND. While all median loading measures were significantly affected by Exo, none was affected by Precision (Table 12). Compared to the No condition, the Shl design significantly reduced median loading of the bilateral measures of AD, but significantly increased ILL-D loading in males (Figure 26). Designs with mechanical arms (Full and Arm) increased loading on AD-ND and ILL-ND (Figures 26d and e) compared to the No condition, although the difference with the Full for the latter muscle was not significant for females ($p = 0.68$). While the Full significantly reduced AD-D loading, this condition yielded a significantly higher loading on the ILL-D (Figures 26a and b). For the RA muscles, the Shl led to a significantly lower median loading than the Arm condition (Figures 26c
and f), although this difference was evident only in females for RA-ND. Regarding the main influence of Gender on AD-D, AD-ND, and RA-D (Table 12), females exhibited ~140, 100, and 160% higher median loading than males, respectively. The effect of Precision on AD-D median loading approached significance ($p = 0.09$); compared to the Low condition, this loading increased by ~9 and 18% in the Middle and High conditions, respectively.

Table 12. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the median loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$)

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Trans.</th>
<th>Exo (E)</th>
<th>Precision (P)</th>
<th>E x P</th>
<th>Gender (G)</th>
<th>E x G</th>
<th>P x G</th>
<th>E x P x G</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD-D</td>
<td>Square root</td>
<td>&lt;0.0001</td>
<td>0.091</td>
<td>0.808</td>
<td>0.024</td>
<td>0.126</td>
<td>0.882</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.245)</td>
<td>(0.047)</td>
<td>(0.029)</td>
<td>(0.688)</td>
<td>(0.056)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>AD-ND</td>
<td>Square root</td>
<td>&lt;0.0001</td>
<td>0.337</td>
<td>0.864</td>
<td>0.045</td>
<td>0.253</td>
<td>0.603</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.480)</td>
<td>(0.022)</td>
<td>(0.025)</td>
<td>(0.557)</td>
<td>(0.040)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>ILL-D</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>0.662</td>
<td>0.999</td>
<td>0.034</td>
<td>0.927</td>
<td>0.797</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.590)</td>
<td>(0.008)</td>
<td>(0.004)</td>
<td>(0.083)</td>
<td>(0.002)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>ILL-ND</td>
<td>Log</td>
<td>&lt;0.0001</td>
<td>0.317</td>
<td>0.989</td>
<td>0.849</td>
<td>0.029</td>
<td>0.917</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.492)</td>
<td>(0.023)</td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.087)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>RA-D</td>
<td>--</td>
<td>0.020</td>
<td>0.295</td>
<td>0.993</td>
<td>0.031</td>
<td>0.308</td>
<td>0.568</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.094)</td>
<td>(0.024)</td>
<td>(0.008)</td>
<td>(0.939)</td>
<td>(0.036)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>RA-ND</td>
<td>--</td>
<td>0.002</td>
<td>0.195</td>
<td>0.961</td>
<td>0.076</td>
<td>0.036</td>
<td>0.529</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.135)</td>
<td>(0.032)</td>
<td>(0.014)</td>
<td>(0.846)</td>
<td>(0.082)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>
Mean static loading across all conditions ranged from 0.82% in the ILL-D to 1.5% in the AD-ND. While all static loading metrics were significantly affected by Exo, only ILL-ND was significantly influenced by Precision (Table 13). Most results for the static loading metrics were similar to those explained earlier for the median values. However, a few differences were found. Compared to the No, the Full and Arm designs significantly increased static loading of the AD-D only among males and females, respectively (Figure 27a). The Shl condition was not significantly different from the No in terms of static loading of the AD-ND ($p = 0.99$; Figure 27d) and ILL-D among males ($p = 0.056$; Figure 27b). For females, the Full condition led to significantly higher ILL-ND
static loading than the No ($p=0.025$; Figure 27e). Regarding the main effect of Precision on ILL-ND static loading (Table 13), compared to the Low condition it increased by $\sim 3$ and $15\%$ in the Middle and High conditions, respectively. For the main effect of Gender on the AD-D and AD-ND (Table 13), females experienced $\sim 190$, and $240\%$ higher static loading than males, respectively.

Table 13. Summary of statistical results for the main and interactive effects of Exo, Precision, and Gender on the static loading EMG measures. Any transformations used on the dependent measures are specified. $P$ values and effect sizes ($\eta^2_p$) in parentheses are reported for each effect, with bold font indicating significant effects ($p<0.05$).
The major objective of this study was to compare three different exoskeletal designs, each intended to support the upper extremity, under varying levels of task precision requirements. This comparison was done in terms of physical demands (as quantified by objective and subjective measures) and quality (as a dimension of performance), and in the context of a simulated drilling task that required arm elevation. As discussed below, exoskeletal designs differentially affected physical demands and quality, with these effects being fairly consistent across the examined precision levels.

Discussion

The major objective of this study was to compare three different exoskeletal designs, each intended to support the upper extremity, under varying levels of task precision requirements. This comparison was done in terms of physical demands (as quantified by objective and subjective measures) and quality (as a dimension of performance), and in the context of a simulated drilling task that required arm elevation. As discussed below, exoskeletal designs differentially affected physical demands and quality, with these effects being fairly consistent across the examined precision levels.
**Influences of Varying Precision Requirements**

Across all three Exo conditions (and the manual condition), increasing Precision led to increases in some metrics of muscle activity. Specifically, there were increases in the total median loading, static and peak loading of the ILL-ND, and median loading of the AD-D. We did not, however, detect significant effects of Precision on the majority of physical demand metrics, potentially because such effect depends on the task “speed.” Here, the task pace was controlled at a level that seemingly was not high enough to detect broader Precision effects on physical demands. However, we set the pace at level that was considered likely to be achievable by most participants and to induce minimal fatigue, and which perhaps was an occupationa realistically pace. Supporting that Precision effects depends on speed, Laursen et al., (1998) found that the influence of Precision on the loading of several muscles loading was dependant on movement speed in a repetitive hand movement task; in contrast to two higher speeds, Precision had little or no effect on muscle loading at a lower speed. The lack of a substantial effect of Precision on muscle loading here is also consistent with the lack of such an effect on perceived exertion.

For all Exo conditions (and the manual condition), increasing Precision led to more errors, and designs with mechanical arms (*Full* and *Arm*) exhibited larger effects (Figure 22). We suggest that this higher sensitivity of designs with mechanical arms to increasing Precision is because the tool was attached indirectly to the wearer’s body, making quality performance more dependant on whole body stablity. Additionally, the asymmetric positioning of the mechanical arms likely led to asymmetric external moments on the torso, potentially causing deviations from the target (i.e., errors). Use of lighter mechanical arms may mitigate this influence, and perhaps attenuate asymmetric loading on the wearer (discussed below). Further, the mounting system employed
here to attach the tool to the mechanical arms (cf. Figure 21) may have contributed to increased errors by impairing visual access. A few participants also reported that their dominant arm (particularly the elbow) occasionally hit the mechanical arm, and this may also have contributed to increased errors. The Shl design appeared to impact quality only at the High precision level; a few participants reported that the onset of support from the device, as the arms were raised, was not “gentle”, an action that may have contributed to increasing errors in the noted condition. Using another vest similar to the current Shl condition, specifically the Levitate Airframe™, performance was reported to increase in a mock overhead precision task (Spada et al., 2017a; Spada et al., 2017b) and in a pilot experiment of a mock painting and welding tasks (Butler, 2016). It is unclear whether these inconsistent results regarding performance vs. precision are due differences in exoskeleton design or the specific simulated tasks demands.

**Influences of Different Exoskeletal Designs**

The Full design reduced the peak values of the total loading metric (Figure 24a) and bilateral measures of RA (Figures 25e and f), but increased the total static and median measures (Figures 24b and c), all bilateral measures of ILL, and lower extremity RPEs (Figures 23e and f). The total static and median measures may have increased due to the relatively heavier weight of this device. This design also attaches portions of the exoskeleton to the wearer’s lower legs (to transmit loads to the ground), which likely accounts for the increased loading on this region. It should be noted, however, that the specific exoskeleton included here was designed for use with work boots. Because offering different size boots was difficult, we added a few additional holes in the original foot straps to allow for using smaller shoes. We then standardized the experiment by using only one type of sports shoes (that were available in our lab). Although considered unlikely, this change
may have contributed to the overall increase in physical loading, specifically the increased RPEs in the lower extremity in this condition.

The Arm design increased static and median values of the total loading metrics (Figures 24b and c) and the non-dominant AD and ILL, while reducing peak loading on the ILL-D (Figure 25b). Likely due to the device weight, there was a general increase in the noted total loading measures. Because the mechanical arm was positioned on the dominant side (and to the dorsal side of wearer; Figure 21b), increased loading on the AD-ND may have stemmed from asymmetric demands imposed to control the tool. More specifically, the increased ILL-ND loading could have occurred to counteract the asymmetric external moment created by the tool and mechanical arm weights (acting on the dominant side). Supporting the increased low back muscle loading observed here, Rashedi et al., (2014) found a similar increase in a study that examined a similar exoskeletal vest in an analogous overhead work task simulation.

The Shl design reduced the values of total peak and median loading metrics (Figures 24a and b), median bilateral measures of AD, and RPEs in the shoulder, upper arm, and the low back, while increasing median loading of the ILL-D for males. Regarding the latter, we speculate that males may have changed their working strategy during the Shl condition and recruited the ILL-D more when generating the required drilling force. This difference in strategy, if present, may also have contributed to the reduced shoulder demands in the Shl condition among males. Contrary to this explanation, however, the Shl design led to lower RPEs of the low back and shoulder for both gender (Table 9). Perhaps due to its relatively lighter weight, the Shl device did not affect the total static metric compared to the No condition (Figure 24c).
Comparison to Results Obtained in Chapter 3

In the earlier study, the same exoskeletal design approaches were compared in terms of physical demands (muscle loading, perceived discomfort, and maximum acceptable frequency of drilling) and quality (one precision level corresponded to the Middle precision level here) using two tool masses (note that the study here used the lighter tool mass from that study). In addition to the difference in work height, the “drilling” exertions were upward (vs. forward here), Shl design was represented by the ShoulderX™ exoskeleton (vs. EksoWorks™ here), the mechanical arm in the Arm condition was connected to the ventral side of the wearer’s body, and task duration was 18 minutes. Of that duration, the initial 15 minutes were used for frequency adjustment and latter 3 minutes continued the repetitive task at the final MAF. During the latter 3 minutes, loadings on several muscles were recorded. In terms of monitored muscles, only the AD and ILL muscles were monitored in both studies.

In terms of quality, the general pattern of outcomes was consistent across the two studies. Yet, the Shl design led to a number of errors similar to those found in the No condition in the earlier study. This can be attributed to the examined precision level in that study. In other words, having the highest precision level here led to increasing the overall difference between the Shl and No conditions.

In terms of physical loading, participants reported earlier that the Full device led to the highest discomfort in the lower extremity. Although only evident here among males, this condition (i.e., the Full) led to the highest perceived exertion in the lower extremity (recall that perceived discomfort was not collected here given the relatively shorter task duration). These findings
suggest that the noted discomfort likely resulted from an increased exertion level in the lower extremity. For loading on the shoulder, direct comparison of results was possible only for the AD muscle (recall that monitoring MD activity was infeasible here). In both studies, the Shl device appeared to most substantially reduce AD peak loading. For both static and median loading of AD-ND, similar results were found: devices with mechanical arms increased loading. This increased loading was potentially needed to control the tool, as discussed earlier (recall that the mechanical arms were positioned in the dominant side). For median loading of AD-D, results were generally similar except that females in the prior study had substantially higher loading in the Full condition compared to the corresponding value obtained here. This difference might be attributed to the difference in the work height between the two studies (recall that work height was higher in Study 2). For static loading of AD-D, results were generally similar, except that females experienced higher loading in the Full condition in the earlier study, while here they experienced the highest loading in the Arm condition. These findings further support that the Arm and Full devices affect shoulder loading differently depending on the work height.

In terms of ILL peak loading, the two studies led to consistent results, in that devices with mechanical arms increased loading of ILL-ND while only the Full increased the ILL-D. For median loading of ILL, both studies found that the Arm condition leads to the highest loading on the ILL-ND and that the Full leads to the highest loading of ILL-D. For static loading of the ILL-D, Exo conditions were not significantly different in the earlier study, while here the Full appeared to increase loading. For static loading of ILL-ND, both studies were generally consistent in finding that the Arm device caused the largest increase in loading. Because of the consistent results found for ILL, the earlier discussion applies here as well. Regarding the Arm condition, we note that
changing the mechanical arm positioning did not seem to change the general pattern of results (regarding both performance and physical loading). Yet, caution should be exercised, because the two studies used different task demands (i.e., confounding effects).

**Limitations**

There are a few limitations in this study that should be indicated. While use of a controlled experiment increases internal validity, the extent of external validity is unknown, as discussed in Alabdulkarim et al., (2017) for a largely similar experimental setup. The current study also compared ekoskeletal designs in the context of only a single simulated task (e.g., one working height, one required drilling force, and one tool mass). As such, it is unclear if the pattern of observed results might change for different task demands. For example, the *Full* design might be more effective particularly when using heavier tools. Additionally, only one exoskeleton was examined from three different design approaches. The extent to which the current results, using the selected devices, generalizes to other exoskeletons from the same design, is unknown. Although we provided training sessions, it is also unknown if same pattern of results will be obtained for more experienced workers. To objectively evaluate physical demands thru measuring muscles activation, only a few muscles were accessible across the examined exkoskeletons. For example, it was infeasible to monitor several other muscles in the shoulder girdle that are important for such overhead work. It is thus unclear if results from the muscles we monitored will generalize to other muscle groups.
Conclusions

In summary, three distinct exoskeletal designs differentially affected physical demands and quality performance, with this effect being fairly consistent across three levels of precision demands. Designs with mechanical arms (Full and Arm) increased loading on the low back and deteriorated quality performance. Designs that mainly supported the shoulder (Shl) appeared to reduce quality at the highest level of precision demands, while reducing overall muscular loading at the upper extremity. Given the growing interest regarding exoskeletons for occupational use, this study highlights a need for future research to simultaneously consider physical demands and task performance and under occupationally relevant factors with more diverse and realistic conditions.

References


Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol, 47*(6), 381.


Chapter 5: Conclusion and Recommendations

The primary objective of this dissertation was to examine the dual influences of different types of ergonomic interventions, in terms of performance and injury risk. Three experiments were completed involving a simulated drilling task. This specific task was considered informative, as it entailed diverse demands (precision, strength, and speed) and permitted quantifying two dimensions of task performance (productivity and quality). Two distinct types of potential ergonomic interventions were investigated: 1) adding *adjustability*, as a commonly advocated approach when considering ergonomics early in the (re)design phase to change *task demands*; and, 2) using *exoskeletons* to potentially enhance *worker capacity*. For the former, the study results were expected to be helpful in estimating the extent of adjustability beyond which dual benefits could be achieved (if any). For the latter, results were anticipated to be valuable for efforts at developing future guidelines for the occupational use of exoskeletons, which consider both of the dual benefits into account, and address three relevant aspects of task design (maximum acceptable frequency, tool mass, and precision requirements). Contributing to such future guidelines is particularly important as currently there are none and existing evidence is very limited. Further, results of this work can inform (re)designing exoskeletons to better achieve the dual benefits. Overall, this research addresses a practical challenge involved in considering occupational proposals of such interventions, specifically justifying their (sometimes considerable) costs. Thus, knowing/estimating the influence of ergonomic interventions on performance can be beneficial to financially justify them, and potentially reduce the risks of WMSDs.
Effects of Workstation Adjustability

The dual influence of three levels of workstation adjustability were examined in Chapter 2. It was hypothesized that increasing workstation adjustability, as a variable amendable to (re)design, would improve performance and reduce ergonomic risk, and that there would be an inverse relationship between performance and ergonomic risk. Supporting our hypothesis, the primary finding from this study was that both productivity and quality significantly improved when increasing adjustability. Yet, improved performance (both productivity and quality) occurred only when a given level of adjustability also succeeded in reducing ergonomic risk. Supporting the inverse association between ergonomic risk and performance, the condition with the highest level of adjustability created the lowest ergonomic risk and the best performance. In contrast, there was no substantial difference in ergonomic risk between the two lower adjustability conditions, in which performance was also comparable. Across all conditions, several significant linear associations were found between risk (e.g., Strain Index score) and performance metrics (e.g., completion time), further supporting the inverse relationship.

Effects of Exoskeleton Design

The dual influence of three distinct passive exoskeletal designs was investigated/compared in Chapters 3 and 4. Three designs were included: full-body (Full) and upper-body (Arm) exoskeletons with attached mechanical arms; and an upper-body (Shl) exoskeleton providing primarily shoulder support. It was hypothesized that devices (with varying levels of maturity) representing these design approaches would lead to different dual effects depending tool mass (~2 and ~5 kg; Chapter 3) and depending on task precision requirements (Chapter 4). Both designs with mechanical arms increased static and median total muscle activity (Chapter 4) while
deteriorating quality performance (Chapters 3 and 4). Comparing these specific designs (i.e., the Full and Arm), the Arm device led to a higher maximum acceptable frequency (a metric of physical demands; Chapter 3) than that of the Full device among females, and the latter notably increased loading on the lower extremity. Increases in the latter metric (i.e., maximum acceptable frequency) can also be interpreted as reflecting potential increases in productivity. Both conditions (i.e., the Full and Arm) increased loading on the low back specifically (Chapters 3 and 4), though this effect was partially alleviated when the design allowed for counterbalancing the load and transferring it to the ground (i.e., the Full device; Chapter 3). The exoskeleton design that mainly supported the shoulder (i.e., the Shl device) reduced median and peak total loading (Chapter 4), with these reductions generally occurring in the shoulder region (Chapters 3 and 4). Yet, this design (i.e., the Shl) increased median and static loading of the dominant upper arm (i.e., TB-R muscle; Chapter 3), and deteriorated quality in the condition with the highest precision requirements (Chapter 4). The influences of increasing precision (Chapter 4) and tool mass (Chapter 3) were fairly consistent across the examined designs, and was evident as increased physical loading, with precision affecting the latter to a lesser degree. As such, no single design was obviously superior in both physical demands and performance.

For (re)designing occupational exoskeletons, specific design recommendations can be made from results obtained in the two studies:

- Given that the weight of the Full mechanical arm was almost double the weight of that of the Arm, more shoulder static loading was evident in that condition (i.e., the Full). Therefore, if the design approach includes a mechanical arm, reducing the weight of this mechanical arm is strongly encouraged. The increased mechanical arm weight was also
found to increase loading on the low back and lead to more deviations from target (potentially reducing quality performance).

- Another potential for improving design was found related to the load transfer and counterbalancing used in the Full design for reducing low back loading. However, the tradeoff that should be recognized/addressed for this mechanism is the increased discomfort in the lower extremity of the wearer. Additionally, because the counterbalancing weights are located out of the visual field of the wearer, caution is needed when designing exoskeletons for use in confined spaces, as there may be safety implications.

- The asymmetric positioning of the mechanical arms resulted in more unloading of the dominant arm. If the design remains asymmetric (a clearly undesirable design characteristic for both performance and physical demands), it is suggested to reduce task demands on the non-dominant arm to the possible extent when using these exoskeletons.

- When using devices with mechanical arms, it was difficult to unload the tool between holes to facilitate recovery. Therefore, the tool was always connected to the wearer when performing the task. In practice, this: 1) can be problematic when workers need to use different tools for a given task; and 2) was found to increase overall static loading for these conditions. Thus, when designing a mechanical arm, a more “efficient” tool (un)loading mechanism would be beneficial.

- While the Shl device was found useful for reducing shoulder demands, this device did not reduce demands on the hand/wrist (unlike the Full and Arm devices). Demands on this specific body part might be the limiting factor for increasing a wearer’s capacity. Thus, it may be useful to expand the targeted body parts in this design approach.
• When using higher support levels in the Shl designs, returning the arm to a neutral posture after performing an arm elevation task required physical effort (to overcome support provided by the devices). This required effort may lead the wearer to “rest” the arms in a non-neutral posture and lead to increased physical loading over longer durations. Thus, there is need to develop a mechanism by which the wearer can obtain the “desired” support and to release “undesired” support.

**Future Directions**

Future research is warranted to investigate/quantify the dual influences of ergonomic interventions. Along with research to address the previously noted limitations, future work should consider factors that can modify the balance between task demands and worker capacity, as imbalances between these two aspects can lead to occupational injuries. Among other influential factors, obesity and aging are growing concerns in the contemporary workforce and can influence the noted balance. Because testing here was highly controlled to maximize internal validity, intervention testing should also be done under more realistic conditions (e.g., field testing). For example, a possible avenue for studying exoskeletons effectiveness is examining them under varying demands (vs. the current simulated task that was designed to have constant demands). Such demands might be imposed by, for examples, different working heights, horizontal distances, workpiece configurations, and number of tools used in a given task. Regarding the latter (i.e., number of tools used), workers may need several tools to complete a given task in practice. It is currently somewhat cumbersome, though, to assemble a tool for use with mechanical arms. Therefore, frequently changing the tool might not be effective, and workers in practice may occasionally avoid using the exoskeletons to skip these preparatory steps (as is the case for many types of personal protective equipment).
Appendices

Appendix A: Participants Instructions (modified from Snook et al., 1995)

Instructions for the drilling task:
The rung has six holes. For an 18-minute duration, you will simulate drilling these holes back and forth at a pace that you control (described next). For a given hole, completing the drilling task requires pushing the probe into the hole above a target force and maintaining this force for a total of 2.5 seconds. We want you to complete the drilling task as quickly and as accurately as possible. Your performance is defined by the number of both “errors” you generate (quality) and holes you complete (productivity). Quality and productivity are equally important. You will hear error beeps when there is contact between the drill “bit” and the surrounding hole (sound 1). We want you to minimize these errors (beeps) to the extent possible. When you generate a sufficient force, you will hear a drilling sound (sound 2). When a hole is completed (the noted 2.5 seconds was elapsed), you will hear a completion sound (sound 3). Immediately after hearing this sound, try to move to the next hole. Start drilling the next hole as soon as you hear the pacing sound (sound 4).

Instructions for setting the pace in the initial 15 minutes:
This task is intended to simulate a paced production line. Please imagine that the shift length in this line is two hours. We want you to select the maximum acceptable pace that you believe would not cause numbness, pain, or excessive discomfort after a 2-hour shift duration in any part of your body. You will adjust the pace by saying out loud “a little faster”, “faster”, or “a lot faster” and this corresponds to three levels of adjustments to your current pace. Slowing the pace is done in a similar manner. To return to a previous pace after you have tried a specific change, you may say “go back”. Knowing your maximum acceptable pace is not an easy task. Only you know your feelings.

If you feel the pace is too fast, do not hesitate to reduce it as indicated above. However, we also do not want you to select a pace that is too slow. If you feel you can increase the pace, please do so as directed above.

Do not be afraid to make adjustments. You have to make enough adjustments so that you get a good feeling for what is too fast and what is too slow. You can never make too many adjustments, but you can make too few.

Remember, this is not a contest. Each person has his/her own maximum acceptable working pace. Here, we want your adjustment to indicate how fast you can perform the drilling task without getting numbness, pain, or excessive discomfort in any body part, but also while keeping errors to a minimum. After the initial 15 minutes, we want you to continue the same task but without changing the last pace that you have converged to for three minutes.
Appendix B: Usability and Discomfort Questionnaire

1. Usability questions (This section was completed immediately after an exoskeleton was used)

For the condition you just completed, how well do you agree with the following statements? Please circle the appropriate number.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>It meets my needs for completing the task</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>It makes the things I want to accomplish easier to get done</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>It is easy to use</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I am satisfied with it</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>It works the way I want it to work</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Overall, I prefer working with it more than without it</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

2. Discomfort

For the condition you just completed, please use the scale and diagram in front of you to rate the level of discomfort you experienced in the following body parts during the task. For body parts with two sides (e.g., left and right legs), please rate the worst side of the two.

<table>
<thead>
<tr>
<th>Body Part</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand/wrist</td>
<td></td>
</tr>
<tr>
<td>Upper Arm</td>
<td></td>
</tr>
<tr>
<td>Shoulder</td>
<td></td>
</tr>
<tr>
<td>Neck</td>
<td></td>
</tr>
<tr>
<td>Low Back</td>
<td></td>
</tr>
<tr>
<td>Thigh</td>
<td></td>
</tr>
<tr>
<td>Lower Leg and Foot</td>
<td></td>
</tr>
</tbody>
</table>

If you rated any body part more than three above, please explain what you think is the source of that discomfort.
(The following questions were presented after all experimental conditions were completed)

a) Do you have any suggestion or concern about the usability and design of any exoskeleton?

b) Please rank the four conditions from highest (# 4) to lowest (# 1) physically demanding conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without device</td>
<td></td>
</tr>
<tr>
<td>Device A</td>
<td></td>
</tr>
<tr>
<td>Device B</td>
<td></td>
</tr>
<tr>
<td>Device C</td>
<td></td>
</tr>
</tbody>
</table>