

ABSTRACT

MA, WENQI. An Application of Quantitative Methods for Motor Ability Level Classification, Performance Prediction and Training Protocol Selection. (Under the direction of Dr. David B. Kaber.)

Over the past four decades, the application of automation technology in industrial operations has expanded dramatically, primarily due to the development of advanced computer technology. However, manual work is still required in manufacturing, especially in assembly processes. To achieve high production rates, assembly operators are required to develop specialized motor skills. Numerous research studies have been conducted to define training protocols for novice operators to support individual motor skill development to maximum achievable levels. The effectiveness of these training approaches has also been validated in prior laboratory research. However, in real-world applications, such as manual assembly on a production line, it is more commonly expected that operators are trained to a uniform performance level (i.e., “normal” or 100% performance) rather than each operator achieving his/her highest skill level in order to prevent bottlenecks or work-in-process inventory accumulation in production operations. Therefore, there is a need to identify appropriate methods to classify novice operators based on their initial motor performance and to assign them to suitable training protocols facilitating different levels of skill development towards 100% performance.

The objective of this research was to develop an algorithm for classification of operator motor ability on the basis of baseline performance in a simple motor-control test. The research was also to specify appropriate virtual reality (VR)-based training methods, on the basis of skill classification, such that operators might achieve desired levels of motor performance in a real-world design assembly task using a computer-mediated environment.

The study followed a two-phase experimental approach. In the first phase, a batch of 21 right-handed participants was recruited for the computer-based motor test performance along with completion of a standardized psychomotor test (in physical form). The results of the standardized test (Purdue Pegboard) were used as a “gold standard” to validate the computer-based motor test for automatically assessing participant motor ability. A statistics-based model was then developed to classify participant motor ability level using as inputs a set of features based on kinematic parameters generated through the computerized motor test. The finalized model achieved a classification accuracy of ~98% using a 75/25 cross validation approach.

In the second phase, a different batch of 36 right-handed participants was recruited to perform the same computer-based motor test. Performance results were input to the classification model to predict motor ability level. Based on the classification results (“high”, “medium” and “low” ability level), each participant received one-of-three pre-designed haptic-VR training protocols, including consistent haptic guidance in motor movements, resistive haptic forces counter to movements, and random haptic disturbances. Prior research has demonstrated varying effects of these methods on training performance with some superiority of haptic disturbances. Results revealed participants identified as “medium” or “low” were able to achieve levels of motor performance comparable to “high” participants through 1-hour training with a VR-based 2D pattern assembly task. Results further validated the accuracy of the motor ability classification algorithm.

The findings of this study verified a quantitative motor skill classification algorithm and an approach to designing and assigning proper training protocols for novice operators to achieve comparable levels of motor performance. This methodology could be applied to

operator training for real-world manual assembly operations and to ensure that a group of workers achieve uniform performance levels.

© Copyright 2014 by Wenqi Ma

All Rights Reserved

An Application of Quantitative Methods for Motor Ability Level Classification, Performance
Prediction and Training Protocol Selection

by
Wenqi Ma

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Industrial Engineering

Raleigh, North Carolina

2014

APPROVED BY:

David B. Kaber
Committee Chair

Yuan-Shin Lee

Chang S. Nam

Peter Bloomfield

Nancy J. Currie

DEDICATION

To my dear mother!

BIOGRAPHY

Wenqi Ma obtained her Bachelor's degree in Industrial Engineering and Technology Management from University of Hong Kong in 2010. In the same year, she travelled to U.S. to study Human Factors and Ergonomics at North Carolina State University under the supervision of Dr. David B. Kaber. Her research interests include rehabilitation engineering, motor skill assessment and training, and motor ability classification.

ACKNOWLEDGMENTS

I would like to firstly express my appreciation to my supervisor, Dr. David Kaber, for his continuous direction and guidance during the past four years. Dr. Kaber has always been generous in providing professional advice and cheerful encouragement. With the inspiring and precious research experience of working with him, I was able to conquer those challenges in completing my study.

Secondly, my sincere gratitude goes to my committee members, Dr. Nam, Dr. Lee, Dr. Bloomfield and Dr. Currie, who have offered time and effort in supporting me with this dissertation. I would like to thank all my committee members for helping me improve my research with professional suggestions.

Also, I am very grateful to my family. To my dear parents, thank you for your love and caring, which have been the greatest motivation for me to overcome any difficulties and achieve a better self. To my beloved husband, thank you for giving me your unconditional support and companionship. It is your enthusiasm that has always cheered me up when facing frustrations.

In addition, I would like to express my gratitude to my esteemed coworkers, Dr. Michael Clamann and Dr. Wooram Jeon. Michael helped me improve my writing and Wooram helped me develop the computer programs applied in this research. I deeply thank both of them for their generous sharing of knowledge and experience.

Last but not least, I would like to thank the National Science Foundation for supporting my Ph.D. program through a grant to NCSU Ergonomics Lab (IIS-0905505). In addition, I thank the NC Ergonomics Center for comparable support of my studies.

TABLE OF CONTENTS

LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
LIST OF ACRONYMS.....	xi
CHAPTER 1: INTRODUCTION.....	1
1.1. Manual assembly training methods	1
1.1.1. Manual assembly requirements and conventional training.....	1
1.1.2. Review of AR technology applied in assembly training	3
1.1.3. Application of VR technology in assembly training.....	5
1.2. Manual assembly learning process	9
1.3. Challenge of applying lean concepts to a manual assembly production line	12
1.4. Research motivation.....	14
CHAPTER 2: LITERATURE REVIEW	16
2.1. Motor skill training protocols	16
2.1.1. VR-based training systems with haptic guidance	16
2.1.2. Robotic-based training systems with haptic guidance	21
2.1.3. Training systems with haptic disturbances	24
2.2. Motor skill assessment.....	25
2.2.1. Clinical-based techniques	26
2.2.2. Robot-aided techniques.....	28
2.2.3. Systems with more dimensions.....	29
2.2.4. VR-based systems.....	31
2.3. Motor performance prediction	36
2.3.1. Theoretical potential to make performance prediction	36
2.3.2. Real-time performance prediction	41
2.4. Quantitative classification approaches.....	43
2.4.1. Design of task	43
2.4.2. Feature selection algorithms	45
2.4.3. Classification algorithms	48
CHAPTER 3: PROBLEM STATEMENT.....	55
3.1. Related work summary	55

3.2. Research objective	58
3.3. Research hypotheses	58
CHAPTER 4: PHASE I EXPERIMENT METHODOLOGY	60
4.1. Participants.....	60
4.2. Equipment and apparatus	60
4.3. Experimental tasks	61
4.3.1. Dice manipulation task	61
4.3.2. PPT.....	64
4.4. Procedures	64
4.5. Construction of model.....	65
4.5.1. Task performance parameters	65
4.5.2. Feature generation.....	67
4.5.3. Motor ability level.....	67
4.5.4. Construction of linear regression model	68
4.5.5. Development of non-linear algorithms	71
4.5.6. 75/25 cross validation method	76
CHAPTER 5: PHASE I DATA ANALYSIS AND RESULTS	79
5.1. Linear model	79
5.2. Non-linear algorithms	81
CHAPTER 6: PHASE II EXPERIMENT METHODOLOGY	83
6.1. Participants.....	83
6.2. Equipment and apparatus.....	84
6.3. Experimental tasks.....	84
6.3.1. Dice manipulation task	84
6.3.2. Basic Block Design (BD) task	84
6.3.3. BD task with haptic features	85
6.4. Procedures	86
6.5. Independent variables	88
6.5.1. CG condition.....	88
6.5.2. RF condition.....	88
6.5.3. HD condition.....	89
6.6. Dependent variables.....	90

6.7. Model application	91
6.7.1. Comparison of dice manipulation task, BD task, and physical assembly task	91
6.7.2. Training conditions assignment	95
6.7.3. Statistical methods of analysis and expected outcomes.....	96
CHAPTER 7: PHASE II DATA ANALYSIS AND RESULTS	100
7.1. Descriptive statistics	100
7.2. Within-group comparison between baseline and final tests	100
7.3. Between-group comparison of baseline and final test performance	102
7.4. Analysis of learning percentages in BD training	103
CHAPTER 8: DISCUSSION.....	105
8.1. Validation of dice manipulation task for assessing motor ability.....	105
8.2. Non-linear classification models.....	106
8.3. Selection of haptic-VR training protocols	109
CHAPTER 9: CONCLUSIONS	113
9.1. Limitations of present study.....	113
9.2. Future work	115
9.3. Research applications.....	117
REFERENCES.....	119
APPENDICES.....	130
Appendix A.....	131
Appendix B	132
Appendix C.....	134
Appendix D.....	136
Appendix E	140

LIST OF TABLES

Table 2.1. Summary of reviewed studies and algorithms	53
Table 2.2. Comparison of reviewed classification algorithms.....	54
Table 4.1. Levels of dice manipulation task difficulty	64
Table 4.2. Summary of task performance parameters	66
Table 5.1. Summary of predictors in the final linear model	80
Table 5.2. Summary of classification accuracy for various algorithms.....	82
Table 5.3. Summary of input variables in final ordinal SVM model	82
Table 7.1. Summary of group mean (SD) statistics in BD completion time (sec.)	100
Table 7.2. Summary of paired t-test results	102
Table 7.3. Summary of one-way ANOVA and Tukey's HSD results.....	103
Table 7.4. Summary of group mean (SD) in learning percentage (%)	104
Table 8.1. Comparison of classification model accuracy.	109

LIST OF FIGURES

Figure 2.1. Classification of participants in target-hitting task (Huegel, 2009)	44
Figure 4.1. Dice manipulation task	62
Figure 4.2. Unfolded virtual die.....	62
Figure 4.3. Forward selection algorithm for developing linear model	70
Figure 4.4. Forward selection algorithm for developing non-linear models	73
Figure 4.5. 75/25 cross validation approach	77
Figure 4.6. Phase I experiment procedures and data flow	78
Figure 5.1. Change of adjusted R^2 and error rate of linear model along with number of predictors.....	80
Figure 6.1. Virtual block design task	85
Figure 6.2. Motion flow chart for dice manipulation task	93
Figure 6.3. Motion flow chart for block design task.....	94
Figure 6.4. Motion flow chart for iPhone assembly task (based on online video, “iPhone 4: The entire assembly process,” 2010)	95
Figure 6.5. Phase II experiment procedures and data flow.....	99
Figure 7.1. Comparison of performance in baseline and final BD tests	101
Figure 7.2. BD training performance	104
Figure 8.1. Final performance of low and medium groups compared with baseline performance of high group.....	112

LIST OF ACRONYMS

2D	2-dimensional
3D	3-dimensional
AIC	Akaike Information Criterion
AMAT	Arm Motor Ability Test
AMPS	Assessment of Motor and Process Skills
ANOVA	Analysis of variance
AR	Augmented reality
ARAT	Action Research Arm Test
BD	Block Design
BN	Bayesian Networks
CLR	Cumulative Logistic Regression
C-SVM	Cost-sensitive SVM
DOF	Degrees of freedom
DTW	Dynamic Time Warping
FIM	Functional independence measure
FMA	Fugl-Meyer Assessment
HMM	Hidden Markov Model
HSD	Honestly Significant Difference
ICC	Intra-class correlation
KNN	k-Nearest-Neighbor
NHPT	Nine-Hole Peg Test
PC	Personal computer
PPT	Purdue Pegboard Test
RBF	Radial Basis Function
RMS	Root Mean Square
SD	Standard Deviation
SEM	Structural Equation Modeling
SHAP	Southampton Hand Assessment Procedure
SVM	Support Vector Machine
UAN	User Action Notation
VE	Virtual Environment
VR	Virtual Reality

CHAPTER 1: INTRODUCTION

Due to technological developments since the advent of the personal computers (PC), reliance on the human workforce has significantly decreased. However, despite the advances in automation and information technology, human work is still necessary in many domains, e.g. surgical operations, rehabilitation training, military operations, etc. As a result of high flexibility, adaptability and creativity of humans, manual work is also heavily relied on in manufacturing, especially in assembly production (e.g., iPhone assembly). Human labor allows processes to accommodate increasing requirements for customized product production. Currently, manual assembly is still the most viable method for producing a variety of products in small to medium batches (Ehrhardt et al., 1994; Alzuheri et al., 2010; Hamrol et al., 2011; Howie et al., 2011).

Assembly operations play a crucial role in the quality and cost of products, as well as in the degree of flexibility of manufacturing enterprises. If human operators assigned to assembly tasks fail to perform effectively, the productivity and efficiency of an entire process can be compromised (Vazquez & Resnick, 1997). However, due to limitations in human physical and cognitive capabilities, manual assembly systems usually suffer from many deficiencies. To overcome such limitations, proper training of assembly line operators is crucial prior to deploying them on a production line.

1.1. Manual assembly training methods

1.1.1. Manual assembly requirements and conventional training

Manual assembly tasks usually require the manipulation and joining of parts to form a whole product. To achieve a simple assembly task involving just two parts, e.g. inserting a

cable into a connector, factors to be considered include reaching for and grasping the cable, moving the cable towards the connector (aiming), determining the relative positions of the cable and connector, and inserting the cable accurately at a predetermined position (Bound et al., 1999; Duan et al., 2010). These factors mainly require a good level of human motor skill. If an assembly task involves a set of subparts, which must be put together in some specific sequence, then higher order cognitive processing must occur as well. Such cognitive work often includes understanding and memorizing procedural information, planning, recalling, and decision making (Duan et al., 2010). Therefore, a typical industrial assembly task requires both long-term memory and procedural knowledge of the way the task should be performed as well as the fine motor knowledge about precise movements and forces to be applied (Gutiérrez et al., 2010). Manual assembly training thus needs to take into account the transfer of both types of task knowledge (physical & cognitive) to novice operators for an efficient approach.

Training methods and equipment designed for manual assembly work efficiency have dramatically improved in recent years. Traditionally, assembly skills were transferred to novice operators through instruction manuals or tutorial videos (Gutiérrez et al., 2010; Srimathveeravalli et al., 2006). These learning methods are usually complemented by the help of an expert worker to demonstrate the assembly procedures (Gutiérrez et al., 2010). Even in today's small-size production companies, this approach is still common due to ease of implementation. However, in some cases, such training methods are not efficient due to limited work time schedules, expert availability or cost constraints. In addition to cognitive understanding and memorization, operators need other skills to achieve assembly tasks, such

as fine motor control. These skills are typically acquired or improved through hands-on practice.

With technological advances in computer hardware and user interfaces, virtual reality (VR)-based training systems and educational software have been developed as complements or alternatives to traditional training methods (Srimathveeravalli et al., 2006; Woll et al., 2011). These systems provide innovative ways of delivering assembly training programs, which allow for hands-on practice within a simulated environment. The following section provides a review of augmented reality (AR) technology followed by a detailed discussion of virtual reality (VR) training systems.

1.1.2. Review of AR technology applied in assembly training

Feiner et al. (1993) described AR training technologies as presenting virtual worlds that enrich rather than replace real-world training methods. AR trainers allow a user to see computer graphics or virtual objects superimposed on a real-world environment, thus enhancing a user's perception with additional information for training. Pathomaree & Charoenseang (2005) proposed an AR system for assembly training. The system was integrated with a haptic device to provide users with force feedback while interacting with a virtual world model displayed based on markers in a real-world scene. The system was able to enhance user perception and reduce assembly task completion times. Step-by-step assembly solutions were presented to users and suggestions were provided to correct assembly actions in case of inappropriate performance. Experiments were conducted to compare assembly task completion time with the AR training system and traditional methods. Results demonstrated the proposed

AR system significantly reduced completion times for both 2-dimensional (2D) and 3-dimensional (3D) assembly tasks.

Woll et al. (2011) developed an automotive mechanic task learning software, called eLearning, in the form of a serious game for mechanic apprentices. The objective of the software was to motivate assembly mechanics to learn specific knowledge (e.g., part lists and process steps) and train specific abilities (e.g., precise motor control and spatial thinking) to achieve car power generator assembly. AR technology was employed to support user experiences of spatial aspects of the task and to support transition of theoretical knowledge into practical skills. Using the AR-based system, users could intuitively interact with virtual components of the virtual assembly as with real assembly components in the real world. Most interestingly, the entire system was presented via a smartphone. The system was evaluated and well received by actual mechanics.

Training results obtained with those systems developed using AR technology are mostly positive. Compared with traditional methods, users can realize practical training and develop cognitive and motor skills. Expert operators are not needed for demonstrations when using AR since most systems have been designed to provide assembly procedures or solutions in real time. However, the application of these systems requires integrated practice with real machines or assembly parts. This is not always feasible due to accessibility and availability factors. In addition, training with real machines may damage delicate AR components and result in increased training cost (Gutiérrez et al., 2010).

1.1.3. Application of VR technology in assembly training

To overcome the drawbacks of AR systems, VR technologies have been developed to provide users with a better understanding of machine and assembly procedures using 3D virtual environments (VE). Such VEs are generated by computers and updated in real time. Users interact with virtual objects through various input and output devices.

Bound et al. (1999) investigated the potential for the training of manual assembly skills using VR and AR technologies as well as conventional training media. Experiments were conducted to compare the immediate impact of a given format of media on task performance with operator training and post-training performance. Task completion time for assembling a water pump was evaluated for participants receiving information from one of five training formats (conventional 2D engineering drawings, desktop VR using a monitor and 2D mouse, desktop VR using stereoscopic glasses to provide 3D images along with a 2D mouse, immersive VR using a head-mounted display, or a context-free AR condition). Results suggested that both VR and AR conditions significantly outperformed the 2D engineering drawing condition in reducing assembly time. Performance under the three VR conditions was comparable, indicating that VR technologies did not differ much for the particular application. However, the VR technology was limited in terms of no haptic feedback, which can play an important role in learning manipulations as part of assembly tasks. The AR technology, on the other hand, provided users with tactile feedback through manipulation of real assembly objects. This difference among the training conditions may account for the superior performance of the AR condition. Despite the VR technology not outperforming the AR technology, Bound et al. concluded that VR was beneficial for training in that it allowed users to experience assembly

manipulation objects without the use of real objects. Hence, users could be trained to assemble a product before the product was physically manufactured. In terms of training applications, VR may be more flexible than AR in that virtual worlds can be separated from the real world.

To make VR a better approximation of physical reality for manual task training, the sense of touch and kinesthetics must be addressed. Such design requirements can be fulfilled through the use of haptic devices, which allow VR users to perceive tactile feedback in VEs. The training effects of VR systems for assembly tasks have been demonstrated to be effective in numerous laboratory research and industry applications (Bound et al., 1999; Edwards et al., 2004; Bhatti et al., 2008; Vo et al., 2009; Gutiérrez et al., 2010). The addition of haptic devices to such systems may further reveal the benefits of VR for assembly skill training.

Adams et al. (2001) conducted an experiment to investigate the benefits of force feedback in VR training of a real task. Three groups of participants received different levels of training before completing a manual assembly task with a LEGO biplane model. A VR system was developed to present a simulation that emulated the real task in a VE. A haptic device was used to simulate tactile feedback generated when interacting with the virtual objects. The participants either received VR training with haptic feedback, VR training without force feedback, or no VR training. Analysis of completion times for the real task (post-training) revealed that participants trained with force feedback performed significantly better than those receiving no training. Although not statistically significant due to the relatively small sample size, the average completion times for participants trained with force feedback were shorter than those trained without. The haptic feedback proved to contribute to the VR assembly training.

Pugmire et al. (2007) developed the “Pick-n-Place” training system, a VR assembly and disassembly application. The application was able to read and display CAD models as well as texture-mapped polygon models to provide realistic environments, such as an assembly bay. Assembly constraints could also be defined to specify how parts could be mated. With this system, users were able to visually inspect the components of an assembly, grab them, and try to connect them. Collision detection was used to help determine interpenetration of parts accompanied with a “thud” sound generated from a stereo sound system. If parts could not fit together or were obstructing one another, both visual and auditory cues were provided to help users adjust their assembly strategy. This research suggests that multimodal systems are a suitable option for providing users with a good sense of a virtual training environment and facilitating assembly skill training.

While a number of computer-based VR systems have been proposed, developed and applied in manufacturing companies, many have not been found to be effective in transferring task sequence knowledge. This outcome is likely due to limitations in realistic representation of real-world tasks or the absence of haptic control devices in such systems. To address this problem and enlarge the knowledge area that can be trained through VR systems, Bhatti et al. (2008) proposed a complete haptic-enabled interactive and immersive VR system to support the learning process of general assembly operators. The prototype system was expected to imitate real physical training scenarios by supporting comprehensive user interactions. Within the designed training environment, users could perform assembly operations with physical restrictions imposed by a haptic device to deliver the feeling of a real-world environment.

Given the system's physically interactive and immersive nature, it facilitated both sequence learning and procedural skill development.

Gutiérrez et al. (2010) said that common VR training systems could only simulate single manual assembly or disassembly operations with rigid components and failed to consider manipulation of tools, which should not be neglected in procedural tasks. On the basis of this contention, they developed a multimodal training system for industrial maintenance and assembly tasks using VR technology. The system was aimed at providing efficient assembly skill transfer by allowing trainees to practice the assembly/disassembly tasks in real time. Users could interact with the virtual scenario in a multimodal way combining haptics, gestures and visual feedback. The system also supported procedural task learning with four strategies to guide users during the training process. Each strategy required different degrees of trainee interaction, from more guided and passive strategies to less guided and active ones. The proposed system outperformed other systems (e.g., the Pick-n-Place system developed by Pugmire et al. (2007)) in terms of flexibility to adapt to task demands and users preferences – flexibility in the available training protocols, flexibility in support of haptic devices, and flexibility in available training protocols.

In summary, during the past decade, VR technology has been applied for manual assembly training and the application has largely improved through development of multimodal systems, including haptic interfaces. All previous research and applications in industry have demonstrated that VR is a powerful tool for training human operators to perform tasks which would otherwise be expensive to duplicate or dangerous for novice operators in the real world. By applying VR systems, companies can design training tasks with more degrees of

flexibility and variety. In addition to receiving information about assembly sequences and procedures, operators can practice with virtual objects generated in VEs. With the use of a haptic device, VR users are provided with tactile feedback in manipulation and assembly tasks. In addition to providing users with a feeling of physical reality through the simulated reality, the visual, auditory and haptic cues as part of the VR can also serve as guidance in assisting trainees in understanding instructions and completing tasks with correct actions. In addition, with VR systems, user performance is easily recorded and analyzed for assessing the effectiveness of training strategies. Considering the advantages of VR systems demonstrated by previous research, the present study will examine VR as a tool for facilitating assembly motor skill training.

1.2. Manual assembly learning process

The flexibility of VR systems allows researchers to effectively design various tasks for delivering manual assembly training. For successful completion of most assembly work, both cognitive and motor skills must be addressed during training. Therefore, it is crucial to understand the human learning process in acquiring assembly skills such that corresponding tasks can be properly designed and provide efficient training for assembly operators.

One of the most widely-accepted theories of human skill learning is Fitts' learning model (Fitts, 1954). The model includes three phases: a cognitive phase, an associative phase, and an autonomous phase. In the first stage, operators learn the basic procedures and properties of the specific task. Secondly, the procedures and relevant knowledge acquired in the first

phase are “chunked” into sequences of actions. Finally, the “chunked” sequences gradually combine into a smooth pattern of activity.

Philbin et al. (1998) suggested that in any training situation, the tasks to be trained must be decomposed into components of cognitive, perceptual and motor demands. These demands must be met during the training process. The cognitive portion primarily consists of the construction of an internal model of the task within trainee memory. During assembly training, this is the stage where trainees gradually develop their cognitive skills. Operators learn the properties of each part and how each particular piece fits within the final model. They then develop their strategies for building the entire assembly. For traditional assembly training, operators need to construct the correct sequence of actions through understanding of engineering drawings and instruction manuals. This part of learning is primarily cognitive in nature and usually requires a long period of time to complete. Using VR systems to train assembly tasks supports both the cognitive and associative stages where plans are developed during task completion. In this way, using VR systems for task training can largely reduce the time needed for the first two stages of learning (Bound et al., 1999).

As previously mentioned, in addition to cognitive skills, operators also need to learn motor behaviors during training of an assembly task. The motor demands of assembly concern dexterous manipulation of parts. Operators must learn how to handle the parts, move them to desired positions, rotate them to correct orientations, and mate them according to requirements; that is, acquire motor skills. With respect to Philbin et al. (1998) model, the perceptual aspect of the training usually plays the role of “glue” to bind the cognitive and motor portions

together. This phase helps promote trainee comprehension of the task environment (Adams et al., 2001).

An assembly task requires operators to repeat the same set of procedures. This characteristic requires development of cognitive skills very early in training. During initial trials, operators are likely to make mistakes due to unfamiliarity with assembly plans or sequences. However, in later iterations of the task, after developing some familiarity with the procedures and having constructed an internal task model in memory, operators are able to work without hesitation and without faults. After reaching this stage of learning, training performance is mainly focused on motor demands. However, motor demands can be difficult to address if there are any limitations in cognitive skill proficiency. Related to this, poor motor control skill development can lead to delays in training progress, dropping of parts and damage to assemblies during training, as well as actual assembly quality problems (e.g., misaligned pieces) in real operations following training.

In summary, current VR systems developed for assembly task training have proven effective in reducing training completion times. This success actually is the result of effectively reducing the time needed to address cognitive skill development. However, the more important component of assembly training, motor demands, can be more difficult to address. Instead of designing tasks that cover both cognitive and motor skill training at relatively similar levels, it may be beneficial to construct a series of tasks that focus more on motor skill development. In this way, the effectiveness of training to fulfill motor demands can be evaluated without disturbance from cognitive demands. In addition, although VR systems may reduce the time

needed to complete training tasks, more variables must be recorded during training to assess user motor skill development in real time.

1.3. Challenge of applying lean concepts to a manual assembly production line

As discussed in the preceding sections, numerous VR systems have been developed to support efficient training protocols so that trainees can acquire assembly skills within the minimum amount of time. In addition, some research has investigated the design of training tasks to assist operators to achieve their highest skill level. However, such an approach does not consider the broader perspective of evaluating the performance level of a group of operators for assembly line work. For production operations, there is a need to train assembly operators to a uniform performance level. Consistent performance across operators on an assembly line is a common assumption in process design and scheduling. However, this assumption is typically violated by actual imbalances in operator performance levels creating bottlenecks in production and reducing overall line productivity.

Along with the development of automation technology, the concept of “lean production” has been introduced to the automated manufacturing production line. The lean concept targets improvement of production rates through elimination of waste generated during production processes, such as material redundancy, work-in-process transportation, operator idling, etc. To achieve this target, the most common method is to maximize production resource utilization through production planning, scheduling of operator workload assignments, working station design, and operator task training. In the ideal situation, a production line should be balanced, meaning that the total workload on the line can be

distributed evenly among workstations, thus reducing idle time. Therefore, all operators will have an equal amount of work while simultaneously achieving the specified production rate of the line (Altuger & Chassapis, 2010; Low Shye Nee et al., 2012).

Recently, the lean concept has also been introduced into the manual production area. However, variability in human operator performance may make achieving the application of lean production challenging (Altuger & Chassapis, 2010; Hamrol et al., 2011). To solve this problem, many algorithms and mathematical models have been developed to optimize operator work scheduling and production plans. One common assumption of these methods is that operators are viewed as interchangeable, i.e., all operators have comparable skill levels in completing production tasks. However, it is more likely that this is not reality for a physical production line (Zee & Slomp, 2005). There are typically individual differences among operators that may have substantial effects on production outcomes. Hence, production control practices implemented based on lean theories are likely to fail in practice without taking into account worker differences in terms of skill, working speed, etc. (Zee & Slomp, 2005; Altuger & Chassapis, 2010).

In summary, as opposed to training workers to the highest possible skill level, in real-world implementation of a production line, it is more commonly expected that all workers operate at comparable level of performance to achieve a uniform production rate. To address this requirement, individual differences in novice operator abilities must be considered during design and selection of training systems and training tasks.

1.4. Research motivation

While previous studies have developed and tested numerous protocols for manual assembly skill training, few have made considerations of individual differences in skill levels and the influence that may occur in training effectiveness. Moreover, the simple target of training a group of operators to their own best ability level using the same protocol is not in accordance with the requirements of standard performance on a production line. Therefore, it is necessary to develop advanced methods by which to classify novice operator skill levels and assign them to appropriate training protocols, integrating contemporary training technologies, based on baseline performance levels. It is expected such methods can be developed and applied in real manual assembly training for production lines (e.g., iPhone assembly) to facilitate operator performance and increase overall production rates.

To achieve these objectives, several issues need to be addressed. First, as discussed above, in manual assembly task training, cognitive demands may be easier to satisfy due to the repetitive nature of assembly work. However, more effort may be required for operators to develop motor skills in order to achieve an assembly task. On this basis, motor skill training will be the main focus of the present research. Secondly, in order to establish an operator's motor skill level, an appropriate assessment approach should be developed, including motor performance tests. Thirdly, since standardized tasks used to assess motor skills differ from assembly training tasks, it is necessary to verify that performance measures obtained in relatively simplistic motor tests can be used to predict performance in complex skilled motor training tasks. Finally, once operator baseline performance metrics have been collected, a

classification algorithm is needed to identify the skill levels of specific operators such that corresponding training protocols can be assigned.

In the next section, a literature review is presented covering four issues relevant to the present research, including motor skill training protocols, motor skill assessment, motor performance prediction, and quantitative skill classification approaches. With the knowledge obtained from related studies, the structure of this research is developed through a problem statement and experiment methodology.

CHAPTER 2: LITERATURE REVIEW

2.1. Motor skill training protocols

2.1.1. VR-based training systems with haptic guidance

Singer (1980) defined motor skill as “the ability to execute a movement in an optimal fashion, or an activity of a person involving a single or a group of movements performed with a high degree of precision and accuracy”. To assist training of operator motor skills, a variety of approaches have been suggested and developed. Todorov et al. (1997) said that the acquisition of new motor skills can be enhanced by using VEs with augmented feedback on operator actions. To assess this claim, they designed a system to teach difficult multi-joint movement in a table tennis environment. The system presented trainees with a fairly realistic computer animation of the environment and a virtual ball. Trainees were asked to attempt to learn the movement by matching the movement pattern of an expert teacher’s paddle also presented in the environment. This training system provided operators with augmented feedback, specifically a minimum set of movement details most relevant to the training task. Experiments were conducted to validate the effectiveness of the system and to promote motor skill acquisition. Results indicated that participants receiving VE training performed significantly better than those receiving real-task practice. The expert teacher’s trajectory provided during VE training served as an effective basis for developing performance in the real-world task. However, practice with only the VE was not sufficient for transfer to the real task. Experiment results also demonstrated that when a critical component of the VE training system was removed, participants showed no transfer to the real task.

The augmented feedback provided in the system of Todorov et al. (1997) focused on visual cues. Similarly, some other work has explored the usefulness of visual cues in motor skill training (Massimino & Sheridan, 1994). With the development of haptic devices, researchers have also explored haptic cues for motor skill training. Feygin et al. (2002) investigated the use of haptic guidance for perceptual-motor skill training. They argued that haptic training might be especially helpful for learning motor tasks since training with a haptic device occurs in physical space, relative to the body. They provided haptic feedback to facilitate kinesthetic perception and memory formation. To validate their theory, experiments were designed to compare haptic training with visual training and combination training involving both haptic and visual information. The findings from the experiment indicated effectiveness of haptic guidance for training performance, especially in temporal aspects of the task. In general, Feygin et al. (2002) study provided support for the role of haptics in motor skill training with VEs. However, one drawback of haptic training was identified; that is, learning occurred passively during training. The authors suggested that more development of haptic training paradigms was needed to promote operator conscious construction of strategies to accomplish motor tasks instead of unconsciously developing movement patterns through passive haptic feedback.

Wang et al. (2006) developed a haptic-enabled Chinese handwriting training system to study the learning and transfer of motor skill. Based on the characteristics of human motor skill learning, four design rules were proposed for training systems. Considering Fitts' three-stage learning process model and previous experiment results (Fitts & Posner, 1967; Solis et al., 2003), the first rule was that the function of the training system should focus on the cognitive

and associative stages of learning in which haptic training could accelerate and improve an operator's progress. Conversely, once an operator reaches the autonomous stage, additional haptic feedback could lead to detrimental effects to task performance. To facilitate developing active strategies during training, the second rule was that error information should be provided to form an active learning environment. In order to assist operators in remembering motor patterns required by a motor task, the third rule was that approaches should be developed to enhance kinesthetic memory. Last but not least, to improve the fidelity and stability of haptic training environments and provide better sensory feedback, the fourth rule was that haptic information should be combined with other sensory modalities, such as visual and audio feedback. Experiment results validated the proposed training rules and the force feedback provided by the haptic device proved to be effective for training novices in Chinese handwriting skill.

With the continuing development of VR systems design and haptic device technology, a number of haptic training programs have been proposed and explored. Even some training tasks that seem to be quite simple and straightforward have been found to provide efficient training and to improve participant motor skill (Petrofsky & Petrofsky, 2004). However, compared to training tasks that merely require operators to perform 3D movements or follow trajectory patterns, assembly tasks also require operators to develop dexterous motor control skills in positioning and mating subassemblies. Edwards et al. (2004) conducted an experiment to explore whether the inclusion of force feedback or auditory cues could improve manipulation performance and subjective usability ratings for an assembly task presented in an immersive VE. Users reported perceptions of increased realism, helpfulness and utility for

assembly training when force feedback was provided. However, contradictory to subjective ratings and expectation, users provided with force feedback actually showed decreased performance with a greater number of collisions of parts. Also, when auditory cues and force feedback were presented together, the performance and usability ratings did not improve, as compared to conditions with only one type of cue. The study results provided an example of detrimental effects of force feedback in haptic training. It should also be noted that the techniques and haptic devices used in this study had many limitations due to technological constraints. These constraints might have diminished the efficacy of the haptic cues. With proper task design and updated haptic devices, haptic training techniques should be further investigated.

As has been discussed thus far, by incorporating haptic feedback in a VE, users can better perceive interaction with virtual objects. However, at the same time, the existence of haptic controls and cues may increase the complexity of a VE and training task performance for users. To assess the performance benefits of haptic-based interaction in virtual assembly, Vo et al. (2009) evaluated user performance in completing similar assembly tasks when provided with only visual feedback or with both visual and haptic cues. A set of experiments was designed and conducted to determine whether completion times and user accuracy would be influenced by force rendering and to distinguish the assembly tasks that would be most affected by haptic feedback. When compared to visual-only methods, statistical analyses revealed that haptics-based interaction was beneficial for improving performance by reducing times for object weight discrimination, permitting higher placement accuracy when positioning virtual objects, and enabling steadier hand motions along 3D trajectories. The use of haptics

for virtual assembly training was demonstrated to benefit performance in some aspects of user behavior. The Vo et al. (2009) study provided additional support for further development and research on haptic-based interaction in VR training simulations.

Howard & Vance (2007) investigated the feasibility of using a desktop haptic VE as a design tool for evaluating assembly operations. The application involved several software packages to explore the benefits and limitations of combining haptics with physical models. One or two-handed force feedback was presented with various haptic devices, including a Phantom from SensAble Technologies, Inc. Test results validated the system's feasibility for simulating assembly operations for engineering evaluation. With similar technology, Nam & Jang (2010) developed a virtual assembly training system in which a Phantom Desktop was used for tool position sensing and force feedback. With the proposed training system, users could touch and move virtual components and use a virtual electric screwdriver. The VE facilitated training of assembly procedures for a product. These studies support application of the Phantom haptic device in developing a virtual assembly training system.

In summary of the above studies, a better understanding of haptic feedback in virtual assembly and updated hardware and software to support haptic-based interaction has led to development of VR assembly training systems integrating haptic devices and use in laboratory research and operator training. The addition of haptic feedback benefits VR training systems by reducing operator learning time, improving task performance, increasing dexterity, and increasing perceptions of realism and the sense of presence in a task simulation (Adams et al., 2001; Williams et al., 2002; O'Malley et al., 2006).

2.1.2. Robotic-based training systems with haptic guidance

In addition to VR-based motor skill training protocols, robotic-based systems have been developed. Krebs et al. (2003) proposed a concept of performance-based progressive robot-assisted motor control therapy. This research was focused on identifying optimal therapy protocols for stroke patients to maximize recovery. In this approach, patients were asked to perform a reaching task with a virtual spring pulling their hand towards a target. Guidance was provided during training with the amount of guidance determined by the spring coefficient and patient performance in speed, time and muscle activation level. Although this study did not yield a statistically reliable result on the effectiveness of the new proposed therapy due to a small sample size, therapists noted that the therapy method benefited patients in terms of reducing arm spasticity and increasing muscle tone.

Similar to the work of Krebs et al. (2003), Reinkensmeyer et al. (2004) conducted a robot-assisted rehabilitation study for gait training. In this research, human motor control adaptation to dynamic environments was modeled as an error corrective learning process. The control gains of the guidance robot were adjusted at each trial based on the degree of patient movement error. Experiment results supported the hypothesis that robotic device assistance of naturalistic movements leads to better functional recovery. Assistance in movement (as needed) also proved to be more effective than providing constant assistance with a fixed gain.

O'Malley & Gupta (2003) designed a Fitts' type targeting task to evaluate human performance under passive robotic assistance as well as shared control (active assistance). In this research, the shared control mode was designed with reference to the study of Steele et al., (2001). Different from the general passive assistance that provides little information pertinent

to the state of the environment, the shared control architecture allows users to effectively “share” control of the interface with an automatic controller or virtual agent. The agent is placed in the perceptual space of the human through use of a haptic display. In this way, both the human and virtual controller exchange power over the control interface and the human may override the virtual controller of the system, if necessary. As compared to passive assistance, shared control was proposed to be an active haptic guidance scheme for training. A shared controller dynamically intervened, through an automatic feedback controller acting upon the system, to modify the system dynamics during guidance (O’Malley et al., 2006). Experiment results demonstrated that active assistance under shared control led to improved human performance of the task comparable to passive assistance. Yet shared control had greater effects in terms of transfer of the skills acquired during training to a real working environment.

On the basis of this research, Li & Patoglu et al. (2009) conducted an experiment on a haptic guidance system designed as a fixed-gain error-reducing shared controller with assistance applied to the dynamics of the manual control task during training. After a month-long training protocol, they showed that use of fixed-gain haptic guidance had a detrimental effect on operator performance in a target-hitting task. This result was independent of the duration of exposure to the shared controller. According to the hypothesis stated by Li and Patoglu et al. (2009), such a negative effect was mainly due to the fact that the control gain was fixed, thus leading participants to become dependent on the guidance during task completion.

In an attempt to improve the effectiveness of haptic training systems and further test the hypothesis of negative performance effects of fixed-gain controllers, Li and Huegel et al.

(2009) introduced a progressive shared control guidance scheme. The scheme was to reduce the dependence of participants on haptic guidance by adjusting control gains based on individual participant performance. Experiments were conducted to compare participant performance in completing a virtual target-hitting task with various types of haptic guidance. Results indicated the proposed progressive shared control scheme performed significantly better than fixed-gain haptic guidance schemes. The superior performance of the progressive shared control group over other haptic guidance groups was attributed to the performance-based progressive control algorithm that reduced dynamic interference and dependence of operators on the haptic guidance and promoted acquisition of motor skills for the task. It was suggested that this scheme could be generalized for haptic guidance based training in a wide range of complex tasks.

Although progressive haptic guidance has shown superior performance in facilitating motor skill learning during training, when compared with other schemes, it has failed to demonstrate superior retention performance. In fact, results of many studies have suggested that haptic guidance might not be the best approach for improving motor skill learning with some studies even showing detrimental results. In those studies that demonstrated positive effects of haptic guidance during motor skill training, learning transfer effects have not been observed due to a lack of experiments involving retention tests. This problem with haptic guidance research can also be found in some research results on augmented feedback in VEs. In studies by Schmidt (1991) and Schmidt and Wulf (1997) it was suggested that when concurrent augmented feedback was provided to VR users too frequently or continuously, motor learning could be degraded. The existence of augmented feedback improved

performance during training yet degraded operator performance when the feedback was not available after training (learning evaluation).

2.1.3. Training systems with haptic disturbances

Some training schemes have been proposed to solve the problem of operator dependence on haptic guidance. Bettini et al. (2001) suggested the implementation of “dead zones” where haptic guidance would become available intermittently only in case large or unsafe operator errors occurred. It was expected that the addition of dead zones would reduce operator dependence on the guidance. However, from the perspective of training, this kind of assistance only provides a safe medium for practice. It lacks the intent of promoting operator learning.

Another type of training scheme was proposed to gradually reduce the amount of haptic guidance during training. Bell & Kozlowski (2002) proposed a performance-based adaptive haptic guidance scheme. The adaptive guidance was designed to augment self-regulation in learning and was tailored to meet differing needs of individual trainees. Individual learning progress was monitored and assessed by a computer. Based on this evaluation, trainees were provided with a scaled amount of haptic guidance. An experiment revealed that the adaptive guidance had a pervasive and substantial impact on the self-learning process and the sequence of trainee study and practice. The proposed adaptive guidance scheme yielded significant improvements in acquiring basic and strategic knowledge and developing performance skills.

More recently, a new approach to haptic-based training was proposed by Lee & Choi (2010) and labeled as haptic disturbances. Their hypothesis was that disturbing the movement of a learner during motor skill training, instead of guiding them, might lead to higher degrees

of motor learning. Haptic disturbances pose a learner with more challenging conditions for motor task execution and this might increase the capacity of the learner to perform the task. Two types of haptic disturbances were explored in this research: repulsive haptic disturbances using feedback forces and noise-like haptic disturbances using random feed-forward forces. The effects of the two methods were experimentally assessed and compared with a conventional visual-only method and progressive haptic guidance. Experiment results revealed that during training the progressive haptic guidance produced the best performance. However, in immediate and delayed retention tests, noise-like haptic disturbances led to the best performance. The results of this study indicate a high potential for application of haptic disturbances in developing motor skill training protocols, expediting the motor learning process, and promoting long-term motor skill learning.

2.2. Motor skill assessment

In order to predict operator motor performance in assembly work, it is necessary to develop methods of motor skill assessment to be applied during training. In conventional protocols used in clinical therapy and rehabilitation, performance is typically measured using gross score scales or subjective observations of clinicians. The development of computer-based training systems has allowed researchers to record participant performance in real time with objective and quantitative measurements. Such systems support more reliable and accurate determination of an individual's skill level and can be used to measure performance improvement during training as well as evaluate effectiveness of training protocols.

While computer-based systems have been researched for operator strategy and motor function training, the design of training tasks and performance assessment methods also needs to be researched. In many previous studies of assembly training protocols, the most common objective measure has been task completion time, which is easy to obtain and straightforward to interpret. However, to evaluate a user's complete motor function, there is a need for more measurements than completion time.

2.2.1. Clinical-based techniques

To assess an individual's hand and arm control, the most common method has been to use clinically-based scales. Examples include standardized tests of hand function proposed by Jebsen et al. (1969), the Action Research Arm Test (ARAT) proposed by Lyle (1981), etc. Although these testing methods have been validated by numerous studies and datasets, the application procedures require clinicians and sometimes subjective observation and judgment. Moreover, the outcomes are usually simple and scalar scores. Therefore, to obtain more detailed and comprehensive information about upper limb function, a number of researchers have explored more quantitative and objective approaches using various technologies.

During the development of constraint-induced therapy, Taub et al. (1993) proposed the Arm Motor Ability Test (AMAT) to assess the effectiveness of an intervention for improving daily living activities among patients. In contrast to other arm ability measures, AMAT was able to assess functional ability, quality of movement and time of performance. The reliability, validity and sensitivity of such tests were evaluated by Kopp et al. (1997) using a previously established test, the Motricity Index (Demeurisse et al., 1980), as a criterion. The AMAT measures were found to be reliable, sensitive and internally consistent. However, Kopp et al.

(1997) also found that the correlation of AMAT measures with upper limb motor impairment measured by the Motricity Index was fair at best. However, due to limitation of the Motricity Index procedure and assumptions, this result was not considered to be conclusive. Hence, Chae et al. (2003) investigated the concurrent validity of AMAT in measuring motor impairment using another criterion, specifically the Fugl-Meyer Assessment (FMA; Fugl-Meyer et al., 1975). Scores of the AMAT & FMA were highly correlated and demonstrated the criterion validity of the AMAT. However, it was also pointed out that AMAT measurements were time consuming and led to significant fatigue for patients. Chae et al. (2003) suggested that the number of testing items in AMAT should be reduced to facilitate clinical implementation.

Light et al. (2002) developed another assessment procedure, the Southampton Hand Assessment Procedure (SHAP) to obtain contextual results on hand function in a clinical environment. The reliability and validity of the proposed assessment procedure was tested against standard medical outcome measurement techniques. Experimental results established the statistical integrity of the SHAP. Compared to existing procedures, the SHAP was superior in terms of allowing clinicians to obtain more information that comprehensively described the prehensile range of the hand. In addition, the SHAP's outcome measure was a contextual rating of functionality that enabled the clinician to initially determine a patient disability, and subsequently monitor their performance throughout a course of treatment or rehabilitation.

While the AMAT and SHAP both proved to be powerful in assessing motor ability, it should be noted that these techniques are still clinical-based testing procedures and require expert decision making in ratings of patient ability. Thus, both methods still do not completely remove subjectivity in assessing upper limb motor function.

2.2.2. Robot-aided techniques

Krebs et al. (1998) applied robotics and automation technology to assist and quantify the extent of neuro-rehabilitation. The prototype robotic system, MIT-MANUS, could move, guide or perturb the movement of a user's upper limb and record motions and mechanical quantities such as the position, velocity and forces applied to the robot. The recorded kinematic and force data profile was later analyzed and validated by standard assessment results generated from the functional independence measure (FIM; Trombly, 1995) as well as the upper limb subsection of the FMA scales. The kinematic data further explained the standard assessment with improved objectivity, repeatability, precision and ease of application. Krebs et al. (1999) also used the MIT-MANUS system to identify the apparent sub-movements composing continuous arm motion in an unloaded task. Kinematic analysis was applied for detailed description of primitive sub-movements. The precise mathematical characterization of sub-movement kinematics provided key information for objectively and reliably deconvolving continuous arm movements into component sub-movements. Following these studies, Krebs et al. (2000) presented an overview of their studies in robot-aided stroke neuro-rehabilitation and recovery. It was stressed that while conventional clinical scales could help assess the impact in the neuro-recovery process, the coarse nature of such scales required extensive and time-consuming trials, and often failed to reveal detailed and important information for optimizing therapy. Alternatively, robot-based motor assessment scales offer the potential benefit of using kinematic measurements to obtain greater insight into the process of recovery from neurological injury. Similar to the work of Krebs et al. (1998, 1999, 2000),

Reinkensmeyer et al. (2000) developed a rehabilitation system called the “ARM Guide” to assess tone, spasticity and lack of coordination for patients after chronic brain injury.

More recently, Zollo et al. (2011) studied quantitative measurement of arm motor control during a robot-aided rehabilitation session. Training tasks consisted of three games of point-to-point movements from the center of a circle to targets at the periphery. Robotic machines (InMotion2 and InMotion3, based on the MIT-MANUS system) were used to aid the therapy and collect kinematic and force data on patient performance. The data was then used to compute performance indices to measure temporal, spatial and force features of motor skill ability. Five quantitative indicators were extracted, including aiming angle, length ratio, jerk index, useful force, and useful work. A traditional clinical impairment assessment was conducted based on clinical scales to provide a validation reference. Linear regression analysis results revealed that arm functions measured by the robotic machines were reliable in assessing functional capability.

The above studies have demonstrated the usefulness of kinematic data analysis generated from robotic-aided systems for objectively assessing upper limb motor function. However, robotics applications are usually constrained to specific movements and kinematic data have been limited to three dimensions.

2.2.3. Systems with more dimensions

Spyers-Ashby et al. (1999) used a multidimensional movement analysis system to record limb tremor over 6 degrees of freedom (DOF). In this study, an electromagnetic sensor (3Space Fastrak; Polhemus, Inc.), was used to collect and quantify the tremor data over three translational directions and three rotations. Salazar-Torres et al. (2004) also used a multi-

dimensional biomechanical device to investigate the excitability of muscle stretch reflexes in order to quantify spasticity. These devices and technologies have enlarged the range of kinematic data obtained on motor performance. However, access to some research devices is limited to experimental studies. Clinical applications require systems that are easier to implement and are more flexible in use (Amirabdollahian et al., 2005).

Still other studies have been conducted on test protocols for motor skill assessment using VR systems. Bardorfer et al. (2001) presented an objective test for evaluating the functional capability of the upper limbs in patients with neurological diseases. The methodology involved creating a VE, using a computer display for visual information and a Phantom haptic device with three active DOFs and six measurement DOFs. The haptic interface was used as a kinematic measuring device and for providing tactile feedback to patients. The test task was designed as a labyrinth puzzle that required patients to move a haptic pointer through the labyrinth as quickly as possible with as few contacts with 3D walls as possible. Various degrees of task complexity were achieved by changing the labyrinth track width and length and wall friction. The preliminary experiment demonstrated the suitability of the new method for upper limb capability assessment, as it provided objective, repeatable and quantitative results. The potential benefits of using this method were identified as high resolution measurements of motor skill and long term stability of measures. In addition, with the haptic device, random force perturbations could be used to trigger and assess patient responses to disturbances. In this case, a patient would perform tracking a predefined trajectory using feedback control. While this method appeared to be sensitive to changes in patient

performance, a statistically reliable comparison of the method with other upper limb assessment methods was not conducted.

2.2.4. VR-based systems

Chuang et al. (2002) also explored the usability and usefulness of a VR-based system in assessing hand functions. Participants were asked to use a dataglove to insert 3D virtual representations of a cylinder and a prism into target holes. The task was designed to assess visual-motor coordination and individuals had to achieve the insertion goal accurately within a set time. For each trial of the experiment, the root mean square (RMS) value of the hand movement trajectory was determined along three axes of translation, which could be used to measure the extent of displacement in manipulation. The RMS values on the three axes, along with the task completion time and accuracy data, represented the VR assessment measurements. To verify the reliability of the performance test, each participant was asked to take the test twice and the measurements were analyzed using an intra-class correlation (ICC) approach, based on a repeated-measures analysis of variance (ANOVA). The results of the analysis indicated a fairly high test-retest reliability of the VR assessment during the movement of the prism, but were not significant for movement of the cylinder block. The study demonstrated that the proposed VR system could provide a real time quantitative 3D task analysis of performance fluctuations among operators. The system was generally consistent in measuring the variables under study. Although the study did not validate the computer-based system measurements with clinical-based scores for standardized peg-movement exercises, it suggested a promising application of VR technology in generating quantitative real time and off-line measures of whole hand function.

With similar technology, August et al. (2005) developed and tested a VR system using a low-cost data glove and computer-based puzzle games. The glove tracked hand and finger movement in space, including x, y, z, yaw, pitch and roll, as well as speed of movement and the duration of activities. Specific movement routines were used for periodic objective assessment of patient performance. It was expected that these measurements would be effective for telerehabilitation applications. Although the technology represents great potential benefits for clinical and home use, the accuracy and reliability of the identical response measures need further validation.

Alamri et al. (2009) proposed a novel approach based on AR technology to enhance patient involvement in rehabilitation exercises and to measure performance without direct supervision of a therapist. Two exercise tasks were designed and a dataglove was used to read participant hand spatial characteristics during task performance. With the data recorded by the system, eight factors were extracted to evaluate the quality of performance. These factors included: task completion time, hand coordination, compactness of task performance, hand steadiness, speed of hand movement, kinetic energy, grasping angles, and finger grip acceleration. In combination, these factors allowed for assessment of functions, including hand steadiness, eye-hand coordination, hand control ability, etc. The proposed system provided therapists with the means to quantitatively measure patient performance and treatment progress. Preliminary experiment results provided support for use of the system as a practical rehabilitation tool.

Related to Alamri et al. research, a few years earlier, Amirabdollahian & Munih, et al. (2005) conducted a project focused on quantifying upper limb skills and performance as well

as characterizing the utility of different computer interfaces in terms of their physical and functional characteristics. To fully quantify user upper limb motor skills, six different assessment modules and three simulation modules were developed. Amirabdollahian, Gomes, & Johnson (2005) presented detailed experiment results and evaluations with one of the assessment modules, a peg-in-hole test. This test was a modified virtual simulation of the validated Nine-Hole-Peg-Test (NHPT; Mathiowetz et al., 1985) used in clinical assessment. As a simplified version of the NHPT, only two holes and one peg were provided in the virtual test. Haptic presentation of the virtual world was facilitated using the Phantom haptic interface. During the test, participants were asked to perform repetitive movements using the haptic device. Performance variables, including position, orientation, velocity and contact/collision reaction forces, were recorded as a vector in the Cartesian coordinate frame for analysis purposes. Although this assessment provided a far greater number of measures compared to the conventional NHPT, extensive data analysis was required to identify objective and reliable performance scales for the VR version of the test. Apart from the virtual peg-in-hole test, the project by Amirabdollahian and Munih, et al. (2005) generated five other assessment modules (linear tracking, circular tracking, target tracking, maximum force, and 2D labyrinth) and three simulations (Handy1 robot simulation, Manus manipulation simulation, and a powered wheelchair simulation). In general, the project developed an assessment package including various test programs focusing on different aspects of upper-limb motor function. The authors considered full clinical evaluations to be necessary for validation of the new system. This was expected to support applications involving clinical and laboratory motor skill analyses.

Emery et al. (2010) developed a VR system to replicate the full version of the NHPT and to allow for objective and quantitative motor ability assessment. The system was developed to exploit the advantages of the classical NHPT (i.e., easy to administer, standardized, and validated) and of VEs (including controlled environment, adjustable parameters, quantitative and objective measurements). The data collected from the virtual test included elapsed time, hand and arm positions, orientations in polar coordinates, forces, and order of peg manipulation. Using raw data from the system, nine parameters were extracted, including reaction time, approach time, mean velocity, maximum velocity, maximum acceleration, sensitivity to orientation, number of zero-crossings of the acceleration signal, number of peg drops, and root mean square of the forces. For validation purposes, the real NHPT test was used. In this particular VR system, haptic feedback was not provided. Thus, participants had to rely on visual cues to determine the relative position of the cursor and pegs in the task. Due to the time required to align the cursor with pegs, the time needed in the virtual test was around three times longer than in the real test. This was likely attributable to participant use of the haptic device and limited training. However, there was a strong correlation between the average completion times for both tests. This result, to some extent, validated the use of the virtual test as a tool to assess motor function.

Feys et al. (2009) investigated the relationship between VR-based movement and motor control measures and clinically valid measures of hand and arm dexterity. A series of movement tasks were designed and tested using the haptic Phantom device within a VE. Force, position, orientation, velocity and contact/collision reaction forces were recorded as a vector in the Cartesian coordinate frame. Based on the raw data from the device, outcome measures

were computed including time needed to execute the task, distance or the trajectory covered during the task, and maximal speed of performance. Clinical assessment consisted of measurements for muscle strength as well as hand and arm functionality. The Motricity Index was used to evaluate the upper-limb force using three measurements (pinch grip, elbow flexion, and shoulder abduction). Upper-limb function was evaluated with the NHPT, the Purdue Pegboard Test (PPT; Tiffin, 1968), the ARAT, and the Tempa (a daily life activity test; Feys et al., 2002). Pearson correlation coefficients were calculated to examine the relation between performance on the VR motor tests and the clinical measures. Study results did not reveal correlation of the severity of muscle weakness and VR motor test performance. This finding indicated that patients with severe muscle weakness might still be able to perform the virtual test tasks quite well due to compensatory movements of the trunk. Arm functional capacity measured with the ARAT and PPT was significantly related to movement control performance during completion of the virtual tasks. However, there were low correlations of the Tempa measurements and virtual task performance. It was suggested that motor control ability measured with the VR haptic system was more closely related to arm and hand motor function than the activities of daily life. This research demonstrated the capability of the Phantom device in training and assessment of upper-limb motor function ability. The measurements obtained in the virtual tasks corresponded with the clinical measurements generated with conventional standardized performance tests.

It should be noted that the previous studies represent only a small portion of the existing and ever-growing research aimed at developing objective and quantitative upper-limb motor function assessment tools. Due to the complex nature of the upper-limb performance,

advancement of different measures and tests is an ongoing and developmental process. The studies reviewed here have direct relevance to the development and test efforts proposed as part of the present research. In summary, a good motor skill assessment methodology should be objective, quantitative, reliable, easy to perform, suitable for routine use and should produce repeatable results. To validate the reliability of such methodologies, cross-validation with other standardized assessment methods should be conducted.

2.3. Motor performance prediction

2.3.1. Theoretical potential to make performance prediction

In the previous section, a number of computer-based motor skill assessment methods have been reviewed. Assessment results are usually used as a basis for identifying a need for treatment or to track training or patient recovery progress in a therapy program. However, the results of such VR-based motor skill tests have not been formally used in practice to predict user performance on other tasks (Mathiowetz et al., 1985). Generally, the tasks designed for skill assessment tests are simple and relatively easy to implement. However, to complete real tasks at a physical workstation, more complex and dexterous motor skills are required. The relationships between performance on low-level skilled motor tests and performance on high-level skilled functional motor tasks need to be further investigated as a basis for effective operator selection or training program development. In each of the following studies, some novel technology is developed for capturing human motor performance information and the technology is used as a basis for quantitatively predicting performance in a real-world task.

Rosen & Goodenough-Trepagnier (1989) made an attempt to develop assessment tests that predict performance in some functional activities. They developed a prototype system for prescribing augmentative communication devices for motor-impaired, non-vocal persons. Virtually all such communication devices require users to sequentially select items from a planar array of “keys” (mechanical switches, touch panels, etc.) using a single actuator (fingertip, u-cuff and stylus, mouthstick, headstick, etc.). Hence, a user’s maximum communication rate is dependent on the speed with which he/she can move between and actuate “keys”. The prototype selection system employed special-purpose assessment instrumentation to measure the speed with which a client could move an actuator from one target to another. The equipment allowed targets (keys) to be varied in size, separation distance, movement direction, actuation force, and travel distance. The system generated regression equations expressing user performance (measured in key actuation time) in terms of any set of values of movement distance, direction, key size, force, and travel distance. For each candidate device, a predicted communication rate for a user could be determined. Preliminary validation tests yielded strong and highly significant correlations between user performance on the assessment test and functional use of the device.

Chase & Casali (1995) proposed an approach for cursor-control device selection by predicting user performance with a given device based on measures of residual motor skills. The system used by Rosen & Goodenough-Trepagnier (1989) utilized custom equipment that essentially isolated and tested the physical aspects of communication tasks. A virtual test keyboard could be quickly reconfigured to approximate performance with a range of physical keyboard devices. In contrast, the approach proposed by Chase & Casali (1995) was based on

a more fundamental assessment of a user's functional skill. It attempted to map a user's performance on a basic functional hand skill test to a complex functional activity. This mapping then provided a basis for device selection for that user. Ultimately, it was expected that clinicians could administer a series of "generic" manual manipulation or dexterity tests and would be able to predict performance in any number of functional activities using different devices.

To explore the feasibility of the approach documented by Chass & Casali (1995), two experimental studies were conducted. Casali (1991) developed a functional assessment test and evaluated its ability to predict user performance with each of several computer input devices using a simple cursor-positioning task. A number of existing functional assessment tests and techniques were evaluated for possible modification or use. However, it was found that no single test was sufficient for the proposed purpose, since most approaches did not appear to be sensitive enough to differentiate between small differences in manual skills that would nonetheless determine an individual's ability to use many computer input devices. Therefore, a new assessment procedure was developed. A variety of input devices were analyzed with respect to their operating characteristics. Using the methods engineering techniques of time and motion study, elementary actions employed in device use were identified. To promote practical utility, the authors said the test should be quick and simple to administer while requiring no specialized or expensive test equipment. Consequently, the test was developed to require manipulation of simple objects of various sizes in ways that mimicked the actions necessary to operate computer input devices. In total, the test included six action categories: point/reach, slide, lift/move, place, reaction time, and repetition speed. To associate

performance in the test with performance using an input device, experiments were conducted with both physically disabled and non-disabled participants. All were asked to complete the assessment test and a benchmark cursor control task using five different input devices. Task completion time was used as primary dependent measure and a multiple regression equation was generated for each device using data from the non-disabled group. The equation described device performance as a function of performance on those subtests, which were thought to represent actions necessary in operation of the device. The resulting R^2 values indicated high predictive utility of the multiple regression equation, suggesting that: (1) the assessment test measured appropriate actions, (2) the correct subtests were chosen for each device, and (3) each subtest sufficiently measured each intended action. However, the same approach failed to yield useful results for the disabled group. There was a high degree of variance in correlations of test scores with functional performance for the disabled group. This result indicated that while the approach could predict functional performance with various input devices based on the assessment test scores, the method did not produce consistently reliable results. In general, the proposed method identified a promising direction for solving performance prediction problems.

Intuitively, it seems reasonable that the performance of an input device on a relatively simple target acquisition task would be a good indicator of performance on relatively more complex tasks. However, the work of Epps (1987) demonstrated that the predictive capability of a simple cursor positioning task was quite limited. Therefore, a more predictive method was needed. Chase et al. (1992) developed and evaluated a method of linking performance in a simple cursor control task to performance in a realistic computer-based task. A variety of

complex tasks were analyzed and decomposed into primitive components using User Action Notation (UAN; Hartson et al., 1990). The lists of primitive components generated were then examined to form a set of primitive cursor actions. In addition to target size and target distance, movement direction and selection mode (i.e., point and click vs. drag) were included. These four factors were then used to develop a primitive set of benchmarks. An experiment was conducted in which participants were asked to complete a task comprised of primitive cursor control actions as well as a complex graphics creation task. Each graphics subtask was analyzed using the UAN to determine the frequency of occurrence of each of the primitive actions. The predicted performance for each graphics task was then calculated by multiplying the user's performance on each primitive action by the number of occurrences of that primitive action within the task, and then summing for all primitive actions that occurred within the complex task. Correlations between an individual's actual times on the benchmark graphics tasks and the predicted times were computed. Results revealed a significantly high correlation between the two time values. The study suggested that the primitive benchmark set could be used to successfully predict performance in a more complex computer task with a given input device.

These studies demonstrate the potential to use physical skill assessment to predict performance in more complex functional tasks. Although most results could not be used to predict a user's exact performance level with a given device, it was possible to predict whether the individual could operate the device. Beyond the above studies, additional research has been conducted on real-time performance prediction methods in order to facilitate adaptive system control.

2.3.2. Real-time performance prediction

Jipp, Bartolein and Badreddin (2009) introduced an approach for dynamically adapting the level of automation of a wheelchair system based on the current level of a user's motor ability. An experiment was conducted to predict participant wheelchair driving behavior based on data obtained during task performance. During the experiment, an industrial PC was mounted on a wheelchair to gather data. Five positional goals were identified and a course was defined consisting of different sections with each section requiring driving from one position goal to another. Response measures were calculated based on data the PC recorded, including task time, average velocity and the distance the wheelchair drove forward (or backward or rotated on the spot) for each section and for the entire course, respectively. To link fine motor abilities with wheelchair control performance, the tremor, the precision, the arm-hand velocity, the wrist-finger velocity and aiming capabilities of users were assessed along with the traditional Motor Performance Test. Univariate general linear model analyses were conducted with the wheelchair movement variables as independent variables and the fine motor abilities as dependent variables. Results demonstrated that some psychomotor abilities are better predicted by task performance time, movement velocities and displacements than other abilities. Precision in control could be predicted best, i.e., by the greatest number of wheelchair variables. In contrast, there were only four significant predictors of wrist finger speed.

On the basis of the above results, Jipp, Bartolein, Badreddin, Abkai and Hesser (2009) presented a new technical system to automatically assess a user's level of psychomotor ability. In addition to the variables collected in the previous experiment, the number of input commands given were also recorded, as was the number of times the wheelchair started to

drive backward from a standing position, the number of directional changes, the means of the translational and rotational inputs given via the joystick, and the variance of the translational and rotational input commands. To predict psychomotor abilities on the basis of the wheelchair movements and the joystick input, two models were developed using the Structural Equation Modeling (SEM) methodology and the Bayesian Networks (BN) methodology. SEM was used as an extension of the general linear model approach enabling simultaneous application of a set of regression equations. The BN method was applied to learn the structure and parameters of a network from real data. The network structure was adjusted depending on the significance level for predicting psycho-motor ability. After the structure of the network was known, the parameters could be learned. Validation results indicated that the BN model outperformed the SEM model in predicting wheelchair users' fine motor abilities and especially precision. With the overall model producing different levels of predictive validity for various sections of the course, and a relatively small study sample size, the current experiment may be biased and the results are limited. However, the study still provided a promising approach in relating low-level motion variables with the high-level motor performance.

In contrast to the previous laboratory studies, Vandamme (2010) conducted research to explore whether the Assessment of Motor and Process Skills (AMPS) tool was appropriate for evaluating the quality of actual work performance. A large set of vocational disabled individuals were recruited for this study. Using AMPS results and medical diagnoses, individual ability to perform in regular unskilled employment was predicted. Subsequently, a skills questionnaire was developed with company managers to be included in a structured interview. The skills questionnaire was administered after a minimum of 2 weeks on the job

for employees. Questionnaire results were checked against the AMPS results to determine whether the AMPS scores for specific skills could predict performance on a shop floor. To investigate the agreement between the predicted outcome and actual outcomes on the shop floor, both proportions and Cohen's kappa were calculated. High prediction accuracy demonstrated that the AMPS was a useful tool for determining employee readiness for regular employment. The AMPS results indicated whether individuals who were searching for unskilled jobs had sufficient skills for regular employment in certain areas. The measurements were useful in determining employability and coaching individuals as they prepared themselves for employment in regular jobs. Although this approach relied heavily on subjective judgment and decisions of AMPS raters and managers, there was a strong potential for predicting actual task performance based on performance on designated motor tests.

2.4. Quantitative classification approaches

2.4.1. Design of task

Based on individual differences in motor skill level, proper training protocols need to be selected for trainees in order to promote desired uniform levels of performance across employees on a production line. In other words, trainees should be classified on the basis of their baseline skill levels in benchmark tests and trained accordingly. In a study conducted by Huegel (2009), participants performed a haptic-enabled virtual target-hitting task and were classified into three types of performers: high performers, low performers or transitional performers (see Figure 2.1 from Huegel (2009)). High performers were defined as being one standard deviation above the mean in initial performance. These performers started out strong,

made modest improvements and generated high scores across all trials. Low performers were defined as subjects whose final target hit count was more than one standard deviation below the mean in initial performance. The low performers improved only slightly across experimental sessions. Transitional performers had characteristics of both groups; they started out performing poorly, like the low performers, but performed as well as high performers in the end.

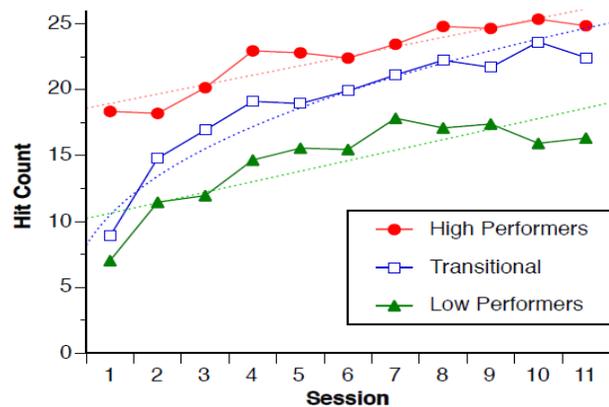


Figure 2.1. Classification of participants in target-hitting task (Huegel, 2009)

Although the study of Huegel (2009) demonstrated that it is possible to classify operator skill level based on their performance in simple tasks, it was not shown whether such results apply to other types of motor tasks. Howie et al. (2011) conducted a study attempting to answer this question using the Neverball video game as training task. Unfortunately, although all participants showed improvement in overall performance metrics as a result of playing the video game, the data did not show evidence that participants could be classified into three types, as done in the research of Huegel (2009). Expert performers were identified

through combination of task metrics, yet no clear split was detected between low performers and transitional performers. It was suggested that the failure to extend the results obtained by Huegel (2009) in the target-hitting task to the Neverball video game was likely due to a limited number of training sessions conducted during the experiment and a lower level of task difficulty.

In summary, these studies indicated that in order to classify trainee skill level, the task used to collect performance measures needs to be carefully designed with a sufficient number of samples and a challenging level of difficulty.

2.4.2. Feature selection algorithms

In addition to designing tasks in order to generate useful classification results, appropriate classification features should be extracted from performance data to serve as input variables in classification models. Some approaches for extracting variables from performance data on low-level skilled tasks have been demonstrated to be useful for evaluating upper-limb motor functions. However, not all features may be important for classification. In fact, some features may be redundant and some may not contribute to discriminating among classes of trainees. Such unimportant features may even degrade the classification results (Gruber et al., 2006).

Price & Sears (2008) examined the functional capabilities of individuals with varying levels of functionality, irrespective of cause of impairment. Fourteen computer-based subtasks were developed for assessment of arm, hand, finger, head and neck motion. A total of 30 measurement metrics were computed from the raw data obtained through a computer interface device, measuring movement capabilities (e.g. smoothness, distance and speed), pause control

(e.g. steadiness and consistency), and the time to task completion. However, a high level of collinearity was detected between various metrics, which significantly compromised validation results. That is, while the set of variables detailed and comprehensively assessed motor function, some metrics were redundant with other metrics. In order to reduce collinearity related concerns for subsequent data analysis, Price and Sears applied a remediation process to remove redundant metrics from their data set. The existence of collinearity among assessment variables in this research suggested a possible detrimental effect of using a variety of classification features.

Based on Price & Sears (2008) study, it is considered advantageous to separate important from unimportant skill classification features. To select features appropriately, researchers have developed different algorithms. Among recent studies, the work of Liu and Motoda (1998) is most notable. In their work on feature selection for knowledge discovery and data mining, several typical feature filter/selection algorithms were reviewed and many of these approaches have been applied to solve classification problems with experimental data. Gruber et al. (2006) presented a flexible online signature verification architecture using a Biometric Smart Pen to collect signal data. Referring to the work of Liu & Motoda (1998), Gruber et al. (2006) applied Las Vegas as well as Quick Branch and Bound algorithms for selecting important features for user skill classification. It was noted that the feature selection was actually performed on a person-specific basis since the features important for signature authentication for one person were not always important for verification of another person's signature. The need for individual performance modeling was also demonstrated by the results of Buchtala et al. (2005). They provided a strong example of individual-related feature

selection. Similarly, using the same interface technology, Dose et al. (2007) presented a quantitative method to distinguish the handwriting of healthy persons from patients. The Relief and Sequential Forward Feature Generation algorithms were applied to determine the level of feature importance, on the basis of the research by Liu & Motoda (1998). More recently, Gruber et al. (2012) applied an approach, the Gini index, to calculate the weight for each feature for selecting the most important ones for handedness classification.

Apart from the typical algorithms, some researchers have developed specific methods with respect to the characteristics of data sets. Hook et al. (2004) introduced a digitizing pen for signature identification. This device could measure the kinematics and dynamics of hand movement during the writing process by recording 3D pressure and pen inclination. Characteristic behavioral patterns were numerically extracted from the signals and a rapid feature classification and matching algorithm was executed to achieve person verification. From the raw data, a total of $n = 110$ features were extracted, including length, mean value, standard deviation, skewness, number of peaks, number of radial loops, sweep of polar angle, etc. These stochastic variables represented a compressed image of the original time series of data, mapped to an n -dimensional feature vector. To select suitable features for classification, the frequency distribution (histogram) of values for each of the 110 features was calculated for a sufficiently large sample taken from the recorded population of signatures. Each histogram provided an empirical probability density for the particular feature. The frequency distributions obtained from the entire study population were then compared with the histograms computed from repeated samples from one person. Features that showed low repeatability, little overall or large individual variance, insufficient specificity and pronounced redundancy were

discarded. In this way, the number of features was reduced to around 50, and the final feature space was expected to be balanced and mutually independent such that the state space could be occupied relatively uniformly, with regional clusters belonging to samples from the same individual.

In general, a variety of algorithms are feasible for solving feature selection problems and numerous studies have demonstrated theoretical and experimental validity of specific methods. In dealing with application problems, it is crucial to choose an algorithm that best matches the data characteristics such that irrelevant features can be identified and removed without discarding other any useful information.

2.4.3. Classification algorithms

In addition to feature selection algorithms, there are also many classification approaches to choose from. The most simple and straightforward algorithm is generalized linear discriminant analysis, which models the boundaries between class regions as generalized linear functions (Fisher, 1936). Such linearized models are easy to formulate, but they assume normality of data and equality of covariance (McLachlan, 1992). Unfortunately, data obtained in real experiments normally fails to satisfy such strict requirements and, consequently, classification results based on linearized models are usually not satisfactory. Therefore, it is more common to apply nonlinear models to solve the classification problems with real data sets (e.g., Spyers-Ashby et al., 1999; Buchtala et al., 2005; Jipp, Bartolein, & Badreddin, 2009; Huntsinger, Roupail, & Bloomfield, 2013).

When dealing with binary responses, Logistic Regression is generally the best classification algorithm choice with the advantages of ease of implementation, efficient

computation, limited assumptions, and flexibility in model component inclusion, etc (Pohar et al., 2004). However, when the number of classification groups is larger than two, the algorithm needs to be applied repeatedly for multiple paired group classification to obtain the complete solutions. In this way, the computation load is increased.

In the case of complex classification problems with large sets of features extracted from experimental data, statistical methods (e.g., Sum Rule, naïve Bayes, k-Nearest-Neighbor (KNN) classifier, etc.), machine learning approaches (e.g. Support Vector Machine (SVM), Radial Basis Function (RBF) networks, Dynamic Time Warping (DTW), Hidden Markov Models (HMM), etc.), and specialized ensemble methods (e.g., dynamic classifier selection), are commonly applied (Gruber et al., 2006). These classifiers all have their own advantages and disadvantages. With differences in data characteristics, the most suitable classifier will vary. In addition, when using a specific classifier, a number of instances of the classifier may be applied, e.g., RBF networks with different learning algorithms, or SVM with different kernel functions, etc. Therefore, to improve the accuracy of classification problem solutions, usually more than one algorithm with several parameter settings would be tested on a single data set (e.g., Kuncheva, 2002; Gruber et al., 2012).

Rohlik et al. (2003) developed a text recognition method for an electronic pen that produced signals corresponding to the movement of the pen on paper. Signals were described by a set of primitives and HMM were used for word recognition. Initially, DTW techniques were tested for classifying words written by different participants. However, results were not sufficient for reliable classification. Moreover, the DTW method could not be used for a large vocabulary, making it unsuitable for possible commercial applications. Therefore, the HMM

approach was chosen. A Baum-Welch algorithm was used to train the model and a backward error-propagation algorithm was used for recognition purposes. Experiment results showed that the HMM-based algorithm could be used for word recognition. However, it was also found that the process of training the HMM could be quite time-consuming and the achieved recognition rate was about 5–10% below the accuracy of state-of-the-art hand-written text recognition systems. Rohlik et al. (2003) were also not able to improve the computational efficiency of the algorithm and concluded that other algorithms need to be applied in further research.

In the research of Dose et al. (2007), Cost-sensitive Support Vector Machines (C-SVM) were used to classify time series data on handwriting using characteristic features as model inputs. SVM uses a hyperplane to separate two classes. For classification problems that cannot be linearly separated in the input space, SVM can find a solution using a nonlinear mapping from the original input space into a high-dimensional “feature space”, where an optimally separating hyperplane is searched. Such hyperplanes are identified as optimal when they have a maximal margin, where margin refers to the distance separating the hyperplane from the closest data point (i.e., the support vector). The transformation is usually realized by nonlinear kernel functions. The C-SVM algorithm allows for non-linear classification analysis and minimizes misclassifications in data sets. The resulting classification rates for Dose et al. study were quite high when using a Gaussian kernel function.

Gruber et al. (2012) also chose SVM algorithm to classify handedness among preschool children. In addition to this binary classification, the SVM was extended to produce a real-value output that could be interpreted as a posterior probability for class membership given a

certain set of inputs. Such probability estimates can be useful when the output of one classifier is used as input for another classifier, or when the reliability of a classification result must be considered. In the Gruber et al. study, several tests were conducted and an ensemble of classifiers was used to handle the subproblems separately instead of using a single classifier for the entire problem. After generating the results for each subtest using the C-SVM with a Gaussian kernel, a C-SVM with a linear kernel was used to combine results of sub-problems and return an overall classification result. This result was also a value that could be interpreted as a posterior probability of a child being right- or left-handed. Validation results indicated good performance of the classification algorithm. In similar research, Gruber et al. (2012) conducted experiments with less complex motor subtests and other classifier paradigms were applied, including KNN, decision trees, naïve Bayes and neural network. However, no improvements in results were achieved. In general, the research by Gruber et al. (2012) demonstrated a reliable and objective handedness classification algorithm.

While numerous studies have developed and applied different types of classification algorithms with various parameter settings, all these approaches have some commonalities. In order to achieve a successful classification model, sufficient data should be provided while feature dimensionality should be constrained. Moreover, all the input data must have *a priori* labels. For example, to establish a model to classify handedness, performance data must be collected from participants with known handedness in order to properly train the model. However, such conditions are not always satisfied in model development. In such cases, clustering methods are often applied, which do not require *a priori* labels of observations. Aharonson & Krebs (2012) conducted a study to search for patterns in motor performance data

acquired during a baseline evaluation, such that prediction of progress during following therapy sessions could be made. This research was aimed at identifying participants that would not respond to therapy before actually conducting sessions. To achieve this target, kinematic and kinetic data were acquired during a robot-mediated evaluation session of children with motor impairments. Since the data set size was relatively small and no information was available to know whether a child would respond to the therapy, i.e., there was no *a priori* label information, a clustering method, K-means algorithm, was applied. The K-means algorithm finds locally optimal solutions minimizing the sum of the squared distance (L2) between each data point and its nearest cluster center, where each cluster is modeled by a spherical Gaussian distribution. One shortcoming of this algorithm is that the number of clusters is typically determined based on user performance. To overcome this limitation, an exhaustive search for the best number of clusters (K) was conducted in the range of 2-20. Results indicated that the clustering method succeeded in identifying those children that ultimately showed no significant improvement during therapy sessions, i.e., non-responders. This research presented an example application of using a clustering method to solve classification problems. In addition, the experimental results supported the hypothesis that participant performance progress in designated therapies could be classified and predicted based on an early baseline evaluation session. In this way, it is possible to identify participants that would not respond to therapy and make proper changes in early evaluation sessions, thus saving time and effort while improving therapy effectiveness.

In summary, a quantity of classification algorithms have been developed and improved to solve practical problems. Table 2.1 presents a summary of the prior studies along with the algorithms applied.

As has been discussed, the various classification models have their own advantages and limitations. Table 2.2 presents a comparison of the classification algorithms reviewed thus far. In general, when choosing an algorithm for classification purposes, it is necessary to identify whether the model assumptions can be satisfied by the obtained data set. Meanwhile, for a specific model, it is crucial to determine appropriate parameter settings in order to achieve high classification accuracies.

Table 2.1. Summary of reviewed studies and algorithms

Study	Feature Selection Algorithms	Classification Algorithms
Hook et al. (2004)	Feature frequency distribution matching	Heuristic algorithm
Gruber et al. (2006)	Las Vegas algorithm; Quick Branch and Bound algorithm	Various methods including RBF, SVM, DTW and HMM
Dose et al. (2007)	Relief and Sequential Forward Feature Generation algorithms	C-SVM with Gaussian kernel function
Gruber et al. (2012)	Gini index	C-SVM with Gaussian and linear kernel functions
Aharonson & Krebs (2012)	None, due to small data sizes	K-means algorithm (clustering algorithm)

Table 2.2. Comparison of reviewed classification algorithms

Algorithms	Relevant Studies	Advantages	Disadvantages
DTW	Rohlik et al. (2003); Gruber et al. (2006)	Well-established for comparing time series	Heavy computational load; not reliable for large samples
HMM	Rohlik et al. (2003)	Proved significantly effective in speech and handwriting recognition; can be easily extended	Strict Markovian assumption; large number of parameters; time-consuming in model training
RBF	Gruber et al. (2006)	Fast training; requires few training samples	Cannot deal effectively with irrelevant attributes
SVM	Dose et al. (2007); Auria & Moro (2008); Gruber et al. (2012)	Flexible model form; robust computation; unique solution delivery	Lack of transparency of results (but not harmful to classification accuracy)
Clustering methods	Aharonson & Krebs (2012)	Do not require <i>a-priori</i> labels	Limited to relatively small data sizes

CHAPTER 3: PROBLEM STATEMENT

3.1. Related work summary

In the previous section, a series of relevant research from four different areas was reviewed. The results of these studies provided significant direction for the development of the present research.

Among various VR systems developed for motor skill training, those integrating different haptic features have led to the greatest effect of tactile feedback on operator performance. With the development of haptic technology, haptic cues provided in virtual training environments have shown great effectiveness for improving trainee performance during training task completion. However, such improvements have not been verified through retention test performance. This means that although performance may be better in the presence of haptic guidance, trainees may not sufficiently develop their own motor skills during training. Related to this, consistent haptic guidance has been shown to have detrimental effects on motor skill training results due to user over-reliance during training. Researchers have developed different solutions to address this problem. The most recent and notable methods are the progressive shared haptic control scheme or the haptic disturbance scheme. The former attempts to decrease user dependence on haptic guidance by progressively reducing haptic cueing based on trainee performance. The latter attempts to motivate trainees to develop their own ability to accomplish tasks by disturbing force feedback during training task performance. Experimental results have shown that both schemes have succeeded in improving trainee motor skill learning. However, haptic disturbances have led to significantly superior

performance in retention tests. In addition, compared to the scheme of progressive shared control, a training protocol with haptic disturbances is much easier to implement. Consequently, in designing different training protocols for this research, the conventional haptic guidance feature, as well as two types of haptic disturbance features (resistive force feedback and noise-like force feedback) were selected.

To objectively and quantitatively assess operator motor skill level, a number of studies have developed various motor tasks. The most common motor assessment tasks include target following, trajectory duplication, and block manipulation. All these tasks are easy to understand; thus, requiring little cognitive skill for accurate performance. However, they all require operators to have dexterous hand-motor control ability and good hand-eye coordination to achieve accurate movement and orientation. In addition, assessment results from these tasks need to be validated by comparison with results obtained with standardized physical methods. In this research, a simple virtual block manipulation task was used to assess the motor skill level of operators. This type of task was selected because of its resemblance to standardized psychomotor tests and manual assembly work. With the Phantom haptic device, it is possible to automatically collect kinematic information and store the data in the form of time series. To verify that the block manipulation task was valid and reliable for operator motor skill assessment, results were compared with an established motor test, the PPT. The PPT results were used as a “gold standard” for skill assessment with the haptic-VR version of the block manipulation tasks.

Although the tasks used in previous research for motor skill assessment have also been used for motor skill training, in some cases, the two tasks are unique. Therefore, it is necessary

to determine whether motor skill evaluation results can be generalized to training task performance, i.e., whether an individual showing high-level motor performance in a low-level skill test will perform as well in a high-level skilled motor task. Compared to the number of studies focusing on motor skill training and assessment, research conducted on motor performance prediction is limited. Several approaches, including objective and subjective methods, have been proposed to predict motor performance using baseline evaluation metrics. Referring to approaches validated by previous research, this study applied linear regression as a basis for preliminary evaluation of the utility of parameters generated from assessment tasks for predicting motor performance in generalized motor tasks.

While a large set of parameters can be computed from kinematic data collected during motor task performance, not all parameters are necessarily important in developing a skill classification model. To identify the most relevant parameters containing important skill information, and to establish a skill classification model, a feature filter/selection algorithm is necessary. Various algorithms have been developed for application to data with different characteristics. To guarantee that irrelevant parameters are not selected without also missing selection of relevant features, algorithms must be chosen carefully.

Similarly, with the selected features, a number of skill classification models can be applied. In comparison with conventional linear regression models, machine learning algorithms are usually superior and do not require normality or equality of covariance of data. In addition, with various parameter settings to choose from, the application of machine learning algorithms is more flexible, especially when processing data collected from physical experiments. In previous studies, researchers have commonly constructed multiple models and

then developed an ensemble of results for more accurate skill classification. Therefore, it was expected that accurate classification results might be obtained by applying a similar approach in this research.

3.2. Research objective

This study was motivated by the need for selection of proper training protocols for operators based on baseline motor skill assessment in order to achieve target levels of motor performance. Two phases of experimentation were conducted in this research. The first phase aimed at developing a quantitative algorithm for accurate classification of operator motor control skill based on baseline performance. The objective of the second phase was to identify haptic-VR simulation training design features for effective skill training to requisite levels and to validate the classification algorithm developed in the first phase.

3.3. Research hypotheses

There were three hypotheses formulated for this research. The first hypothesis (H1) was that a customized haptic-VR manipulation task would prove to be a valid and reliable task for motor skill assessment as compared with a standardized physical psychomotor task. The use of the VR system was primarily motivated by the capability to automatically collect data on user motor performance in a realistic 3D representation of a test task and to automate the processing of such data for skill assessment.

The second hypothesis (H2) was that a non-linear classification model assessing motor control skill would yield superior skill classification accuracy as compared with a linear statistical model.

The third hypothesis (H3) was that specific haptic features (e.g., haptic disturbances) could be designed and selected for operator training to desired motor skill levels based on motor skill classification using baseline performance data. No prior research has successfully demonstrated the selection of haptic-VR training protocols for motor skill development based on *a priori* skill classification using simplified motor tests.

CHAPTER 4: PHASE I EXPERIMENT METHODOLOGY

4.1. Participants

Twenty-one participants (11 male, 10 female, average age = 30.2, SD = 9.2) were recruited from both university and off-campus population for this phase. All participants were right-hand dominant, as verified with a questionnaire. The participants were also required to be in good health and have 20/20 or corrected vision (if vision was corrected, they were required to wear the corresponding prosthetic to reach 20/20 acuity during the experiments). All participants were expected to have no prior experience in similar experiments or training and no background information on the research goals or experimental hypotheses. In addition, participants were expected to be able to follow all instructions given for performing the motor tasks. The experiment tasks mainly involved repetitive elbow-wrist movements in a seated posture. Any participant with current or chronic wrist disorders (e.g., carpal tunnel syndrome) was excluded from the study. The above screening criteria were summarized in a questionnaire (see Appendix A) for potential applicants. Only those satisfying all requirements were contacted and scheduled for participation.

4.2. Equipment and apparatus

A VR simulation of the block manipulation task was presented to participants on a PC integrated with a stereoscopic display using a NVIDIA® 3D Vision™ Kit, including 3D goggles and an emitter. Stereoscopic rendering of the task simulation was facilitated by an OpenGL quad-buffered stereo, high-performance video card (NVIDIA® Quadro™). A

SensAble Technologies Phantom Omni[®] Haptic Device was used as the haptic control interface. The Omni included a boom-mounted stylus that supported 6 DOFs movement and 3 DOFs force feedback. All data on participant performance with the Omni was recorded automatically by the simulation software.

4.3. Experimental tasks

All participants were asked to complete two tasks in the experiment. The first one was the computer-based virtual block manipulation and the second was a standardized physical psychomotor test (PPT).

4.3.1. Dice manipulation task

This VR task was previously developed for a National Science Foundation research project and was demonstrated to be effective for familiarizing experiment participants with use of a haptic device with a stylus control interface (Ma et al., 2012). The task involved manipulation of a die using the haptic-VR workstation described above. The training VE included a single virtual die placed near the left-side of a virtual work surface and a square near the right-side of the work surface (see Figure 4.1). A 2D image of a single side of a die (stimulus) was presented at the top of the screen in the display area. The goal of the task was to move the die as quickly and accurately as possible to the target square with the top surface of the die matching the stimulus.

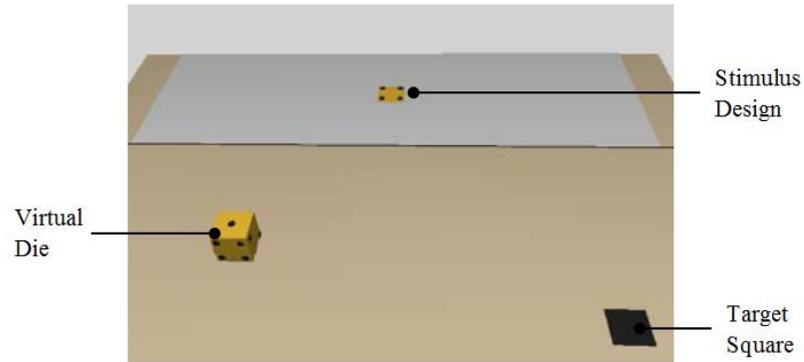


Figure 4.1. Dice manipulation task

The virtual die was modeled after a standardized die, i.e., the opposite sides always sum to seven, and the sides with numbers one to three were arranged clockwise around a common corner after folded. Figure 4.2 shows the configuration of the virtual die when unfolded.

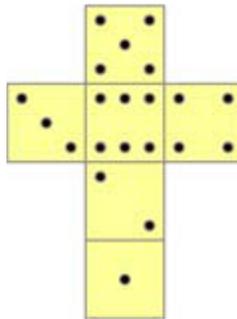


Figure 4.2. Unfolded virtual die

In this task, the “target side” referred to the side of the die facing up when it was placed on the work surface. The number of dots appearing on the target side determined the specific

orientation. Thus, sides with 1, 4 or 5 dots were orientation independent, while sides with 2, 3 or 6 dots were orientation dependent (i.e., a side with 4 dots appeared identical no matter how it was rotated, yet a side with 3 dots would slant upward or downward with different rotation). Based on the orientation and rotation required, the dice manipulation task included four levels of difficulty:

- (1) In the 1st level, the target side of the die appeared face-up at the initial position and required no specific orientation to match the stimulus. The starting position of the die was randomly determined within a defined X-Y plane on the left side of the work surface. Therefore, at level 1, participants only needed to grasp the die block and place it within the target square without rotation.
- (2) At level 2, the target side was again face-up at the initial position, yet the die required a specific re-orientation of 90 degrees. Participants had to rotate the top surface of the die prior to placement.
- (3) In the 3rd level, the target side was offset by 90 degrees at the starting position and no specific orientation was required. To achieve this level of task performance, participants needed to rotate the die to locate the target side while moving the block to the target square. However, the final orientation was always correct, i.e., the target side had 1, 4 or 5 dots.
- (4) At level 4, the target side was offset by 180 degrees at the starting position and the target side required a specific orientation. Therefore, participants had to rotate the block to the opposite side during movement and place the block with a specific orientation. The four levels of task difficulty are summarized in Table 4.1.

Table 4.1. Levels of dice manipulation task difficulty

Design	Orientation Requirement	Target Side at Initial Position
Level 1	Flexible	Face-up
Level 2	Specific	Face-up
Level 3	Flexible	Adjacent
Level 4	Specific	Opposite

Performance differences among the above four levels of dice task difficulty were previously confirmed by Ma et al. (2012). Consequently training under each successive level was expected to contribute to greater motor skill in the dice manipulation.

4.3.2. PPT

The standardized PPT (Model 32020; Lafayette Instrument[®]) was presented to all participants. Only the “right-hand” portion of the test was used in this study. The administration of the task followed instructions by “Lafayette Instrument for the Purdue Pegboard Test” (Lafayette Instrument Company Inc., 2002).

4.4. Procedures

The entire experiment was set up and conducted in an independent and quiet lab room without disturbances. A brief introduction to the study was provided for participants prior to the beginning of experiment. The participants were told that they were going to complete a set of computer-based motor task trials and one standardized psychomotor test with their right hand only. All screening criteria for recruiting participants were confirmed before the experiment began and the participant was asked to sign an informed consent form (see Appendix B).

Following the briefing, participants were directed to sit in front of the VR workstation to receive haptic device training. The 2nd level of the dice manipulation task was used to demonstrate how the haptic control could be used to reach, grab, move and release the die. Participants were asked to use only their right hand to perform the task and to complete the task as quickly and accurately as possible. Upon participant request, an additional practice trial was provided.

After the device training, each participant was asked to perform 40 trials of the dice manipulation task. Participants were required to perform under the four levels of difficulty in ascending order, from the easiest to the most difficult. Each of the first two levels was repeated in 5 trials in order to facilitate participant familiarity with the task. Levels 3 and 4 were each repeated in 15 trials in order to generate substantial data on participant motor performance for baseline skill evaluation.

A mandatory 5-minute break was provided after completion of the dice manipulation task. Participants were then asked to perform three trials of the PPT. The total time for the experiment was approximately 30 minutes. Detailed experiment instructions are provided in Appendix D.

4.5. Construction of model

4.5.1. Task performance parameters

During the 40 trials of the dice manipulation task, kinematic response data were automatically captured with the haptic device and recorded in the form of time series with a frequency of 10 Hz. The raw data included the elapsed task time, and stylus displacement and

orientation in the Cartesian coordinate frame. Based on the raw kinematic data, a set of 29 task performance parameters were computed for each trial, as summarized in Table 4.2. These parameters were determined based on the results of prior research (Amirabdollahian et al., 2005; Emery et al., 2010; Zollo et al., 2011; Ma et al., 2012) investigating performance in virtual object manipulation tasks.

Table 4.2. Summary of task performance parameters

Parameter	Explanation
Reaction time	Time from the start of the task to actual performance.
Approach time	Time taken to move from the pointer origin position to the initial position of the die.
Approach moving speed (3)	Average speed along x-, y-, and z-axes while moving the pointer to the initial position of the die.
Approach moving acceleration (3)	Average acceleration along x-, y-, and z-axes while moving the pointer to the initial position of the die.
Approach rotating speed (3)	Average rotation speed along yaw, roll, and pitch while moving the pointer to the initial position of the die.
Approach rotating acceleration (3)	Average rotation acceleration along yaw, roll and pitch while moving the pointer to the initial position of the die.
Manipulation time	Time taken from grabbing the die to release at the target square.
Manipulation moving speed (3)	Average speed along x-, y-, and z-axes while manipulating the die.
Manipulation moving acceleration (3)	Average acceleration along x-, y-, and z-axes while manipulating the die.
Manipulation rotating speed (3)	Average rotation speed along yaw, roll, and pitch while manipulating the die.
Manipulation rotating acceleration (3)	Average rotation acceleration along yaw, roll, and pitch while manipulating the die.
Deviation distance	Distance between the center of target square and the center of the placed die.
Deviation angle	Degrees of rotation of the placed die relative to the target square.

4.5.2. Feature generation

After computing the 29 parameters for each of the 40 trials of the task, a 40×29 matrix was obtained. For each parameter, four statistical features – mean, SD, minimum, and maximum – were generated to describe participant performance across the 40 trials. In addition, the 40 data points corresponding to one parameter were fitted using a learning model in order to compute the learning percentage (k-value), according to the following formula (Konz & Johnson, 2004):

$$Y_n = Y_1 \cdot n^b$$

where Y_n is task performance in the n^{th} trial. Taking the natural log of both sides, the formula becomes:

$$\ln(Y_n) = \ln(Y_1) + b \cdot \ln(n)$$

After solving coefficient b , the learning parameter, using linear regression, the learning percentage (k) can be calculated as:

$$k = 2^b$$

Therefore, with five features corresponding to each parameter, in total, a set of 145 (i.e., 29×5) features were generated as potential predictors for developing the motor performance level classification models.

4.5.3. Motor ability level

The average score obtained from the three trials of the PPT was calculated for each participant and used as “gold standard” in order to determine participant motor ability level. The instructions from “Lafayette Instrument for the Purdue Pegboard Test” provided

normative data collected on 225 healthy adults, 15-40 years old (female mean score = 16.64, SD = 2.10; male mean score = 15.65, SD = 1.71). However, participants in the present study did not show similarly good performance (mean score = 14.52, SD = 1.42). Similar results were also obtained in Clamann's (2014) research experiment (mean score = 14.67, SD = 2.15). It is possible that the normative data set previously developed with the PPT is generally representative of persons with elevated levels of motor skill. Considering differences among the study samples, statistics on the current data set were used to determine performance level ranges. On this basis, each individual was assigned to one of three motor performance levels, i.e., high, low and medium. High performers were identified as those achieving an average score at least one SD higher than the overall average PPT score, i.e., higher than 15.94. Low performers were identified as those persons achieving an average score at least one standard deviation lower than the overall average PPT scores, i.e., lower than 13.10. Medium performers were identified as those persons generating an average score in between the high and low performer levels, i.e., between 13.11 and 15.93. In this way, the 21 participants were labeled with three levels, including seven participants in each level. The level information for each individual served as an *a priori* label for developing the classification models.

4.5.4. Construction of linear regression model

To evaluate the utility of the generated features for assessing motor performance, the average PPT scores were related to features using a linear regression approach. All 145 features were available to serve as predictors, which was much larger than the size of responses (i.e., 21 average PPT scores). Thus, in order to guarantee the validity of the linear model, a forward selection algorithm was applied. Linear regression analysis includes the assumptions of

linearity between responses and predictors, independence and homoscedasticity of errors, and normality of error distribution (Neter, Wasserman, & Kutner, 1990, pp. 118-128). For selection of features to fit the model to the data set, adjusted R-squared was used as selection criterion. Figure 4.3 shows a flow diagram of the forward selection algorithm applied in this study. Features leading to a greater model fit were selected and used as predictors to construct the linear model. To verify the reliability of the dice manipulation task in assessing participant motor skill (H1), adjusted R-squared values were computed based on the final linear model.

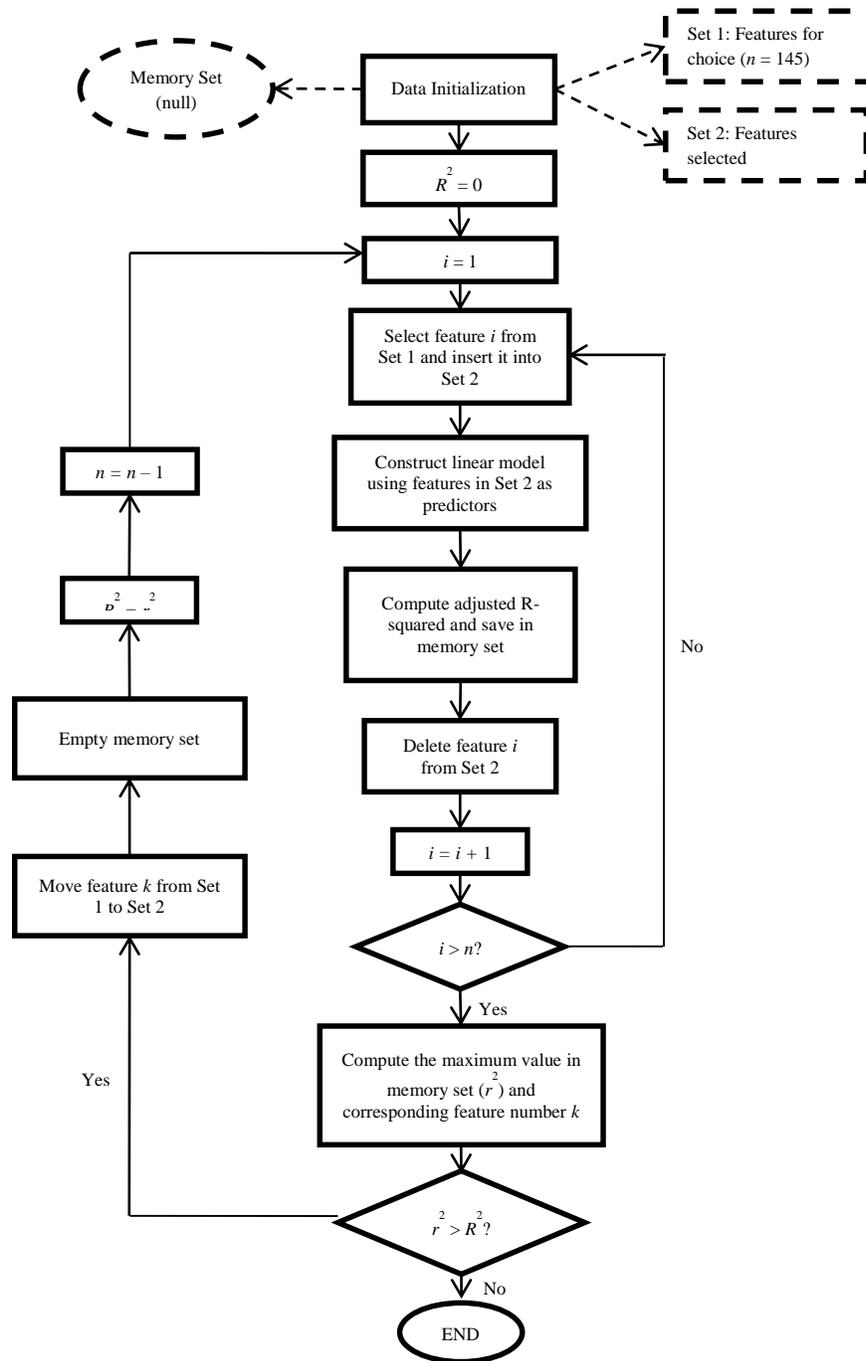


Figure 4.3. Forward selection algorithm for developing linear model

4.5.5. Development of non-linear algorithms

Linear regression models have been applied in many prior research studies (e.g., Chase & Casali, 1995; Jipp et al., 2009). However, when compared with non-linear models, this approach has strict assumptions regarding the normality of data, linearity of condition means, etc. As summarized in Section 4.5.1, some parameters generated based dice manipulation task performance were related to human response time (e.g., reaction time). Prior research has demonstrated these responses generally follow a lognormal distribution (Bonto-Kane, 2009). Consequently, such data poses violations of the assumptions of linear regression. On this basis, it was reasonable to expect that non-linear models might generate results with greater accuracy than the linear regression model.

In the previous section, various classification algorithms were discussed in association with experimental applications. Based on the results of these studies, the SVM algorithm, with the advantages of flexible model form, robust computation and unique solution delivery (Auria & Moro, 2008), was applied as the main classification algorithm in this research. Previous research has commonly used linear kernel functions. However, it has been suggested that SVM with a radial basis function (RBF) kernel yields good classification results with relatively little “effort” (Staelin, 2002). Considering the influence of data characteristics on classification results, it would be difficult to predict the best-performing kernel function. Therefore, three different kernel functions, i.e., the RBF, linear, and polynomial, were tested to obtain a model with the highest classification accuracy. Similar to the linear classification model development, a forward selection algorithm was used to identify those features from the complete set of 145 features (refer to Figure 4.4) providing the greatest predictive utility in the context of the

model. Since the SVM algorithm does not generate a coefficient similar to adjusted R-squared for use in evaluating model fitness, classification rate of error was used as a criterion instead. The rate of error was computed by applying a 75/25 cross validation method (as described in Section 4.5.6).

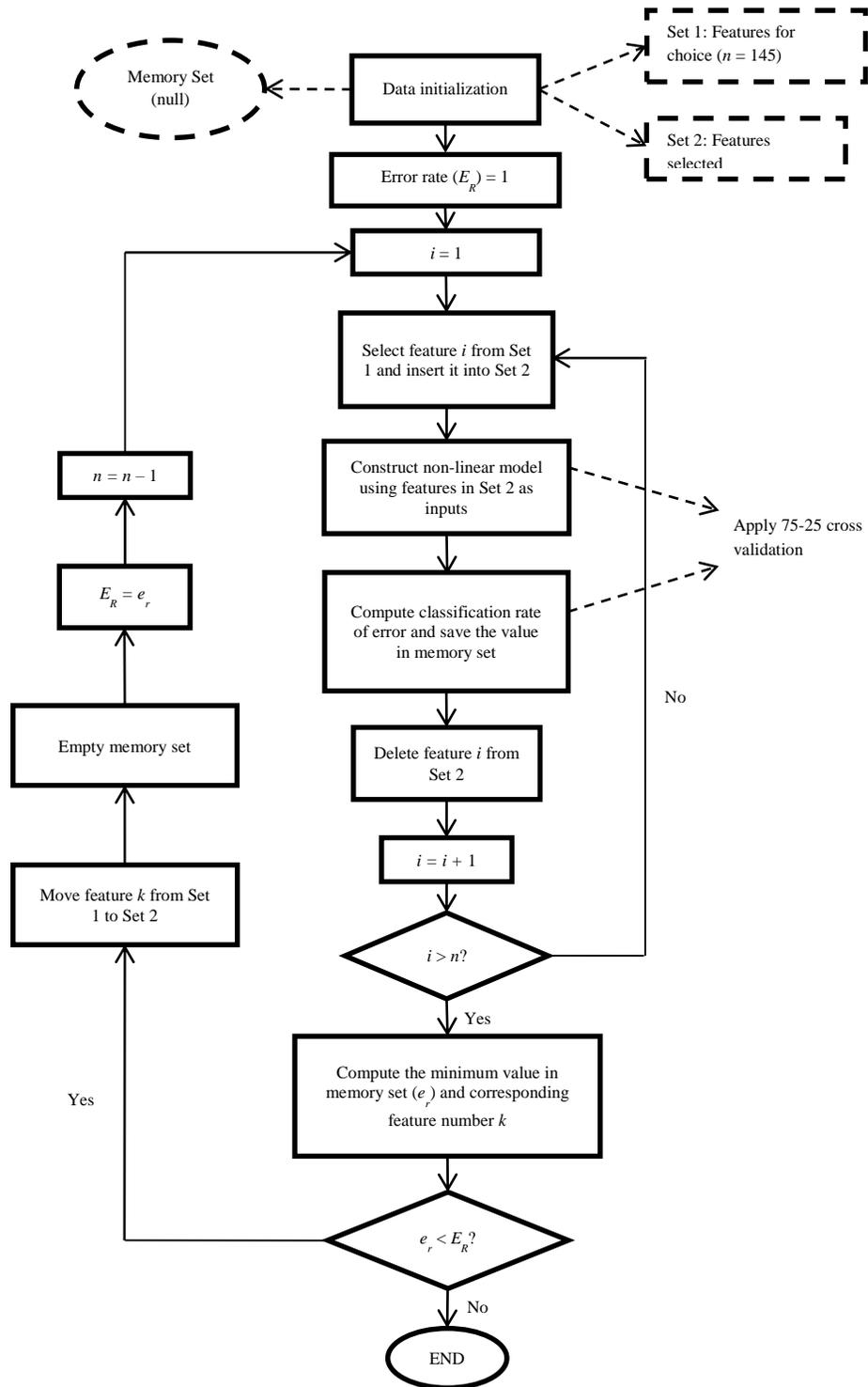


Figure 4.4. Forward selection algorithm for developing non-linear models

In this study, the motor performance level determined for each participant (following the method in Section 4.5.3) was used as an *a priori* label and it contained ordering information (i.e., high > medium > low). However, the original SVM algorithm does not utilize such information in developing classification models. Although this problem can still be solved using the SVM methodology, discarding the ordering information potentially degrades the predictive performance of the resulting classifier (Frank & Hall, 2001). To make use of such information, in addition to the original SVM models, a simple algorithm presented by Frank and Hall (2001) was associated with SVM in order to develop ordinal classification models. This algorithm was demonstrated to be effective for improving model accuracy through extensive experimental results (Frank & Hall, 2001).

In applying this method, the original dataset (with attributes of high, medium, and low for each participant) was converted into two new datasets:

New dataset 1: Converting the class attribute “low” to 0 and all others to 1.

New dataset 2: Converting the class attribute “high” to 1 and all others to 0.

Subsequently, binary SVM were applied to the two new datasets separately, generating classification models with probabilities of $\Pr(\text{Target} > \text{low})$ (i.e., the likelihood new participant performance would be higher than a low performer) and $\Pr(\text{Target} > \text{medium})$ (i.e., the likelihood new participant performance would be higher than a medium performer). The probabilities that a new participant belonged to each of the three groups were then computed as:

$$\Pr(\text{Target is low}) = 1 - \Pr(\text{Target} > \text{low})$$

$$\Pr(\text{Target is medium}) = \Pr(\text{Target} > \text{low}) - \Pr(\text{Target} > \text{medium})$$

$$\Pr(\text{Target is high}) = \Pr(\text{Target} > \text{medium})$$

The class with the highest estimated probability was assigned to the new participant.

Experimental results of Frank and Hall's (2001) study showed that for datasets with fewer class values, there was relatively little improvement in SVM classifier accuracy as a result of exploiting the ordering information. In the present research, only three classes were involved and it was possible that SVM classification accuracy would not be markedly better for ordinal classification models. Therefore, both original and ordinal SVM models with three different kernel functions were developed.

As a complementary approach, a cumulative logistic regression (CLR) model (Huntsinger et al., 2013), also known as ordered logistic regression, was also developed. The CLR model is widely used in social sciences to estimate the relationship between an ordered categorical dependent variable and a set of independent variables (Huntsinger et al., 2013). In applying the CLR algorithm, the underlying value is estimated as a linear function of the independent variables and a set of cut points. The probability of observing a specific outcome corresponds to the probability that the sum of the estimated linear function and random error is within the range of estimated cut points for the estimated outcome. For this type of analysis, the error term is assumed to be logistically distributed. The coefficients of the linear function and cut points are estimated using maximum likelihood.

Although the CLR model has been demonstrated as a powerful tool for solving various problems (Chen et al., 2008; Liu et al., 2008; Huntsinger et al., 2013), few prior studies have applied it in dealing with human performance data. Therefore, in order to obtain a classification model with the highest accuracy, the original and ordinal SVM models, as well as CLR model,

were all developed. The one with lowest prediction rate of error was selected for application in the Phase II experiment.

4.5.6. 75/25 cross validation method

To evaluate the classification accuracy of the various models, a 75/25 cross validation method was applied. The entire experimental data set was randomly partitioned such that 75% of it (i.e., 16 data points) was used to train the model while the remaining 25% (i.e., 5 data points) was used to test the model accuracy. To reduce bias, the process was repeated 500 times such that all data points were used as both training and model verification observations. The average rate of error was computed as an evaluation criterion for each replication in order to compare model accuracy. Figure 4.5 presents a data flow diagram to describe the cross validation process. To verify the hypothesis that a non-linear model would yield superior classification accuracy to a linear model (H2), the rates of error were computed and compared between the various models.

All computation and algorithm development for this phase of the study was conducted using *R*, a software for statistical computing and graphics. Figure 4.6 presents a diagram summarizing the procedures and data flow for the Phase I experiment.

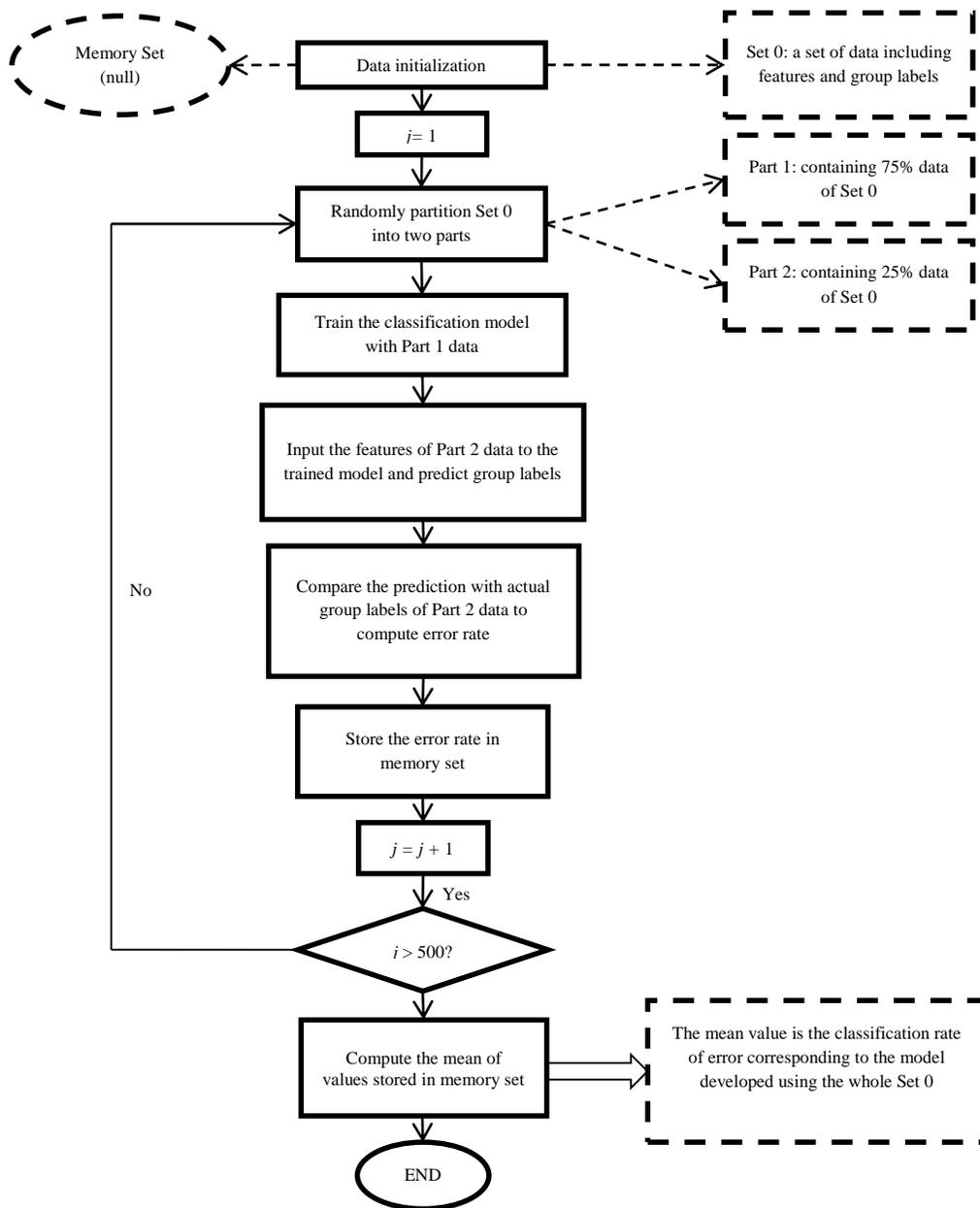


Figure 4.5. 75/25 cross validation approach

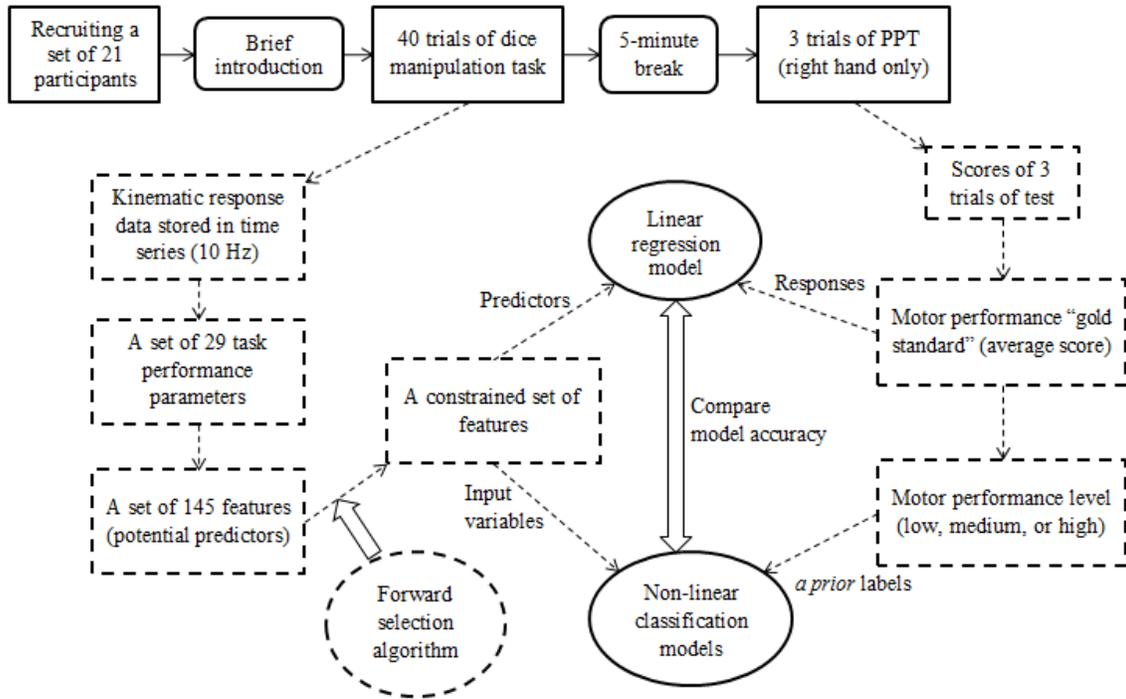


Figure 4.6. Phase I experiment procedures and data flow

CHAPTER 5: PHASE I DATA ANALYSIS AND RESULTS

5.1. Linear model

By applying the forward selection algorithm and using the adjusted R-squared as a criterion, a total of 15 features were selected as predictors for inclusion in the linear classification model. As a complementary approach, the selection algorithm was also applied using the Akaike Information Criterion (AIC) as model fitness indicator. Both processes generated the same set of features. For each of the 15 intermediate linear models, error rate was computed using the 75/25 cross-validation method, as described in Section 4.5.6. The changing trend of the adjusted R-squared value and rate of error, along with the increase in number of predictors during the selection process, is presented in Figure 5.1. For the first 11 predictors, the model accuracy improved as more predictors were involved in the model. However, beginning with the model with 12 predictors, the rate of error started to rebound, as additional predictors were added. In this case, involvement of more predictors actually degraded the model performance in prediction. Therefore, the final linear model was chosen as the one with the first 11 predictors, which had a prediction error rate of 8.88% and adjusted R-squared value of 0.9966.

The finalized linear model is summarized in Table 5.1, with each of the 11 predictors specified in terms of the corresponding parameter and feature (e.g., minimum of approach time).

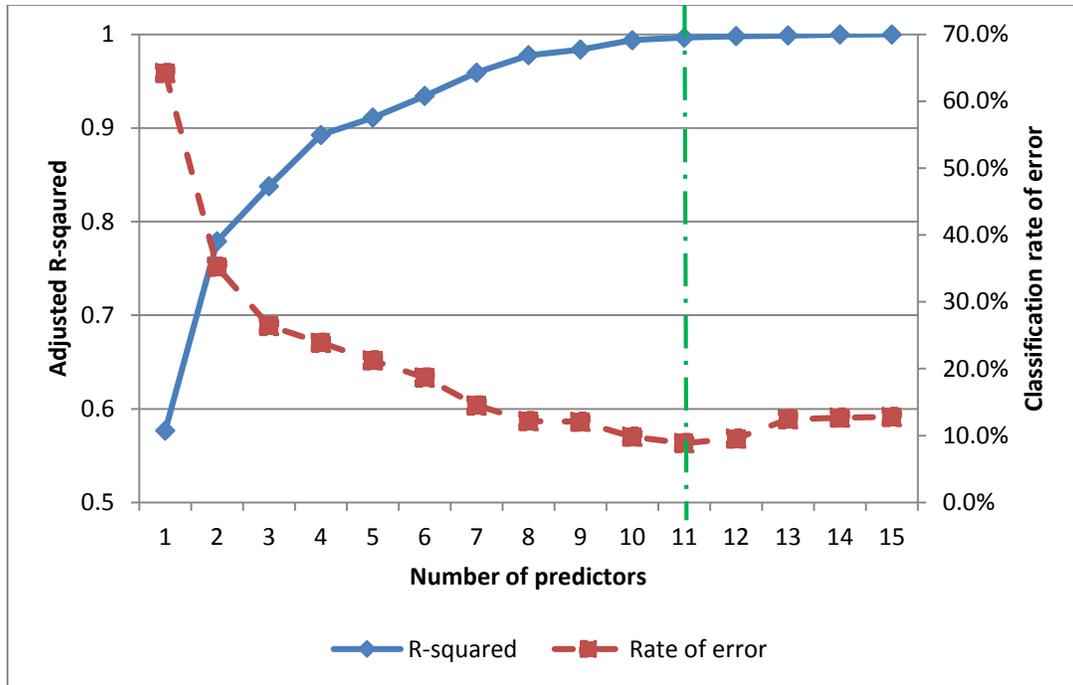


Figure 5.1. Change of adjusted R² and error rate of linear model along with number of predictors

Table 5.1. Summary of predictors in the final linear model

No.	Parameter	Feature	t-value	p-value
1	Approach speed along y-axis	Minimum	20.91	< 0.001*
2	Approach acceleration along z-axis	SD	26.76	< 0.001*
3	Approach rotation acceleration in roll	Maximum	-5.70	< 0.001*
4	Approach rotation acceleration in yaw	Maximum	10.86	< 0.001*
5	Manipulation acceleration along z-axis	SD	-20.51	< 0.001*
6	Approach acceleration along x-axis	k-value	17.40	< 0.001*
7	Manipulation acceleration along z-axis	k-value	-4.68	0.0016*
8	Approach acceleration along x-axis	Minimum	-8.98	< 0.001*
9	Manipulation acceleration along y-axis	Mean	-6.68	< 0.001*
10	Approach time	Minimum	-6.28	< 0.001*
11	Approach speed along y-axis	k-value	-2.94	0.0188*
Adjusted R-squared				0.9966
Overall significance: F(11, 9) = 511.2, p < 0.001*				

5.2. Non-linear algorithms

SVM models using the RBF, linear, and polynomial kernel functions were developed. In addition, in order to explore the use of ordinal information contained in the performance level labels, both the CLR and ordinal SVM algorithms were applied. The rates of error for these non-linear models are summarized in Table 5.2, along with the selected number of features for each model. By comparing prediction rates of error generated by the various models, it was found that ordinal SVM with the linear kernel function produced the lowest error rate, i.e., the highest prediction accuracy. However, it can also be noted that this model included the second highest number of features among all models (with only the SVM model using a linear kernel function including more). Related to this, the ordinal SVM model using the RBF kernel yielded a higher rate of error reduction per model feature (on average, 0.59 error points) than the ordinal SVM model using the linear kernel (on average, 0.49 error points per feature). However, the inclusion of additional features in the model using the RBF kernel (i.e., more than nine) led to a reduction in the total error rate for the model. Therefore, the ordinal SVM model using the linear kernel function was selected for application in the Phase II experiment for classification of motor skill level of new study participants based on their performance in dice manipulation task.

Table 5.2. Summary of classification accuracy for various algorithms

Algorithm	Kernel function	Rate of error (%)	Number of features	Error rate reduction (%) (compared to linear model)
SVM	RBF	4.56	10	4.32
	Linear	2.44	15	6.44
	Polynomial	26.24	4	-17.36
Ordinal SVM	RBF	3.61	9	5.27
	Linear	1.96	14	6.92
	Polynomial	23.44	5	-14.56
CLR	N/A	7.85	8	1.03

Table 5.3 summarized the 14 input variables used in constructing the ordinal SVM model with linear kernel function.

Table 5.3. Summary of input variables in final ordinal SVM model

No.	Parameter	Feature
1	Approach acceleration along z-axis	SD
2	Deviation angle	k-value
3	Approach speed along y-axis	Mean
4	Approach acceleration along z-axis	Minimum
5	Reaction time	Minimum
6	Manipulation acceleration along z-axis	Mean
7	Manipulation rotation speed in pitch	k-value
8	Approach speed along x-axis	SD
9	Manipulation speed along x-axis	Minimum
10	Approach rotation acceleration in roll	Maximum
11	Manipulation rotating speed in yaw	Mean
12	Deviation distance	SD
13	Approach time	Minimum
14	Manipulation acceleration along z-axis	k-value

CHAPTER 6: PHASE II EXPERIMENT METHODOLOGY

6.1. Participants

A group of 36 participants (25 males, 11 females, average age = 24.4, SD = 5.8) was recruited from both university and off-campus population for the second phase of the study. None of the persons who participated in the first phase of experiment were included in the Phase II study sample. The age range was a bit narrower for Phase II participants, as compared to Phase I. As in the first phase of the research, all participants were right hand dominant, verified through a questionnaire. Similar to the other requirements for participants in the first phase, Phase II participants were required to have 20/20 or corrected vision (if vision was corrected, they were required to wear the corresponding prosthetic to reach 20/20 acuity during the experiments). All participants were required to have no prior experience in similar experiments or training and no background information on the present research goals or experimental hypotheses. In addition, participants were required to be in good health and to be able to follow instructions during performance of the experiment tasks. The tasks used in this phase of the experiment involved repetitive elbow-wrist movements in a seated posture. Any participant with current or chronic wrist disorders (e.g., carpal tunnel syndrome) was excluded from the experiment. All relevant participant inclusion information was obtained by the same questionnaire, as used in the Phase I experiment (see Appendix A).

6.2. Equipment and apparatus

The equipment used for the second phase of experiment was identical to that applied in the first phase.

6.3. Experimental tasks

6.3.1. Dice manipulation task

In this phase of the research, participants performed the same version of the dice manipulation task as described in the first phase. Results of the task performance were used as inputs to the ordinal SVM model (with linear kernel function) in order to assign participants to one of three motor skill classification groups (i.e., high, medium, or low).

6.3.2. Basic Block Design (BD) task

A VR version of the BD task, based on the Wechsler Adult Intelligence Scale (WAIS; an adult IQ test), was presented to participants. The physical form of this task involves using small wooden blocks with white and red colored sides to assemble patterns presented in a design book. Participants are asked to complete the pattern assembly task as quickly and accurately as possible in each trial. This task was simulated with high fidelity in a VR application (see Figure 6.1, from Ma et al. (2012)). The features of the VR-BD task included a virtual tabletop divided into two parts, including a display area and a work area. The display area presented the stimulus design pattern to be replicated by a participant. The work area was used to arrange the blocks. The work area and blocks were presented at approximately 70% of actual size to allow the design pattern and workspace to be viewed on a 21-inch stereo monitor. All BDs were constructed with the aid of a target grid, which appeared as a 2×2 or 3×3

collection of squares in the work area (see Figure 6.1), depending on the design stimulus. In this experiment, in order to eliminate the potential influence of cognitive workload on participant performance, a grid was presented over the stimulus design in order to segment the design pattern into separate squares. In this way, participants were provided with the test solution and performance was expected to be dictated by motor skill.

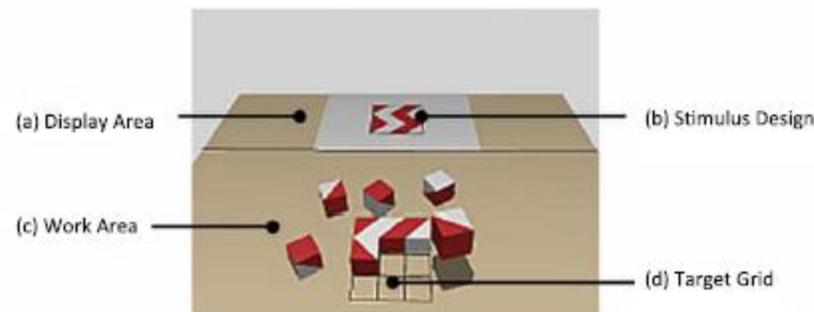


Figure 6.1. Virtual block design task

6.3.3. BD task with haptic features

On the basis of the general form of the BD task, different haptic features were presented under each of three training conditions (Section 6.5 provides detailed descriptions of the conditions) offering different levels of assistance to participants according to their skill classification results. In addition, the stimulus designs used in this part of training were based on the Wechsler Abbreviated Scale for Intelligence (WASI; The Psychological Corporation, 1999). This version of the BD task is identical to the WAIS version except for the design

patterns. As in basic BD task, a grid was superimposed on the stimulus design patterns during each trial in order to reveal the solution to construction for participants.

6.4. Procedures

This phase of experiment was conducted in the same lab room used for the Phase I experiment. No disturbances were allowed during the experiment. During the introduction, participants were told that they would complete a set of computer-based motor tests, followed by three training sessions and a final test. All test and training tasks were required to be performed with the right hand only. The screening criteria for recruiting participants were confirmed before the start of experiment and each participant was asked to sign an informed consent form (see Appendix C).

After the initial briefing, participants were directed to the VR workstation and they began the experiment with 40 trials of the dice manipulation task. The administration of the task was the same as in the first phase of experiment. Participants were allowed to take a short break after completing this part of the study.

The second baseline test involved performance of the basic VR-BD task with no haptic features applied. Design 8 from WAIS version of the task (Wechsler, 1997) was used to demonstrate the requirements of BD and use of the haptic control interface. Participants were told to complete the task as quickly and accurately as possible with their right hand only. If necessary, they were allowed an additional trial involving construction of Design 8 for practice. Designs 9-14 from the WAIS were used for formal testing. These designs are the most complex

designs as part of the set of 14 designs. A mandatory 5-minute break was provided after completion of the baseline BD task.

Based on their motor performance in the dice manipulation task, participants were classified and assigned to one of the three training conditions. All participants were required to complete three training sessions of VR-BD task with additional haptic features. Each session included 6 trials with different stimulus design patterns. Design 7 from the WASI was used for practice to help participants become familiar with the haptic feature presented as part of their training condition. A second practice trial with Design 7 was permitted upon participant request.

An earlier experiment by (Kaber et al., 2014) involving training in the VR-BD task revealed participants to exhibit signs of fatigue after 5-6 consecutive trials when using their non-dominant hand. It was suggested that better training could be achieved by providing several sessions with short breaks in between. Moreover, Ma et al. (2012) showed that participants were able to reach asymptotic performance in the VR-BD task after 3-5 training sessions using the non-dominant hand. Since participants were required to use the dominant hand in the present research, it was assumed that three sessions would be sufficient to promote participant asymptotic performance without signs of fatigue. The 6 trials in each training session involved participant construction of Designs 8-13 from the WASI, posed in random order. These designs were selected as they are the most complex among the set of 13 WASI designs. Between training sessions, participants were asked to take a break for at least 5 minutes in order to further prevent the potential for wrist and forearm fatigue.

Following completion of the training, participants were required to complete an additional basic VR-BD test, identical to that as part of baseline testing. Designs 9-14 from the WAIS were also used for this post-test. The entire experiment lasted for about 100 minutes. Throughout the entire experiment, participants were not provided with information on their design completion time in order to avoid inducing temporal stress/workload. The detailed experimenter instructions are presented in Appendix E.

6.5. Independent variables

The independent variable for this study was the training condition assigned to participants. There were three types of training based on the haptic feature presented during the VR-BD task, including consistent haptic guidance (CG), resistive haptic force feedback (RF), and haptic disturbances with random force feedback (HD).

6.5.1. CG condition

In the CG condition, the haptic device presented an assisting force to facilitate accurate participant hand movement during block manipulation. A snap force with a consistent magnitude of 0.5 N was also presented with a direction from the current cursor position to the center of the targeted square location in a design pattern.

6.5.2. RF condition

In the RF condition, resistive forces were presented to inhibit participant hand movements to correctly orient and locate blocks for design pattern reconstruction. The resistive force was also presented with a consistent magnitude of 0.5 N. Different from the CG

condition, this force was directed from the center of a targeted location in a design pattern to the current cursor position.

6.5.3. HD condition

In the HD condition, noise-like force was presented to disturb participant performance. The disturbing noise was randomly generated between 0 and 0.6 N and presented at a random frequency between 0 and 0.5 Hz. The magnitude of the noise-like force and the frequency were both generated by the RAND() function in a C++ program. Applying such computation, the direction of the noise-like force was also randomized. The force would last for 1 second each time it was presented.

Lee and Choi's (2010) study suggested that, in order to promote motor task training effectiveness, the frequency of disturbing force feedback should depend on the type and difficulty of the task. They found that low frequency disturbances did not yield a training effect; however, high frequency disturbances led to participant fatigue in task performance. However, when exposed to the noise-like force with a regular frequency, participants were likely to only perform between disturbances and pause when the noise force was present. In this way, the expectation that participants could develop strong motor skills through fighting with the disturbing forces was violated. To avoid such issues in the present study, the noise-like force was designed to be presented with a random frequency, such that participants would not be able to predict the time when a disturbance would occur.

Under the RF and HD conditions, in order to achieve accurate block orientations and placements, more effort was required by participants to maintain the steadiness of the haptic

stylus. This requirement increased the potential for fatigue under these conditions and further motivated the breaks between training sessions.

6.6. Dependent variables

The time needed to complete the baseline and final BD tests were the primary response measures for this experiment. The completion time of all BD task trials was automatically recorded by the VR-BD program. It should be noted that the original scoring method, developed for the WAIS version of the BD task, was not applied in this study. This method was developed for the native (physical) version of the BD test, which is easier to perform than the VR version due to the use of the haptic stylus. Consequently, the task time limit, and scores typically assigned for various completion times, were not applicable to evaluate the VR-BD task performance. In addition to this, the original BD scoring method is a step-wise index dependent upon both task speed and accuracy; therefore, the BD is a confounded measurement. In some cases when completion time improves (i.e., decreases), the overall task score may not change as a result of the time still falling within a defined scoring range. Moreover, with the implementation of the visual grid for segmentation of each stimulus design patterns, participants in this study were able to achieve 100% accuracy in completing the task, which further limited the applicability of the original BD score.

Beyond the test response time measures, participant learning percentage (k-value) for each training condition was calculated based on the trend of BD task completion time across the three training sessions. The set of k-values were used as a secondary measurement to explore potential learning and rate of improvement in task performance.

6.7. Model application

6.7.1. Comparison of dice manipulation task, BD task, and physical assembly task

The BD training condition assignment was based on the motor performance assessment results obtained in the dice manipulation task. Here it was assumed that the motor skill level evaluated based on the dice manipulation task performance could serve as a basis for predicting motor performance in the BD task. In order to theoretically verify this assumption, fundamental motion unit analysis was conducted on each task (see Figure 6.2 and 6.3). Since both tasks were presented as part of VEs and were controlled with a haptic stylus using only the right hand, the two-hand process chart for general motion unit analysis (Niebel & Freivalds, 2008) was not appropriate. Instead, the process of each task was mapped using a flow diagram and by identifying fundamental motion units. By comparing the flow diagrams for the two tasks, any differences among the tasks were identified.

The BD task involved motion planning and block selection for manipulation. (Task trials involved either four or nine blocks.) However, the dice task involved manipulation of only one die in each trial. In manipulating a single object (block or die), the general movements are the same as in manipulating many blocks. In addition, the functions of “grasp” and “release” are achieved in the same manner across both tasks by using the haptic stylus. Therefore, the two tasks were comparable in workstation design, required movements and motor workload. Since the planning and selection steps in the BD task involved greater cognitive work than motor control, it was reasonable to expect that motor performance in the BD task could be predicted based on motor skill as evaluated using the dice manipulation task.

Moreover, for some real assembly tasks such as iPhone assembly, which mainly deal with relatively small parts and seldom require special tools, the main work steps are similar to those observed in the dice manipulation and BD tasks. The general work steps for the iPhone assembly were extracted from an online video (“iPhone 4: The entire assembly process,” 2010) and are presented in Figure 6.4, including fundamental motion units. From this additional analysis, it is clear that the general procedures involved in the phone assembly are quite similar with those in completion of the BD task. Although there are some differences in the way the “Grasp” and “Release” motions are completed between the VR-based BD task and real physical assembly task, the critical components of moving and rotating objects are the same. In addition, in the iPhone assembly task, a frame is provided on to which workers fasten semi-finished assemblies. The frame also identifies specific positions in which parts should be placed as part of the complete assembly. The iPhone frame was simulated in the VR-based BD task by presenting a target grid to indicate desired block positions for assembly of design patterns. Therefore, it was expected that the evaluation results based on the BD task could be potentially applied to similar physical production tasks.

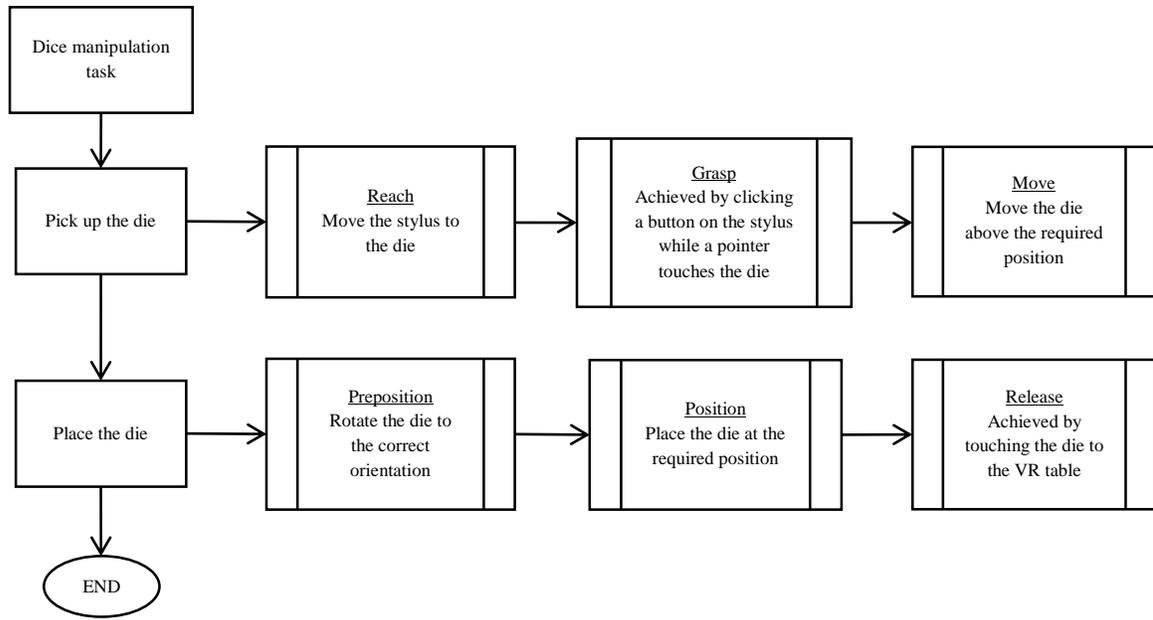


Figure 6.2. Motion flow chart for dice manipulation task

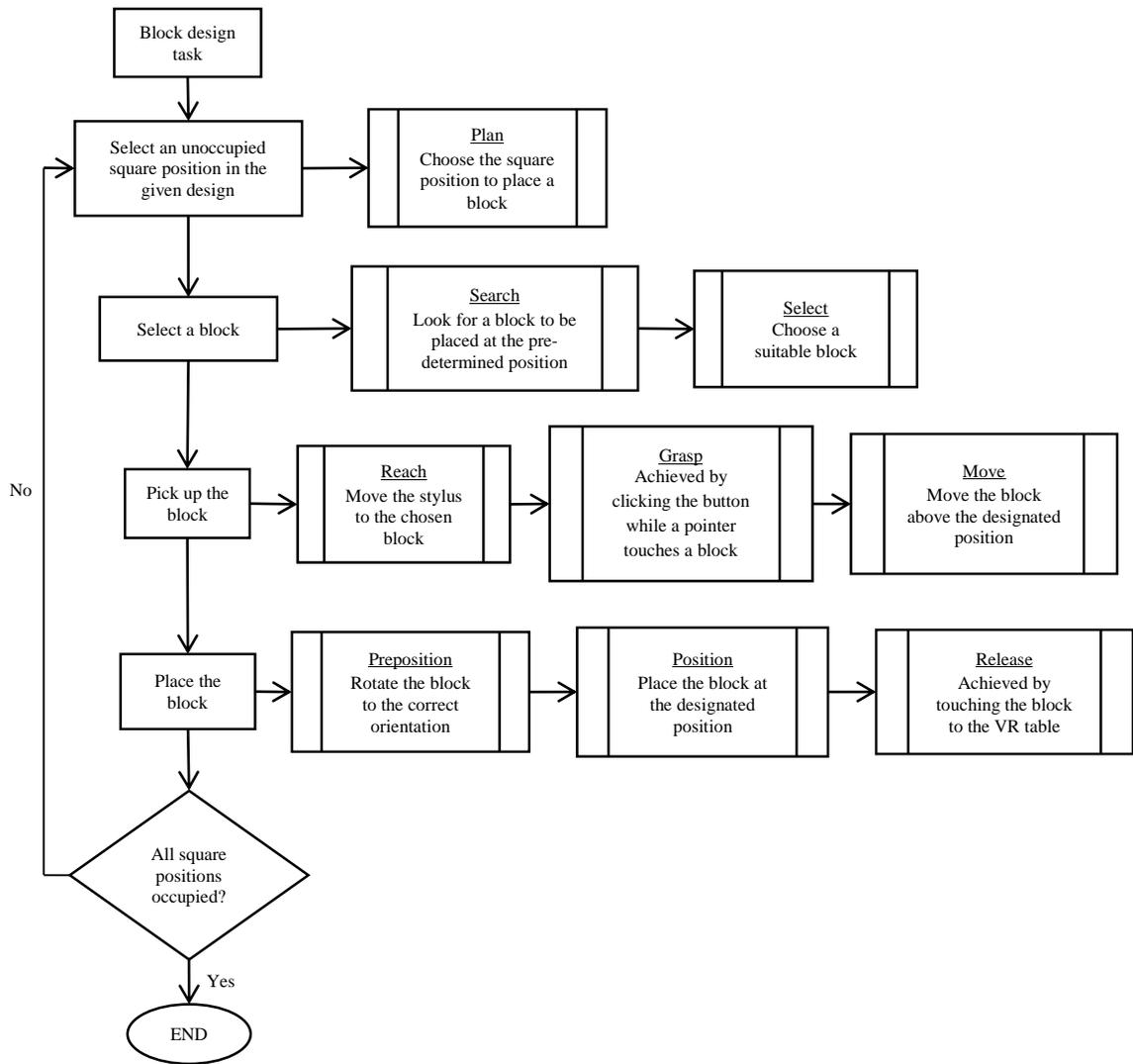


Figure 6.3. Motion flow chart for block design task

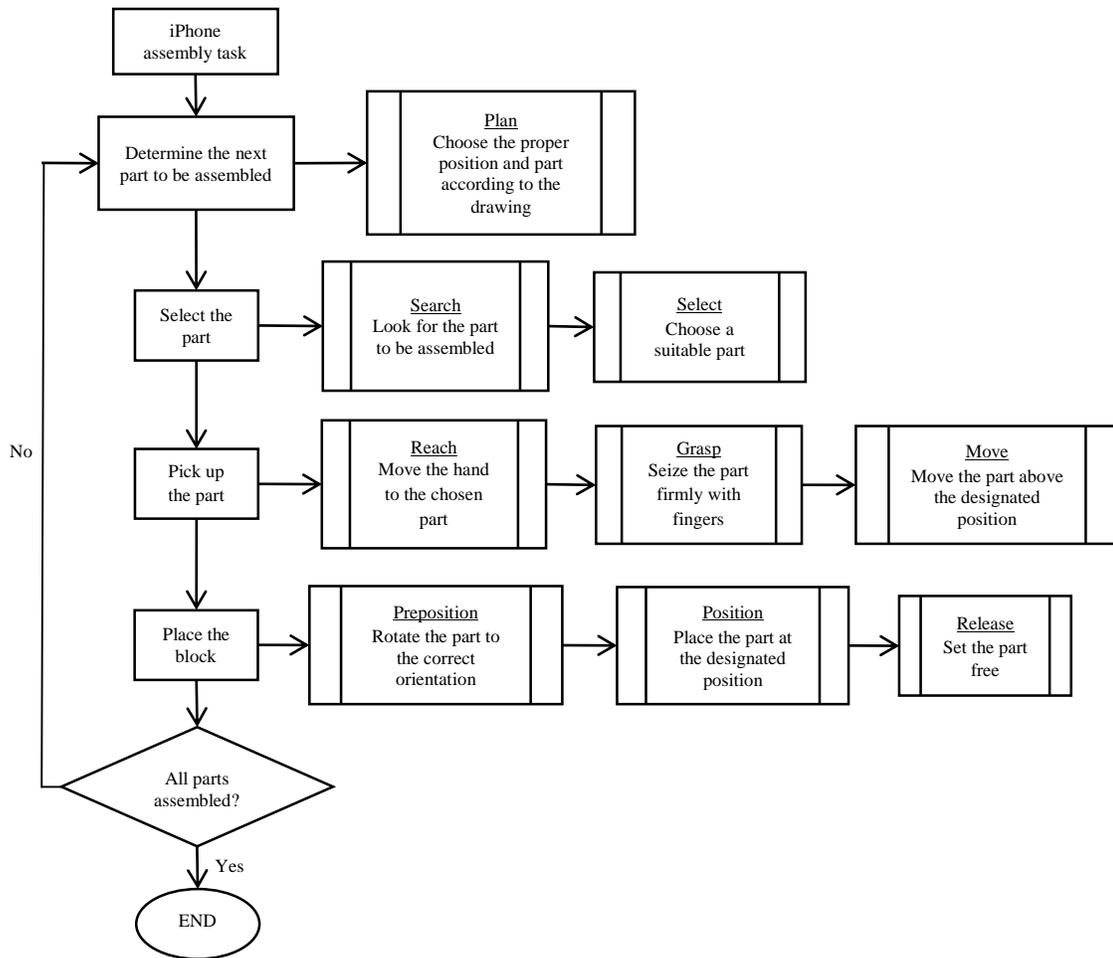


Figure 6.4. Motion flow chart for iPhone assembly task (based on online video, “iPhone 4: The entire assembly process,” 2010)

6.7.2. Training conditions assignment

During the 40 trials of the dice manipulation task, participant performance was automatically recorded by the haptic device and software in the form of time series with a frequency of 10 Hz. Observations were made on the set of features selected in the Phase I experiment as part of the ordinal SVM model with linear kernel function. Using these features

as inputs to the model, participants in the Phase II study were classified into one of the three groups, i.e., high performers, medium performers, or low performers.

In previous research, it was demonstrated that consistent haptic guidance resulted in participant over-reliance on given aids, degrading training effects due to participants losing motivation to develop their own skills to achieve the task. In contrast, VR training with haptic disturbances produced significantly greater effects in terms of promoting motor skill learning. Experimental results of Lee & Choi (2010) showed that participants receiving training with resistive haptic disturbance and noise-like haptic disturbance performed a skill retention test with significantly better results, as compared to those completing training under haptic guidance (with test scores ~19% and ~31% higher, respectively). Comparing the two types of haptic disturbances (HD and RF), noise-like force feedback was shown to outperform resistive force feedback (with retention test scores ~13% higher, yet not statistically significant).

Based on these results, participants demonstrating an initial high performance in the dice testing application were assigned to the CG training condition. For participants demonstrating an initial low performance, three training sessions were conducted with the haptic device presenting random force-feedback, i.e., the HD condition. For participants demonstrating a transitional medium ability in the dice test, the RF training condition was provided. The presentation of haptic forces was designed, as described in Section 6.5.

6.7.3. Statistical methods of analysis and expected outcomes

In Section 6.7.1, the assumption that motor performance assessments, obtained based on dice manipulation task performance, could be used for predicting motor skill level in the BD task was theoretically verified through fundamental motion unit analysis. To further

validate this assumption, data collected on baseline VR-BD test task performance were analyzed. A one-way unbalanced ANOVA test was applied to compare task completion time between the three skill groups (low, medium and high performers). The test was unbalanced as the number of participants assigned to the low, medium and high performing groups, based on the dice manipulation trials, was not uniformly distributed (6 high performers, 17 medium performers, and 13 low performers). It was expected that with accurate classification, the three groups of participants would show significantly different completion times on the baseline test. Tukey's Honestly Significant Difference (HSD) method was applied to further detect differences between any pair of skill groups.

Paired t-tests were applied to compare participant performance between the baseline and final VR-BD tests within each training group (i.e., CG, RF, HD). It was assumed that all three training conditions would show an effect of reduced task completion time.

To validate the utility of the designated haptic features for facilitating expected levels of training proficiency, based on the classification developed in the first phase of the study, results of the final VR-BD test were analyzed in order to determine the motor skill level of the three groups of participants after training. A one-way unbalanced ANOVA test was used to identify any group differences in terms of final BD task completion times. Again, Tukey's HSD method was applied for multiple comparisons between pairs of skill/training groups. The expectation for this process was that matching of the haptic feature conditions in VR task training to participant motor skill classification, based on the simple dice manipulation performance, would lead to comparable levels BD/assembly task performance among skill groups. Demonstration of this expectation represented the major contribution of the present

research. Figure 6.5 presents a diagram summarizing the procedures and data flow for the Phase II experiment.

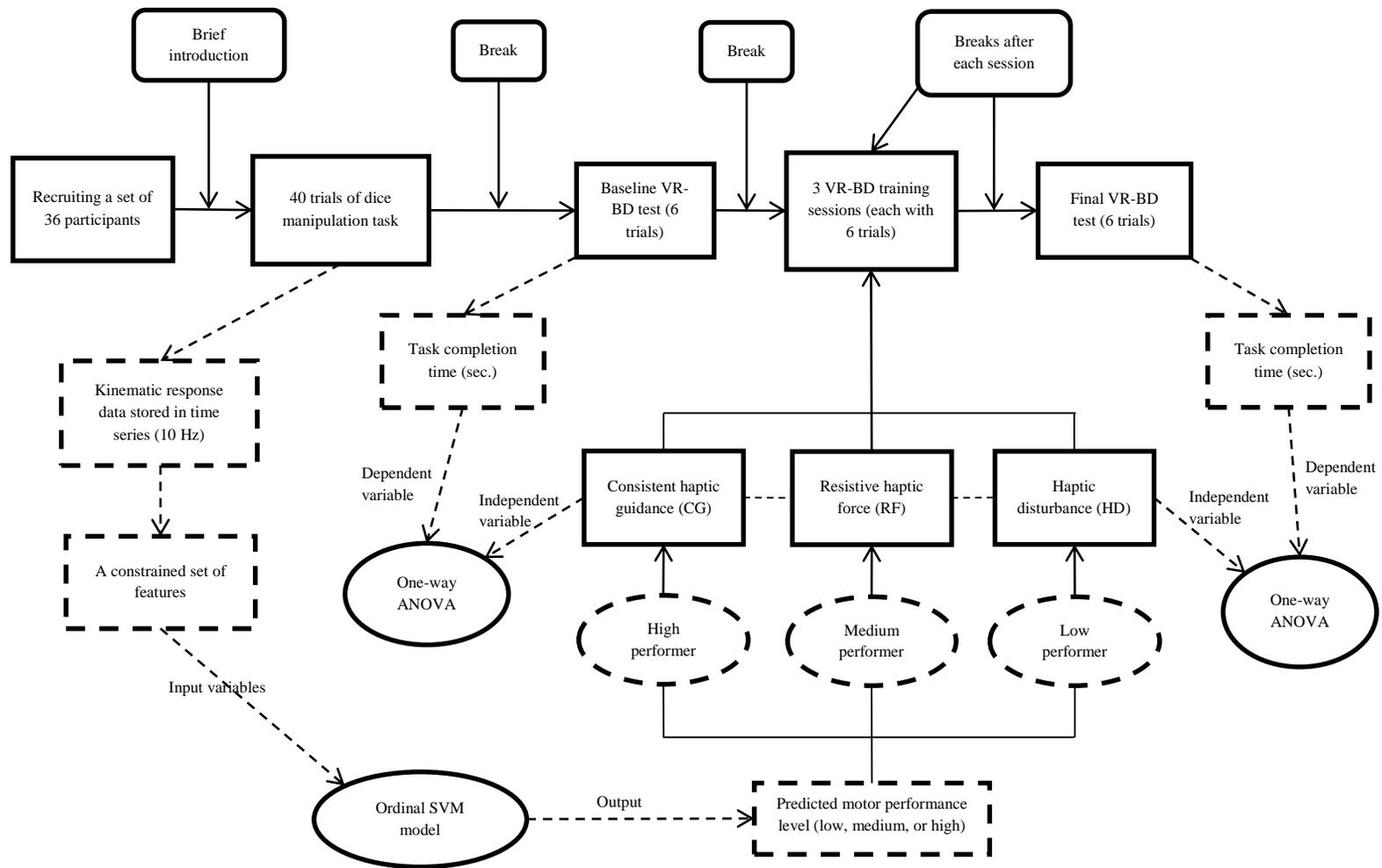


Figure 6.5. Phase II experiment procedures and data flow

CHAPTER 7: PHASE II DATA ANALYSIS AND RESULTS

7.1. Descriptive statistics

Applying the ordinal SVM model developed in Phase I, 36 participants were classified as 6 high performers, 17 medium performers, and 13 low performers. For each participant, the completion time in the six trials of the same BD test or training session was added up for further analysis. Table 7.1 shows a summary of group-based general statistics (group mean and SD) in BD session completion time.

Table 7.1. Summary of group mean (SD) statistics in BD completion time (sec.)

	N	Baseline	Training 1	Training 2	Training 3	Final
High	6	511.3 (45.0)	371.8 (78.6)	322.3 (62.7)	304.7 (47.6)	312.3 (36.7)
Medium	17	657.1 (163.6)	633.1 (146.3)	559.1 (139.1)	502.2 (140.9)	435.5 (102.7)
Low	13	830.7 (231.1)	871.9 (289.9)	678.1 (204.4)	650.2 (192.4)	479.2 (141.2)

7.2. Within-group comparison between baseline and final tests

In Figure 7.1, a bar chart was present showing session completion time of three condition groups in baseline and final VR-BD tests. The error bar indicated the standard error of corresponding sample mean. From the bar chart, it was clear that all three condition groups achieved improvement in their final test performance as compared to in baseline test.

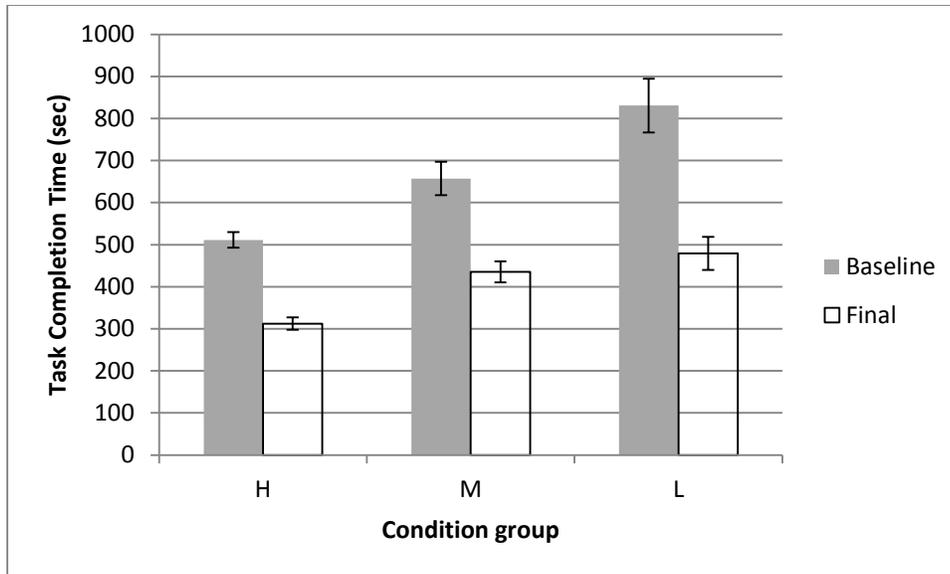


Figure 7.1. Comparison of performance in baseline and final BD tests

To further analyze the within-group differences, a series of paired t-test were conducted to compare within-group differences between baseline and final test results. Table 7.2 presents a summary of the t-test results, including t-statistics, p-values and statistical powers. All t-tests conducted here were one-tailed which assumed that participants would show improvement (i.e., reduction in completion time) in final test. All three t-tests gave extremely small p-values, suggesting a significant decrease in completion time from baseline to final test for all three condition groups. Also, the high statistical power supported the confidence of such results.

Table 7.2. Summary of paired t-test results

	Paired t-test	Statistical power (1 – β)
High	t(5) = 14.28, p < 0.001*	1
Medium	t(16) = 7.75, p < 0.001*	1
Low	t(12) = 11.14, p < 0.001*	1

7.3. Between-group comparison of baseline and final test performance

One-way ANOVA unbalanced tests were conducted to check between-group differences in baseline and final VR-BD test completion time. Tukey's HSD method was applied to generate multiple pairwise group comparisons. Statistical results were presented in Table 7.3, including F-values, overall p-values, statistical powers, and p-values of pairwise group comparisons. In baseline test, significant difference was verified between three groups and low performers showed to spend significantly longer time than high and medium performers in completing the test. On average, high performers spent less time than medium performers. However, such difference was not proved significant.

Participants also showed significant between-group differences in performing final test. However, the difference was only significant between low performers and high performers while the other two pairs of groups were comparable.

Table 7.3. Summary of one-way ANOVA and Tukey's HSD results

	One-way ANOVA	Tukey's HSD	Statistical power (1 - β)
Baseline	F(2, 33) = 7.13, p = 0.002*	Low vs. High: p = 0.003* Low vs. Medium: p = 0.036* Medium vs. High: p = 0.22	0.84
Final	F(2, 33) = 4.58, p = 0.018*	Low vs. High: p = 0.013* Low vs. Medium: p = 0.546 Medium vs. High: p = 0.068	0.68

7.4. Analysis of learning percentages in BD training

Performance of three condition groups in completing BD training was described in Figure 7.2. Again, error bar was included to show the standard error of each corresponding sample mean. From the line chart, all three groups showed a trend of decrease in completion time, indicating improvement throughout training sessions. By fitting the training session completion time (i.e., sum of six trials in the same session) in the learning model, learning percentage was calculated for each participant (computing method described in section 4.5.2). Table 7.4 provides a summary of group-based mean and SD statistics. One-way ANOVA test was conducted but no significant difference was detected between groups ($F(2, 33) = 0.59$, p -value = 0.56, $1 - \beta = 0.15$). However, on average, low performers showed the lowest learning percentage, indicating the fastest learning potential.

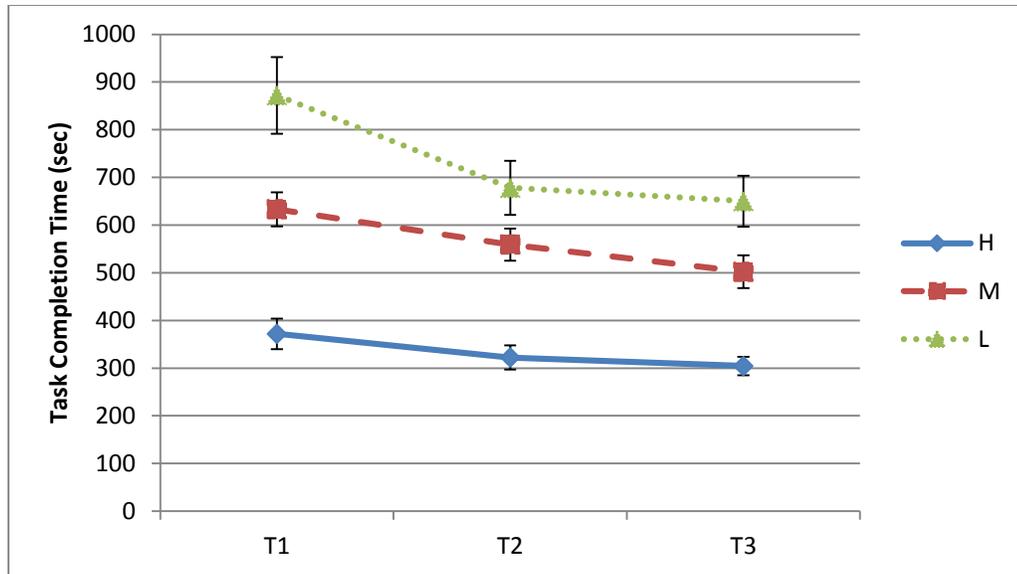


Figure 7.2. BD training performance

Table 7.4. Summary of group mean (SD) in learning percentage (%)

High	88.6 (6.5)
Medium	86.6 (7.5)
Low	84.0 (12.1)

CHAPTER 8: DISCUSSION

8.1. Validation of dice manipulation task for assessing motor ability

The first hypothesis (H1) as part of this research was that human performance in the haptic-VR dice manipulation task would prove to be a reliable and efficient indicator of motor skill, as compared with standardized psychomotor tasks. The results of the Phase I experiment supported this hypothesis. Dice manipulation task performance records were used to generate a set of statistical features. Linear regression analysis revealed subsets of these features to have utility for predicting average PPT scores. The PPT scores were considered to be a “gold standard” for indicating motor performance capability. The resulting linear regression model produced an adjusted R-squared value of 0.9966, indicating a good fit of performance in the dice manipulation task to the standardized psychomotor test (PPT) results. Therefore, the computer-based dice manipulation task was proved to be a valid and reliable method for assessing motor skill. Most importantly, the dice manipulation task can be administered with efficiency as compared to other physical psychomotor test (PPT, BD task) and use of a haptic control interface with the VR task simulation allows for spatial and temporal data on participant motions to be captured automatically, which supports rich kinematic parameter calculations.

However, despite a significant model fitness level ($p < 0.001$), the linear regression approach only achieved 91.12% accuracy in prediction (i.e., an error rate of 8.88%) of PPT performance. Among the eleven dice manipulation performance features selected as predictors in the final linear model of PPT scores, there was no feature related with performance accuracy

(i.e., deviation distance and deviation angle). Considering the fact that performance of PPT required both speed (in order to insert as many pins as possible within limited time) and accuracy (to insert each pin properly into the little hole) in assessing motor skill level, lack of accuracy related predictors was a potential reason for the reduced prediction power of the linear model. It is possible that knowledge-based selection of predictors vs. statistical or data-driven selection might serve to further increase the predictive utility of the linear regression model for describing motor skill.

8.2. Non-linear classification models

The classification accuracy achieved with the linear model (91.12%) might be considered high accuracy when dealing with small sample-size problems, such as classifying the motor skill levels of members of a hockey playoff team (12 persons). However, large sample-size problems, such as classification of iPhone assembly worker skill, exist and the present research focuses on this application. The largest iPhone assembly factory, Foxconn's Longhua facility, employs approximately 300,000 workers in assembly activities (Perlin, 2013). An average shift involves 7000 to 8000 workers. For such large-scale, real-world applications, even a 1% improvement in classification accuracy would likely represent practically significant differences in terms of training costs and worker skill development. The second hypothesis (H2) as part of this research stated that a non-linear classification model would achieve higher classification accuracy than a linear model. The Phase I results also supported this hypothesis. The original SVM algorithms applying the RBF and linear kernel functions both achieved reductions in the prediction rate of error that was obtained with the

linear regression modeling approach. However, the SVM model using a polynomial kernel function degraded the classification accuracy with a significantly high error rate of 26.24%. This finding indicated that for the specific data collected in this study, a polynomial kernel function would not perform as good as it might with other data sets. Previous research demonstrated the capability of polynomial kernel functions in studies concerning natural language processing (e.g., Goldberg & Elhadad, 2008), facial expression pattern recognition (e.g., Wang et al., 2013), and disease diagnosis (e.g., Mo & Xu, 2010).

To further improve the predictive model accuracy, CLR and ordinal SVM models were developed to exploit ordinal information contained in the participant performance level labels. Based on performance in the PPT task, each participant was assigned an initial motor skill classification (low, medium or high). The CLR algorithm made use of this information and produced superior classification results as compared to linear regression model; however, it did not perform better than the original SVM model using RBF and linear kernel functions. Although the CLR method has proved to be a popular and useful approach for classification of ordinal data in prior studies, it was not a superior fit for the specific dataset analyzed in the present study.

On the contrary, the ordinal SVM model produced a decreased error rate as compared to the original SVM model with the same kernel function. However, this improvement in accuracy was not significantly large (as summarized in Table 8.1), which was likely influenced by the relatively small number of classification groups. These results were consistent with the conclusion of the study by Frank and Hall (2001), which indicated that the degree of improvement in ordinal model performance over the original SVM model would increase with

the number of classes. Therefore, while model prediction accuracy was not substantially improved in this study, as a result of applying ordinal SVM, it is possible that an ordinal model would show superior performance to non-ordinal SVM in addressing similar motor skill classification problems with more levels of classification.

When comparing the variables used in the ordinal SVM algorithm with the predictors included in the linear regression model (see Table 5.1 and 5.3), it was found that the two sets had some overlap but not a perfect matching of variables. This finding suggested that greater model prediction power might also be attributable to some specific features generated based on participant motor performance data. Related to this, it is important to note that two parameters included in the ordinal SVM model were indicators of participant accuracy in the dice manipulation task, specifically the deviation distance and deviation angle of the die from the target. This information was utilized by the ordinal SVM model, which showed significant superiority to linear regression model in motor skill prediction accuracy. This observation further supports the suggestion of incorporating both speed and accuracy related parameters in constructing such classification models, based on expert knowledge of the nature of the tasks being performed for skill classification.

Table 8.1. Comparison of classification model accuracy.

	Two-sample t-test	Statistical power (1 – β)
RBF: Original vs. Ordinal SVM	t = 1.13, p = 0.129	0.1930
Linear: Original vs. Ordinal SVM	t = 1.30, p = 0.096	0.2321
Polynomial: Original vs. Ordinal SVM	t = 1.19, p = 0.117	0.2074
Ordinal SVM with linear kernel vs. Linear model	t = 11.24, p < 0.001*	1.0

* - Significant at p < 0.01 level.

8.3. Selection of haptic-VR training protocols

The ordinal SVM algorithm with linear kernel function, developed through the Phase I research, was used to classify participants in the Phase II experiment, based on their performance in dice manipulation task. Statistical analysis on performance in a VR version of the WAIS BD task revealed significant differences ($p = 0.002$) among the three groups of participants (low, medium and high performers). This finding suggested that performance of the dice manipulation task was able to serve as a reference for predicting motor ability level in the BD task. In addition, these results further verified the capability of the non-linear classification model for identifying individual participant motor skill level.

The third and final hypothesis (H3) of this research stated that training protocols, presenting specific haptic features, could be designed and selected to bring participants to desired, comparable motor skill levels. Results of the Phase II experiment were in partial support of this hypothesis. Within-group comparisons of baseline vs. post-training (VR-BD) test performance revealed highly significant differences for all three groups ($p < 0.001$). This

finding indicated that there were significant training effects of all three haptic feature conditions for improving performance of the VR-BD task.

During training as part of the Phase II experiment, those participants classified as low performers were assigned to the most challenging VR haptic feature condition with the expectation of producing the greatest improvements in motor skill test performance. Similarly, those participants classified as high performers were assigned to the least challenging VR haptic feature condition with the expectation of producing the lowest degree of improvement in motor skill test performance. Medium performers were assigned to what was considered as the moderately challenging haptic feature condition. Although the participants initially classified as low and medium performers revealed significant differences in baseline motor test performance ($p = 0.036$), after exposure to equivalent amounts of training, they demonstrated comparable levels of motor skill based on VR-BD test performance ($p = 0.546$). Furthermore, there was no significant difference among medium and high performance post-training test performance ($p = 0.068$), as a result of prescription of the specific haptic training conditions based on initial skill classification. However, there was a significant gap between low performer post-training test performance and high performer outcomes ($p = 0.013$); thus, the inference of partial support of H3.

Although prior studies have shown VR training with haptic guidance to potentially degrade motor task learning, as compared to the use of haptic disturbances, while holding fixed the amount of training, high performers in this study still achieved some improvement in motor skill with guidance. This situation likely led to the gap in final test performance between the low and high performers. It is possible that specific designs of haptic guidance conditions (e.g.,

trajectory tracking, force boundaries, target attraction forces, etc.) may vary in terms of motor control learning effects. This observation is also reflected in contradictory results generated by prior studies. For example, Feygin et al. (2002) and O'Malley et al. (2006) showed positive training results while Edwards et al. (2004) and Li et al. (2009) demonstrated negative results using different haptic conditions.

Related to this explanation of the results, an additional one-way unbalanced ANOVA test was conducted to compare the final (post-training test) performance of medium and low performers with the initial performance of high performers (as shown in Figure 8.1). The test result ($F(2, 33) = 1.19, p = 0.316, 1 - \beta = 0.219$) showed no significant difference among the three groups. Therefore, through the three training sessions as part of the Phase II experiment, medium and low performers were successfully brought to a similar level of motor skill, as compared to high performers, before exposure to any training. This result further supports the contention that the significant difference detected in final test performance among the low and high performers was mainly caused by “overtraining” of the high performers. Had the high performers not been exposed to any haptic VR training whatsoever, it is likely that the post-test results would have been comparable for all groups. It is therefore reasonable to assume that by reducing the amount of training for high performers, equivalent levels of BD task performance among different skill groups could be achieved. In general, the problem of specific training condition prescription for operators, based on baseline motor skill classification is complex, and specific haptic VR features should be associated with values of specific motor performance features extracted from initial classification task performance.

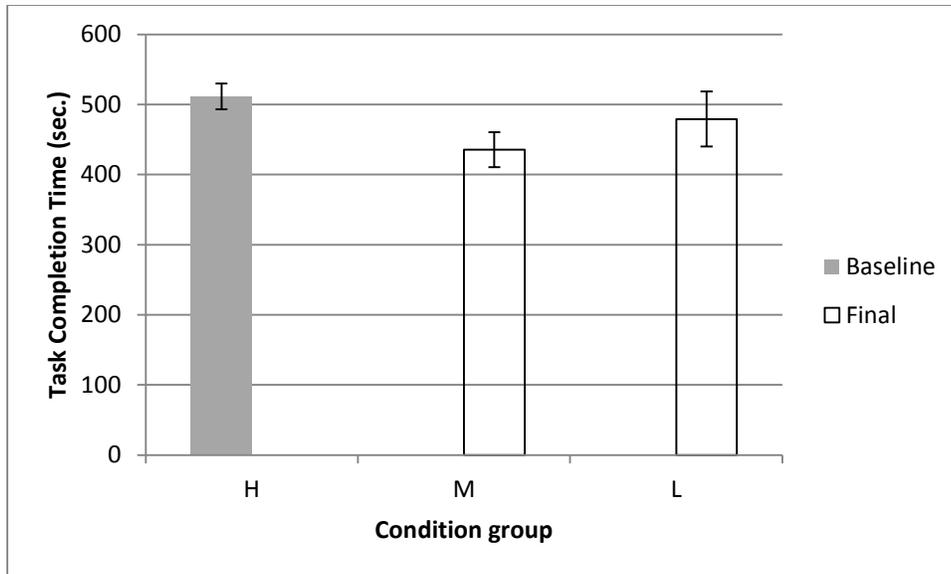


Figure 8.1. Final performance of low and medium groups compared with baseline performance of high group

CHAPTER 9: CONCLUSIONS

In conclusion, the objectives of the present research were to develop a quantitative algorithm for classifying operator motor skill levels, based on baseline motor test performance, and to identify appropriate protocols for effective motor skill training to desired levels. To achieve these objectives, a quantitative classification algorithm was developed to predict individual motor performance level based on performance in a simple customized motor task. Three haptic-VR training protocols were designed and assigned to participants based on their identified skill levels. The customized motor task was verified as a reliable and efficient means for assessing motor performance, as compared to standardized psychomotor testing. The developed algorithm incorporated a non-linear classification methodology exploiting ordinal information contained in training and verification data sets. This model was validated and proved accurate for classification purposes. The specific training conditions (VR haptic features) designed for the various levels of baseline motor skill provided effective skill training to desired levels, especially for those individuals identified as moderate and low performers, in terms of motor control.

The following sections identify limitations of the present work, future study directions, and potential industry applications of the technology developed here.

9.1. Limitations of present study

In this research, the only aspect of training that was manipulated through experimentation was the design of haptic guidance or challenge features in a VR environment. All other dimensions of the training regimen were held fixed across participant groups,

including the timing and extent of exposure to training conditions. In order to overcome any motor skill gaps among low and medium, low and high or medium and high performers identified at the outset of a training program, the amount of training can also be manipulated, in addition to the level of difficulty of the training condition. For example, based on the results of this study, low performers could be provided with more training sessions with haptic disturbances while less or no training with haptic guidance could be provided for high performers. Such prescriptions would be aimed at achieving comparable levels of performance among motor skill groups through training.

Another potential limitation of the present study concerned the stylus-shape of the haptic control device, which led to differences in performance with physical blocks in BD task. In using the haptic device, users needed to translate the rotation of the stylus according to the orientation of the virtual blocks. This process likely increased cognitive workload, especially for low performers. Also, differences in hand-hold postures might also influence the effectiveness of using the present VR-BD task simulation for training in physical assembly skills.

In addition, since the PPT was not performed as part of the Phase II experiment, the classification model developed in Phase I was only verified using the baseline BD test results. However, this process did not directly validate the error rate computed in the Phase I data analysis. It is possible that the error rate (1.96%), obtained based on Phase I experiment data, might have been different for the Phase II experiment, given that a completely different group of participants was tested.

A final limitation of this work was the lack of assembly task as an evaluation test. The developed classification algorithm was only validated by performance in VR-BD task. Although the similarity between VR-BD task and physical assembly work was verified using fundamental unit analysis, it was not experimentally confirmed in this study that motor skill classification via VR-based task performance was applicable for identifying individual performance level in real assembly task performance. Therefore, the present research did not demonstrate that motor skills trained through a haptic-VR simulation could also be transferred to improve performance in an assembly task.

9.2. Future work

In the present study, only one variable, concerning the VR training protocol, was manipulated. It would be interesting to expand the current design of experiment, including manipulation of the amount of training, as an additional control parameter. Furthermore, the three groups of participants could be divided and exposed to all training conditions. That is, low, medium and higher motor performers could all be exposed to haptic guidance, resistive force-feedback and haptic noise. It is also possible that the duration of training might interact with the type of training in a manner other than linear increases in skill with greater exposure time and level of training task difficulty. Comparison of performance of participants assigned to different durations of training under the various haptic VR conditions could lead to identification of an optimal number of training sessions as well as specific training conditions that should be combined and provided for low performers, medium performers, etc. to achieve

specific levels of training or psychomotor and assembly task performance. It is likely that some types of haptic feedback would be better for training specific groups of participants than others.

In addition, the main experiment as part of this study only used the WAIS BD test to evaluate improvement in motor performance. With respect to future research, the effectiveness of haptic VR training for motor skill development should be verified through complimentary testing in a physical or simulated version of an assembly task. Moreover, apart from completion time, more measurements should be used to assess test task performance, such as task accuracy measures, in order to obtain more robust results. In addition to the haptic feedback conditions controlled in this study, there are other factors in designing training protocols that may influence participant learning, such as the shape of the haptic control device, the difficulty of training tasks, etc. It would be worthwhile to investigate such potential effects of other factors in future research.

Lastly, the simple VR-based dice manipulation task, which was validated as a predictor of motor performance in psychomotor tasks as part of the present study, should also be evaluated for predicting motor skill in other types of visual-spatial and fine-motor control tasks. It is possible that performance in the fundamental motions as part of dice orientation and discrete movement might also be predictive of skill in graphomotor (drawing) production or tasks involving constructional praxis (3D object manipulation). Such validation could support use of the dice manipulation task for predicting performance in other domains than assembly work.

9.3. Research applications

The primary application of the present research is worker training for manual assembly that requires fine handwork with small objects (e.g., iPhone assembly). A correspondence was drawn between fundamental motion requirements of simplistic psychomotor and discrete movement task performance with the basic requirements of iPhone assembly in order to demonstrate potential utility of motor classification based on simplistic skill classification for predicting assembly performance. The motor skill classification algorithm developed as part of this research makes it possible to specify individual ability level with a simple baseline motor performance test.

Based on the present research results, if an objective of a real-world assembly operation is uniform performance across operators and workstations in order to ensure production line balancing, baseline operator motor skill classification should be related to individual training protocol design or selection. For example, persons with high motor skills may need little, if any, targeted training in order to reach the desired performance level for a production line. In the present study, it was observed that high motor performers needed no training, or a small amount of training, for psychomotor task performance to be comparable to low and medium performers. Medium and low performers may need exposure to moderate to highly challenging training conditions over time in order to achieve comparable levels and, in some cases, performance equal to a high skill group. In the present study, it was observed that fixed duration training with varying difficulty was effective for enhancing low and medium skill performance levels. In general, the simplistic dice manipulation task studied here, and the non-linear model

for motor skill classification, allow for indexing of individual baseline motor skill level to specific psychomotor task training conditions towards performance improvement.

Finally, results of this work supported the use of disturbances (or noise) in trainee motor control for promoting skill learning rather than guidance in training tasks. In general, if the duration of training is held fixed, perturbations in motor performance via a task interface are more effective for improving performance levels than trainee assistance in task completion. The present study suggested that uniform performance among groups of workers could be achieved by applying such a methodology. In this way, individual differences among assembly operators could be largely eliminated, leading to potential increases in assembly line production rates. In addition, the motor performance analysis, classification and training approach developed here could also be applied in training of military tasks or other sophisticated manual control operations requiring uniform performance levels among soldiers or operators.

REFERENCES

- Adams, R., Klowden, D., & Hannaford, B. (2001). Virtual training for a manual assembly task. *Haptics-E*, 2(2), 1–7.
- Aharonson, V., & Krebs, H. I. (2012). Prediction of response to robot-aided motor neuro-rehabilitation of children with cerebral palsy. *Biomedical Signal Processing and Control*, 7(2), 180–184.
- Alamri, A.; Jongeun Cha; Eid, M.; El Saddik, A. (2009). Evaluating the post-stroke patients progress using an Augmented Reality Rehabilitation system. *2009 IEEE International Workshop on Medical Measurements and Applications (MeMeA)*, 89–94.
- Altuger, Gonca; Chassapis, C. (2010). Manual assembly line operator scheduling using hierarchical preference aggregation. *Proceedings - Winter Simulation Conference*, 1613–1623.
- Alzuheri, A.; Lee, Luong; Ke, X. (2010). Ergonomics design measures in manual assembly work. *2010 Second International Conference on Engineering System Management and Applications (ICESMA 2010)*.
- Amirabdollahian, F., Gomes, G. T., & Johnson, G. R. (2005). The Peg-In-Hole: A VR-Based Haptic Assessment for Quantifying Upper Limb Performance and Skills. *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*, 422–425.
- Amirabdollahian, F., Munih, M., Kouris, F., Laudanna, E., Stokes, E., & Johnson, G. (2005). The I-Match Project: A VR Based System to Allow Matching of an Optimum Interface to a User of Assistive Technology. *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*, 526–529.
- August, K.; Bleichenbacher, D.; Adamovich, S. (2005). Virtual reality physical therapy: a telerehabilitation tool for hand and finger movement exercise monitoring and motor skills analysis. *2005 IEEE 31st Annual Northeast Bioengineering Conference (IEEE Cat. No. 05CH37674)*, 73–74.
- Auria, L., & Moro, R. A. (2008). *Support Vector Machines (SVM) as a Technique for Solvency Analysis*.
- Bardorfer, A., Munih, M., Zupan, A., & Primožic, A. (2001). Upper limb motion analysis using haptic interface. *IEEE/ASME Transactions on Mechatronics*, 6(3), 253–260.

- Bell, B., & Kozlowski, S. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology*, 55, 267–306.
- Bettini, A., Lang, S., Okamura, A., & Hager, G. (2001). Vision-assisted control for manipulation using virtual fixtures. *IEEE International Conference on Intelligent Robots and Systems*, 2, 1171–1176.
- Bhatti, A.; Young Bing Khoo; Creighton, D.; Anticev, J.; Nahavandi, S.; Zhou, M. (2008). Haptically enabled interactive virtual reality prototype for general assembly. *2008 World Automation Congress*, (1).
- Bonto-Kane, M. V. (2009). *Statistical Modeling of Human Response Times for Task Modeling in HCI*. North Carolina State University.
- Bound, A.C.; Haniff, D.J.; Baber, C.; Steiner, S. J. (1999). Virtual reality and augmented reality as a training tool for assembly tasks. *1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)*, 32–36.
- Buchtala, O., Klimek, M., & Sick, B. (2005). Evolutionary optimization of radial basis function classifiers for data mining applications. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics*, 35(5), 928–947.
- Casali, S. P. (1991). A physical skills based strategy for choosing an appropriate interface method. *Extra-Ordinary Human-Computer Interaction*, Cambridge University Press.
- Chae, J., Labatia, I., & Yang, G. (2003). Upper limb motor function in hemiparesis: concurrent validity of the Arm Motor Ability test. *American Journal of Physical Medicine and Rehabilitation*, 82(1), 1–8.
- Chase, J.D.; Casali, S. P. (1995). Introducing a methodology for selecting cursor-control devices for computer users with physical disabilities. *Proceedings of the 3rd Annual Mid-Atlantic Human Factors Conference*, 115–120.
- Chase, J.D.; Casali, S.P.; Hartson, H. R. (1992). The predictability of cursor control device performance based on a primitive set of user object-oriented cursor actions. *Proceedings of the Human Factors Society 36th Annual Meeting*, 306–310.
- Chen, C. W., Yang, Hsien Chueh Peter; Chen, C. Y., & Chang, Alex Kung Hsiung; Chen, T. H. (2008). Evaluation of inference adequacy in cumulative logistic regression models: An empirical validation of ISW-ridge relationships. *China Ocean Engineering*, 22(1), 43–56.

- Chuang, T.-Y., Huang, W.-S., Chiang, S.-C., Tsai, Y.-A., Doong, J.-L., & Cheng, H. (2002). A virtual reality-based system for hand function analysis. *Computer Methods and Programs in Biomedicine*, 69(3), 189–196.
- Clamann, M. P. (2014). *Adaptive Haptic Forces in a Virtual Environment Improve Fine Motor Skill Training*. North Carolina State University.
- Demeurisse, G., Demol, O., & Robaye, E. (1980). Motor evaluation in vascular hemiplegia. *Eur Neurol*, 19, 382–389.
- Dose, M., Gruber, C., Grunz, A., & Hook, C.; Kempf, J.; Scharfenberg, G.; Sick, B. (2007). Towards an Automated Analysis of Neuroleptics' Impact on Human Hand Motor Skills. *2007 4th Symposium on Computational Intelligence in Bioinformatics and Computational Biology (IEEE Cat. No.07EX1576)*, 494–501.
- Duan, F., Gao, Q., & Arai, T. (2010). Assembly skill transfer system for cell production. *2010 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 253–258.
- Edwards, G. W., Barfield, W., & Nussbaum, M. A. (2004). The use of force feedback and auditory cues for performance of an assembly task in an immersive virtual environment. *Virtual Reality*, 7(2), 112–119.
- Ehrhardt, Ina; Herper, Henry; Gebhardt, H. (1994). Modeling strain of manual work in manufacturing systems. *Winter Simulation Conference Proceedings*, 1044–1049.
- Emery, C., Samur, E., Lambercy, O., Bleuler, H., & Gassert, R. (2010). Haptic/VR assessment tool for fine motor control. *Haptics: Generating and Perceiving Tangible Sensations*, 186–193.
- Epps, B. W. (1987). A comparison of cursor control devices on a graphics editing task. *Proceedings of the Human Factors Society 31st Annual Meeting: Rising to New Heights with Technology*, 1, 442–446.
- Feiner, S., Macintyre, B., & Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of the ACM*, 36(7), 53–62.
- Feygin, D., Keehner, M., & Tendick, R. (2002). Haptic guidance: experimental evaluation of a haptic training method for a perceptual motor skill. *Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002*, 40–47.
- Feys, P., Alders, G., Gijbels, D., De Boeck, J., De Weyer, T., Coninx, K., ... Eijnde, O. B. (2009). Arm training in Multiple Sclerosis using Phantom: Clinical relevance of robotic

outcome measures. *2009 IEEE International Conference on Rehabilitation Robotics*, 576–581.

Feys, P., Duportail, M., Kos, D., Van Asch, P., & Ketelaer, P. (2002). Validity of the TEMPA for the measurement of upper limb function in multiple sclerosis. *Clinical Rehabilitation*, *16*(2), 166–173.

Fisher, R. A. (1936). THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS Table I. *Annals of Human Genetics*, *7*(2), 179–188.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381–391.

Fitts, P., & Posner, M. (1967). *Human Performance*. Belmont, CA: Brooks/Cole.

Frank, E., & Hall, M. (2001). A simple approach to ordinal classification. In *EMCL '01 Proceedings of the 12th European Conference on Machine Learning* (pp. 145–156).

Fugl-Meyer, A. R., Jaasco, L., Leyman, L., Olsson, S., & Steglind, S. (1975). The post-stroke hemiplegic patient. *Scand. J. Rehab. Med.*, *7*, 13–31.

Goldberg, Y., & Elhadad, M. (2008). splitSVM : Fast , Space-Efficient , non-Heuristic , Polynomial Kernel Computation for NLP Applications. *Proceedings of the 46st Annual Meeting of the Association of Computational Linguistics (ACL)*, 237–240.

Gruber, C., Hook, C., Kempf, J., Scharfenberg, G., & Sick, B. (2006). A Flexible Architecture for Online Signature Verification Based on a Novel Biometric Pen. *2006 IEEE Mountain Workshop on Adaptive and Learning Systems*, 110–115.

Gruber, T., Meixner, B., Prosser, J., & Sick, B. (2012). Handedness tests for preschool children: A novel approach based on graphics tablets and support vector machines. *Applied Soft Computing Journal*, *12*(4), 1390–1398.

Gutiérrez, T.; Rodríguez, J.; Vélaz, Y.; Casado, S.; Suescun, A.; Sánchez, E. J. (2010). IMA-VR : A Multimodal Virtual Training System for Skills Transfer in Industrial Maintenance and Assembly Tasks. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 428–433.

Hamrol, A.; Kowalik, D.; Kujawin'ska, A. (2011). Impact of selected work condition factors on quality of manual assembly process. *Human Factors and Ergonomics in Manufacturing*, *21*(2), 156–163.

- Hook, C., Kempf, J., & Scharfenberg, G. (2004). A novel digitizing pen for the analysis of pen pressure and inclination in handwriting biometrics. *Biometric Authentication*, 283–294.
- Howard, B., & Vance, J. (2007). Desktop haptic virtual assembly using physically based modelling. *Virtual Reality*, 11(4), 207–215.
- Howie, N., Purkayastha, S. N., Byrne, M. D., & O'Malley, M. K. (2011). Motor Skill Acquisition in a Virtual Gaming Environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 55(1), 2148–2152.
- Huegel, J. (2009). *Progressive haptic guidance for a dynamic task in a virtual training environment*. Doctoral Dissertation, Rice University, Houston, TX.
- Huntsinger, L. F., Roupail, N. M., & Bloomfield, P. (2013). Trip Generation Models Using Cumulative Logistic Regression. *Journal of Urban Planning and Development*, 139(3), 176–184.
- iPhone 4: The entire assembly process. (2010). *Youtube*. Retrieved from <https://www.youtube.com/watch?v=QgaP4zLyFAY>
- Jebsen, R. H., Taylor, N., & Trieschmann, R. B. (1969). An objective and standardized test of hand function. *Archives of Physical Medicine and Rehabilitation*, 50, 311–319.
- Jipp, M., Bartolein, C., & Badreddin, E. (2009). Predictive validity of wheelchair driving behavior for fine motor abilities: Definition of input variables for an adaptive wheelchair system. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 39–44.
- Jipp, M., Bartolein, C., Badreddin, E., Abkai, C., & Hesser, J. (2009). Psychomotor Profiling with Bayesian Networks Prediction of User Abilities based on Inputs of Motorized Wheelchair Parameters. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 1680–1685.
- Kaber, D., Tupler, L. a., Clamann, M., Gil, G.-H., Zhu, B., Swangnetr, M., Jeon, W., Zhang, Y., Qin, X., Ma, W., Lee, Y. S. (2014). Evaluation of an Augmented Virtual Reality and Haptic Control Interface for Psychomotor Training. *Assistive Technology*, 26(1), 51–60.
- Konz, S., & Johnson, S. (2004). *Work design: Occupational ergonomics* (6th ed.). Scottsdale, AZ: Holcomb Hathaway.
- Kopp, B., Kunkel, A., Flor, H., Platz, T., Rose, U., Mauritz, K. H., ... Taub, E. (1997). The Arm Motor Ability Test: reliability, validity, and sensitivity to change of an instrument

for assessing disabilities in activities of daily living. *Archives of Physical Medicine and Rehabilitation*, 78(6), 615–620.

- Krebs, H. I., Aisen, M. L., Volpe, B. T., & Hogan, N. (1999). Quantization of continuous arm movements in humans with brain injury. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8), 4645–4649.
- Krebs, H. I., Hogan, N., Aisen, M. L., & Volpe, B. T. (1998). Robot-aided neurorehabilitation. *IEEE Transactions on Rehabilitation Engineering*, 6(1), 75–87.
- Krebs, H., Palazzolo, J., Dipietro, L., Ferraro, M., Krol, J., Rannekleiv, K., ... Hogan, N. (2003). Rehabilitation robotics: Performance-based progressive robot-assisted therapy. *Autonomous Robots*, 15(1), 7–20.
- Krebs, H., Volpe, B., Aisen, M., & Hogan, N. (2000). Increasing productivity and quality of care: robot-aided neuro-rehabilitation. *Journal of Rehabilitation Research and Development*, 37(6), 639–652.
- Kuncheva, L. I. (2002). Switching between selection and fusion in combining classifiers: an experiment. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics : A Publication of the IEEE Systems, Man, and Cybernetics Society*, 32(2), 146–56.
- Lafayette Instrument Company, I. (2002). *Purdue Pegboard Test User Instructions*. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco* (Vol. 15).
- Lee, J., & Choi, S. (2010). Effects of haptic guidance and disturbance on motor learning: Potential advantage of haptic disturbance. *2010 IEEE Haptics Symposium*, 335–342.
- Li, Y., Huegel, J. C., Patoglu, V., & O'Malley, M. K. (2009). Progressive shared control for training in virtual environments. *World Haptics 2009 - Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 332–337.
- Li, Y., Patoglu, V., & O'Malley, M. K. (2009). Negative efficacy of fixed gain error reducing shared control for training in virtual environments. *ACM Transactions on Applied Perception*, 6(1), 1–21.
- Light, C. M., Chappell, P. H., & Kyberd, P. J. (2002). Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: Normative data, reliability, and validity. *Archives of Physical Medicine and Rehabilitation*, 83(6), 776–783.

- Liu, M., Yu, L., Zhang, X., & Guo, S. (2008). Cumulative logistic regression-based measurement models of road traffic congestion intensity. *Journal of Beijing Jiaotong University*, 32(6), 52–56.
- Liu, H.; Motoda, H. (1998). *Feature selection for knowledge discovery and data mining*. Kluwer Academic Publishing.
- Low Shye Nee; Goh Wen Juin; Fow Jun Yan; Chan Yee Theng; Kamaruddin, S. (2012). Lean production in manual assembly line — A case study. *2012 IEEE Symposium on Humanities, Science and Engineering Research*, 815–820.
- Lyle, R. C. (1981). A performance test for assessment of upper limb function in physical rehabilitation treatment and research. *International Journal of Rehabilitation Research*, 4, 483–492.
- Ma, Wenqi; Kaber, David; Gil, Guk-Ho; Clamann, Michael; Jeon, Wooram; Zhu, B. (2012). Comparison of Virtual Reality-Based Visual and Haptic Aiding for Psychomotor Training. *In Review*.
- Massimino, M. J., & Sheridan, T. B. (1994). Teleoperator Performance with Varying Force and Visual Feedback. *Human Factors*, 36(1), 145–157.
- Mathiowetz, V., Weber, K., Kashman, N., & Volland, G. (1985). Adult Norms for the Nine Hole Peg Test of Finger Dexterity. *The Occupational Therapy Journal of Research*, 5, 24–33.
- Mathiowetz, V.; Volland, G.; Kashman, N.; Weber, K. (1985). Adult norms for the Box and Block Test of manual dexterity. *American Journal of Occupational Therapy*, 39(6), 386–391.
- McLachlan, G. (1992). *Discriminant Analysis and Statistical Pattern Recognition* (1st ed.). John Wiley & Sons, Inc., Hoboken, New Jersey.
- Mo, Y., & Xu, S. (2010). Application of SVM Based on Hybrid Kernel Function in Heart Disease Diagnoses. *2010 International Conference on Intelligent Computing and Cognitive Informatics*, 462–465.
- Nam, Y., & Jang, T. (2010). Virtual assembly training system using a haptic device. *IADIS Multi Conference on Computer Science and Information Systems (MCCSIS 2010)*, 431–433.

- Neter, J., Wasserman, W., & Kutner, M. H. (1990). *Applied linear statistical models: Regression, analysis of variance, and experimental designs* (3rd ed., pp. 118–128). Irwin (Homewood, IL).
- Niebel, B., & Freivalds, A. (2008). *Methods, Standards, and work design* (12th ed.). McGraw-Hill Science.
- O'Malley, M. K., & Gupta, a. (2003). Passive and active assistance for human performance of a simulated underactuated dynamic task. *11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2003. HAPTICS 2003. Proceedings.*, 348–355. doi:10.1109/HAPTIC.2003.1191308
- O'Malley, M. K., Gupta, A., Gen, M., & Li, Y. (2006). Shared Control in Haptic Systems for Performance Enhancement and Training. *Journal of Dynamic Systems, Measurement, and Control*, 128(1), 75–85.
- Pathomaree, N., & Charoenseang, S. (2005). Augmented reality for skill transfer in assembly task. *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, 500–504.
- Perlin, R. (2013). Chinese Workers Foxconned. *Dissent: a quarterly of politics and culture*. Retrieved from <http://www.dissentmagazine.org/article/chinese-workers-foxconned>
- Petrofsky, J. S., & Petrofsky, D. (2004). A simple device to assess and train motor coordination. *Journal of Medical Engineering & Technology*, 28(2), 67–73.
- Philbin, D.A.; Ribarsky, W.; Walker, N.; Hubbard, C. E. (1998). Training in virtual environments: analysis of task appropriateness. *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*, 210.
- Pohar, M., Blas, M., & Turk, S. (2004). Comparison of Logistic Regression and Linear Discriminant Analysis : A Simulation Study. *Metodološki Zvezki*, 1(1), 143–161.
- Price, K., & Sears, A. (2008). Performance-based functional assessment: An algorithm for measuring physical capabilities. *ASSETS'08: The 10th International ACM SIGACCESS Conference on Computers and Accessibility*, 217–224.
- Pugmire, D., Modi, D., & Monroe, L. (2007). Pick-n-Place: A Virtual Reality Assembly Tool. *NW Highlights, Los Alamos National Laboratory*.
- Reinkensmeyer, D., Aoyagi, D., Emken, J., Galvez, J., Ichinose, W., Kerdanyan, G., ... Edgerton, V. (2004). Robotic gait training: toward more natural movements and optimal

training algorithms. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 7(26), 4818–4821.

Reinkensmeyer, David J.; Leonard E.; Averbuch, Michele; Mckenna-Cole, Alicia; Schmit, Brian D.; Rymer, W. Z. (2000). Understanding and treating arm movement impairment after chronic brain injury: progress with the ARM guide. *Journal of Rehabilitation Research and Development*, 37(6), 653–662.

Rohlik, O., Mautner, P., Matousek, V., & Kempf, J. (2003). HMM based handwritten text recognition using biometrical data acquisition pen. *Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium (Cat. No.03EX694)*, 2, 950–953.

Rosen, M.J.; Goodenough-Trepagnier, C. (1989). The tufts-MIT prescription guide: Assessment of users to predict the suitability of augmentative communication devices. *Assistive Technology*, 1(3), 51–61.

Salazar-Torres, J.; Pandyan, A.D.; Price, C.I.M.; Davidson, R.I.; Barnes, M.P.; Johnson, G. R. (2004). Does spasticity result from hyperactive stretch reflexes? *Disability and Rehabilitation*, 26(12), 756–760.

Schmidt, R. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. *Tutorials in Motor Neuroscience*, 59–75.

Schmidt, R. A., & Wulf, G. (1997). Continuous concurrent feedback degrades skill learning: Implications for training and simulation. *Human Factors*, 39(4), 509–525.

Singer, R. N. (1980). *Motor learning and human performance: An application to motor skills and movement behavior* (3rd ed.). Macmillan Publications.

Solis, J.; Vizzano, C.A.; Bergamasco, M. (2003). Validating a skill transfer system based on reactive robots technology. *Proceedings. RO-Man 2003. The 12th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Cat. No. 03TH8711)*, 175–180.

Spyers-Ashby, J. M., Stokes, M. J., Bain, P. G., & Roberts, S. J. (1999). Classification of normal and pathological tremors using a multidimensional electromagnetic system. *Medical Engineering & Physics*, 21(10), 713–723.

Srimathveeravalli, G., Gourishankar, V., & Thenkurussi, K. (2006). Haptic Attributes and Human Motor Skills. *2006 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 453–459.

- Staelin, C. (2002). *Parameter selection for support vector machines* (Vol. 354).
- Steele, M., Gillespie, R. B., & Arbor, A. (2001). Shared control between human and machine: Using a haptic steering wheel to aid in land vehicle guidance. *Human Factors and Ergonomics Society 45th Annual Meeting*.
- Taub, E.; Miller, N.E.; Novack, T.A.; Cook, E.W. 3rd; Fleming, W.C.; Nepomuceno, C.S.; Connell, J.S.; Crago, J. E. (1993). Technique to improve chronic motor deficit after stroke. *Archives of Physical Medicine and Rehabilitation*, 74(4), 347–354.
- Tiffin, J. (1968). *Purdue Pegboard Examiner Manual*. Chicago, IL.
- Todorov, E., Shadmehr, R., & Bizzi, E. (1997). Augmented feedback presented in a virtual environment accelerates learning of a difficult motor task. *Journal of Motor Behavior*, 29(2), 147–158.
- Trombly, C. A. (1995). *Occupational Therapy for Physical Dysfunction*. Baltimore, MD: Williams and Wilkins.
- Vandamme, D. (2010). Assessment of motor and process skills: assessing client work performance in Belgium. *WORK: A Journal of Prevention, Assessment & Rehabilitation*, 35(2), 201–208.
- Vazquez, M., & Resnick, M. (1997). Supporting Flexible Assembly through Human Factors. *Proceedings of the Human Factors and Ergonomics Society*, 1, 561–564.
- Vo, D. M., Vance, J. M., & Marasinghe, M. G. (2009). Assessment of haptics-based interaction for assembly tasks in virtual reality. *World Haptics 2009 - Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 494–499.
- Wang, D., Zhang, Y., & Yao, C. (2006). Machine-mediated Motor Skill Training Method in Haptic-enabled Chinese Handwriting Simulation System. *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 5644–5649.
- Wang, F., He, K., Liu, Y., Li, L., & Hu, X. (2013). Research on the selection of kernel function in SVM based facial expression recognition. *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*, (2), 1404–1408.
- Williams, L. E. P., Loftin, R. B., Aldridge, H. A., Leiss, E. L., & Bluethmann, W. J. (2002). Kinesthetic and visual force display for telerobotics. *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, 2, 1249–1254.

- Woll, R., Damerou, T., Wrasse, K., & Stark, R. (2011). Augmented reality in a serious game for manual assembly processes. *2011 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities*, 37–39.
- Zee, D. van der, & Slomp, J. (2005). Simulation and gaming as a support tool for lean manufacturing systems: a case example from industry. *Proceedings of the 2005 Winter Simulation Conference (IEEE Cat. No.05CH37732C)*, 10.
- Zollo, L., Rossini, L., Bravi, M., Magrone, G., Sterzi, S., & Guglielmelli, E. (2011). Quantitative evaluation of upper-limb motor control in robot-aided rehabilitation. *Medical & Biological Engineering & Computing*, 49(10), 1–14.

APPENDICES

Appendix A

North Carolina State University
Edward P. Fitts Department of Industrial and Systems Engineering
The Cognitive Ergonomics Laboratory

Background Survey Form

Title of Study: Motor ability level classification model development
Subject No: _____

1. Your name: _____
2. Your contact information: Tel.: _____ Email: _____
3. Your age: _____
4. Your gender: Female Male
5. Your dominant hand (the one you use to write) is: left hand right hand
6. Do you have 20/20 or corrected vision? Yes No
7. Do you have prior experience with haptic devices or motor training tasks?
 Yes No
8. Do you have a diagnosed motor control disability? Yes No

Appendix B

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

Title of Study: MOTOR ABILITY LEVEL CLASSIFICATION FOR MODEL DEVELOPMENT

Principal Investigator: Wenqi Ma

Faculty Sponsor (if applicable): Dr. David B. Kaber

General Information

You are being asked to take part in a research study. Your participation in this study is voluntary. You have a right to be a part of this study, to choose not to participate or to stop participating at any time. The purpose of research studies is to gain a better understanding of how to develop quantitative classification models for determining operator motor skill levels based on baseline task performance. You are not guaranteed any personal benefits from joining this study. Research studies like this may also pose risks to participants. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something listed in this form it is your right to ask for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact Wenqi Ma at 919-515-7210.

Purpose of this study

The purpose of this study is to learn more about how to classify the participants based on their motor performance in baseline tasks requiring simple motor skills. A quantitative model will be developed in this study and will provide assistance in further research about appropriate training protocol selection. The tasks to be performed in this study involve virtual dice manipulation training and the Purdue Pegboard Test. The dice manipulation task will be presented using a virtual reality (VR) simulation with a hand-held haptic stylus. The VR simulation system also includes a 3-D visual display that is viewed through active LCD goggles. Your performance on the motor tasks using the VR system is expected to provide insights into the motor ability and quantitative approaches for classifying motor ability levels. The Purdue Pegboard Test will be completed using the right hand and the average score in the test will provide a “gold standard” for assessing actual motor ability levels.

Procedure

If you agree to participate in this study, you will be asked to:

- 1) Read and sign this consent form, once you understand the research and agree to participate.
- 2) Perform 3 trials of the right-hand portion of the Purdue Pegboard Test (screen test). This session will last about 5 minutes.
- 3) Perform a baseline motor test involving 40 trials of virtual dice manipulation. This session will last about 20 minutes.
- 4) Fill out a payment form.

Risks

Potential risks from this research include: (1) general fatigue due to attending to the VR displays during the baseline motor test; and (2) hand and wrist fatigue due to controlling the haptic device. Rest periods will be arranged after the test session.

Benefits

You may derive benefits from this research including improved motor skills through exposure to a series of motor training tasks. You may also derive some indirect benefits including understanding of human factors research methods and knowledge of concepts such as virtual environment and haptic device design, etc.

Confidentiality

The information in this study records will be kept confidential. Data will be stored securely in the Cognitive Ergonomics Lab in the NC state Department of Industrial and Systems Engineering. The data will only be made available to researchers conducting the study. No reference will be made to you in oral or written reports of the study, which could link you to the research. Any identifying information collected during the study will only be used for recruitment and will be removed from data. A code number will be matched to your name and a master list with the code number and any identifying information will be kept separately from all other survey and response data collected as part of the experiment. This code list and all other data will be destroyed at the close of the study.

Compensation

For completing this experiment, you will be paid at a rate of \$15/hour.

Contact

If you have any questions at any time about the study or the procedures, you may contact the researcher, Wenqi Ma, at 458 Daniels Hall, NC State University main campus, or 919-515-7210. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Debra Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919-515-4514).

Consent to Participate

“I have read and understood the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may withdraw at any time.”

Participant’s signature _____ **Date** _____

Investigator’s signature _____ **Date** _____

Appendix C

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

Title of Study: MOTOR ABILITY LEVEL CLASSIFICATION FOR TRAINING PROTOCOL SELECTION

Principal Investigator: Wenqi Ma

Faculty Sponsor (if applicable): Dr. David B. Kaber

General Information

You are being asked to take part in a research study. Your participation in this study is voluntary. You have a right to be a part of this study, to choose not to participate or to stop participating at any time. The purpose of research studies is to gain a better understanding of how to select suitable training protocol based on operators' baseline performance. You are not guaranteed any personal benefits from joining this study. Research studies like this may also pose risks to participants. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something listed in this form it is your right to ask for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact Wenqi Ma at 919-515-7210.

Purpose of this study

The purpose of this study is to learn more about how to select suitable training protocols for operators with different motor abilities in order to facilitate uniform target task performance. The tasks to be performed in this study involve virtual dice manipulation training and block design pattern assembly. Both tasks will be presented using a virtual reality (VR) simulation with a hand-held haptic stylus. The VR simulation system also includes a 3-D visual display that is viewed through active LCD goggles. Your performance on the motor tasks using the VR system is expected to provide insights into the motor ability and quantitative approaches for classifying motor ability levels.

Procedure

If you agree to participate in this study, you will be asked to:

- 5) Read and sign this consent form once you understand the research and agree to participate.
- 6) Perform a baseline motor test involving 40 trials of virtual dice manipulation. This session will last about 30 minutes.
- 7) Complete three training sessions. You will be assigned to one of three training conditions according to your performance in the baseline motor test. Each training condition will involve a similar block design (BD) task, but with different levels of automated assistance. Each training session will involve six BD trials. This phase of the experiment will take about 45 minutes.
- 8) Perform a basic BD test involving six trials. This session will last about 15 minutes.
- 9) Fill out a payment form.

Risks

Potential risks from this research include: (1) general fatigue due to attending to the VR displays during the training trials; and (2) hand and wrist fatigue due to controlling the haptic device. Rest periods will be arranged between training sessions.

Benefits

You may derive benefits from this research including improved motor skills through exposure to a series of motor training tasks. You may also derive some indirect benefits including understanding of human factors research methods and knowledge of concepts such as virtual environment and haptic device design, etc.

Confidentiality

The information in this study records will be kept confidential. Data will be stored securely in the Cognitive Ergonomics Lab in the NC state Department of Industrial and Systems Engineering. The data will only be made available to researchers conducting the study. No reference will be made to you in oral or written reports of the study, which could link you to the research. Any identifying information collected during the study will only be used for recruitment and will be removed from data. A code number will be matched to your name and a master list with the code number and any identifying information will be kept separately from all other survey and response data collected as part of the experiment. This code list and all other data will be destroyed at the close of the study.

Compensation

For participating in this study you will be paid \$15 for each hour you spend in the experiment. If you withdraw from the study prior to its completion, you will receive compensation for all time you have contributed in the experiment.

Contact

If you have any questions at any time about the study or the procedures, you may contact the researcher, Wenqi Ma, at 458 Daniels Hall, NC State University main campus, or 919-515-7210. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Debra Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919-515-4514).

Consent to Participate

“I have read and understood the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may withdraw at any time.”

Participant’s signature _____ **Date** _____

Investigator’s signature _____ **Date** _____

Appendix D

Phase I Experiment Instructions

This document includes instructions and scripts to administer the Phase I experiment as part of the dissertation, “*An application of quantitative methods for motor ability level classification, performance prediction and training protocol selection.*” The instructions are provided in tabular format, with sections to be read aloud in *italics*. Supplemental information (e.g., materials, scoring methods, etc.) are provided for each section.

Orientation

Introduction	<p><i>Thank you for participating in this experiment. There are two parts to this experiment. You will begin with a computer-based motor skill test called the Dice Manipulation Task. After finishing this test, you will complete a standardized psychomotor test, the Purdue Pegboard Test. The entire experiment will last about an hour.</i></p> <p><i>Please understand that this experiment is not to test your personal ability or skills. The goal of this study is to collect sufficient data to develop a model for classifying human motor skill in order to determine the ideal type of training for patients, workers, etc.</i></p>
Informed consent	<p>[Prepare the informed consent form and a pen.]</p> <p><i>This form summarizes information you need to know about the experiment. Please read it. If you have any questions, please feel free to ask me. Please note that in order to participate in this study you should be at least 18 years old, have 20/20 vision (with or without correction; glasses or contact lenses are OK). You will receive \$15/hour for your participation. If you consent to participate, please sign and date the form.</i></p> <p>After signing the form, seat the participant at the testing station for the Dice Manipulation Task and begin with the instructions.</p> <p><i>Before we start I need you to turn off your cell phone to avoid distractions during the experiment. You will be able to take a break between the two tests.</i></p>

Part I. Dice Manipulation Task

Introduction and training	<p>Start the dice manipulation task program. Enter the participant number. Use the Level 2 task difficulty for the tutorial.</p> <p><i>In this part of the experiment, you are going to complete 40 trials of the Dice Manipulation Task using a computer-based program. The task you'll be performing uses a 3D virtual environment, so you will need to wear active light shutter goggles for all trials.</i></p> <p>Ask the participant to don the 3D glasses. Lift the haptic stylus before starting the trial.</p> <p><i>Now, I am going to ask you to use this stylus to manipulate the die in the virtual environment. Please hold the stylus as you would hold a pen and use your index finger to press the dark button. [Show the position of the button on the stylus.] Look at the ball-shaped pointer on the virtual table. [Point to the pointer on the screen.] Please move the stylus to control the pointer and position it at the surface of the die. Feel the resistance force when touching the die. You can click the dark button to hold the die. The button can be released after the die leaves the table and the die will remain attached to the pointer. You can change the die orientation by rotating your wrist. The die will be automatically released when touching the table.</i></p> <p>[Demonstrate how to use the haptic stylus to manipulate the die.]</p> <p><i>The goal of this task is to move the die to the target square shown in black [point to the target square on the screen] with the top surface matching the stimulus [point to the stimulus on the screen]. Please complete each trial as quickly and accurately as possible.</i></p> <p><i>Now please try to complete this trial. You should familiarize yourself with the operation of the device as much as possible. Tell me when you finish the task. We will NOT record your performance at this time.</i></p>
Formal test	<p>Now we will start the formal test sessions.</p> <p><i>The dice manipulation program will automatically record your performance in each trial from now on. Tell me when you finish the trial. Please complete each trial as quickly and accurately as you can.</i></p>

	<p>[Ask the participant to hold the stylus.]</p> <p><i>Are you ready?</i></p> <p>Enter the task difficulty level. Press the OK button to start the trial.</p> <p>Each participant will complete 5 trials at Level 1, 5 trials at Level 2 and 30 trials at Level 3 and 4 alternatively. In total, 40 trials will be completed.</p>
End of session	<p><i>Now you have finished all 40 trials of the dice manipulation task. Please take a short break. We will start the second part of the experiment in 5 minutes.</i></p>

Part II. Purdue Pegboard Test

Introduction and training	<p>[Make ready the test board, pins, and stopwatch.]</p> <p><i>Now we will begin the second part of the experiment. Your task in this part is to insert as many pins as possible in the board within a limited time. You will use your right hand to pick up one pin at a time from the cup and place each pin in the board starting from the top hole.</i></p> <p>[Provide a demonstration of inserting a pin.]</p> <p><i>Now you may insert a few pins in the board for practice. You should familiarize yourself with the task as much as possible. We will NOT record your score at this time.</i></p> <p><i>During the test time, if you drop a pin, please do not stop to pick it up. Simply continue by picking up another pin from the cup.</i></p> <p>Correct any errors made in placing the pins and answer any questions. After the participant has inserted three or four pins and appears to understand the operation, stop the training.</p> <p><i>Please stop. Now, please remove all pins from the board and put them back into the cup.</i></p>
Formal test	<p><i>Now we will begin the formal session. When I say “Start”, place as many pins as possible in the right-hand row of the board, starting from the top</i></p>

	<p><i>hole. Work as rapidly as you can until I say “Stop”. You will have 30 seconds to finish this trial.</i></p> <p><i>Are you ready? Start.</i></p> <p>[Start timing. After 30 seconds, stop the participant.]</p> <p><i>Stop.</i></p> <p>Count the number of pins inserted correctly and record the number as the score for this trial.</p> <p>Take out the pins and put them back into the cup. Repeat the test two more times.</p>
--	---

Departure

Departure and thanks	<p><i>Now you have finished the experiment. Please fill in the payment form. Your compensation for the experiment will be \$XX.</i></p> <p>[Let the participant fill out the payment form.]</p> <p><i>The data collected today will be used to develop a model for classifying human motor performance. The model will be applied in a further study to assign participants to proper motor skill training protocols. If you are interested in additional information about this study or have any questions, please contact me or Dr. Kaber. Our contact information is listed in the consent form that you will take home with you.</i></p> <p>[Give a copy of the consent form to the participant.]</p> <p><i>Thank you for participating in this study.</i></p>
----------------------	---

Appendix E

Phase II Experiment Instructions

This document includes instructions and scripts to administer the Phase II experiment as part of the dissertation, “*An application of quantitative methods for motor ability level classification, performance prediction and training protocol selection.*” The instructions are provided in tabular format, with sections to be read aloud in *italics*. Supplemental information (e.g., materials, scoring methods, etc.) are provided for each section.

Orientation

Introduction	<p><i>Thank you for participating in this experiment. There are three parts to this experiment. You will begin with a computer-based motor skill test called the Dice Manipulation Task. After finishing this test, you will take three training sessions. You will complete another computer-based Block Design test for the last part. The entire experiment will last about two hours.</i></p> <p><i>Please understand that this experiment is not to test your personal ability or skills. The goal of this study is to verify a motor ability classification model developed in prior experiment. The model was expected to determine the ideal type of training for patients, workers, etc.</i></p>
Informed consent	<p>[Prepare the informed consent form and a pen.]</p> <p><i>This form summarizes information you need to know about the experiment. Please read it. If you have any questions, please feel free to ask me. Please note that in order to participate in this study you should be at least 18 years old, have 20/20 vision (with or without correction; glasses or contact lenses are OK). You will receive \$15/hour for your participation. If you consent to participate, please sign and date the form.</i></p> <p>After signing the form, seat the participant at the testing station for the Dice Manipulation Task and begin with the instructions.</p> <p><i>Before we start I need you to turn off your cell phone. You will be able to take breaks, but to avoid distractions, it needs to be off during the experiment.</i></p>

Part I. Dice Manipulation Task

<p>Introduction and training</p>	<p>Start the dice manipulation task program. Enter the participant number. Use the Level 2 task difficulty for the tutorial.</p> <p><i>In this part of the experiment, you are going to complete 40 trials of the Dice Manipulation Task using a computer-based program. The task you'll be performing uses a 3D virtual environment, so you will need to wear active light shutter goggles for all trials.</i></p> <p>Ask the participant to don the 3D glasses. Lift the haptic stylus before starting the trial.</p> <p><i>Now, I am going to ask you to use this stylus to manipulate the die in the virtual environment. Please hold the stylus as you would hold a pen and use your index finger to press the dark button. [Show the position of the button on the stylus.] Look at the ball-shaped pointer on the virtual table. [Point to the pointer on the screen.] Please move the stylus to control the pointer and position it at the surface of the die. Feel the resistance force when touching the die. You can click the dark button to hold the die. The button can be released after the die leaves the table and the die will remain attached to the pointer. You can change the die orientation by rotating your wrist. The die will be automatically released when touching the table.</i></p> <p>[Demonstrate how to use the haptic stylus to manipulate the die.]</p> <p><i>The goal of this task is to move the die to the target square shown in black [point to the target square on the screen] with the top surface matching the stimulus [point to the stimulus on the screen]. Please complete each trial as quickly and accurately as possible.</i></p> <p><i>Now please try to complete this trial. You should familiarize yourself with the operation of the device as much as possible. Tell me when you finish the task. We will NOT record your performance at this time.</i></p>
<p>Formal test</p>	<p><i>Now we will start the formal test sessions.</i></p> <p><i>The dice manipulation program will automatically record your performance in each trial from now on. Tell me when you finish the trial. Please complete each trial as quickly and accurately as you can.</i></p>

	<p>[Ask the participant to hold the stylus.]</p> <p><i>Are you ready?</i></p> <p>Enter the task difficulty level. Press the OK button to start the trial.</p> <p>Each participant will complete 5 trials at Level 1, 5 trials at Level 2 and 30 trials at Level 3 and 4 alternatively. In total, 40 trials will be completed.</p>
End of session	<p><i>Now you have finished all 40 trials of the dice manipulation task. Please take a short break. We will start the next part of the experiment in 5 minutes.</i></p>

Part II. Block Design Baseline Test

Introduction	<p>Start the test program. Enter the participant number. Use Trial 8 task for the tutorial.</p> <p><i>In this test, you will perform a block design task using the haptic control device as in last part.</i></p> <p>Ask the participant to put on the 3D goggles. Lift the haptic stylus before starting the trial.</p> <p><i>The function of the stylus will be the same as in last part when manipulating the block. During this training, your task is to construct designs using a set of 4 or 9 virtual blocks. As you can see, these blocks [point to the blocks on the screen] are identical. Some sides are whole white, some are whole red, and some are half and half. Please move the given blocks into the grids [point to the grids on the virtual desk as desired destination] to make the combined top surface look exactly like the given stimuli [point to the stimuli design on the screen].</i></p> <p><i>Now please try to complete this trial as accurately and quickly as possible. Tell me when you finish the task. We will NOT record your performance at this time.</i></p>
--------------	--

Formal test	<p><i>Now we will start the formal test sessions.</i></p> <p><i>The block design test program will automatically record your performance in each trial from now on. Tell me when you finish the trial. Please complete each trial as quickly and accurately as you can.</i></p> <p>[Ask the participant to hold the stylus.]</p> <p><i>Are you ready?</i></p> <p>Enter the trial number. Press the OK button to start the trial.</p> <p>Each participant will complete 6 trials in the test. The trials are Trial 9-14 from WAIS, given in successive order.</p>
End of session	<p><i>Now you have finished all the block design test. Please take a short break. We will start the next part of the experiment in 5 minutes.</i></p>

Part III. Block Design Task Training

Motor ability level prediction	<p>During the break, process the dice manipulation trajectory data following steps below.</p> <ol style="list-style-type: none"> 1. Open the txt file recording the trajectory. The file will be named as “haptic_DATE_TIME_SUBJECT#_trajectory.txt” (e.g., haptic_2013_821_0905_trajectory.txt). 2. Transfer all data to the file named “raw.txt”. Save and close the file. 3. Run R-scripts “parameter.R”, “feature extraction.R”, “level classification.R” successively. The output letter will indicate the motor ability level of the participant (H – high, L – low, M – medium). The participant should take one of the three training programs according to predicted level. High group will take training with assistive force (AS), low group will take training with random force feedback (RN), and medium group will take training with resistive force feedback (RE).
Introduction	<p>Start the training program. Enter the participant number. Use Trial 7 task for the tutorial.</p>

	<p><i>In this part of the experiment, you are going to complete three sessions of block design training, each including six trials. Again, you will be performing the task in a 3D virtual environment with the same goggles.</i></p> <p>Ask the participant to put on the 3D goggles. Lift the haptic stylus before starting the trial.</p> <p><i>Different from the block design task you did in last part, in completing the training task, additional force will be given during your performance. Also, the stimulus designs will be different.</i></p> <p>For participants taking AS: <i>When you move the block to the corresponding target location, you will feel a snap force which will help you get it done easily.</i></p> <p>For participants taking RE: <i>When you move the block to the corresponding target location, a resistive force will be given and you have to overcome the force to allocate the block.</i></p> <p>For participants taking RN: <i>When you move the block to the corresponding target location, a random force will be given to distract your performance. You will have to struggle with it to complete the task.</i></p> <p>[Demonstrate the force feature by positioning one block and return the stylus to the participant.]</p> <p><i>The other requirements are the same as in last part. Now please try to complete this trial as accurately and quickly as possible. You need to familiarize yourself with the force feature as much as possible. Tell me when you finish the task. We will NOT record your performance at this time.</i></p>
Formal test	<p><i>Now we will start the formal test sessions.</i></p> <p><i>The block design training program will automatically record your performance in each trial from now on. Tell me when you finish the trial. Please complete each trial as quickly and accurately as you can.</i></p> <p>[Ask the participant to hold the stylus.]</p> <p><i>Are you ready?</i></p> <p>Enter the trial number. Press the OK button to start the trial.</p>

	Each participant will complete 6 trials during one training sessions. The trials should be Trial 8-13 from WASI, given in random order. Three training sessions should be completed. A 5-minute mandatory break should be given between training sessions.
End of training	<i>Now you have finished three training sessions. Please take another short break. We will start the last part of the experiment in 5 minutes.</i>

Part IV. Block Design Final Test

Introduction	<p>Start the test program. Enter the participant number.</p> <p><i>In this part, you are going to complete six trials of block design tasks as in previous sections. However, this time, there will be no additional force features involved. In fact, it will be exactly the same as you did in the baseline test.</i></p>
Formal test	<p><i>Now we will start the formal test sessions.</i></p> <p><i>The block design test program will automatically record your performance in each trial from now on. Tell me when you finish the trial. Please complete each trial as quickly and accurately as you can.</i></p> <p>[Ask the participant to hold the stylus.]</p> <p><i>Are you ready?</i></p> <p>Enter the trial number. Press the OK button to start the trial.</p> <p>Each participant will complete 6 trials in the test. The trials are Trial 9-14 from WAIS, given in successive order.</p>

Departure

Departure and thanks	<p><i>Now you have finished the experiment. Please fill in the payment form. Your compensation for the experiment will be \$XX.</i></p> <p>[Let the participant fill out the payment form.]</p> <p><i>The data collected today will be used to verify the classification model developed in prior study. The model is expected to be applied in classifying assembly workers to select proper training methods accordingly. If you are interested in additional information about this study or have any questions, please contact me or Dr. Kaber. Our contact information is listed in the consent form that you will take home with you.</i></p> <p>[Give a copy of the consent form to the participant.]</p> <p><i>Thank you for participating in this study.</i></p>
----------------------	---