Fall 2014

Acquisition, retention and transfer of heavy equipment operator skills through simulator training

Chung Yin So
Purdue University

Graduate School Form 30
(Revised 08/14)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Chung Yin So

Entitled
Acquisition, Retention and Transfer of Heavy Equipment Operator Skills through Simulator Training

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Vincent Duffy
Co-chair

Steven Landry

Robert Proctor
Co-chair

Mark Lehto

Phillip Dunston
Co-chair

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Vincent Duffy

Approved by: Abhijit Deshmukh 12/01/2014

Head of the Department Graduate Program Date

ACQUISITION, RETENTION AND TRANSFER OF HEAVY EQUIPMENT OPERATOR SKILLS THROUGH SIMULATOR TRAINING

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Chung Yin So

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2014

Purdue University

West Lafayette, Indiana

This is dedicated to my dear grandparents, my parents, my sister and my husband.

Thanks for all the unconditional love, support, and encouragement you have given me.

ACKNOWLEDGEMENTS

Throughout the course of my Ph.D. studies, I have received support and encouragement from many individuals. First, I would like to thank my advisors, Drs. Robert Proctor and Phillip Dunston, for their invaluable advice and continuous support, especially the opportunity to get involved in the project of virtual reality construction equipment simulator training. In particular, I would like to thank you Dr. Proctor for inspiring me to complete an additional Master's degree in Cognitive Psychology on top of my Ph.D. degree in Industrial Engineering. I would like to thank Dr. Vincent Duffy for serving as my co-advisor in Industrial Engineering and his support and advice on the dissertation. All the guidance, advice, and support received from them have made my Ph.D. studies a fruitful and rewarding journey. I would also like to thank Drs. Steven Landry and Mark Lehto, for serving on my exam committee.

I express many thanks to all the lab members of Team Proctor with whom I have worked for several years. Their friendship and support have made my time at Purdue enjoyable and bearable.

I would like to acknowledge Dr. Julian Sanchez, the Director of John Deere Technology Innovation Center, for his willing collaboration and providing inspiration on simulator training, in particular, offering resources and providing connections with the experienced construction operators who took part in my research.

Special thanks to Dr. Alan H.S. Chan, my mentor at City University of Hong Kong, who has inspired me and trained me to be a good researcher and influenced me to be enthusiastic about Human Factors and Ergonomics.

Without the love and support of my family, I would not be where I am today. I especially want to thank my parents, Wai Yin (Rebecca) Chan and Hon Fat (Danny) So for their endless love and encouragement throughout the course of my life. I would also like to give a big thank you to my sister, Chung Kei (Do Do) So, for always being so patient and understanding. Finally, I want to say thank you to my beloved husband, Chun Wing Lam, for being my source of strength and encouragement throughout my postgraduate studies.

TABLE OF CONTENTS

Page

LIST OF TABLES xii

LIST OF FIGURES xiv

ABSTRACT xvii

CHAPTER 1. OBJECTIVES AND SIGNIFICANCE 1

CHAPTER 2. LITERATURE REVIEW 5

2.1. Task Switching 6
2.2.1. Cognitive Task-Switching 7
2.2.2. Motor Control Task-Switching 9
2.2.3. Does the Task-Switching Cost in Cognitive and Simple Motor Tasks Hold for Complex Perceptual Motor Tasks?. . . 10
2.2. Acquisition, Retention, and Transfer of Training 12
2.3. Training Principles 13
2.3.1. Specificity of Training 14
2.3.2. Contextual Interference. 16
2.3.3. Task Difficulty. 17
2.3.4. Variability of Practice 19
2.4. Transfer Taxonomy. 20
2.5. Transfer of VR Training 21
2.5.1. Surgical Simulators 22
2.5.2. Flight Simulators 23
2.5.3. Construction Equipment Simulators 24
2.6. Methodology for Collecting Qualitative Data 29
2.6.1. NASA Task Load Index 29
2.6.2. Task Analysis 30
2.6.2.1. Hierarchical Task Analysis 31
2.6.3. Verbal Protocol Analysis 33
2.7. Summary 35

Page

CHAPTER 3. RESEARCH FRAMEWORK 37

CHAPTER 4. SIMULATORS. 41

4.1. Experiment Apparatus. 41
4.1.1. Hydraulic Excavator Simulator 41
4.1.2. Loader Simulator. 44

CHAPTER 5. STUDY 1: TRAINING FOR SIMPLE TASKS WITH TWO MACHINES 50

5.1. Objectives. 51
5.2. Hypothesis 52
5.3. Method. 52
5.3.1. Participants. 52
5.3.2. Experimental Setup 52
5.3.3. Design 53
5.3.4. Experimental Task 54
5.3.4.1. Controls Familiarization of Hydraulic Excavator Simulator 55
5.3.4.2. Controls Familiarization of Wheel Loader 55
5.3.4.3. Reading Task 55
5.3.5. Procedure 56
5.4. Results. 57
5.4.1. Control Group. 57
5.4.1.1. Practice Effect 57
5.4.1.2. Total Execution Time. 60
5.4.1.3. Total Number of Errors 60
5.4.1.4. Workload Measures 62
5.4.2. Initial Performance on Excavator (Session 1). 62
5.4.3. Effects of Intervening Task. 63
5.4.3.1. Unrelated Reading Task 63
5.4.3.2. Practicing on Loader 64
5.4.3.3. Workload Measures 64
5.5. Discussion 65
5.6. Conclusion 66

Page

CHAPTER 6. STUDY 2: TRAINING FOR COMPLEX TASK WITH TWO MACHINES 67

6.1. Objectives. 67
6.2. Hypotheses 68
6.3. Method. 70
6.3.1. Participants. 70
6.3.2. Experimental Setup 70
6.3.3. Design 71
6.3.4. Experimental Task 72
6.3.5. Procedure 74
6.4. Results. 75
6.4.1. Retention Test on the Excavator Simulator 75
6.4.1.1. Initial Test (Session 1) 77
6.4.1.2. Group 1 vs. Group 2 77
6.4.1.3. Group 1 vs. Group 3 78
6.4.2. First Three Sessions on Excavator Simulator. 81
6.4.3. Performance on Loader Simulator of the Two Intervening Groups 83
6.4.4. Workload Measures 85
6.5. Discussion 89
6.6. Conclusion 92

CHAPTER 7. STUDY 3: HIERARCHICAL TASK ANALYSIS OF TRUCK LOADING TASKS 94

7.1. Objectives. 94
7.2. Comparison Between Loader and Excavator Simulators 96
7.2.1. Controls and Functions of Both Machines. 96
7.2.2. Truck Loading Scenarios 97
7.3. Task Analysis of Truck Loading Tasks 102
7.3.1. Develop Preliminary HTAs 102
7.3.2. Verify Task Analysis With Experienced Operators. 104
7.3.2.1. Participants 104
7.3.2.2. Experimental Setup 104
7.3.2.3. Procedure 106
7.3.3. HTAs Revision 107

Page

7.4. Comments From the Operators 107
7.5. Discussion 116
7.6. Conclusion 118

CHAPTER 8. STUDY 4: VERBAL PROTOCOL ANALYSIS BY EXPERTS. . . 120

8.1. Objectives. 122
8.2. Method. 123
8.2.1. Participants. 123
8.2.2. Experimental Setup 124
8.2.3. Design 124
8.2.4. Experimental Task 124
8.2.4.1. Simple Bucket Loading (B1) 126
8.2.4.2. Filling a Trench (B2) 126
8.2.4.3. Truck Loading (B3) 127
8.2.4.4. Moving a Load With Wide Forks/Fork Lifting (F) . . 127
8.2.5. Procedure 127
8.3. Data Analysis: How to use Verbal Protocol Analysis to Develop HTA 130
8.3.1. Phase 1: Transcription of Verbal Data and Familiarizing Yourself With Your Data 131
8.3.2. Phase 2: Cleaning Data. 132
8.3.3. Phase 3: Identifying Verbs of Actions (Generating Initial Codes) 135
8.3.4. Phase 4: Rank Verbal Reports by Number of Actions 136
8.3.5. Phase 5: Searching for Themes 136
8.3.6. Phase 6: Reviewing Themes and Codes 144
8.3.7. Phase 7: Naming Themes Into Goals and Subgoals 144
8.3.8. Phase 8: Tracing the Plans of the Subgoals 145
8.3.9. Phase 9: Building the HTA Diagram 146
8.4. Results. 146
8.4.1. HTAs of Four Loader Tasks 146
8.4.2. Workload Measures 148
8.4.3. Post Questionnaires 152
8.5. Discussion 153
8.5.1. Comparisons of B1, B2, B3 and F HTA Results. 153
8.5.2. Experts' Verification Versus Verbal Protocol Analysis . . . 155

Page

8.5.3. Difficulty Levels 157
8.6. Conclusion 157

CHAPTER 9. STUDY 5: SKILL TRANSFER AND RETENTION ON A MACHINE 158

9.1. Objectives. 158
9.2. Hypotheses 159
9.3. Method. 160
9.3.1. Participants. 160
9.3.2. Experimental Setup 161
9.3.3. Design 161
9.3.4. Performance Measures 163
9.3.5. Procedure 163
9.4. Results. 165
9.4.1. Initial Practice. 165
9.4.2. Performance Measures 166
9.4.2.1. Productivity on Truck Loading Task 167
9.4.2.1.1. Effects on First and Last Trials in Each Session 169
9.4.2.2. Workload Measures 172
9.4.2.2.1. Transfer Test Versus Retention Test 172
9.4.2.2.2. Workload Measures of the Four Loader Tasks. 174
9.4.2.3. Task Difficulty Ranked by the Novices 174
9.5. Novices vs. Experts. 175
9.6. Discussion 178
9.7. Conclusion 184

CHAPTER 10. GENERAL DISCUSSION 185

10.1. Summary of Main Findings 185
10.2. Research Questions 188
10.2.1. How much does training on one machine transfer (positively or negatively) to other machines?. 188

Page

10.2.2. Does insertion of training on various machines facilitate (or inhibit) learning and retention on a previously practiced machine? 188
10.2.3. When should an alternative machine be introduced in the training if skills on multiple machines are required of an operator? 190
10.2.4. What is contributing to positive or negative transfer when switching between machines?. 190
10.2.5. Is there positive or negative transfer due to switching tasks within a machine?. 191
10.2.6. Can the complex perceptual-motor operator skills acquired during simulator training be retained for at least a week over which there is no interaction with the simulator or related equipment? 192
10.3. Practical Implications 192
10.4. Limitations and Future Direction 194
10.4.1. Variety of Machines 194
10.4.2. Training Modules 195
10.4.3. Training Time 195
10.4.4. Transfer to Real Machine. 196
10.4.5. Research-Friendly Simulators 196
10.5. Conclusion 197

LIST OF REFERENCES 199

APPENDICES

Appendix A: NASA-TLX 220
Appendix B: Consent Form (IRB Protocol #1110011339) for Studies 1, 2 and 5 221
Appendix C: Preliminary Questionnaires for Studies 1, 2 and 5. 222
Appendix D: Study 1. 223
Appendix E: Study 2. 227
Appendix F: Consent Form (IRB Protocol #1304013518) for Studies 3 and 4 Conduced at John Deere Sites. 233
Appendix G: Preliminary Questionnaire for Studies 3 and 4 234
Appendix H: Study 4. 235
Appendix I: Study 5 253

Page

VITA. 260

PUBLICATIONS 262

LIST OF TABLES

Table		Page
2.1.	Characteristics of Different Types of Transfer	20
2.2.	Standard Instructions to Participants for Making Their Verbal Protocol Reports (adapted from Ericsson & Simon, 1993)	35
4.1.	The Simulation Modules Incorporated in the Current Version of the Hydraulic Excavator Personal Simulator, as Ordered in the Instruction Manual	45
4.2.	The Simulation Modules Incorporated in the Wheel Loader.	48
7.1.	Comparisons of Machine Constraints and Control Configurations Between Excavator and Loader	99
7.2.	The Demographic Information of the Operators.	105
7.3.	Summary of the Operators (Numbers 1–14) Reporting Comments on the HTA of Truck Loading Task on Excavator Simulator	108
7.4	Summary of the Operators (Numbers 1-14) Reporting Comments on The HTA of Truck Loading Task on Loader	110
7.5.	Responses Collected in the Post Questionnaires in Study 3	115
8.1.	The Demographic Information of the Operators in Study 4	123
8.2.	Phases for Conducting Thematic Analysis as Outlined by Braun and Clarke (2006) and Modified for Developing HTA Diagrams From Verbal Protocols	130
8.3.	Operators With the Most and Least Utterance in the Four Experimental Tasks	133

Table Page

8.4. Eight Ranked Verbal Reports With Highlighted Actions for the Filling a Trench Task. 137

8.5. An Illustration of Sorting the Different Codes Into Potential Themes From the Verbal Protocols of Filling a Trench Task 139

8.6. Responses Collected via the Post Questionnaires in Study 4 154

LIST OF FIGURES

Figure		Page
3.1.	Structure of the dissertation	39
4.1.	Setups for (a) the excavator simulator and (b) the loader simulator	42
4.2	Screen shot from the simulated Caterpillar 320CL hydraulic excavator	43
4.3.	SAE pattern of the joystick controls and pedals of the simulated hydraulic excavator	46
4.4.	Screen shot from the simulated John Deere 544K 4WD Loader	47
4.5.	Controls of the simulated loader: (a) joystick and pedals, (b) Sealed-Switch Module (SSM) controls, and (c) steering wheel	49
5.1.	Mean execution time per 30-trial module of the controls familiarization task on the excavator simulator across the three sessions for all groups	58
5.2.	Mean execution time per 30-trial module of the controls familiarization task on the excavator simulator across the three sessions for the control group	59
5.3.	Two-way interaction plot of session × block on total execution time of each module	61
6.1.	Exocentric view screen captures from the truck loading scenarios for (a) the excavator simulator and (b) the loader simulator	73
6.2.	Productivity of truck loading task on the excavator across the 5 sessions for the three experimental groups	76
6.3.	Group 1 vs. Group 2: Two-way interaction plots of (a) session × trial, (b) session × group, and (c) trial × group on productivity of each module	79

Figure Page

6.4. Group 1 vs. Group 3: Two-way interaction plots of (a) session × trial and (b) session × group on productivity of each module 82

6.5. Two-way interaction plots of (a) session × trial and (b) trial × intervening task on productivity of each module across the first three tests on excavator 84

6.6. Productivity of truck loading task on the loader across the 2 sessions (three-way interaction plot) for loader group and long-loader group . . . 86

6.7. Two-way interaction plots of session × group on mean TLX rating across the five sessions. 88

7.1. The motion constraints of (a) excavator and (b) loader 98

7.2. Preliminary HTAs for (a) excavator digging a trench and loading a truck and (b) loader transferring soil from stockpile to truck 103

7.3. Finalized HTA for excavator digging a trench and loading a truck. . . . 113

7.4. Finalized HTA for loader transferring soil from stockpile to truck. . . . 114

8.1. Screen shot from the four training modules 125

8.2. Feedback indicators shown on the screen of loader simulator 129

8.3. HTA for bucket loading module (B1) 147

8.4. HTA for filling a trench module (B2) 149

8.5. HTA for truck loading module (B3) 150

8.6. HTA for moving with a wide fork module (F) 151

9.1. Four experimental groups in Study 4 162

9.2. Productivity of truck loading task on the loader across the 2 sessions for the four experimental groups 168

9.3. Two-way interaction plot of session × method on productivity of truck loading task 170

Figure Page

9.4. Three-way interaction plot of session × trial × method on productivity of truck loading task. 171

9.5. Two-way interaction plots of session × group on mean TLX rating across the two sessions 173

9.6. Two-way interaction plots of trial × group on productivity (red line indicates the benchmark) 177

9.7. Two-way interaction plots of feedback × group on the perceived % of time spent on each type of feedback 179

9.8. The performance on the simulated excavator of all trials in two sessions (initial test and 2-week retention test) in So et al.'s (2013) study 183

ABSTRACT

So, Chung Yin. Ph.D., Purdue University, December 2014. Acquisition, Retention and Transfer of Heavy Equipment Operator Skills Through Simulator Training. Major Professors: Robert W. Proctor, Phillip S. Dunston, and Vincent Duffy.

Initiatives and collaborations among heavy construction equipment manufacturing companies and training technology firms to develop and employ simulators for varied training purposes are becoming commonplace. However, human factors research on simulator training for operators of construction equipment is still sparse. For simulator training to be effective, it is necessary to understand how skills are learned using the simulator, how those skills are transferred to other tasks, devices, and real scenarios, and how well skills are retained after simulator training.

This research is on skill development, specifically as it applies to operator training for two specific types of heavy construction equipment: excavator and wheel loader. It aims at decomposing the complexity of equipment operation and distinguishing the skills to be acquired for each machine. It consists of five studies, three conducted with students at Purdue and two with expert operators at John Deere.

Study 1 investigated whether operation of a simulated hydraulic excavator is influenced by an intervening task performed between initial practice on the excavator and a subsequent retention test using a controls familiarization task (which involves just knowing the control functions). Two intervening tasks were inserted: practicing on

a simulated loader, and reading an unrelated text intended to distract the participants. Performance on the simulator was compared against that of a group of participants who practiced on the simulated excavator throughout. The results showed no performance cost attributable to inserting practice on the simulated loader while learning the controls on the simulated excavator. The learning trends, however, prompted the question of whether the same results would bear true for learning a more complex perceptual-motor task.

Study 2 was intended to verify whether the alternating equipment sequence yields the same outcome for a more complex truck loading task that involves multiple operations. Besides the two experimental groups (control and loader groups) in Study 1, an additional group which was given practice on the two machines (but with a different practice schedule from the original loader group) was added to address the question of whether the duration of practice on an alternative machine affects performance on the previously learned machine. The number of sessions was also increased, from three to five, to examine the possible influence when participants continue to switch between the machines. Those participants who engaged in intervening practice on the simulated loader showed a smaller performance improvement on learning the truck loading task on the simulated excavator than did the control group who practiced on the simulated excavator for all five sessions. This outcome confirms that the controls familiarization tasks on both machines studied in the preliminary study may have been too simple for the full effects of switching between the machines to be evident. This finding of continued skill improvement upon return to the previously practiced machine inspires consideration of concurrent the previously practiced machine inspires consideration of concurrent simulator-based training rather than the practice of learning to operate only one machine at a time.

Study 3 analyzed skill transfer using hierarchical task analysis (HTA) to investigate the degree of overlap in specific task components by studying the similarity and dissimilarity of the truck loading task performed in Study 2 on excavator and wheel loader simulators. After the modification and verification by operators of the initial HTAs, the finalized HTAs revealed that the lack of positive transfer found in performing the truck loading task alternately with the excavator and loader was likely due to the differences between loader and excavator in terms of the controls, physical constraints, and the explicit goals and subgoals of the task. In addition, comparing the number of levels of subgoals of HTAs did not evidence any level-of-difficulty differences between tasks.

Studies 4 and 5 investigated whether there is a cost when switching between different types of training modules within the same machine. Study 4 was conducted with experienced operators, who provided information on how the four selected tasks on the loader should be performed and classified the perceived difficulty level of each. Verbal protocol analysis was used to decompose the tasks of the four training modules on the loader simulator: 1) Simple Bucket Loading (B1), Filling a Trench (B2), Truck Loading (B3), and Fork Lifting (F). A nine phase, systematic method for deriving the HTAs from the think-aloud protocols was also developed in this study, which successfully generated the four HTAs. The findings show that 1) the HTA of the Fork Lifting module is significantly different than those of the three bucket loading tasks, and 2) although all three bucket loading tasks shared a similar mechanism, the operators ranked B1 as the easiest, followed by B2 and then B3 due to the corresponding accessibility of the dump targets, and fork lifting was ranked as the most difficult task. The results were used to justify the hypotheses for Study 5.

Study 5 sought to verify whether an alternating practice sequence within the same machine, i.e. training with an alternative tool (a wide fork) and returning to the original learned tool (a bucket) on a loader simulator, yields better skill transfer and retention (after a one-week interval). Four groups of undergraduate students were tested. Two groups were given two tasks involving bucket loading to practice in the first two sessions, whereas the other two groups were given a bucket loading task in the first session and the fork lifting task in the second session. The transfer and retention tasks both involved a bucket loading task that had not been performed in Sessions 1 and 2. The results showed that the groups who were assigned to practice on two tasks involving the manipulation of buckets performed better in the skill transfer test when the new task was introduced that also involved manipulation of the bucket. The results support the *specificity of training principle* (for which the practice conditions match the test conditions and thus facilitate retention or transfer) but not the *progressive difficulty training principle* (for which difficulty impedes performance in the learning stage but facilitates retention). It is suggested that, when training perceptual-motor tasks, tasks practiced during the learning phase should match the transfer task. Manipulation of task difficulty may play a role only if the tasks share task-relevant cognitive processes and mental models.

The overall findings of this research provide: 1) better understanding of skill development for the operation of construction equipment, and 2) evidence as to how the trainees can better utilize their time when training on a single machine and concurrently on multiple machines. The findings add to the general body of knowledge on perceptual-motor skill acquisition and to that on training in a specific domain via a specific technology. The findings are expected to generalize to heavy equipment training in related domains, such as forestry and mining, and domains requiring instrument handling skills and robotic arms, such as surgery and orbital space vessel external operations.

CHAPTER 1. OBJECTIVES AND SIGNIFICANCE

Virtual-reality (VR) simulators allow cost-effective, safe, and efficient training of operators in risk-free environments by real eliminating fuel costs, equipment damage, and emissions. With the increasing quality of three-dimensional graphics and decreasing costs of personal computers, it has become possible to employ affordable simulator-based training more widely than was the case just over a decade ago. The use of simulator training is therefore appealing across many industries, including aviation, mining, rail and power. Heavy construction equipment simulators have been used customarily to provide an alternative to a portion of the field training that involves costly, logistically difficult and hazardous tasks (Dunston, Proctor & Wang, 2014). Nowadays, commercially available training simulators of construction equipment are modeled after specific models of real equipment, and the equipment manufacturers promote these simulators, which feature different lessons and tasks intended to develop skills in basic machine controls, proper operator technique, and safe job site operation.

Although industry training programs employ established curricula that introduce equipment functions and typical task objectives, there is no firm evidence that these curricula are informed by a systematic scientific analysis of the tasks performed by operators. VR-based training systems in construction have increasingly received research attention (e.g. Dunston et al., 2014; Tichon and Diver, 2012; Wang &

Dunston, 2005), yet scant confirmation of the principles and standard curricula for efficient use of construction equipment operator-training systems is found in the literature. By interviewing trainers and course managers, Tichon and Diver (2012) studied the usability and usefulness of integrating simulator training into an existing civil construction training program for helping disadvantaged job-seekers become ‘job ready’. The study reported numerous advantages, including the possibility of providing immediate expert feedback, the opportunity to practice dangerous or potentially costly conditions without tying up real machines, and the ability to learn from one’s own mistakes. Since their evaluation only reports subjective feedback from the trainers and course managers, the effectiveness of simulator training within computer-generated civil construction sites has yet to receive a thorough, objective testing. It is also notable that much of the research focus has been on the technical aspects of prototype systems (e.g., Dopico, Luaces, & González, 2010; Torres 2004; 2005), with only a few studies conducted on learning or transfer of training for construction equipment (e.g., Hildreth & Heggestad, 2010; Hildreth & Stec, 2009; Visser, Tichon, & Diver, 2012; So, Proctor, Dunston, & Wang, 2013).

Design of effective training programs requires understanding the tasks performed by operators and the required skills. Construction-equipment operation entails navigating and maneuvering vehicles, and also cutting, moving, and processing material. A skilled construction operator must have a thorough understanding of multiple machines’ capabilities, the principles behind their operation, and countless of hours of practice (Ober, 2010). Thus, it is crucial to determine effective training for these various machines as well as whether and how skills at operating one machine may transfer to the others. For the present project, experiments were designed to provide to address the following research questions: 1) How much does training on one machine transfer (positively or negatively) to other machines? 2) Does insertion of training on various machines facilitate (or inhibit) learning and retention on a previously practiced machine? 3) When should an alternative machine be introduced in the training if skills on multiple machines are required of an operator? 4) What is contributed to positive or negative transfer when switching between machines? 5) Is there positive or negative transfer due to switching tasks within a machine? 6) Can the complex perceptual-motor operator skills acquired during simulator training be retained for at least a week over which there is no interaction with the simulator or related equipment?

Overall, the findings of this research were expected to provide: 1) better understanding of skill development for the operation of construction equipment, and 2) evidence as to how the trainees may better spend their practice time for (a) single machine and (b) multiple machines training. The findings of this research add to a general body of knowledge (i.e., perceptual-motor skill acquisition) as well as to the body of practice for training in a specific domain via a specific technology (i.e., VR-based simulators for training construction equipment operators). The findings are expected to generalize to heavy equipment training in related domains (such as forestry and mining) and domains requiring instrument handling skills (such as surgery, dentistry, and orbital space vessel external operations).

This dissertation is organized into the following sections. Chapter 2 provides review of research related to skill development and transfer, task switching paradigm, training principles, VR training applications, hierarchical task analysis, verbal protocol analysis and cognitive workload. Chapter 3 presents the research framework and the goals of each study. Chapter 4 describes the two simulators employed in this research. Chapters 5 through 9 present the details and results of five studies devised to address the questions posed above. Chapter 10 concludes with the explanation of contributions and final remarks related to this research.

CHAPTER 2. LITERATURE REVIEW

Versatility is required for operators of specialized equipment. For example, pilots are expected to perform as well or better when they return to a particular model of aircraft after flying a second one (Lyall & Wickens, 2005). Likewise, skilled operators of heavy construction equipment may become proficient at operating several machine types, such as excavators, loaders, graders, and dozers, and be able to switch between them (Dunston et al., 2014). Since practice to obtain such skills is both time-consuming and costly, it is therefore essential to determine effective training for these machines, as well as whether skills at operating one machine type transfer to the others. A question of interest is whether the similarities and differences promote transfer (positive or negative) and retention as an individual moves from practice on one machine or task to another and back again. Regarding simulator training, it is important to understand how skills are learned using the simulator, how skills are transferred to other tasks, devices, situations or real scenarios and how much skill is retained after simulator training.

Research has found that practice schedules on motor control tasks may differentially influence performance and learning (e.g., Baddeley & Longman, 1978; Schmidt & Lee, 2011). Also, the task-switching phenomena observed in cognitive tasks (e.g., Chamberland & Tremblay, 2011; Waszak, Hommel, & Allport, 2003) alone may not be enough to explain the task-switching implications in the motor control domain. Whether switching between two complex perceptual-motor tasks with different task sets is always detrimental to speed and accuracy and may lead to a switch cost is questionable and worth investigation. Most experiments have "task switching" referenced only as switching on a trial-to-trial basis (e.g., Meiran, 1996), which is different from the current interest of this research, in which novices were given a task to practice for a few trials before switching to another task or machine. Indeed, alternating task practice in motor performance can be introduced through practice schedule manipulation to create different task-switching demands across experimental conditions. Practice schedule (e.g., blocked or mixed), motor learning schedule (e.g., massed or distributed practice), contextual interference, skill transfer and retention theory and training principles such as task difficulty, variability of practice and specificity of practice, may be considered when it comes to motor control task switching. Thus, some empirically valid principles of training identified by Healy and Bourne (2012) are also discussed here.

2.1. Task Switching

Task-switching paradigms for revealing cognitive processes and mental resources involved in decision making or allocating attention have been investigated widely over the last three decades (e.g., Chamberland & Tremblay, 2011; Sohn & Carlson, 2000). Some seminal studies explained switch costs in terms of the anticipatory components of executive mental control (Jersild, 1927; Spector & Biederman, 1976). In contrast, Allport et al. (1994) posited that switch costs originate from task-set inertia, relating to the proactive interference between conflicting stimulus-response mappings for successive tasks. In choice-reaction task-switching experiments in the cognitive domain, Monsell (2003) suggested that switch costs have been explained by the need to reconfigure cognitive processes during each decision-making process. Other researchers (e.g., Logan & Gordon, 2001; Rogers & Monsell, 1995) pointed out that the costs reflect the time necessary for task-set reconfiguration, and still others argued that neither task-set intertia nor reconfiguration alone best explains the switch cost phenomenon (e.g. Ruthruff, Remington, & Johnston, 2001).

Two types of switch costs—local and global—are typically studied. Local switch costs refer to the RT difference between switch and nonswitch trials within mixed blocks (Meiran, 1996; Rogers & Monsell, 1995). Local switch costs are thought to require executive processes to deactivate the task set relevant on the previous trial and to activate the currently relevant task set (Monsell, 2003). Global switch costs refer to the RT difference between nonswitch trials in a condition in which only a single task is performed and a condition in which subjects alternate between two different tasks (i.e., a mixed block) and are thought to measure the set-up cost associated with maintaining and scheduling two mental task sets, as well as the added load associated with maintaining multiple task sets in working memory (Kray & Lindenberger, 2000).

2.2.1 Cognitive Task-Switching

In a typical task-switching experiment, subjects are asked to make a decision about a stimulus that requires two alternative mental computations. Jersild's (1927) study was to alternate between arithmetic tasks by asking the students to add or subtract a number and then report the sum or difference verbally. More recent studies examined task switching on memory tasks, such as verbal categorization tasks to judge whether the words rhymed or not; spatial categorization tasks to judge whether the two patterns were identical; and spatial categorization tasks to report the order of the items (e.g., letters, dots) presented in a sequence (Chamberland & Tremblay, 2011). Other studies include number comparisons and tone discriminations (Sigman & Dehaene, 2006), picture naming and word reading (Waszak, Hommel, & Allport, 2003), and color naming and word reading for Stroop color-word stimuli (Gilbert & Shallice, 2002). General findings show that switching tasks leads to slower response times and more errors than performing a single task repeatedly.

Monsell (2003) identified the following four basic phenomena of task switching:

1. *Switch cost*: A longer time is needed to initiate a response on a 'switch trial' than on a 'non-switch' or task-repetition trial, often by a substantial amount.
2. *Preparation effect*: If advance knowledge of the upcoming task is provided allowing preparation time for it, the average switch cost is usually reduced.
3. *Residual cost*: Although preparation may reduce a switch cost, a further increase in the preparation interval does not further reduce the time cost of a switch. Such "residual" cost is resistant to be eliminated by the further lengthening of the preparation interval (e.g., Nieuwenhuis & Monsell, 2002).

4. *Mixing cost*: Although performance recovers quickly after a switch, but responses remain slower than performing the same task throughout the block.

Chamberland and Tremblay (2011) attempted to investigate the extent to which the cost of switching between tasks is universal in cognitive tasks by exploring the differential impact of two types of switches: switching by processes (categorization and serial memory) and switching by content (verbal and spatial target stimuli). Their results revealed that high-level cognitive activities such as serial memory might not be negatively affected by task switching. Indeed, if serial memory is involved, shifting to another task, to some extent, may be more beneficial than just performing on the same task.

2.2.2 Motor Control Task-Switching

Task-switching experiments have been mainly focused on cognitive tasks and have not made connections to complex motor tasks until recent years (Bernardin & Mason, 2011). Most of such studies are found in bimanual coordination tasks. For example, Bernardin and Mason (2011) conducted a bimanual coordination task-switching study to investigate the consequences of an unexpected environmental perturbation on reaction time and movement time. They tied their results to the perturbation paradigm, which requires subjects to reorganize their movements in mid-execution due to a size or location change of the target object (e.g. Mason 2008). The most robust finding in bimanual coordination tasks revealed that mirror-symmetric (in-phase) bimanual movements usually resulted in higher accuracy and consistency than nonsymmetric (anti-phase) movements (e.g., Donchin & Cardoso De Oliveira, 2004; Swinnen & Wenderoth, 2004).

Most of the studies so far are still limited to simple motor tasks that involve only reaching, grasping, tapping, etc. In addition, past research incorporated only a short experimental time, i.e., less than a second or only a few seconds per trial. It becomes especially challenging to study the switch cost when executing complex perceptual-motor tasks, which usually involve multiple goals and require more than one set of motor skills and decision making, and which may also take an appreciably longer amount of time.

2.2.3 Does the Task-Switching Cost in Cognitive and Simple Motor Tasks Hold for Complex Perceptual Motor Tasks?

To answer this question, the difference between cognitive and perceptual-motor skills first needs to be understood. Cognitive skills are used in problem solving for intellectual tasks, where a subject's knowledge is more critical to success than their physical prowess (VanLehn, 1996); thus subject's prior knowledge plays a role in the learning of a cognitive skill. Perceptual-motor skills rely on hand-eye coordination, analytical reasoning, working memory abilities and practice (e.g., Rosenbaum, 2001).

Although most task-switching studies involve cognitive tasks, activities in the natural world often involve a task switch that requires motor execution. The operation of construction equipment requires performance of tasks needing complex perceptual-motor skills that are known to improve over years of practice. For example, truck loading from stockpiled aggregates requires multiple skill sets where the operator needs to repeat the steps of driving to the aggregate pile, loading the bucket, driving out from the pile and toward the truck, and releasing the bucket to dump the aggregates into the truck. It has been well established in the verbal domain that task switching slows down cognitive operations related to decision-making and stimulus categorization and increases errors; it seems probable to assume that switching from one complex movement task to another may also cause a switch cost. However, the task switching literature may not be the best fit to explain the effects of skill acquisition of complex perceptual motor skills for the following reasons:

1. Reaction time is often short (measured in milliseconds).
2. Most experiments have investigated task-switching between trials, and switch costs focus on switching on a trial-to-trial basis.
3. Motor switching tasks are limited to simple motor execution.

It is questionable whether current findings concerning the task-switching cost in both cognitive tasks and simple motor tasks may apply in the same manner to complex perceptual-motor tasks for the following reasons: Complex perceptual-motor tasks involve a much longer response time measured in minutes; the tasks may be altered by sessions or blocks, but not necessarily trial-to-trial; and the tasks may involve multiple goals requiring more complex cognitive and motor skills. Whether the current findings concerning the task-switching cost in both cognitive tasks and simple motor tasks may apply in the same manner to complex perceptual-motor tasks is still unclear and worth investigation.

2.2. Acquisition, Retention, and Transfer of Training

There are three fundamental cognitive components of training: acquisition, retention and transfer, and three corresponding goals: efficiency, durability, and generalizability (Healy & Bourne, 2012; Healy, Kole, & Bourne, 2014). Critical questions relating to effective training are 1) how much time and effort are required to achieve a criterion of performance, 2) how can transfer of training to related equipment and tasks be ensured, and 3) what training methods promote retention of the trained knowledge and skills during periods of disuse?

Acquisition refers to acquiring new knowledge and skill, depends upon repeated exposure to and practice of the knowledge and skills to be learned. Group curves for skill acquisition typically approximate a “power law of practice” (Newell & Rosenbloom, 1981). This law formalizes the relationship between trials of practice and time to make a correct response as a power function,

$$R = aN^{-b},$$

where R is response time on trial N, a is response time on trial 1, and b is the rate of change.

Retention refers to the decline in performance or failure to retain information over time, sometimes without opportunity to rehearse or refresh acquired knowledge or skills. The relationship between response time and retention interval has been expressed as a power law by Wickelgren (1974),

$$R = d + fT^{-g},$$

where R is response time, T is the retention interval, d is the criterion of original learning, f is a scaling parameter, and g is the rate of forgetting. Later it was also named 'the power law of forgetting' as the inverse of the power law of acquisition (Rubin & Wenzel, 1996; Wixted & Carpenter, 2007).

Transfer refers to the acquisition of one task affecting performance on another. The effect of training on one task can be either positive (facilitation) or negative (interference) on performance of another task (Taylor et al., 2007). More discussion of transfer is presented in Sections 2.4 and 2.5.

2.3. Training Principles

Healy, Schneider, and Bourne (2012) have identified several positive elements to promote skill retention and transfer, including applying deliberate practice, using distributed practice, employing a mixed practice schedule, adding sources of contextual interference, introducing an external focus of attention, applying errorless learning, introducing task difficulty and increasing the variability of practice. They reviewed and organized the principles of training into three categories based primarily on underlying cognitive processes and training requirement:

1. Principles relating to Resource and Effort Allocation: The learner is required to allocate cognitive resources and effort to acquire specific aspects of the knowledge or skills.
2. Principles relating to Context Effects: The knowledge and skills acquired are bound (context specific), to some degree, to the circumstances in which they are acquired.

3. Principles relating to Task Parameters: training can vary by manipulating different task dimensions such as spacing, feedback, task difficulty.

Some principles relating to 2) context effects and 3) task parameters relevant to the research interest here are reviewed and discussed below.

2.3.1. Specificity of Training

Specificity of training holds when the conditions of practice match the conditions of retention or transfer. The implication is that the conditions of practice should closely match performance to optimize transfer. The theoretical explanation of this principle originated in Thorndike's "*identical elements theory*" (Thorndike & Woodworth, 1901), where they explored how learning was transferred in one context to another context that shared similar characteristics in tasks involving perception and memorization. They proposed that transfer of learning depends on the proportion to which the learning task and the transfer task are similar. The commonality of most transfer theories is advocating that transfer of training is proportional to the similarity between any two tasks (Pavlov, 1935/1955; Henry, 1958).

In verbal learning, for example, Osgood (1949) proposed a model for meaningful similarity and focused interest on its relation to direction and amount of transfer produced. He also proposed three "empirical laws" to account for all transfer phenomena in both serial and paired-associate learning tasks:

1. When stimuli are identical and response similarity is varied, the amount of negative transfer will decrease as response similarity increases.
2. When responses on two lists are identical and stimulus similarity is varied, positive transfer increases as stimulus similarity increases.
3. When both stimulus and response similarity are varied simultaneously, negative transfer will increase as stimulus similarity increases.

The specificity of training principle can also be explained in terms of the *procedural reinstatement principle* (Lohse & Healy, 2012). According to procedural reinstatement principle, when the mental procedures that are acquired during learning can be used during testing, such duplication of test procedures facilitates retention and transfer (e.g. Healy, Wohldmann, & Bourne, 2005). Healy et al. (2012) found that this principle is similar to the following principles that were derived primarily from studies of list learning:

1. *Encoding specificity* (e.g., Tulving & Thomson, 1973): When retrieval cues elicit the original encoding operations, the memory for information is optimized.
2. *Transfer-appropriate processing* (e.g., Morris, Bransford, & Franks, 1977): When the test evokes the procedures used during prior learning, the memory performance is optimized.

3. *Context-dependent memory* (e.g., Kole, Healy, Fierman, & Bourne, 2010): Being tested with a new context other than that tested in the original learned context, the memory for information is worse.

A general conclusion from this procedural reinstatement principle is that specificity occurs when training tasks are based primarily on procedural information, or skill, whereas, generality occurs when training tasks are based primarily on declarative information, or facts (Healy, 2007). Alternatively speaking, retention is strong but transfer is limited for skills (procedural information) learning whereas, retention is poor but transfer is robust for facts (declarative information) learning (Healy et al., 2012).

2.3.2. Contextual Interference

Contextual Interference (CI) refers to the resulting interference when performing different variations of a skill in a practice environment (Magill & Hall, 1990). Such effects have been found in verbal skills (Battig, 1979), motor skills (Shea & Morgan, 1979), and logical rules (Schneider, Healy, Ericsson, & Bourne, 1995). The CI effect can be manipulated by how a practice session is organized. For example, blocked and random schedules are the two most commonly studied practice structures. A blocked practice schedule consists of performing the same task until all of the trials for that particular task are completed before switching to the next task, whereas, a mixed practice schedule frequently changes from one task to another, such that immediate repetitions of any single task are infrequent (Schmidt & Lee, 2011). Typically, participants practicing with a blocked schedule exhibit better performance during acquisition (initial practice) compared to those who practice with a mixed schedule (e.g. Lee & Simon, 2004; Shea & Morgan, 1979). But, in most cases the mixed-practice schedule elicits better performance on a retention or transfer test, and thus better learning, than the blocked schedule (e.g., Battig, 1979; Lee & Magill, 1983).

Task switching in motor performance can be introduced through practice schedule manipulation to create different task-switching demands across experimental conditions. Studies above have shown that different practice schedules for motor control tasks may differentially influence performance and learning; whether switching between two complex perceptual-motor tasks with different task sets is always detrimental to speed and accuracy is worth further investigation.

2.3.3. Task Difficulty

The degree of the contextual interference effect could be a function of the difficulty of the task as noted by Battig (1979), where a greater level of item or task difficulty could produce greater amounts of processing (i.e. contextual interference). One question arises as to which stimulus set (the easier or more difficult one) in a cognitive task should be trained first. Pellegrino et al. (1991) found that initial training on a difficult subset of stimuli was beneficial relative to initial training on an easy subset of the stimuli in a visual discrimination task. Research has suggested that manipulation of task difficulty during training may have facilitating effects during retention and transfer testing (e.g., Schneider, Healy, & Bourne, 2002). However, others have noted that training the difficult task first does not necessarily yield the optimal strategic skills. For example, in a Morse code reception task, participants should be given initial training on easy stimuli, which allowed participants to adopt a more effective unitization strategy for representing codes. For motor skills, Maxwell, Masters, Kerr, and Weedon (2001) introduced 'errorless learning' in a golf putting task in which participants begin with the easiest task (where fewer or no errors are made) and move on to more difficult tasks.

For some complex skills, it is not appropriate or possible to start training at the full complexity level of the transfer task. For example, learning to fly a plane requires understanding of the controls and their functions, the mechanics of the plane, safety violations and the concept of air dynamics. Thus, a strategy is to start with a simple version of the task and gradually increase its difficulty as learning progresses, a technique called 'simplification' by Wightman and Lintern (1985). Briggs and Naylor (1962) examined this technique training flight dynamics in aircraft control using a three-dimensional compensatory tracking system. They concluded that progressive-part training (practice trials on separate dimensions followed by each of the three possible pairings) will be superior to pure-part (involving sessions on each of three separate dimensions) and simplified-whole (from easy to hard) for the acquisition of skill in a complex, multidimensional task, since the progressive-part method utilizes a training task of high similarity to the transfer task and also provides an opportunity to develop efficient timesharing behavior.

The overall success of *progressive difficulty training* compared to training that initiates training the task at its full difficulty level is conclusively established (Healy & Bourne, 2012; Wickens, Hutchins, Carolan, & Cumming, 2011), where conditions which cause difficulty during learning facilitates and enhances later retention and transfer. However, not all sources of difficulty are desirable. Some researchers argued that introducing difficulties during training is facilitative only when the training and retention tasks share task-relevant cognitive processes (McDaniel & Butler, 2011; McDaniel & Einstein, 2005). For example, in a memory task, memory performance will be enhanced when the processes engaged in the initial learning match the processes of the critical task.

2.3.4. Variability of Practice

The principle of variability of practice predicts that training individuals on several tasks (variable practice) often yields better performance on a transfer test than does training individuals on a single task (constant practice). The benefit of variability of practice was first explained by Schmidt (1975)'s Schema theory for discrete motor tasks. In Schmidt's schema theory, *schemas* are generalized rules that generate the spatial and temporal muscle patterns to produce a specified movement within one movement class. Thus, increasing variability of practice on a particular task builds a more effective generalized motor program which could produce similar but different movement. These findings were also found in both motor tasks (e.g., Schmidt, 1975) and non-motor tasks (e.g., Goode, Geraci & Roediger, 2008). However, not all forms of variable practice are effective. For example, Wohldmann, Healy, and Bourne (2008) suggested that varying task parameters within a single motor program enhances transfer, but varying the motor programs themselves has no benefit. In their study, they found that practicing to move a single defective mouse to a variety of targets would enhance transfer to moving that same mouse to new targets.

2.4. Transfer Taxonomy

Transfer refers to an influence of prior knowledge and skills gained in earlier settings on learning and performance in other newly encountered settings. That is, knowledge and skills are passed on from one domain or task to another. To delineate transfer, different taxonomies have emerged (see Table 2.1), concerned with distinguishing different types of transfer. Barnett and Ceci (2002) suggested that the content of transfer (i.e., what is transferred) can be decomposed into three dimensions: (a) the specificity–generality of the learned skill, (b) the nature of the performance change assessed, and (c) the memory demands of the transfer task. The latter factor both captures and extends the near-versus-far-transfer distinction.

Table 2.1. Characteristics of Different Types of Transfer

Type	Characteristics	References
Near	Overlap between situations, i.e. transfer to a more similar context.	Barnett & Ceci (2002)
Far	Little overlap between situations, i.e. transfer to a less similar or dissimilar context.	
Positive	Previously learned information facilitates performance of the new task.	Smode, Beam, & Dunlap, (1959); Cree & Macaulay (2000).
Negative	Previously learned information impedes the recall of previously learned information.	
Vertical	Previously learned knowledge is essential to acquire new knowledge.	Ormrod (2004); Singley & Anderson (1989)
Horizontal	Previously learned knowledge is not essential but helpful to learn new knowledge.	

According to Valverde (1973), transfer may occur when two activities are similar, either in substance or procedure. Anything which the trainee can learn can be transferred, including skills, facts, learning sets, self-confidence, interests, and attitudes.

Transfer may be specific, as when elements of one learning situation occur in identical or similar form in another. In general, trainers desire positive transfer between contexts and tasks to occur and not negative transfer. For example, pilots may be required to switch from flying one aircraft in a mixed fleet to flying another, for which the control-display configurations differ, and then switch back again (Lyall & Wickens, 2005). The pilots are expected to perform as well or better when they return to the first aircraft after flying the second. Likewise, the more experienced operators of heavy construction equipment must become proficient at operating several machine types, including excavators, loaders, graders, and dozers, and may be called upon to switch between them. Conventional training occurs with one equipment type at a time, but VR simulators enable concurrent training. It is crucial, thus, to determine effective training for these machines, as well as whether skills at operating one machine transfer positively or negatively to the others.

Overall, transfer research has attracted much attention in various domains since the beginning of the 20th century and many studies with empirical findings and theoretical interpretations have continued to be conducted in the fields of education and pedagogy (Bransford, Brown, & Cocking, 1999; Soini, 1999), linguistics (e.g., Jiang & Kuehn, 2001; Kecskes & Papp, 2000; Odlin, 1989) and VR training (e.g. Boyle & Lee, 2010; Lehmann, 2005; Valverde, 1973). A review on VR training and transfer is presented in the next section.

2.5. Transfer of VR Training

The value of any training medium depends upon how effectively transfer of training is achieved from the training device to the operational task. For example, consideration should be given to the extent specific flying tasks should be trained in the decision of employing a flight simulator for a pilot training program (Valverde, 1973). The groundbreaking development of VR allows users to participate in a virtual world reproduced by readily available computers, enabling safe, convenient, and planned repetitive training. Training simulators, in general, consist of basic functions of the controls, virtual reality content representing realistic situations, virtual reality interface devices, and the capability of monitoring and reporting the practice results. Much human-factors research has been conducted on simulator training relating to fidelity of flight simulators and design of effective training routines (Koonce & Bramble, 1998), fidelity of driving simulators (Boyle & Lee, 2010), sports expertise (Beauchamp, Harvey, & Beauchamp, 2012; Williams & Ward, 2003), industrial tasks (Duffy, Ng, & Ramakrishnan, 2004; Lin, Duffy, Yu, & Su, 2002) and surgical procedures (Lehmann, 2005; Tan & Sarker, 2011). Research on simulator training on construction equipment is sparse. In this section, research on flight simulators, surgical simulators and construction equipment is reviewed, as these skill domains all involve the complex manual operation of equipment that may be classified as instruments or tools.

2.5.1. Surgical Simulators

Sutherland et al. (2006) categorized 30 studies into four categories of simulation (computer, video, model, and cadaver) and compared them with no training and standard training. They concluded that none of the methods of simulated training has yet been shown to be better than other forms of surgical training. Some studies have proven learning curves and training improvement with simulators (e.g., Seymour et al., 2002; Grantcharov et al., 2003). However, the studies trying to address the important question whether skills acquired during simulator training can be transferred to a real situation do not provide uniform results (e.g. Ahlberg et al., 2002; Torkington et al., 2001).

One of the most common examples of simulated medical training is laparoscopic simulation-based training. For example, Lehmann (2005) investigated the transfer of basic psychomotor skills from VEST to conventional video training. The results demonstrated that skills can principally be transferred from one device to the other and there is an adaptation period when switching to the new device.

2.5.2. Flight Simulators

Cumulative research has shown that the use of flight simulators combined with aircraft training produces more performance improvements in real aircraft than aircraft-only training (e.g., Jacobs, Prince, Hays, & Salas, 1990; Orlansky & String, 1977; Pfeiffer & Horey, 1987). It has been reported that motion feedback improves in-simulator flight performance and increases the realism of pilot behavior and performance (e.g., Bürki-Cohen, Soja, & Longridge, 1998; Pool, Mulder, Van Paassen, & Van der Vaart, 2008). However, some researchers have argued that motion feedback does not imply improved learning, as humans are well able to integrate the available information to maximize their performance. For example, Martin (1985) showed that the use of direct concurrent motion stimuli—a tactual seat pan by providing motion cues with tactile pressure to the buttock and upper thigh areas—aids the pilot in the simulator by providing additional information during the simulator training, but the way these stimuli are perceived and processed by the pilot does not necessarily correspond to real flight (Gundry, 1976). In fact, Schmidt and Wulf (1997) found that augmented feedback that enhances performance during training can interfere with performance in a transfer condition, because the learner has become reliant on the supplementary information.

The transfer of training paradigm is probably the most valid means of investigating the training effectiveness of motion (Advisory Group for Aerospace Research and Development, 1980). Two types of transfer of training motion experiments can be distinguished for flight simulator training, true and quasi-transfer. In a true transfer experiment (i.e., simulator-to-real machine transfer), a group of participants is exposed to simulator training with motion. A second group is exposed to the same training without motion. After training, the performance of both groups is evaluated in a real aircraft. A positive training effect of motion is confirmed when the motion-trained group performs better in the aircraft than the no-motion-trained group. Quasi-transfer of training (i.e., simulator-to-simulator transfer) follows the same procedure as true transfer, except the transfer session is conducted in the simulator. A quasi-transfer design has been advocated because it avoids the cost, hazard, and scheduling hindrances (e.g., bad weather) of true transfer and offers the possibility of testing dangerous disturbances such as engine failures (Caro, 1976; Taylor, Lintern, & Koonce, 1993).

2.5.3. Construction Equipment Simulators

The effectiveness of simulation in training construction equipment operators has been documented in the literature for the last decade, but most of the research has focused on technical aspects of prototype systems rather than on learning or transfer of training. For example, Torres (2004; 2005) developed a haptic interface-based simulator of a semiautomatic hydraulic excavator 2D arm in a virtual environment. Dopico et al. (2010) have applied real-time simulation techniques from multibody system dynamics to develop a full 3D physics-based excavator simulator which could deliver realistic real-time behavior and simulate common scenarios for real excavators: slipping on slope terrains, stabilizing the machine with the blade or the outriggers, using the arm for support or impulsion to avoid obstacles, etc. Kamezaki, Iwata, and Sugano (2008) quantified the effect of simulation training for operators of double front work machines and found substantial improvements in task completion time and positional accuracy. Later, they proposed a new conceptual design for an operator support system and evaluated it using their newly developed simulator (Kamezaki, Iwata, & Sugano, 2009a; 2009b). Their experimental results showed that the support system improves the work performance, including decreasing the operational time for completing a task, the number of operation errors, and the mental workload for the operators.

Research has been conducted on examining the effectiveness of simulators with motion and zero-motion platforms. Hildreth and Stec (2009) sought to verify skill development and transfer from motion and zero-motion wheel loader simulators. They compared anxiety levels with those experienced with training on real equipment. They measured the loading cycle time and production rate as well as levels of operator confidence and anxiety before and after training. No statistically significant difference was found between on-machine and simulation-based training, but among the two simulation types, full-motion simulation-based training was found to increase production rate and confidence, while decreasing cycle time and anxiety. Hildreth and

Heggestad (2010) examined the rate at which skills are developed, the degree to which simulator skills transfer to actual equipment, and the degree to which operator anxiety when operating the physical equipment is decreased. They reported no statistically significant difference in operator performance and anxiety level between those trained using full motion simulation and those trained using static simulation. They argued that while 20 minutes of simulation training was sufficient time to become familiar with the controls and operation, it is not a sufficient amount of training to produce a field ready operator. This short training duration did not progress trainees beyond developing fundamental skills.

Current commercially available training simulators of construction equipment are modeled to feature different lessons and tasks intended to develop skills in basic machine controls, proper operator technique, and safe job site operation. Some, but still little, was found on the effectiveness of the training modules offered in virtual training systems for construction equipment training. Bhalerao (2009) focused on basic control familiarization with a comparison between explicit classroom instruction on control functions and hands-on exploration on a computer-based Virtual Reality excavator simulator and concluded that incorporation of the classroom instructional session is more efficient with regard to learning the basic controls. Following that line, Su, Dunston, Proctor, and Wang (2013) investigated the effect of training practice schedule and contextual interference on construction equipment operating skill development through a VR excavator simulator and concluded that a mixed practice schedule and a blocked practice schedule of coordination skills for training made no difference with regard to training efficiency and the trainee's confidence level. The findings suggested that there is a need to understand the task complexity and task difficulty for construction equipment training prior to designing task schedules. Consequently, more in the way of systematic experimentation on use of these virtual-reality systems is needed to demonstrate what factors affect acquisition and retention of skills as well as transfer of those skills to operation of real equipment.

Another study investigated whether training on one control configuration will transfer, positively or negatively, to another configuration (Lopez-Santamaria, 2011; Proctor et al., 2013). In this experiment, transfer between standard control configurations of a hydraulic excavator and a backhoe, both controlled by joysticks operated with the left and right hands but with different control mappings, was studied. Participants performed two sessions on the simulator, being divided randomly into four groups that differed in terms of which sequence of control configurations was used for the two sessions. Two groups practiced on the same control configuration for both sessions (either the hydraulic excavator or the backhoe loader), whereas the other two practiced one control configuration (hydraulic excavator or backhoe loader) in the first session and switched to the alternative configuration in the second session. The main result was that the switch in control configuration affected performance in general, but the enduring costs were not large.

I and my colleagues have evaluated part-task training in comparison to whole-task training to determine whether this approach accomplishes its goal of more effective training (So et al., 2013). In particular, the study examined whether part-task training produces better learning and retention than whole-task training of a trench-and-load task performed on the hydraulic excavator simulator (using the Society of

Automotive Engineers [SAE] excavator control configuration). The trench-and-load task requires the operator to perform three relatively distinct subtasks: (a) position the excavator between a dump truck and trenching area; (b) dig soil from the trench; (c) dump the soil into the truck. These task components were performed in sequence, enabling comparison of part-task training on the components to whole-task training. The experiment involved three phases: training, immediate test, and retention test (return in 2 weeks). Participants were randomly assigned to one of the two training-method groups: part task and whole task. The results showed that the part-task group began with a lower production rate than the whole-task group, which is to be expected since the whole-task group already had practiced the complete task (though with different scenarios) in the practice session. By the end of the first session, though, the production rates of the two groups did not differ. On returning two weeks later, both groups showed an initial dip in production rate, but with the exception of the first trial, the performance curves trended as if they were continuations of those from the immediate test. The part-task group obtained higher productivity rates than the whole-task group in the retention test. The benefit of part-task training for better retention was found.

In summary, VR-based training systems in construction have increasingly received research attention in the last decade (e.g., Dunston et al., 2014; Wang & Dunston, 2005), yet there is meager confirmation of the training principles and standard curricula for efficient use of construction equipment operator-training systems in the literature. Although industry training programs employ established curricula that introduce equipment functions and typical task objectives, there is no published firm evidence that these curricula are informed by a systematic analysis of the tasks performed by operators. It is also notable that the heavier focus of research on construction equipment operator-training systems has been on the technical aspects of prototype systems, (e.g., Dopico et al., 2010), with only a few studies conducted on learning or transfer of training for construction equipment (e.g., Hildreth & Heggestad, 2010, So et al., 2013; Su et al. 2013).

2.6. Methodology for Collecting Qualitative Data

In addition to obtaining performance measures, it is common in human factors studies to obtain subjective measures of workload using the NASA-TLX (Hart, 2006. Hierarchical task analysis and verbal protocol analysis are methods that can be employed to understand the structure of tasks and how experts perform them. In the present research, they provide means for understanding skill development for the operation of construction equipment and to identify the skills to be acquired for each task or machine. The remainder of this chapter provides overview explanations of these three methods.

2.6.1. NASA Task Load Index

The NASA Task Load Index (NASA-TLX) has been employed extensively as a measure of subjective cognitive load over the past 20 years. Its use has spread far beyond its original application to aviation (for a review, see Hart, 2006). It is a multi-dimensional scale designed to obtain workload estimates from one or more operators while they are performing a task or immediately afterwards (Hart & Staveland, 1988). It consists of six subscales: Mental, Physical, and Temporal Demands, Frustration, Effort, and Performance (see Appendix A).

Some studies have generated a global estimate of cognitive workload by summing up the subscales of the NASA-TLX (e.g., Byers, Bittner, & Hill, 1989). For example, a recent study by Stinchcombe and Gagnon (2013) explored the effect of complexity on cognitive workload under different driving scenarios. They reported the summed workload measure, and their results indicated that all participants exhibited greater workload regardless of age when information-processing demands were increased, through the addition of traffic, and buildings.

2.6.2. Task Analysis

Task analysis originated in Time and Motion Study, combining the concepts of the Time Study work of Frederick W. Taylor (1911) with the Motion Study work of Frank Gilbreth and Lillian Gilbreth (1917, 1919). The original intent was to break down complex tasks into small and simple steps to increase the efficiency of work and reduce errors by careful observation to detect and eliminate redundant or wasteful motion and measurement of precise time taken. The rapid growth in technologies involving conditional situations with different choice, skill and knowledge selections gave rise to Hierarchical Task Analysis (HTA), pioneered by Annett and Duncan (1967), in which the task is analyzed in terms of goals, subgoals, and the actions for accomplishing the goals. Hoffman and Militello (2007) pointed out two distinctions between HTA and other forms of task analysis. First, the tasks being analyzed by this method cannot be described as single sequences of activities, but involve contingencies or conditionality. Second, the tasks can be analyzed in terms of both sequences of actions and goals (or functions). In this research, hierarchical task analysis and verbal protocol analysis were together adopted to facilitate understanding of skill development for the operations of construction equipment and to distinguish the skills to be acquired for each task or machine.

While task analysis has been strongly associated with job analysis and work design, in the era of industrialization, there is another type of task analysis — Cognitive Task Analysis (CTA) — which was invented largely as a result of computerization and which emerged in the 1980s. Hoffman and Militello (2007) defined CTA as "a methodology for the empirical study of workplaces and work patterns, resulting in: (a) descriptions of cognitive processes and phenomena accompanying goal-directed work, (b) explanations of work activity in terms of the cognitive phenomena and processes, and (c) application of the results to the betterment of work and the quality of working life by creating better work spaces, better supporting artifacts (i.e. Technologies), and by creating work methods that enhance human satisfaction and pleasure, that amplify human intrinsic motivation, and that accelerate the achievement of proficiency" (pp. 59). Examples of CTA methods existing today include retrospective interview techniques, real-time observations, think-aloud problem solving, etc. (More details of CTA methods are discussed in Chapter 4 of Hoffman & Militello's, 2007 book.)

2.6.2.1. Hierarchical Task Analysis

HTA was developed in response to the need for a systematic basis for understanding the component skills required in complex non-repetitive operator tasks, especially process control tasks found in industrial work practices (Annett & Duncan, 1967). HTA has since been extended to depict many other types of tasks, for example, preparation for and delivery of anesthesia (Phipps, Meakin, Beatty, Nsoedo, C., &

Parker, 2008). As noted by Phipps, Meakin, and Beatty (2011), "It is particularly useful as a general task analysis method because it provides a flexible, exhaustive and systematic means of identifying the behaviours that occur during a task (Patrick 1992)" (p. 741).

HTA begins by decomposing complex tasks into a hierarchy of goals and subgoals. The way in which a goal can be achieved is conceived of as an operation, and an operation includes 1) the actions that can lead to goal fulfillment, 2) conditions that will activate the goal, and 3) conditions that will fulfill the goal. The analysis is intended to consider both how the task should be performed and how it is actually carried out by operators (Annett, 2004). Because the task is decomposed into subgoals, performance can be analyzed at a number of different levels (Stanton, 2006). Through the contingencies and timelines from the HTAs, researchers can assess work demand by studying the plans in an HTA which set out how operators must respond to events in order to meet the demands of the task and by examining whether several events occurred at once which required the attention of the operator (Shepherd, 2001, pp. 164).

Annett (2004) outlined seven procedural steps in conducting HTA with typical purposes of designing a new system, troubleshooting and modifying an existing system, and developing operator training:

Step 1: Decide on the purpose of the analysis (e.g., designing a new system, troubleshooting an existing system, developing operator training)

Step 2: Get agreement between stakeholders and determine task goals and performance criteria

Step 3: Identify Sources of task Information and select means of data acquisition (e.g. direct observation, walk-through, protocols, expert interviews)

Step 4: Acquire data and draft decomposition table/diagram

Step 5: Recheck validity of your decomposition with stakeholders

Step 6: Identify significant operations in the light of the purpose of the analysis

Step 7: Generate and Test Hypothetical Solutions to the Performance

2.6.3. Verbal Protocol Analysis

Verbal protocol analysis, as recommended by Ericsson and Simon (1993), is a method for collecting and analyzing verbal data about cognitive processing. The main assumption of verbal protocol analysis is that it is possible to instruct subjects to verbalize their thoughts in a manner that does not alter the sequence of thoughts mediating the completion of a task, and that such utterances can therefore be accepted as valid data on thinking. The general finding that a task analysis can identify, a priori, the specific intermediate products that are later verbalized by subjects during their problem solutions, provides the strongest evidence that concurrent verbalization reflects the processes that mediate the actual generation of the correct answer (Ericsson, 2003).

The verbal protocol methodology can be divided into two different experimental procedures: concurrent and retrospective (Ericsson & Simon 1993; Kuusela & Paul, 2000; Ryan & Haslegrave, 2007 a, b). The concurrent think-aloud protocol is collected during the decision task, whereas, the retrospective think-aloud protocol is gathered after the decision task. Concurrent verbal reports are produced under specific instructions to the participant to 'think aloud' as they are performing a set of specified tasks, for example, doing a mental calculation, solving a problem, making a decision. Such verbal protocols are sometimes known as 'thinking aloud protocols' (Lewis, 1982). Subjects are asked to say whatever they are looking at, thinking, doing, and feeling as they go about their task. This enables observers to examine first-hand the process of task completion, rather than only its final product. Observers of such a test are asked to take notes of what the users say, without attempting to interpret their actions and words. This method is thought to be more objective in that participants merely report how they go about completing a task rather than interpreting or justifying their actions (Ericsson & Simon, 1993, see standardized instruction in Table 2.2). Verbal protocol becomes the most direct tool available in examining the on-going processes and intentions as and when learning happens (Gu, 2014). In addition, recent studies using eye tracking techniques to validate the think-aloud method have also shown encouraging evidence supporting the usefulness of the method (e.g.,Guan, Lee, Cuddihy, & Ramey, 2006). Research using verbal protocol analysis has been continuously reported for topics including road user behavior (Cornelissen, Salmon, McClure, & Stanton, 2013), operation in a nuclear power plant

(Lee, Park & Seong, 2012), clinical decision making (Hoffman, Aitken, & Duffield, 2009), and execution of a manual materials-handling task (Ryan & Haslegrave, 2007).

Table 2.2. Standard Instructions to Participants for Making Their Verbal Protocol Reports (adapted from Ericsson & Simon, 1993)

Verbal Protocol Procedure	Instructions to the Participant
Concurrent report	I am interested in what you are thinking about as you work. I am going to ask you to think aloud as you work on the task. I want you to tell me everything you are thinking from the moment you start the task until you have completed the task. I would like you to talk constantly from the time you start until you complete the task. I don't want you to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. It is important for you to keep talking. If you are silent for a long period I will ask you to talk. Do you understand what I want you to do?
Retrospective report	I want to see how much you can remember about what you were thinking from the time you started the task until the time you completed the task. I am interested in what you can actually remember rather than what you think you must have thought. If possible I would like you to tell me about your memories in the sequence in which they occurred while you were working. Please tell me if you are uncertain about any of your memories. Just report all that you can remember thinking about during the task.

2.7. Summary

Whether the current findings concerning the task-switching cost in both cognitive tasks and simple motor tasks may apply in the same manner to perceptual-motor tasks is still unclear and worth investigation. In the cognitive experimental literature on training, many training principles have been identified and supported by empirical research (Healy, et al., 2012; Wulf & Shea, 2002). However, all these principles do not necessarily apply for all tasks under all circumstances, but give inconsistent or contradictory results in different contexts (Healy et al., 2014; Travlos, 2010). This dissertation mainly focuses on skill acquisition, retention and transfer between machines and between tasks within a machine.

CHAPTER 3. RESEARCH FRAMEWORK

The ultimate goal of training is to optimize efficiency, durability and generalizability. Research has shown that practice schedules, specificity of practice, variability of practice and task difficulty may differentially influence learning, retention, and transfer (see Section 2.3). The manipulation of any of these principles might facilitate one aspect but impede the others (Healy et al., 2014; Travlos, 2010).

The present research has the goal of assessing three common training principles and related theories on skills retention and transfer: specificity of training (Identical Elements Theory), variability of practice, and task difficulty (Progressive Difficulty Training). Two of these training principles—specificity and task difficulty—are explored here within the context of the operation of construction equipment. This research mainly focuses on skill acquisition, retention, and transfer between two machines and between tasks within a single machine as demonstrated on VR-based training simulators. Whether introducing an alternative type of construction equipment or a different task to practice during training has positive or negative effects on learning, retention, and transfer is addressed. To understand skill development for the operations of construction equipment and to distinguish the skills to be acquired for each task or machine, HTA and verbal protocol analysis are employed. TLX ratings are also gathered to measure subjective cognitive load associated with each task. The structure of the dissertation is shown in Figure 3.1. This dissertation research is divided into two parts:

Part 1 involves experiments with two machines (hydraulic excavator and front end loader):

Studies 1 & 2 – Experiments conducted on Purdue campus

Study 3 – Interviews at John Deere site

Part 2 involves experiments with a single machine (front end loader):

Study 4 – Experiment conducted at John Deere site

Study 5 – Experiment conducted on Purdue campus

Part 1 consists of three studies. Studies 1 and 2 sought to verify whether an alternating practice sequence (inserting practice on a simulated loader while also learning on a simulated excavator) yields better skills transfer and retention for both a simple response-selection task and a complex task that involves multiple operations, based on the principle of specificity of training (when the conditions of practice match the conditions associated with retention or transfer). Study 3 aims at conceptualizing and analyzing transfer using HTA through the degree overlap of specific task components to provide theoretical explanations and ultimately postulate some guidelines that allow prediction of possible transfer across different tasks or machines.

Part 2, containing Studies 4 and 5, sought to verify whether an alternating practice sequence (training with an alternative tool and returning to the original learned tool) yields better skill transfer and retention (after a

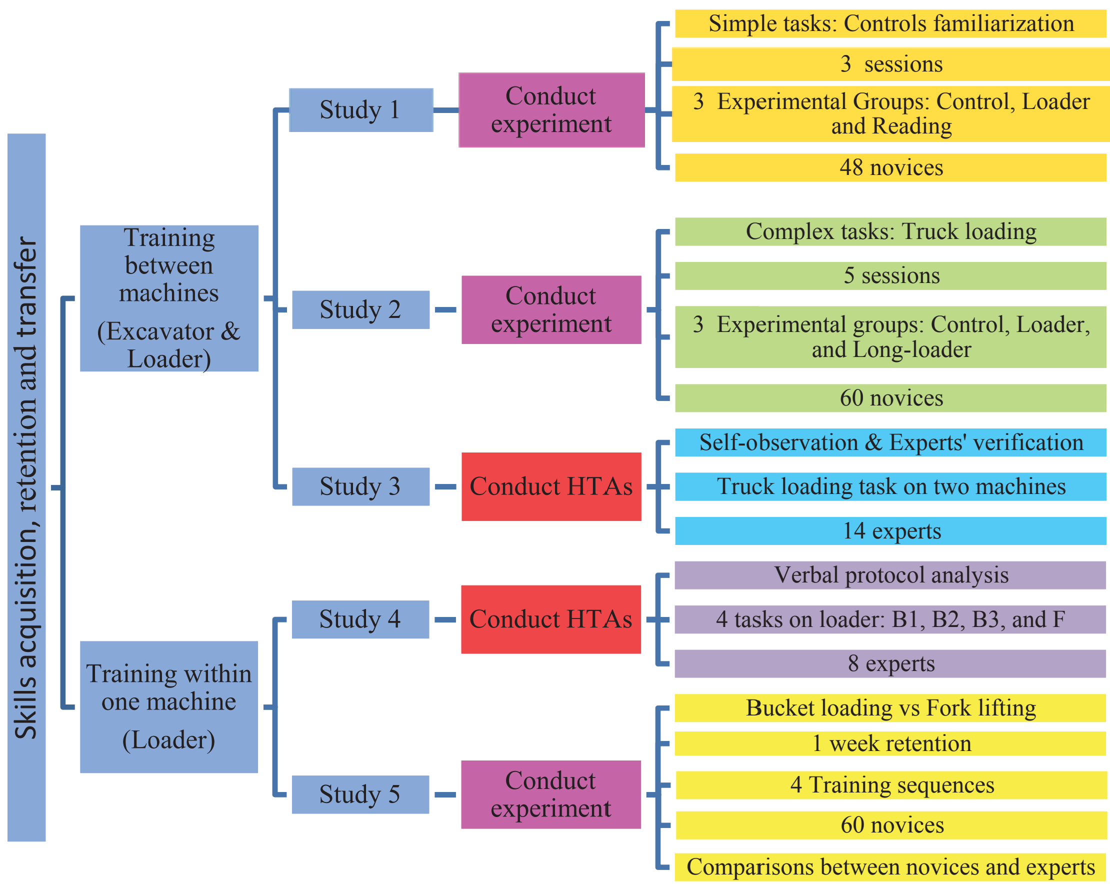


Figure 3.1. Structure of the dissertation.

one-week interval). These two studies also examined the principles of specificity of training and task difficulty. The results of Study 4, conducted with the experienced operators, provide information on how the four selected tasks on the loader should be performed and classify the difficulty level of each task to bolster the hypotheses for Study 5. The experiment conducted for it investigated whether the specificity of training and progressive difficulty training principles, for which difficulty should impede the learning stage (tests on the first session) but facilitate retention, holds for construction equipment training.

CHAPTER 4. SIMULATORS

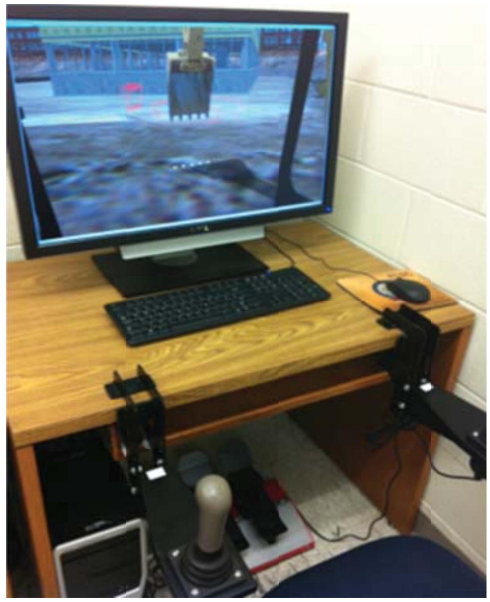
This chapter presents the details of the two construction-equipment simulators used in this research, as well as the setup and the controls of the simulators.

4.1. Experiment Apparatus

The simulators used in this study are Simlog's PC-based Hydraulic Excavator Personal Simulator Version 2.0 and John Deere's PC-based 4-Wheel Drive (4WD) Loader Simulator. Both simulators utilize real-time 3D software supported by a personal computer. Each system is installed on desktop computers equipped with 30-in. LCD Dell monitors, with speakers to each side (see Figure 4.1). Participants are presented with a virtual scene from the perspective of a person in the machine cabin, controlling the virtual machine through some combination of actions with joystick(s), pedals, a steering wheel (for the loader) and the Sealed-Switch Module (SSM) controls (for the loader), which mimics the way in which the actual construction equipment is controlled. Both simulators were designed for training purposes and include modules intended to allow trainees to develop skill at operating the simulated piece of equipment.

4.1.1. Hydraulic Excavator Simulator

Simlog's PC-based Hydraulic Excavator Personal Simulator simulates a Caterpillar 320CL hydraulic excavator (Figure 4.2). The Hydraulic Excavator



(a)

(b)

Figure 4.1. Setups for (a) the excavator simulator and (b) the loader simulator.

Figure 4.2. Screen shot from the simulated Caterpillar 320CL hydraulic excavator.

Simulator is designed to train and orient an entry-level operator on basic machine operation and skill, and to provide specific training exercises applicable to the hydraulic excavator. The training curriculum progressively takes a student from basic control orientation to complex machine tasks by presenting a series of instruction modules. The simulation modules incorporated in the current version of the Hydraulic Excavator Personal Simulator are summarized in Table 4.1.

For each simulation module, key performance indicators measure how well (or how poorly) the simulated work was performed. Typical examples are the time to complete the simulated task, the amount of material dug or loaded, and equipment collisions. Once each trial ends, the values of these performance indicators are displayed in a "Results" window until the user activates the horn to start the next trial.

The simulated hydraulic excavator consists of a stick, boom, bucket and cab on a rotating platform sitting atop an undercarriage with tracks. There are two joysticks to execute control functions. Each joystick can move in four directions up (forward), down (back), left, and right. There is a button on the top of each joystick. The left top button is called “horn button”, which is used to end a trial of a virtual task. The right top button is used to shift control function from bucket motion to carrier driving in some specific virtual tasks. The two joystick axes control the core functionality of the simulated hydraulic excavator according to the SAE pattern (see Figure 4.3).

4.1.2. Loader Simulator

John Deere’s PC-based 4-Wheel Drive (4WD) Loader Simulator simulates a John Deere 544K 4WD Loader (Figure 4.4). This training simulator features real-world situations, jobsite hazards, safety violations, hand signals, equipment damage,

Table 4.1. The Simulation Modules Incorporated in the Current Version of the Hydraulic Excavator Personal Simulator, as Ordered in the Instruction Manual

Simulation Modules	**Objectives of the modules**
Controls Familiarization	to master the controls of the hydraulic excavator
Excavator Positioning	to learn to position the tracks and bucket of the hydraulic excavator
Raking the Green	to learn to position the bucket so as to follow a trajectory that takes the form of a straight line
Over the Moon	to learn to position the bucket so as to follow a trajectory that takes the form of an arc
Bench Climbing/Descending	to climb and descend a bench safely and to place the hydraulic excavator in the proper parked position
Trench Crossing	to safely drive to an open trench, to safely cross the trench, and to place the hydraulic excavator in the proper parked position
Single-Pass Dig and Dump	to learn the basics of digging and dumping with the hydraulic excavator
Trenching	to expand upon the basics of digging and dumping by excavating a small trench
Trench and Load	to expand upon the basics of digging and dumping by excavating a small trench and loading the spoil into a dump truck
Bench Loading with Truck Spotting	to dig material from a bench, to spot an empty articulated dump truck for loading, and to load the truck from the bench
Bench Loading with Truck Spotting - Boulders	to dig heavy boulders from a bench, to spot an empty articulated dump truck for loading, and to load the truck from the bench
Ramp Building	to build a ramp to the top of the bench, using the available material, to climb the ramp safely, and to place the hydraulic excavator in the proper parked position

budget-based scoring, and replica machine controls. Nine highly detailed and realistic lessons teach proper operator technique, machine controls, and safe operation in a virtual jobsite. The simulation modules incorporated in the wheel loader are summarized in Table 4.2.

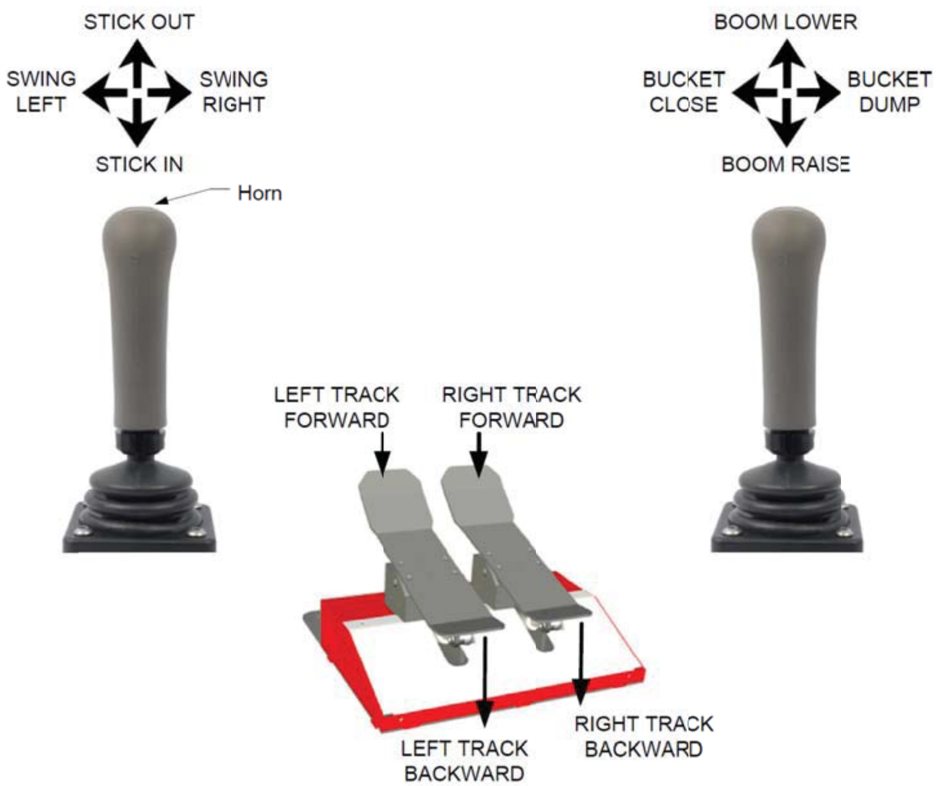


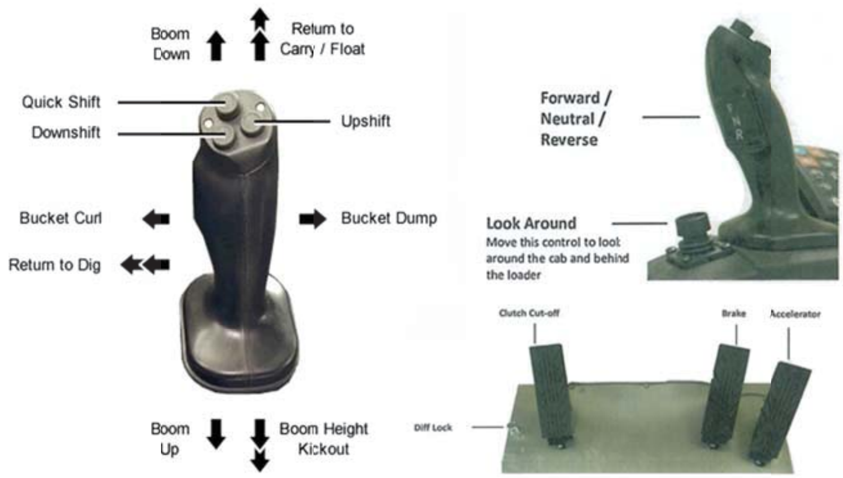
Figure 4.3. SAE pattern of the joystick controls and pedals of the simulated hydraulic excavator.

Figure 4.4. Screen shot from the simulated John Deere 544K 4WD Loader.

Table 4.2. The Simulation Modules Incorporated in the Wheel Loader

Simulation Modules	**Objectives of the modules**
Controls Familiarization	to master the controls of the wheel loader
Bucket Loading	to accurately approach an aggregate stockpile, and to position the boom height and bucket angle to achieve maximum bucket fill
Driving on a Jobsite	to safely maneuver a 4WD Loader through a jobsite while carrying a full bucket of aggregate.
Moving a Load with Narrow Fork	to detach the 4WD Loader bucket and attach narrow/utility forks.
Unloading a Flatbed with Forks	to unload bundles of 20' iron pipe off of a flatbed trailer
Moving a Load with Wide Forks	to transport a tall load (port-a-potty) through the jobsite
Feather bedding into a Trench	to lightly dump aggregate into a trench while following hand signals.
Truck Loading	to quickly and accurately load dump trucks
Loading onto a Lowboy Trailer	to load a 4WD Loader onto a lowboy trailer for transport

After completing each lesson, operators receive immediate feedback based on their performance. Operators are scored against a budget and other skilled operators to help identify strengths and weaknesses. The simulated loader is wheeled and has a wide front mounted bucket connected to the end of two boom arms to scoop up loose material, such as dirt, sand or gravel, and carry it from one location to another. The loader consists of one joystick, pedals, a steering wheel and the Sealed-Switch Module (SSM) controls (see Figure 4.5). Some SSM controls are the same as the joysticks control, which provides an additional option for the operators to choose their most convenient way to execute their desired machine movements.



(a)

(b)

(c)

Figure 4.5. Controls of the simulated loader: (a) joystick and pedals, (b) Sealed-Switch Module (SSM) controls, and (c) steering wheel.

CHAPTER 5. STUDY 1: TRAINING FOR SIMPLE TASKS WITH TWO MACHINES

Transfer refers to re-use of prior knowledge gained in earlier settings to affect learning and performance in other newly encountered settings. That is, knowledge and skills are passed on from one domain or task to another. Transfer of training, in particular, refers to how learning which responses to make to stimuli in one situation influence the responses in another (Adams, 1987). Negative transfer is said to occur when newly learned information degrades or impedes the recall of previously learned information (Smode, Beam, & Dunlap, 1959). In contrast, positive transfer arises when the previously learned information facilitates performance of the new ask. In general, trainers desire positive transfer between contexts to occur but not negative transfer.

A skilled operator of construction equipment needs to become proficient in the use and control of various classes or types of machines. Also, the advent of training simulators means that training programs can provide ready access to learning the operation of multiple machine types concurrently. It is thus important to establish effective training methods across these machine types and to determine the extent to which skills operating one machine transfer to another. An issue of concern is whether learning to operate a single piece of equipment is best if all practice is on that equipment, or whether intermixed training on a related piece of equipment can be of value (or possibly a hindrance). Accumulating evidence suggests that switching between tasks leads to longer response times (RTs) and more errors than performing a single task repeatedly (e.g., Monsell, 2003). The complexity of an operator's perceptual-motor tasks raises a question of whether switching tasks across different machines (e.g., loader vs. excavator). Whether the findings of local and global switch costs from cognitive switch tasks hold in this context is questionable in that the motor tasks involve higher execution complexity and multiple movements (e.g., Kray & Lindenberger, 2000). Research has found that practice schedules on motor control tasks may differentially influence performance and learning; whether switching between two complex perceptual-motor tasks with different task sets is always detrimental to speed and accuracy is worth further investigation because simulators make concurrent training more readily accessible.

5.1. Objectives

This study investigated whether operation of a simulated hydraulic excavator is influenced by an intervening task performed between initial practice on the excavator and a subsequent retention test using a controls familiarization task which involves just knowing the control functions. Two intervening tasks were examined: practicing on a simulated loader and reading an unrelated text (intended to distract the participants). Performance was compared against that of a group of participants who practiced on the simulated excavator throughout. The reading task allowed evaluation of the extent to which directing attention to a task other than excavator training during the intervening period affected subsequent performance on the excavator.

5.2. Hypothesis

According to the specificity of training principle and identical elements theory of transfer, the amount of positive transfer, or benefit, that training in one situation will have on another is determined by the number of elements that the two situations have in common. It is hypothesized that:

> *Hypothesis 1*: *Positive transfer from the loader to the excavator should occur because the tasks being trained on both machines are similar (i.e., controls familiarization) and the elements of operation are similar (e.g., operation of the bucket with joysticks).*

5.3. Method

5.3.1. Participants

Forty-eight undergraduate students (38 males and 10 females, distributed evenly across the three groups), ages 19–34 years (*M* = 19.9; *SD* = 2.5), participated for experimental credits toward an Introductory Psychology course requirement participated for experimental credits toward an introductory psychology course requirement according to Institutional Review Board (IRB) Human Subject Protocol #1110011339 (see Appendix B). All were right-handed, physically capable of operating the simulator, and had no experience operating construction equipment.

5.3.2. Experimental Setup

The simulators used in this study were Simlog's PC based Hydraulic Excavator Personal Simulator, which simulates a Caterpillar 320CL hydraulic excavator, and John Deere's PC-based 4-Wheel Drive (4WD) Loader Simulator, which simulates a

John Deere 544K 4WD Loader. Both systems were installed on desktop computers equipped with 30-in. LCD Dell monitors, with speakers to each side (see Figure 4.1). Participants were presented with a virtual scene from the perspective of a person in the machine cabin. They controlled the virtual machine through some combination of actions with joystick(s), pedals, a steering wheel (for the loader) and the Sealed-Switch Module (SSM) controls (for the loader), which mimics the way in which the actual construction equipment is controlled. The simulated hydraulic excavator consists of a stick, boom, bucket and cab on a rotating platform sitting atop an undercarriage with tracks. The simulated loader, however, is wheeled and has a wide front mounted bucket connected to the end of two boom arms to scoop up loose material, such as dirt, sand or gravel, and carry it from one location to another.

5.3.3. Design

The experiment involved three sessions: 1. skill acquisition on the controls of the excavator simulator; 2. performance of an intervening task; 3. retention test on controls of the excavator simulator. All sessions used training modules provided as part of the simulator software. In Session 1, all three groups were given an introductory lesson on the basic parts and controls of the excavator simulator, followed by an assessment test using the Controls Familiarization module on the excavator simulator. In Session 2, participants were all given an introductory lesson on the basic controls of the loader (similar setting as the introductory lesson for the excavator in Session 1) and then received one of the intervening task conditions (randomly determined): (a) continued practice on the Controls Familiarization module on the excavator simulator (control group), (b) practice on the Controls Familiarization module of the loader simulator; and (c) reading a book unrelated to equipment operation for the same length of time. In Session 3, all participants performed the same Controls Familiarization module as in Session 1, except that no introductory lesson was given.

The factors and levels studied were: two sessions (initial, retention), four blocks (1 to 4) within each session, and three intervening tasks (practice on loader, reading, and continuing practice on excavator). Sessions and blocks were within-subject factors, and intervening task was a between-subjects factor. Several performance measures were recorded on the excavator simulator, including execution time (elapsed time since the beginning of the trial) and the total number of errors in each trial. In addition to obtaining performance measures, the subjective measures of workload were gathered using the NASA-TLX (see Appendix A). Participants rated workload with the TLX after each session of the experiment. Several analyses of variance (ANOVAs) were used to compare the execution time and total number of errors obtained by either the loader group or reading group in the retention test (Session 3) to those from the 2nd and 3rd sessions of the excavator (control) group.

5.3.4. Experimental Task

Participants were given one of the intervening task conditions (randomly determined) in Session 2: (a) continued practice on the control familiarization module on the excavator simulator (control group), (b) practice on the control familiarization module of the loader simulator; and (c) reading a book unrelated to equipment operation for the same length of time. The details of each experimental task are illustrated below.

5.3.4.1. Controls Familiarization of Hydraulic Excavator Simulator

Each trial (simulation exercise) began with a view from inside the operator cabin, along with an instruction displayed at the top of the simulation window. The participant had to read the instruction, recall the appropriate control action, and then activate the instructed function. A summary of results, built into the system software, showing the total number of errors and execution time appeared after the function was activated correctly. The participant had to activate the correct control action before the next trial began. A 5-min break was given between sessions.

5.3.4.2. Controls Familiarization of Wheel Loader

In this module, the participants learned how to react quickly with accurate responses. Participants needed to respond to each control prompt with the correct joystick movement, foot pedal press, or SSM button press. A green checkmark was shown for a correct response and a red X for an incorrect response and did not allow re-correction of the mistake. Different from the excavator simulator, the loader simulator did not display the summary report after each trial, the total execution time and accuracy for 30 trials in total were reported and automatically recorded in the database.

5.3.4.3. Reading Task

To engage the participants in the reading task, the book *1001 Great Stories, Volume 1*, edited by Douglas Messerli (2005), was selected. This book consists of short stories with a variety of themes.

5.3.5. Procedure

Participants were informed of the study's aim and that the goal was to master the controls of the hydraulic excavator. A preliminary questionnaire (see Appendix C) obtaining demographic information was administered before the session began. The first session of experimentation – skill acquisition – involved two parts. Part 1 started with an introductory lesson on the excavator, which was a 3-page instruction presented on the screen. It described the parts and basic functions of the excavator and the corresponding operation of the joystick and pedal controls. Participants were given 10 minutes to study the instruction. In Part 2, participants were seated at the excavator simulator and tested with the Control Familiarization module 4 times (30 trials each). Participants answered the NASA-TLX questionnaire with regard to the control familiarization task just performed on the excavator at the end of the session.

To make the intervening task and the time approximately equivalent in Session 2, the participants were all given a 10-min introductory lesson on the controls and parts of a loader before being assigned to one of the three tasks. To ensure that the participants had processed the information in the introductory lesson, they were not told which intervening task condition they would receive until the lesson was completed. For the control group, who continued practicing on the excavator simulator, the participants performed the control familiarization module another 4 times (30 trials each). Similarly, the loader group performed the control familiarization modules of the loader simulator 4 times (30 trials each). For the group assigned with reading, participants were given the book to read for 15 minutes. They were allowed to start reading any story; they were asked to process all pages and told that skipping pages was not allowed. The number of pages read was recorded at the end of the session. Participants again answered the NASA-TLX questionnaire at the end of session, this time with regard to the just-completed intervening task.

In the last session, all participants were returned to the excavator simulator and performed the control familiarization module 4 times (30 trials each). The NASA-TLX questionnaire was filled out at the end in relation to the last session on the excavator.

5.4. Results

Four participants with total execution time for the whole session over 1000 s (one each from the loader group and excavator group, and two from the reading group) were excluded in the analysis. Their long execution times were mainly due to an extreme number of errors they made during the experiment, which greatly slowed their performance. It was deemed that these deletions were few enough to not compromise appearance of the effects that were being investigated. Figure 5.1 shows the mean execution time per 30-trial module of the control familiarization task on the excavator simulator across the three sessions for all groups.

5.4.1. Control Group

5.4.1.1. Practice Effect

A learning curve was plotted for the control group that practiced the Controls Familiarization modules on the excavator simulator across the three sessions (initial, intervening, retention). A total of 15 participants were included in this analysis. To obtain the learning curve, the mean total execution time of each module (which consisted of 30 trials) is plotted in Figure 5.2 as a function of blocks. Because group curves for skill acquisition typically approximate a "power law of practice" (Newell &

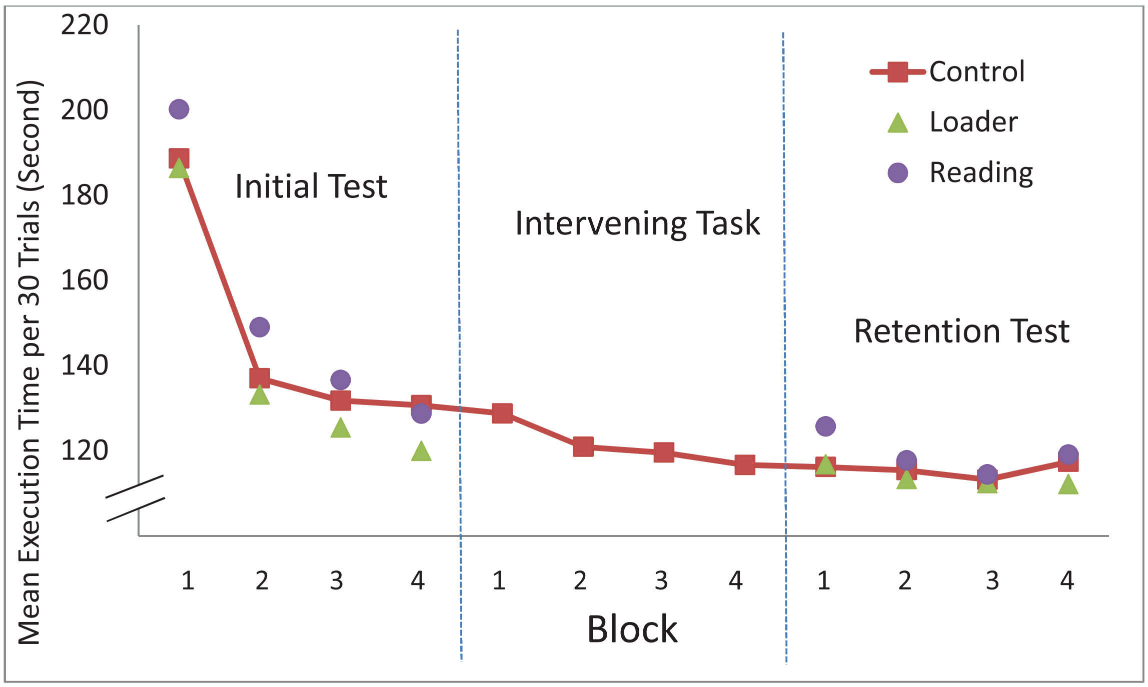


Figure 5.1. Mean execution time per 30-trial module of the controls familiarization task on the excavator simulator across the three sessions for all groups.

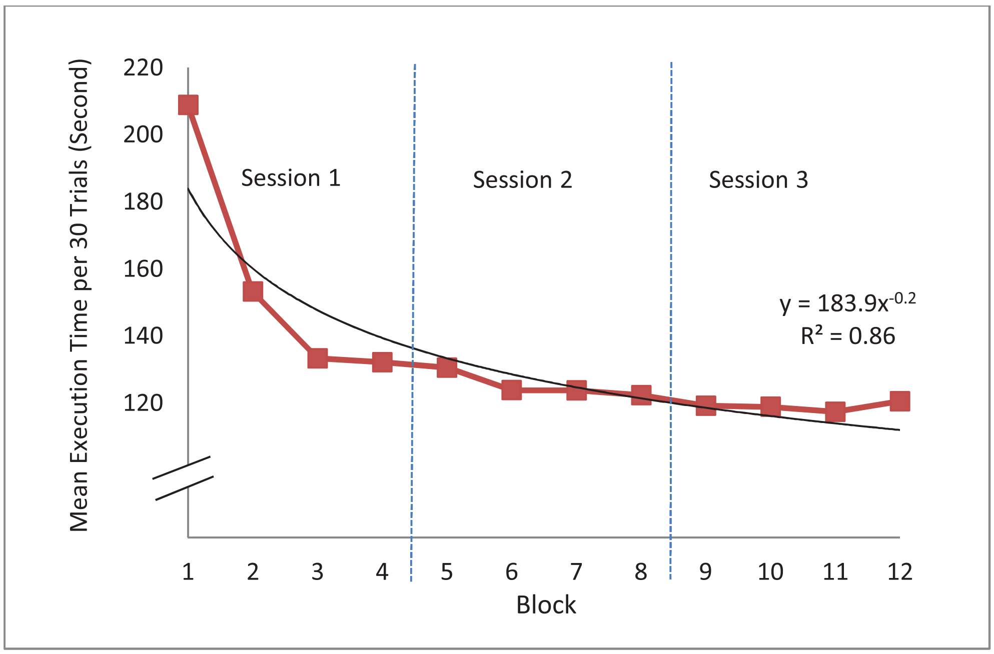


Figure 5.2. Mean execution time per 30-trial module of the controls familiarization task on the excavator simulator across the three sessions for the control group.

Rosenbloom, 1981), a power function was fit using the least-squares method. This function conforms well to the data: $y = 183.9\ x^{-0.2}$, $R^2 = .85$; $F(1, 11) = 61.97$, $p < .001$, where y is the total execution time per module and x is the block number (four for each of the three sessions).

5.4.1.2. Total Execution Time

A repeated-measures ANOVA was used to test the effects of session (initial, intervening, retention) and block (1 to 4) on total execution time per module on the excavator simulator (see Appendix D). An initial analysis for a gender difference yielded $F < 1.0$, so gender was not included in the ANOVA. The results showed a main effect of session, $F(2, 28) = 75.55$, $p < .001$, $\eta_p{}^2 = .854$, with total execution time being shortest for the retention test of Session 3 (115 s), next for Session 2 (121 s) and then for Session 1 (146 s). Trial was also a significant factor, $F(3, 42) = 68.56$, $p < .001$, $\eta_p{}^2 = .830$, showing a decrease in execution time across the blocks within a session. The session × block interaction (Figure 5.3) was also found to be significant, $F(6, 84) = 39.43$, $p < .001$, $\eta_p{}^2 = .738$. This interaction mainly reflects that the majority of learning occurred in Session 1.

5.4.1.3. Total Number of Errors

A repeated-measures ANOVA was used to test the effects of session (initial, intervening, retention) and block (1 to 4) on total number of errors per 30-trial module. The results (see Appendix D) showed a main effect of session, $F(2, 28) = 6.51$, $p < .005$, $\eta_p{}^2 = .317$, with total errors decreasing from 3.42 in Session 1 to 2.18 and 1.86 in Sessions 2 and 3, respectively. Trial was a significant factor, $F(3, 42) = 5.01$, p

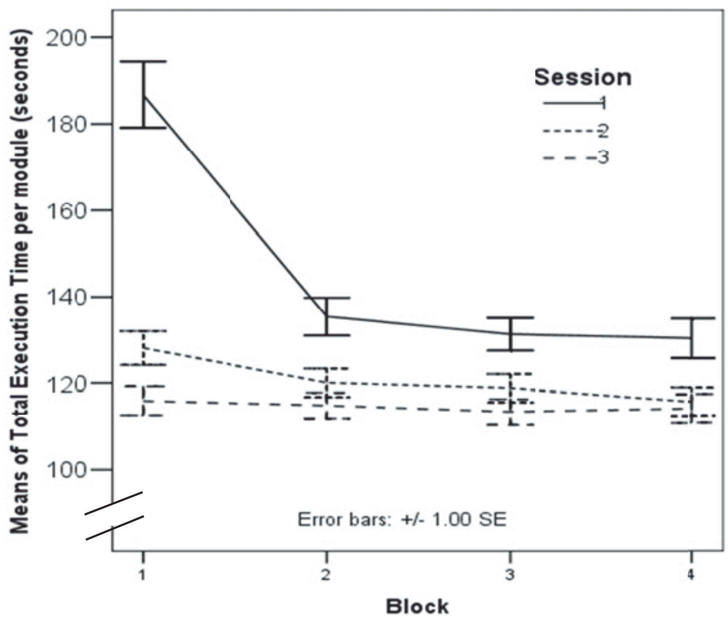


Figure 5.3. Two-way interaction plot of session × block on total execution time of each module.

$< .005$, $\eta_p^2 = .263$, showing a reduction in errors across the blocks within a session. No session × block interaction was found, $F(6, 84) = 1.93$, $p = .085$, $\eta_p^2 = .121$.

5.4.1.4. Workload Measures

The six different subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration were analyzed using a repeated-measures ANOVA, with the workload measures and session (initial, intervening, retention) as within-subject factors. The scale is on 1-10 with 0.5 increments, with 1 indicating low workload and 10 indicating high workload. The Huynh-Feldt correction for violations of sphericity was applied, but because that correction did not change the significance level of the results, those with sphericity assumed are reported here. The results of ANOVA show a main effect of session, $F(2, 28) = 5.79$, $p = .008$, $\eta_p^2 = .293$, for which the overall workload decreasing from 4.40 in Session 1 to 3.43 in Session 2 and 3.42 in Session 3. The main effect of measure was also significant, $F(5, 70) = 8.53$, $p < .001$, $\eta_p^2 = .379$, where participants found rated the tasks as requiring higher mental demand ($M = 4.68$ out of 10) and effort ($M = 4.78$) but lower physical demand ($M = 2.60$) and they were very satisfied with their performance ($M = 2.33$). No session × measure interaction was found, $F(10, 140) = 1.04$, $p = .412$, $\eta_p^2 = .069$.

5.4.2. Initial Performance on Excavator (Session 1)

All participants had the same training process before they were assigned to one of the three experimental groups. The data collected from the 44 participants were tested to check the consistency in performance across groups for Session 1. Separate mixed design ANOVAs were used to test the effects of blocks and intervening groups on the total execution time and the total number of errors per module. The results showed a significant block effect, $F(3, 123) = 146.23$, $p < 0.001$, $\eta_p{}^2 = .781$, but neither an intervening group main effect, $F(2, 41) = 1.27$, $p = .292$, $\eta_p{}^2 = .058$, nor interaction with block, $F(6, 123) = .753$, $p = .609$, $\eta_p{}^2 = .035$, was significant. Thus, all three groups performed at approximately the same level before they were introduced to the intervening task. The ANOVA on the total number of errors per module also showed the block effect, $F(3, 123) = 6.865$, $p < .001$, $\eta_p{}^2 = .143$, with a significant reduction in errors from the 1st block ($M = 5.20$) to the 4th block ($M = 3.27$), but no group main effect, $F(2, 41) = 2.629$, $p = .084$, $\eta_p{}^2 = .114$, or interaction with block, $F(6, 123) = .590$, $p = .738$, $\eta_p{}^2 = .028$. No significant differences were found in the workload measures among the three groups, $F(2, 41) = 1.10$, $p = .344$, $\eta_p{}^2 = .051$, indicating that mental workload estimates for the three groups were similar in Session 1. The equivalence of results across groups in this session allows any later effects on the performance in the retention test to be attributed to the effects of the intervening task.

5.4.3. Effects of Intervening Task

5.4.3.1. Unrelated Reading Task

During the reading task, the participants read on average 26.0 (*S.D.* = 0.85) pages of the story book provided to them. Several ANOVAs were used to compare the execution time and total number of errors per module obtained by the reading group in the retention test (Session 3) to those from the 2nd and 3rd sessions of the excavator group. No significant differences were found between the execution time in the retention test of the loader group and the 2nd session, $F(1, 27) = .11$, $p = .741$, $\eta_p^2 = .004$, or 3rd session, $F(1, 27) = 1.54$, $p = .225$, $\eta_p^2 = .054$, of the control group.

The result pattern was different for the total number of errors per module, for which an ANOVA showed that the intervening task was a significant factor. The reading group made more errors ($M = 5.25$) in the retention test (Session 3) compared to the second ($M = 2.18$) and third sessions of the control group ($M = 2.43$), $F(1, 27) = 9.21$, $p < .005$, $\eta_p^2 = .254$ and $F(1, 27) = 8.19$, $p < .005$, $\eta_p^2 = .233$, respectively.

5.4.3.2. Practicing on Loader

The loader condition was compared to the control condition in a manner similar to the analyses just reported, using execution time and total number of errors as measures. No significant differences were found between the execution time in the retention test of the loader group and the 2nd session, $F(1, 28) = .084$, $p = .775$, $\eta_p^2 = .043$, or 3rd session, $F(1, 28) = .074$, $p = .787$, $\eta_p^2 = .003$, of the control group.

Similar results were found for the total number of errors per module for comparison with the 2nd session, $F(1, 28) = 1.81$, $p = .190$, $\eta_p^2 = .061$, and 3rd session, $F(1, 28) = 1.26$, $p = .272$, $\eta_p^2 = .048$, of the control group. The data analysis of the performance and errors on the loader simulator were reported in Appendix D.

5.4.3.3. Workload Measures

The six different subscales of NASA TLX questionnaire were analyzed using a repeated-measures ANOVA (see Appendix D), with the workload measures and session (initial, intervening, retention) as within-subject factors and experimental group as a between-subjects factor. Group was not found significant, $F(2, 41) = .184$, p

= .833, η_p^2 = .009, and no interactions with group were found, indicating that no significant differences were found among the three experimental groups for any subscale. The main effect of measure was significant, $F(5, 205) = 21.59$, $p < .001$, η_p^2 = .345, indicating that participants rated the tasks as requiring higher mental demand (*M* = 4.52), temporal demand (*M* = 4.48), and effort (*M* = 4.45) but lower frustration (*M* = 3.07), physical demand (*M* = 2.49), and they were very satisfied with their performance (*M* = 2.49).

5.5. Discussion

Participants practiced a training module for a simulated excavator, which requires prompt operation of a correct control action in response to a visual command. Those who practiced for three sessions showed continuous improvement in performance and a reduction in rated mental effort. The main concern of this study was whether practice on a simulated loader that intervened between sessions on the excavator would influence performance of the task requiring operation of the excavator controls. The results did not show effects of having received the intervening training on the loader. Total execution time and number of errors on the excavator were not different from those of the group who maintained practice on the excavator, which is not consistent with Hypothesis 1. Improvement in the total execution time was observed in the excavator retention test for the group that was diverted to practice on the simulated loader. The lack of significant difference from the group who practiced on the excavator for all three sessions suggests that switching from one machine to another does not inhibit the original performance and may even facilitate the learning on the original task. This conclusion is similar to one reached by Lyall and Wickens (2005) for transfer of commercial airline pilots from one model of aircraft to another.

One limitation of the study is that the control familiarization tasks on both machines may be too simple, requiring only selection of a control action in response to a stimulus, for effects of switching between the machines to be evident. To confirm this conclusion that switching from one machine to another is not detrimental, a follow-up study was conducted using more complex tasks. In Study 2, instead of using the Controls Familiarization modules as the assessment tests, using truck loading modules that require not only navigating and maneuvering the vehicles but also fine motor skills to handle the implement (i.e., bucket) may reveal differences not evident in the present study.

5.6. Conclusion

The general finding of this study is that no performance cost on the controls familiarization task is attributed to inserting practice on a simulated loader while also learning on a simulated excavator. The practical implication is that trainees can move from excavator to loader training on controls familiarization without negative impact on learning the basic excavator controls. Because the controls familiarization task is restricted to selection of a control action, Study 1 does not rule out the possibility of transfer between machines occurring when the tasks involve actual operation of the simulated machinery.

CHAPTER 6. STUDY 2: TRAINING FOR COMPLEX TASK WITH TWO MACHINES

Study 1 investigated whether performance on controls familiarization with the simulated hydraulic excavator was influenced by learning the controls of the simulated loader between an initial practice session on the excavator and a subsequent retention test. Participants were asked to perform controls familiarization tasks. Each trial began with a view from inside the operator cabin, along with a single written instruction displayed in the simulation window. The participant had to read the instruction, select the appropriate control action, and execute the correct control function. Performance was compared against that of participants who practiced on the simulated excavator throughout. The performance measures for the excavator showed no cost (or benefit) attributable to inserting practice on the simulated loader between the initial and final sessions on the excavator. The absence of an effect of switching between the machines on learning to perform the controls familiarization tasks in Study 1 could have been due to the simplicity and similarity of the tasks, which involved selecting an appropriate response to the written instruction.

6.1. Objectives

Study 2 sought to verify this result using Truck Loading modules that require navigating and maneuvering the vehicles and fine motor skills to handle the implement through large ranges of motion, not just selection of a signaled response, as it may reveal differences not evident in Study 1. Also, the number of sessions was increased from 3 to 5, to examine performance when participants continued to switch between the two machines in sessions 4 and 5. This increase was done to examine how the duration of the inserted practice on the alternative machine affects performance on the initially learned machine.

6.2. Hypotheses

In this experiment, the ‘Trench and Load’ task on the excavator requires participants to dig and dump by excavating a small trench and then loading the spoil into a dump truck. For the loader, the module named ‘Truck Loading’ was used for assessment. Similar to the excavator, the ‘Truck Loading’ task requires participants to drive to an aggregate pile, get a full bucket of aggregate, and then approach the dump truck and dump the aggregate into the truck bed. From comparing the controls of the two pieces of equipment and the task natures, the following hypotheses were formulated:

> *Hypothesis 2*: By the principles of specificity of training, the transfer is best when the transfer test matches the task being practiced during training, i.e., in what is typically called a retention test. *Therefore, the control group will continuously benefit from practicing the same task throughout all sessions, showing a significant decrease in subjective mental workload.*

Hypothesis 3: In terms of control configuration, the simulated excavator consists of left and right joystick(s) and left and right pedals, whereas the simulated loader consists of one joystick, and the Sealed-Switch Module (SSM), a steering wheel, an accelerator and a brake. Although both tasks share a similar goal of loading a bucket and filling the truck bed, the loader involves driving with the steering wheel for every bucket loading cycle whereas the excavator is stationary without the need to move for every bucket load. *Therefore, the dissimilarities in both the controls and the task procedures between the excavator and the loader will lead to a switch cost when returning to the original practiced machine.*

Hypothesis 4: In a comparison between a practice sequence that employs equal alternating practice first from loader (L) to excavator (E) and one wherein the alternate loader practice is of double duration before returning to the excavator, a larger negative impact and larger interference to skill improvement on the excavator will result from the longer loader practice. *Therefore, the long-loader group with the practice sequence, E>L>L>E>E, performing on the loader in Sessions 2 and 3, will show a larger negative impact and larger interference on performance with the excavator than the loader group with the practice sequence, E>L>E>L>E*. This is due to a longer lag time between the initial test on the excavator and the retention test performed upon returning to the excavator. A longer period of diversion to practice on an alternate type of simulated equipment will result in a greater negative impact on performance when returning to practice on the original type of simulated equipment.

6.3. Method

6.3.1. Participants

Sixty undergraduate students (48 males and 12 females, distributed evenly across the three groups), ages 18–26 years (*M* = 19.8; *SD* = 1.6), participated for experimental credits toward an introductory psychology course requirement according to IRB Human Subject Protocol #1110011339. All were right-handed, physically capable of operating the simulator, and had no experience operating construction equipment.

6.3.2. Experimental Setup

The setup was the same as presented in the Chapter 5 (Study 1). The two simulators were Simlog's PC-based Hydraulic Excavator Personal Simulator, which simulates a Caterpillar 320CL hydraulic excavator, and John Deere's PC-based 4-Wheel Drive (4WD) Loader Simulator, which simulates a John Deere 544K 4WD Loader. Participants were presented with a virtual scene from the perspective of a person in the machine cabin. They controlled the virtual machine through the same interface mechanisms described in Study 1.

6.3.3. Design

This study investigated whether performing a complex task on a simulated hydraulic excavator is influenced by an intervening task performed on a simulated loader between initial practice on the excavator and a later retention test. Four modifications were made from Study 1:

1. The original controls familiarization test was replaced with a more complex truck loading task that involves multiple operations.
2. The number of sessions was increased, from three to five, to examine the possible influence when participants continue to switch between the machines.
3. The intervening reading task group was not included.
4. Besides the two experimental groups (control and loader groups), an additional group which was given practice between the two machines (but a different practice schedule from the original loader group) was added to address the question of how the duration of insertion of practice on an alternative machine matters to the performance on the previous learned machine.

The factors and levels studied were: three sessions (initial, 1st retention, 2nd retention), two trials within each session, and three intervening tasks (control, and two loader groups). Sessions and blocks were within-subject factors, and intervening task was a between-subjects factor. Several performance measures were recorded on the excavator simulator, including execution time (elapsed time since the beginning of the trial) and the total number of errors in each trial. The total percentage of truck being filled per trial was recorded on the loader simulators. Participants were asked to rate workload with the NASA-TLX after each session of the experiment.

6.3.4. Experimental Task

All sessions used training modules provided as part of the simulator software. For the excavator, the module named ‘Trench and Load’ was used (see Figure 6.1a). For each trial of this module, the excavator bucket is empty, and the excavator is positioned some distance away from the trench to be dug, with an empty articulated dump truck parked next to the trench. Participants were asked to drive to a position in line with the marked trench and to dig and dump by excavating the small trench (area indicated by green stakes) and loading the spoil into the dump truck. For the loader, the module named ‘Truck Loading’ was used for assessment (see Figure 6.1b). Similar to the excavator, participants were asked to drive to the material source, in this instance an aggregate pile, get a full bucket of aggregate, and then approach the dump truck and dump the aggregate into the truck bed. The execution time of the truck loading module as hard-coded into the software is fixed at 7 minutes on the loader. Consequently, the trench and load task on the excavator simulator, which allows stopping the module at any time by pressing the horn, was fixed at 7 minutes to make the time in each trial equivalent. The results available from the excavator system software—the volume transferred to the truck and execution time—were recorded for further analysis.

(a)

(b)

Figure 6.1. Exocentric view screen captures from the truck loading scenarios for (a) the excavator simulator and (b) the loader simulator.

6.3.5. Procedure

A preliminary questionnaire obtaining demographic information was administered before the first session began. Participants were informed of the study's aim and that the goal was to perform a truck loading task to obtain maximum productivity on the simulated hydraulic excavator. In Session 1, participants were given 5 min to study a three-page instruction presented on the screen. It described the parts and basic functions of the excavator and the corresponding operation of the joystick and pedal controls. Participants were then seated at the excavator simulator and tested with the Controls Familiarization module once (30 trials). Next, participants were given 2 minutes of free-play to try the trench and load module, during which they could ask questions. Then, they were tested twice with the 7-minute truck loading module.

In Session 2, participants were divided into three groups according to their practice sequences:

Group 1: Control group (E>E>E>E>E)

Group 2: Loader group. (E>L>E>L>E)

Group 3: Long-loader group (E>L>L>E>E)

For the two loader groups, the participants were given an introductory lesson on the basic controls of the loader, followed by the controls familiarization module (30 trials). Next, they were given 2 min of free-play to try the truck loading module on the loader simulator, and they could ask questions during the free-play. Then, they were tested with the truck loading module on the loader simulator twice (7 min x 2 trials).

While the two loader groups were given an introductory lesson on the basic controls of the loader, the control group was given a 10-minute reading task in order to make the tasks and the time approximately equivalent in Session 2 (the participant's choice of short stories from the book *1001 Great Stories, Volume 1*, edited by Douglas Messerli, 2005) before they continued practice on the trench and load module of the excavator simulator. Similar to Session 1, the control group performed the trench and load task twice in this session.

Session 3 was the retention test for the truck loading task on the simulated excavator for the control and loader group, in which all participants performed the same trench and load module as in Session 1, without the introductory lesson. The long-loader group continued to practice the truck loading task twice in Session 3.

In Session 4, the loader group was seated at the loader simulator again to perform the truck loading module twice, whereas the control group continued on the excavator simulator to perform the trench and load module twice. The long-loader group was seated at the simulated excavator for the first retention test.

In Session 5, all participants again returned to the simulated excavator and performed the same truck loading module twice, with no introductory lesson. A 5-min break was given between sessions.

6.4. Results

6.4.1. Retention Test on the Excavator Simulator

Productivity (m^3/hr) in the trench-and-load task with the excavator was calculated from the total volume transferred from the trench to the truck, divided by the total execution time. Figure 6.2 illustrates the mean productivity performance on the

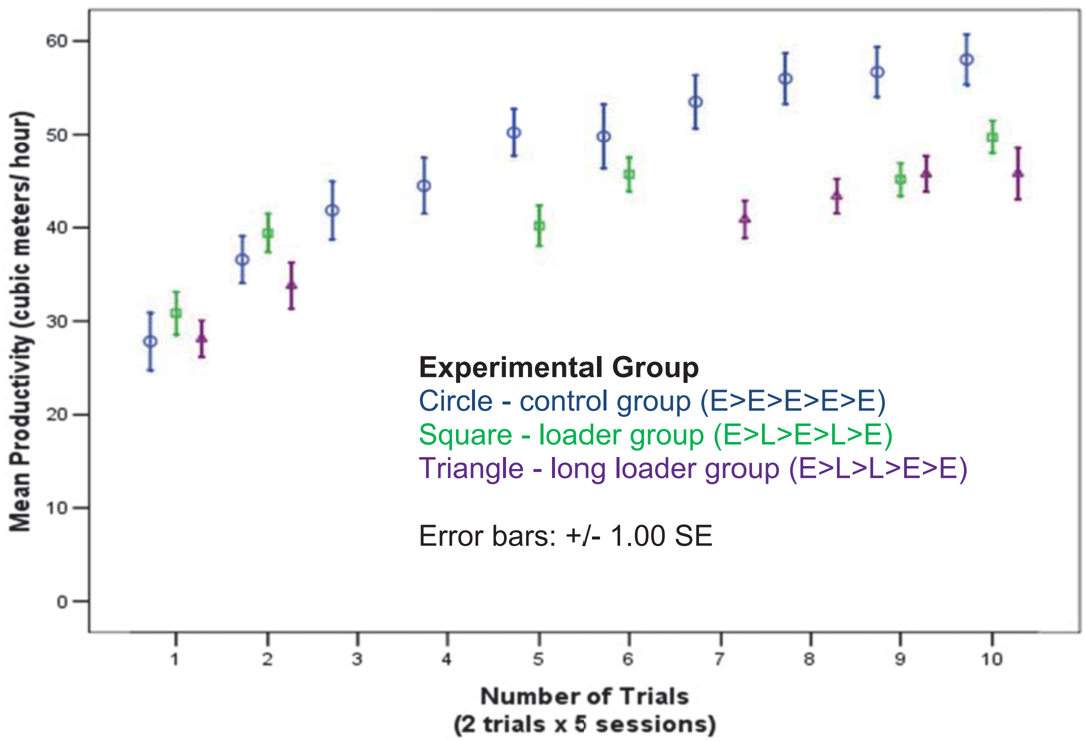


Figure 6.2. Productivity of truck loading task on the excavator across the 5 sessions for the three experimental groups.

excavator across the 5 sessions for all groups. The three groups did show an improvement in productivity across the sessions, but at different rates. Statistical analyses were conducted to investigate the effects of the intervening tasks by comparing the productivity of the control group with the two loader groups. The insignificant gender effects yielded $F < 1.0$, so gender was not included in the ANOVA.

6.4.1.1. Initial Test (Session 1)

All participants had the same training process before they were assigned to one of the three experimental groups. The data collected from the 60 participants were tested to check the consistency in performance across groups. A one-way ANOVA was conducted to test the effect of the three intervening groups on mean productivity in the first session. The results showed a significant trial effect, $F(1, 57) = 68.37$, $p < 0.001$, $\eta_p^2 = .545$, with a significant increase in productivity from trial 1 (29.03 m^3/hr) to trial 2 (36.65 m^3/hr), but neither an intervening group main effect, $F(2, 57) = .829$, $p = .442$, nor interaction with block, $F(2, 57) = 1.12$, $p = .333$, was significant. Thus, all three groups performed at approximately the same level before they were introduced to the intervening task.

6.4.1.2. Group 1 vs. Group 2

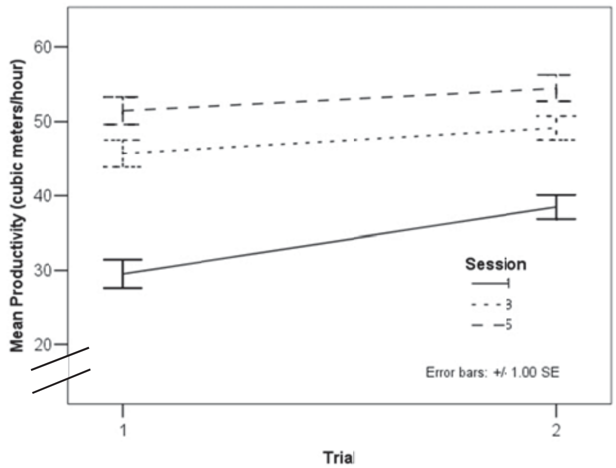
A mixed-design repeated-measures ANOVA (Appendix E) was used to test the effects of session (Session 1 – initial test on excavator simulator, Session 3 – first retention test, and Session 5 – second retention test) and trial (Trials 1 and 2, 7 min each) on productivity per module on the excavator, with group (control group and loader group) as a between-subjects factor. The ANOVA showed a main effect of session, $F(2, 76) = 91.38$, $p < .001$, $\eta_p{}^2 = .706$, with productivity increasing across the three sessions. Trial was also a significant factor, $F(1, 38) = 65.67$, $p < .001$, $\eta_p{}^2 = .633$, showing an increase in productivity from trial 1 to trial 2 within a session. The session × trial interaction (Figure 6.3a) was significant, $F(2, 76) = 9.26$, $p < .001$, $\eta_p{}^2 = .196$, reflecting that the majority of learning occurred in Session 1.

Intervening task only approached the .05 level, $F(1, 38) = 3.47$, $p = .07$. However, both the session × group (Figure 6.3b) and trial × group (Figure 6.3c) interactions were significant, $F(2, 76) = 10.95$, $p < .001$, $\eta_p{}^2 = .224$, and $F(1, 38) = 5.04$, $p < .05$, $\eta_p{}^2 = .117$, respectively. The former (Figure 6.3b) suggests that the control group had a greater improvement from Session 1 to Session 3 and continued to improve to a larger extent than did the loader intervening group. Also, the interaction revealed that the two groups performed at approximately the same level before being introduced to the intervening task ($p > .05$). The trial × group interaction revealed that loader group showed greater improvement in the second trial within each session compared to the control group.

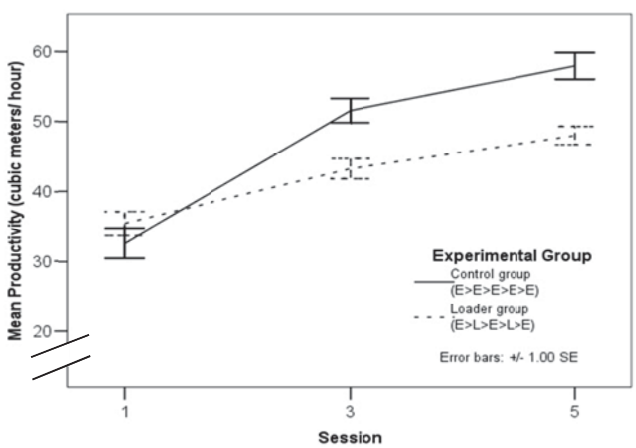
6.4.1.3. <u>Group 1 vs. Group 3</u>

Next, the control group (#1) and long-loader group (#3) were compared in a manner similar to the analysis conducted above. A mixed-design repeated-measures ANOVA was used to test the effects of session (Session 1 – initial test on excavator simulator, Session 4 – first retention test, and Session 5 – second retention test) and trial (Trials 1 and 2, 7 min each) on productivity per module on the excavator, with group (control group and long-loader group) as a

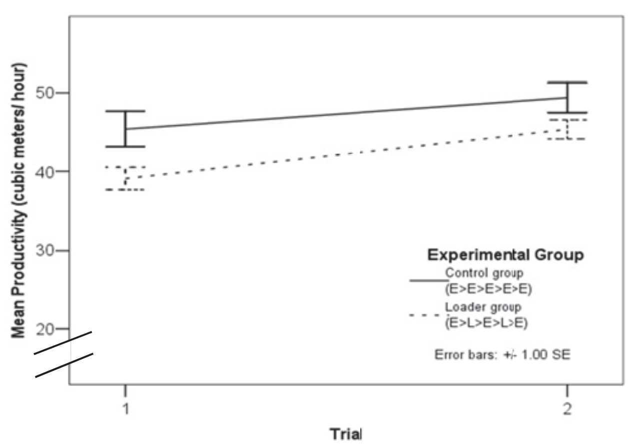
Figure 6.3. Group 1 vs. Group 2: Two-way interaction plots of (a) session × trial, (b) session × group, and (c) trial × group on productivity of each module.


Mean Productivity (cubic meters/hour)
60
50
40
30
20
Session
1
3
5
Error bars: +/- 1.00 SE
1
2
Trial

(a)


Mean Productivity (cubic meters/ hour)
60
50
40
30
20
Experimental Group
Control group
(E>E>E>E>E)
Loader group
(E>L>E>L>E)
Error bars: +/- 1.00 SE
1
3
5
Session

(b)


Mean Productivity (cubic meters/ hour)
50
40
30
20
Experimental Group
Control group
(E>E>E>E>E)
Loader group
(E>L>E>L>E)
Error bars: +/- 1.00 SE
1
2
Trial

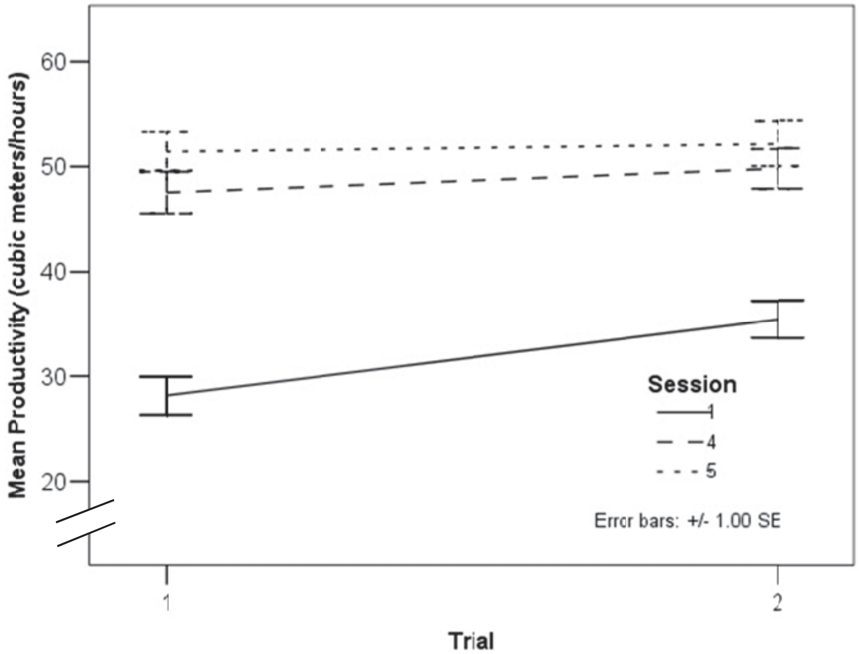
(c)

between-subjects factor. The results (see Appendix E) showed a main effect of session, $F(2, 76) = 106.70$, $p < .001$, $\eta_p^2 = .737$, with productivity increasing across the three sessions. Trial was also a significant factor, $F(1, 38) = 25.16$, $p < .001$, $\eta_p^2 = .398$, showing an increase in productivity from trial 1 to trial 2 within a session. The session × trial interaction (Figure 6.4a) was significant, $F(2, 76) = 5.20$, $p < .005$, $\eta_p^2 = .120$, reflecting that the majority of learning occurred in Session 1.

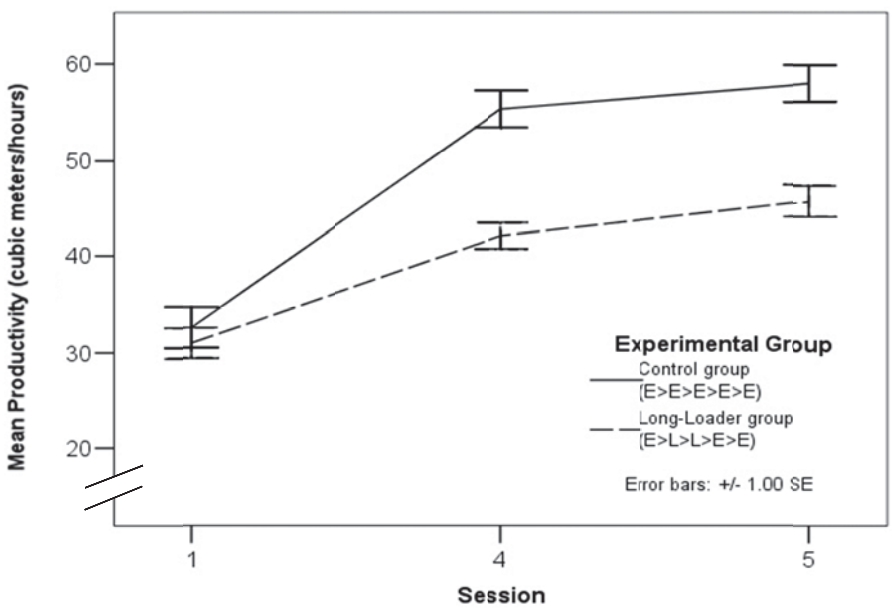
Different from what was found in the comparison between control group (#1) and loader group (#2), where intervening group only approached significance with $p = .07$, intervening group, here, was significant, $F(1, 38) = 9.61$, $p = .004$, $\eta_p^2 = .202$. Session × group (Figure 6.4b) was the only interaction found to be significant, $F(2, 76) = 8.64$, $p < .001$, $\eta_p^2 = .185$, suggesting that the control group had a greater improvement from Session 1 to Session 4 and continued to improve to a larger extent than did the loader intervening group. Also, the interaction revealed that the two groups performed at approximately the same level before being introduced to the intervening task ($p > .05$). Unlike, the trial × group interaction found between the control group and Group 2 (Figure 6.3c), the long-loader group did not show greater improvement in the second trial within each session compared to the control group, no interaction between trial and group was found.

6.4.2. First Three Sessions on Excavator Simulator

To examine whether there is a cost when switching from the alternative machine back to the previously learned one, the performances when the participants were practicing the truck loading task on the excavator for the first, second, and third



(a)



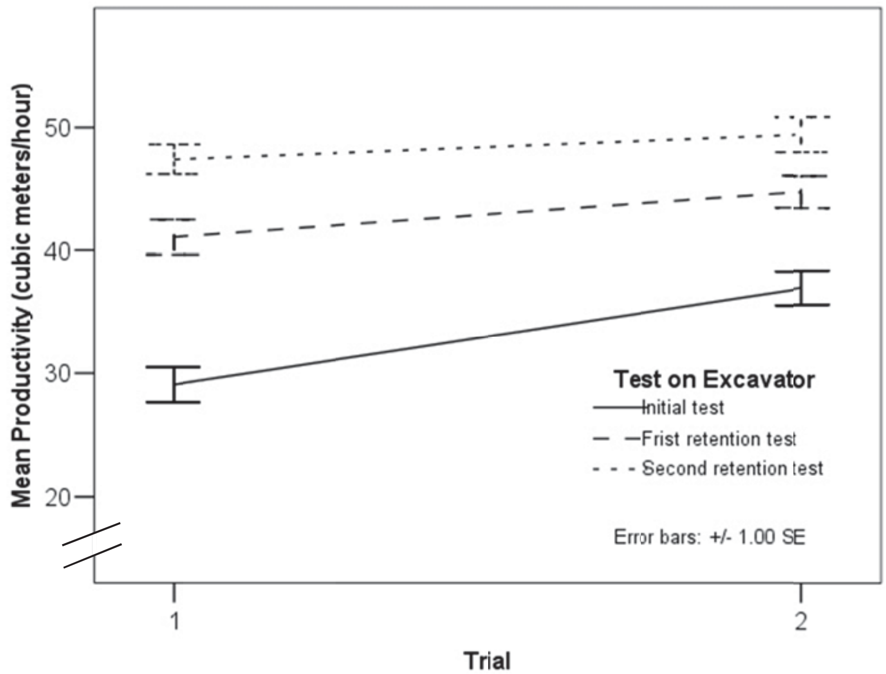
(b)

Figure 6.4. Group 1 vs. Group 3: Two-way interaction plots of (a) session × trial and (b) session × group on productivity of each module.

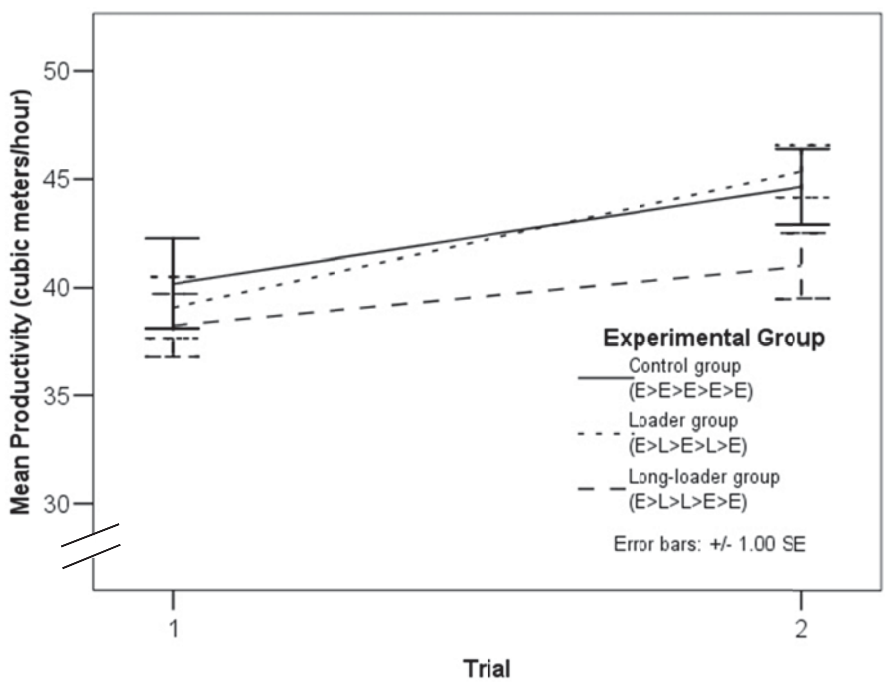
times were also compared. In other words, performance for the control group in Sessions 1, 2 and 3 was compared to performance of the loader group in Sessions 1, 3, and 5 and to that of the long-loader group in Sessions 1, 4, 5. An ANOVA (see Appendix E) with group (control, loader and long-loader) as a between-subjects factor and test (1st, 2nd, and 3rd) and trial (1 and 2) as within-subject factors was conducted on productivity. The ANOVA showed a main effect of session, $F(2, 114) = 108.42$, $p < .001$, $\eta_p^2 = .655$, with productivity increasing across the three sessions. Trial was also found significant, $F(1, 57) = 64.52$, $p < .001$, $\eta_p^2 = .531$, showing an increase in productivity from Trial 1 to Trial 2 within a session. The session × trial interaction was significant as well, $F(2, 114) = 8.88$, $p < .001$, $\eta_p^2 = .135$, corroborating the indications of Figure 6.5a, reflecting that the majority of learning occurred in Session 1 (i.e., the first session on the excavator simulator). Group was not significant, $F(1, 38) = .470$, $p = .628$, but interacted with trial, $F(1, 38) = .3.59$, $p <.05$, $\eta_p^2 = .112$ (Figure 6.5b), indicating a significant difference in Trial 2 between loader and long-loader group ($p > .05$). The insignificant interaction between session and group is shown in Appendix E.

6.4.3. Performance on Loader Simulator of the Two Intervening Groups

The total percentage of each 12-cubic-yard dump truck being filled per 7-min trial was recorded on the loader simulator. Productivity (m^3/hr) in the trucking loading task with the loader was calculated as the total volume transferred from the pile to the truck (converted into m^3), divided by the total execution time (converted to hr). A mixed-design repeated-measures ANOVA was used to test the effects of session (1st



(a)



(b)

Figure 6.5. Two-way interaction plots of (a) session × trial and (b) trial × intervening task on productivity of each module across the first three tests on excavator.

and 2nd time practice on the loader simulator) and trial (Trials 1 and 2, 7 min each) on productivity, with group (loader group and long-loader group) as a between-subjects factor. The results (see Appendix E) showed that both the main effects of session, $F(1, 38) = 109.09$, $p < .001$, $\eta_p^2 = .742$ and trial, $F(1, 38) = 68.12$, $p < .001$, $\eta_p^2 = .642$, are significant. However, group is not a significant factor and none of the interactions were found significant. This could also be seen in a three-way interaction plot shown in Figure 6.6, where the productivity by the two groups followed a similar increasing trend. These results indicated that the two groups practicing on loader showed increasing in productivity from the first session practicing on loader to the second session returning on the loader and also from trial 1 to trial 2 within a session. As shown by the insignificant group effect, practicing on the excavator between the two loader sessions for the loader group (E>L>E>L>E) did not affect their returning performance on the loader, indeed continued to improve the performance on the loader in trials 3 and 4 (Bonferroni pairwise comparisons tests, *ps* > 0.001).

6.4.4. Workload Measures

For the mental workload measures, the six subscales of the NASA TLX for the 5 sessions were analyzed using a mixed design ANOVA with the workload measure and session as within-subject factors and experimental group as a between-subjects factor. The scale for the Performance measure was reversed before the analysis, so that a higher score meant higher workload, but the original data are shown in the plots. The ANOVA (Appendix E) showed a main effect of session, $F(4,228) = 48.08$, $p < .001$, $\eta_p^2 = .458$, with mental workload decreasing across the sessions (*p*s <.001). The main

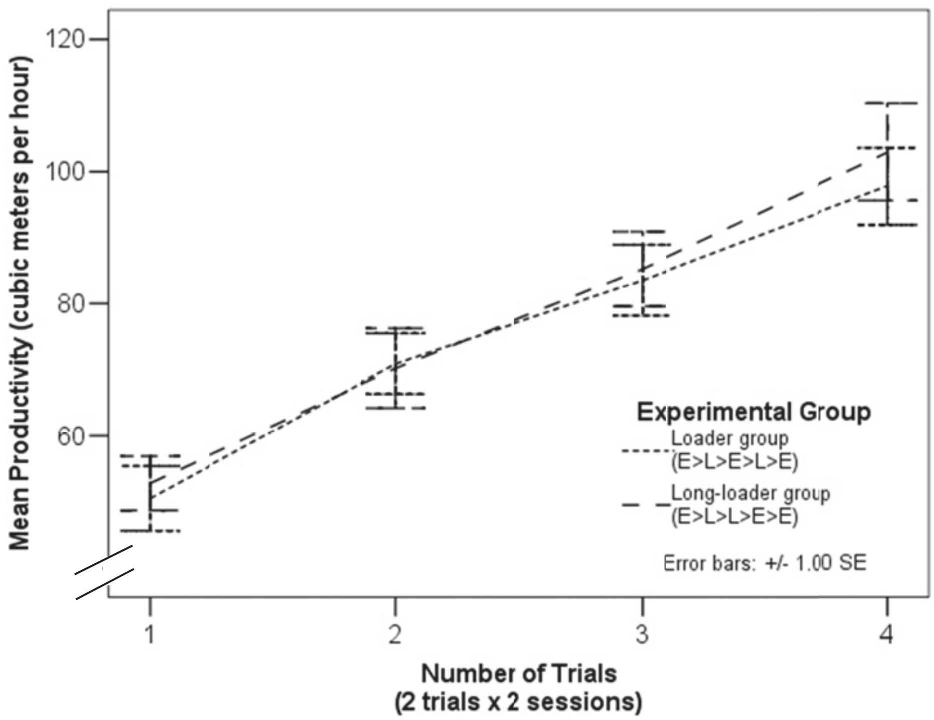


Figure 6.6. Productivity of truck loading task on the loader across the 2 sessions (three-way interaction plot) for loader group and long-loader group.

effect of measure was also significant, where participants rated the tasks as requiring higher effort and mental demand, but lower temporal demand, physical demand and less frustration. Overall, the inverted scale of performance level (the lower, the least dissatisfaction with their performance) received the lowest rating, indicating that the participants were very satisfied with their performance throughout the sessions.

The major interests are the main effect of group and the interaction between session and group, both of which were significant: group, $F(2,57) = 9.481$, $p < .001$, $\eta_p^2 = .250$; session × group, $F(8,228) = 8.27$, $p < .001$, $\eta_p^2 = .225$. Both groups that practiced with the loader showed higher workload than did the control group ($ps < 0.001$). The session × group interaction is plotted in Figure 6.7. The plots of the six subscales (see Appendix E) all follow a similar pattern as reported above, explaining the insignificant 3-way interaction. Three observations were made for each experimental group:

1. The overall workload for the control group continuously decreased across sessions.
2. The loader group showed an increase in workload for sessions 2 and 4, in which they were practicing on the loader instead of the excavator. The workload demand decreased in session 4 compared to session 2 on the loader simulator. The workload demand continued to diminish the next time they returned to the excavator simulator.

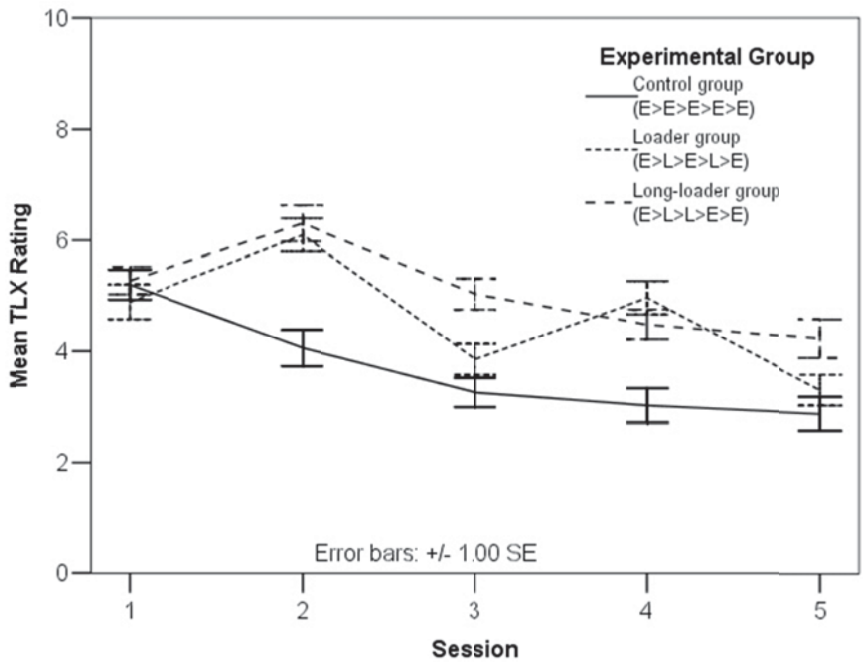


Figure 6.7. Two-way interaction plots of session × group on mean TLX rating across the five sessions.

3. The long-loader group also showed an increased in workload demand in session 2 when the loader practice was introduced. In session 3, in which they continued to practice on the loader, the workload was lower than in session 2. The workload continued to decrease in sessions 4 and 5 when the participants returned to the excavator, but the workload was significantly higher than for the control group.

6.5. Discussion

Participants were given training on loading a truck on either one or two simulated pieces of construction equipment, to examine whether learning to operate a single piece of equipment is best if all practice is on that equipment, or whether intermixed training on a related piece of equipment can be of value given a fixed training period. In this study, those participants who engaged in intervening practice on the simulated loader showed a smaller performance improvement on learning the truck loading task on the simulated excavator than did the control group who practiced on the simulated excavator for all five sessions. This outcome confirms that the control familiarization tasks on both machines studied in the preliminary study may have been too simple for the full effects of switching between the machines to be evident.

Another possible reason for the discrepancy of the results for the two studies is that the controls familiarization modules on the two pieces of equipment are very similar and rely on response selection for task performance rather than on perceptual-motor control of the equipment. Both present to the trainee a virtual scene from the perspective of a person in the cabin of a construction machine. The participants are asked to read the instruction, retrieve the appropriate control action, and then merely activate the correct function. This task similarity may allow practice on the loader to benefit performance with the excavator. In terms of what is called the procedural reinstatement principle (Healy et al., 2012), the procedures acquired in performing on the loader may have been sufficiently similar to those on the excavator to allow complete transfer between machines. Alternatively, it may be that the proportion of learning that occurs after the first session is so small that response time is at a floor.

By comparing the performance of the three groups for the first, second, and third times they carried out the truck loading tasks on the simulated excavator, no differences were found in all the three test sessions. This result suggests that, although there was no benefit of practicing on the loader, there was also no negative transfer. In other words, the skills learned previously on the excavator simulator were retained even after they learned and practiced on the loader simulator. The results also supported *Hypothesis 2* that the control group did continuously benefit from practicing on the same task throughout all session with significant dropping of mental workload measures.

Compared to the Controls Familiarization modules of Study 1, the truck loading modules in this study required more complex perceptual-motor skills to navigate and maneuver the vehicles and fine motor skills to handle the implement through large ranges of motion. In contrast to Study 1, the truck loading modules showed a significant interaction between session and group. Although no negative transfer was found, as proposed in *Hypothesis 3*, the reasons for the differences in performance of the control and the two loader groups in the retention tests may still hold. First, although the chosen tasks for both machines share a similar goal of loading a dump truck by means of filling the bucket and transferring material to the truck bed, they involve different subgoals/steps to achieve this goal (Proctor, Dunston, So, & Wang, 2012): 1) Performing the task with the loader involves driving to move from the stockpile to the truck bed each cycle, whereas performing it with the excavator does not. 2) The excavator is stationary and only requires being driven when the trenching position is no longer optimal to fill the bucket. 3) The excavator has higher degrees of freedom because the bucket location is controlled by both stick and boom, whereas the loader is only controlled by the boom. 4) An excavator operator is required to move the boom, stick and bucket concurrently in order to control the bucket movement efficiently. Second, the control configurations are not the same. The simulated excavator consists of a stick, boom, bucket and cab on a rotating platform sitting atop an undercarriage with tracks, controlled by two joysticks (left-hand and right-hand) and pedals (for driving). The simulated loader, however, is wheeled, has a bucket connected to the end of a pair of boom arms, and travels with its load from one location to another, all controlled by one joystick, accelerator and brake pedals, and a steering wheel.

Comparing the loader and long-loader group, the group effect between the control group and loader group was marginally significant (p = .07), suggesting that the 20-min intervening task may not be long enough to show a significant main effect of performance cost. The results also partially support *Hypothesis 4* that the long-loader group performing on the loader in Sessions 2 and 3 before returning to the excavator practice had larger interference with learning, but not negative impact between the initial and retention test. This could be confirmed by the results that the long-loader group did show continuous improvement, but at a significantly smaller amount compared to the control group. Because the participants still continued to improve in the retention test for both intervening (loader) groups, it leads to another interesting question of whether the differences in improvement between the two groups could be reduced by alternating the practice sequence. Also, by examining the performance on the loader, it was found that practicing on the excavator between the two loader sessions for the loader group (E>L>E>L>E) did not deteriorate their returning performance on the loader, but indeed they continued to improve their performance when returning to the loader. This result suggests that the truck loading task on these two machines does share some components which may assist/ facilitate their learning from one machine to the other. The next study aims at understanding how the truck loading tasks are performed by interviewing experienced operators and having them verify the HTA of the truck loading tasks on each machine.

6.6. Conclusion

The main finding of this study is that no cost or benefit was found from inserting practice on a simulated loader while also learning on a simulated excavator for a complex task—truck loading. The group whose practice on loading a truck with an excavator was broken up by intervening practice of the same task with a loader continued to show improvement when returning to the excavator, showing neither positive nor negative transfer compared to the control group. An implication of these findings for training is that if a trainer wants to maximize learning to operate a machine during a finite time period, practice should be devoted to that machine, whereas if the trainer wants to provide experience with two machines, this can be done without the practice on one machine having a negative effect on the learning of the other. An inference of the present study is that similarity in the overall goals of the tasks, e.g., truck loading, is less important than similarities among the subgoals that comprise the tasks as performed on the respective equipment types. Detailed task analyses should reveal common elements that define the essential similarities at various levels in the overall task structures.

CHAPTER 7. STUDY 3: HIERARCHICAL TASK ANALYSIS OF TRUCK LOADING TASKS

Studies 1 and 2 revealed no performance cost (diminished subsequent performance) attributable to inserting practice on the loader while also learning on the excavator for a more complex task - truck loading. Given a fixed amount of total training time, the two groups whose practice was intervened by the practice of a similar task with a loader continued to show improvement when returning to the excavator, but at a significantly smaller amount compared to the control group. In other words, practice with the loader between the excavator sessions did not alter the excavator learning, as performance picked up at the level of the prior excavator session and continued to improve. The question of what caused the loader group to improve less when returning to the excavator compared to the control group becomes of interest. Are there any relationships between the requirements of the tasks, and operation of the equipment, which influence performance of one subsequent to performance of the other, i.e., what is the nature of transfer across machine types?

7.1. Objectives

In this study, Hierarchical Task Analysis (HTA), a well-accepted and developed form of task analysis (Annett, 2004; Stanton, 2006), is employed for illustrating the complexity of equipment operation and distinguishing the skills to be acquired for each machine. HTA analyzes both the task goals and the sequences of actions involved in executing tasks, and it can be extended to make explicit the cognitive demands and design requirements (Phipps et al., 2011). The HTA should clarify the extent of similarity and differences between the two cases, which might in turn indicate effective common (i.e., transferable) methods to facilitate skill development (Proctor et al., 2012).

The purpose of this study is to conceptualize and analyze transfer through the degree of overlap of specific task components by studying the similarity and dissimilarity of the truck loading task performed in Study 2 on excavator and wheel loader simulators. Tasks performed through operation of different equipment types but having similar goals are analyzed for the purpose of this phase of investigation. To study the similarity and dissimilarity of the truck loading task on the excavator and wheel loader simulators, a detailed comparative analysis of truck loading tasks for these two machines using two approaches are conducted:

1. Direct observation:
    i. Studying in-depth the controls and motion constraints of the two machines
    ii. Developing preliminary HTAs of truck loading task
2. Knowledge elicitation from experienced operators:
    i. Interviewing experienced operators to elaborate the HTAs and elicit common knowledge shared by the two machines
    ii. Having other experienced operators verify the final HTAs

7.2. Comparison Between Loader and Excavator Simulators

Both machines perform the same general task of digging and placing soil in a different location, but they do so with different mechanisms and motion constraints. Chapter 3 already provided the details of the two simulators used in this research, including experimental setup, controls and functions, and the features of the training modules incorporated in the two simulators. In this section, a detailed comparative analysis examining the similarities and dissimilarities between the controls and the motion constraints of the two machines, as well as the truck loading scenarios presented on these two machines are discussed.

7.2.1. Controls and Functions of Both Machines

The simulated hydraulic excavator consists of a stick, boom, bucket and cab on a rotating platform sitting atop an undercarriage with tracks, controlled by two joysticks (left-hand and right-hand) and pedals (for driving). The left joystick controls the stick movement and the rotation of the cab, whereas the right joystick controls the boom and bucket movement. The simulated loader, however, is wheel-mounted, turns by means of a hydraulically actuated pivot point in the loader frame between the front and rear axles (i.e., articulation), and has a wide front mounted bucket connected to the end of two boom arms to scoop up loose material, such as dirt, sand or gravel, and carry it from one location to another. Similar to the excavator, the boom and bucket are controlled by the right joystick, but the loader is driven by accelerator and brake pedals, a steering wheel and pressing the FNR button to reverse direction. The loader operators are required to have their left hand stay on the steering wheel to direct to the right location while having the right hand holding the joystick to control the bucket motions.

The FNR button (which reverses direction) is attached in the front of the joystick controlled by index finger and middle finger for reserve direction.

As illustrated in Figure 7.1, the excavator has higher degrees of freedom because the bucket location is controlled by both the stick and boom, whereas the loader is only controlled by the boom. Thus, to control the bucket movement efficiently, an excavator operator needs to move the boom, stick and bucket, concurrently. In terms of the number of joints of movement of the buckets, the excavator can move and rotate in 4 directions, where it is allowed to 1) rotate in a close or dump position, 2) reach in or extend out by its stick, 3) raise up or down by its boom and 4) swing left and right with the rotating platform, whereas the loader can only do 1, 3 and 4. A comparison showing the similarity and dissimilarity of the machine constraints and controls is summarized in Table 7.1.

7.2.2. Truck Loading Scenarios

In the excavator scenario (Figure 6.1a), the operator starts from the position away from the trench. The operator drives the excavator to a parked position in line with the trench, then loads the bucket by extending and angling the bucket for executing a smooth pass through the soil. After the bucket is filled, the bucket is also curled toward the machine to ensure that the soil is contained and swung over to the truck bed (on the left). Because the excavator and truck are on the same ground level, the bucket must be raised to an appropriate height to clear the sides of the truck bed. Once over the truck bed, the bucket is uncurled to release the soil before the machine is rotated back to place the bucket above the trench for the next digging pass.

(a) (b)

Figure 7.1. The motion constraints of (a) excavator and (b) loader.

Table 7.1. Comparisons of Machine Constraints and Control Configurations Between Excavator and Loader

Characteristics	Excavator	Loader
Machine constraints		
Number of joint rotations	Four	Three
Cab on a rotating platform	Yes (allows 360 degree turn)	No (allows 80 degree turn only)
Driving with tracks/wheels	Tracks	Wheels
Controls configuration		
Number of joysticks	Two (left and right)	One (right-handed)
Bucket close	Right - left (joystick - direction)	Right – left
Bucket dump	Right – right	Right – right
Boom up	Right – backward	Right – backward
Boom down	Right – forward	Right – forward
Stick in	Left – backward	/
Stick out	Left – forward	/
Swing left	Left – left	/
Swing right	Left – right	/
Return to dig	/	Right – push left twice
Return to carry	/	Right – push front twice
Boom height kickout	/	Right – pull back twice
FNR	/	attached on the front of the joystick
Driving controls	Two pedals (left and right)	Steering wheel, accelerator and brakes

In the loader scenario (Figure 6.1b), the operator starts from the position away from the stockpile. The operator has to first drive to the stockpile to load the bucket by driving squarely towards the stockpile while lowering the bucket to ground level for cutting from the base of the stockpile. Then the operator drives the bucket into the stockpile and next simultaneously lifts and curls the bucket upward to contain the soil. After the bucket is filled, the operator backs the loader away from the stockpile while turning to bring the truck into view. Next the loader is driven squarely towards the truck bed while raising the bucket to dump over the side of the truck bed and then uncurling the bucket. After emptying the bucket, the cycle is then completed with backing away from the truck while lowering the bucket once again to travel height.

At this stage of comparison, the following can be seen. 1) Performing the task with the loader involves driving to move from the stockpile to the truck bed each cycle, whereas performing it with the excavator does not. 2) The excavator is stationary and only requires being driven when the trenching position is no longer optimal to fill the bucket. 3) Loading bucket with an excavator is a downward trenching motion, whereas loader fills the bucket upward.

In summary, through this stage of direct observation of the truck loading task with these two machines, it is noted that:

1) The chosen tasks for both machines share a similar goal of loading a dump truck by means of filling the bucket and transferring material to the truck bed.

2) The simulated excavator consists of a stick, boom, bucket and cab on a rotating platform sitting atop an undercarriage with tracks, controlled by two joysticks and pedals. The simulated loader, however, is wheeled, has a bucket connected to the end of a pair of boom arms, and travels with its load from one location to another, all controlled by one joystick, accelerator and brake pedals, and a steering wheel.

3) The excavator has higher degrees of freedom (number of rotating joints). Hence, the excavator can move and rotate the bucket in four directions, whereas loader can only move in three directions.

4) Performing the task with the loader involves driving to move from the stockpile to the truck bed each cycle, whereas the excavator is stationary, thus no driving is involved during bucket filling. In other words, driving and controlling the bucket movement are in parallel when performing truck loading with a loader, whereas driving is a sequential task using the excavator only when the trenching position is no longer optimal to fill the bucket.

The next section will start to develop HTAs, for modeling the tasks in the form of goals, subgoals, and sub-operations. By identifying elements (layout in the HTAs) that tasks have in common, the HTAs can suggest where benefits of training may transfer.

7.3. Task Analysis of Truck Loading Tasks

The task analysis of the truck loading tasks on two machines: excavator and loader, conducted here involve four major steps. First, two preliminary HTAs were developed by researcher's self-observation. Second, the preliminary HTAs were examined by experienced operators. Third, after consolidation of the comments from the experienced operators, revised HTAs were developed. Lastly, additional independent experienced operators were involved in a final stage of confirming the HTAs.

7.3.1. Develop Preliminary HTAs

The preliminary HTAs of truck loading tasks for a hydraulic excavator and those for a loader presented in Figure 7.2 (a) and (b), respectively, are based on the direct observation of the truck loading task through studying the user manuals and training videos for the two pieces of equipment. Figure 7.2 (a) and (b) and the discussion on the development of the preliminary HTAs here have been published in Proctor et al. (2012). The HTAs here followed the methodology stated in *Step 4: Acquire Data and Draft a Decomposition* in Annett et al.'s (2004) article. The HTA diagram employed the method of notation from Annett, Cunningham, and Mathias-Jones (2000). The overall goal (0) is at the top of the hierarchy, with the main subgoals located immediately underneath. Some of these subgoals are decomposed into a second level of subgoals. The boxes in the diagram are numbered in an outline structure, i.e., with subgoals inheriting the number of their parent goal plus a period and new ordinal number. Also, the 'Plan' specified in the ovals shows the conditions under which each of the subgoals are triggered. The symbol ">" is used when subtasks

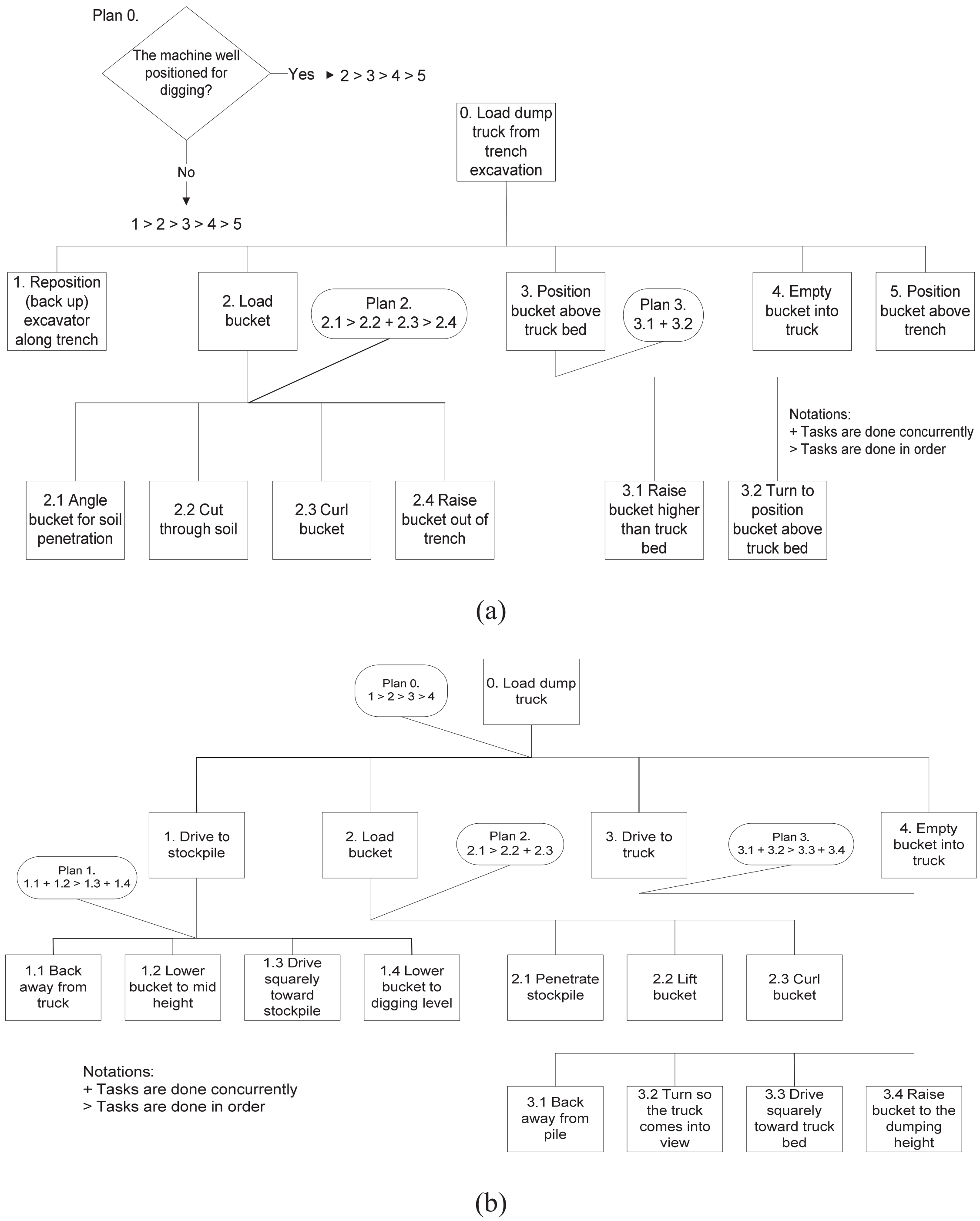


Figure 7.2. Preliminary HTAs for (a) excavator digging a trench and loading a truck and (b) loader transferring soil from stockpile to truck.

are performed sequentially, "+" is used if subtasks are performed in parallel, "/" to represent if either/or subtask is needed to perform, and ":" to represent multiple operations in which timing and order are not important.

7.3.2. Verify Task Analysis With Experienced Operators

The next step was to refine these analyses by having experienced heavy equipment operators evaluate the analyses and provide feedback in order to revise the hierarchies to capture the task structure more accurately.

7.3.2.1. Participants

Through contacts with the User Experience group at the Moline Technology Innovation Center of Deere & Company, a total of 14 machine evaluators from John Deere Dubuque Works, Dubuque, Iowa, experienced in the operation of the wheel loader, excavator or both, were invited in participating this study. The participants were invited based on the availability of their work schedules and followed IRB Human Subject Protocol #1304013518 (Appendix F). The first 11 participants were assigned to comment on the HTAs while the last three participants were assigned to verify the final HTAs. The demographic information of the machine operators is summarized in Table 7.2.

7.3.2.2. Experimental Setup

This study was conducted at the Virtual Reality Lab of John Deere Dubuque Works, in Dubuque, Iowa. The two simulators used in the study were John Deere's PC-based Excavator Simulator, equipped with 60-in. Mitsubishi DLP TV monitor, which simulates a John Deere 200D excavator (for more details of the John Deere's excavator simulator, please refer

Table 7.2. The Demographic Information of the Operators

Operator #	Age	Loader experience (years)	Excavator experience (years)	No. of equipment known	Average years of experience
1	37	15	2	4	10
2	38	2	1	3	2
3	35	3	3	5	5.5
4	40	30	30	5+	30
5	27	1	1	4	2
6	45	2	2	2	2
7	25	1	1	4	1
8	24	1	1	4	1
9	31	3	5	4	12
10	30	5	3	4	3
11	36	11	11	4	11
12	28	2	1	3	6
13	36	20	15	4	20
14	26	5	1	3	2

http://www.deere.com/en_US/services_and_support/training_and_safety/excavator_simulator.page), and John Deere's PC-based 4-Wheel Drive Loader Simulator, equipped with 60-in. Samsung LED TV monitor, which simulates a John Deere 544K 4WD Loader (similar model illustrated in Chapter 4). The truck loading module on the loader simulator presented to the machine evaluators is the same as the one used in Study 2 (which presented to the college students). However, the truck loading module on the John Deere's excavator simulator is slightly different from the Simlog PC-based excavator simulator, for which, the one presented on the John Deere excavator simulator is loading from the bench instead of loading at the same level of the truck. Because the HTAs which developed in this study are both based on the tasks used in Study 2, the machine evaluators were made aware that although the truck loading module presented on the John Deere excavator simulators was from the bench, the initial HTA was based on both the truck and the excavator being on same ground level.

7.3.2.3. Procedure

After signing the consent form, the operators were informed of the study's aim and that the goal was to refine the HTA for a truck loading task on both excavator and loader simulators. A preliminary questionnaire (see Appendix G) obtaining demographic information was administered before the first session began. In Session 1, the machine evaluators were sitting on the simulators presenting the truck loading task. They were first explained with the goal of the truck loading task and asked to try out the truck loading module on the simulator to become familiarized with it. In Session 2, they were given and explained the preliminary HTA diagrams. They were asked to comment and revise the given HTA from their experience and their understanding of the tasks. In Session 3, a post questionnaire was administered, consisting of questions regarding the difficulty of conducting truck loading task using both simulators:

1. Which machine is more difficult for truck loading task?
2. What is the major difficulty you encountered with the excavator simulator or loader simulator, or both (in terms of the nature of the controls and tasks)?
3. What features of the simulator (both excavator and wheel loader) do you think were counterproductive to your learning during training?

7.3.3. HTAs Revision

After the data collection was completed with 11 machine evaluators (#1-11), the comments on the two HTAs were compiled summarized in Tables 7.3 and 7.4. The modifications were then added accordingly to the preliminary HTAs for the excavator and loader for final stage of verification with additional operators. If different approaches were suggested among the operators, the ambiguity was evaluated by three independent experts (#12, 13 and 14) in order to verify the resulting task analyses.

The procedure for the three 'Independent Operators' was similar to that for the previous participants in this study, where they first signed the consent form, were informed of the study's aim, answered a preliminary questionnaire and participated in Sessions 1 – 3. In Session 2, however, they were given the 'modified HTAs' instead of the preliminary version and asked to confirm and verify the accuracy of the refined hierarchies. When different approaches were suggested among the previous operators, the 3 "Independent Operators' were asked to judge which is the most common way that operators do. The verification comments are presented in the fourth column (most right) of Tables 7.3 and 7.4. The finalized HTAs are shown in Figures 7.3 and 7.4.

7.4. Comments From the Operators

After the experienced operators commented on and revised the given HTA, a post questionnaire was administered, consisting of questions regarding the difficulty of conducting truck loading task using both simulators. The questions and answers are summarized in Table 7.5.

Table 7.3. Summary of the Operators (Numbers 1–14) Reporting Comments on the HTA of Truck Loading Task on Excavator Simulator

Modification	Modification based on Preliminary HTA	Operator #	Modification before verification	Final modification after verification with operators #12, 13, and 14
	Step 2			
M1	'Position bucket above trench' is missing, which consists of 'Position bucket (Swing) above trench' and 'Lower bucket for soil penetration' before Step 2.	1,3,8	A separate goal (2. 'Position bucket above trench') was added. Two subgoals (2.1 'Swing above trench' and 2.2 'Lower bucket') were added under it.	Yes
	Step 3			
M2	3.2 cut through soil + 3.3 Curl bucket	1,3,4,5,6,7,8,9,10,11	Require confirmation which one is the most accurate description.	The operators have agreed on grouping them into three main subgoals (i.e., 3.1 'Lower bucket', 3.2 'Fill bucket' and 3.3 'Raise bucket' and the subgoals of each goal are verified as illustrated in the final HTA.
M3	3.2 Cut through soil + 3.3 Curl bucket + 3.4 Raise bucket out of Trench	4,6,7,9,10,11		
M4	3.2 Cut through soil + 3.3 Curl bucket + 3.4 Raise bucket out of trench + add 'Arm in to pull the bucket towards the operator'	9		
M5	3.3 reword truck bed to truck box	1,4,7	Reworded	Yes

(continued on next page)

Table 7.3 continued.

Modification	Modification based on Preliminary HTA	Operator #	Modification before verification	Final modification after verification with operators #12, 13, and 14
	Step 4			
M6	4.1 Raise bucket higher than truck bed + 4.2 Turn to position bucket above truck bed , then add 'arm out' + ‘curl bucket’	2,9,10,11	Two additional subgoals (4.3 ‘Arm out’ and 4.4 ‘Curl bucket’) were created, following this sequence: 4.1+4.2>4.3+4.4.	Yes
	Step 5			
M7	Expand Step 5: Arm out + dump	7,9,11	Require confirmation which one is the most accurate description.	Arm out + dump is sufficient, Swing movement is not necessary.
M8	Expand Step 5: Arm out + dump + add 'Swing'	11		

Table 7.4. Summary of the Operators (Numbers 1–14) Reporting Comments on the HTA of Truck Loading Task on Loader

Modification ID	Modification	Operator ID	Modification before verification	Final modification after verification with operators #12, 13, and 14
	Step 1			
M1	Move 1.1 ‘Back out from the truck’ and 1.2 ‘Lower bucket to mid height’ to last step	1, 2,3, 4, 6, 7, 9, 11	Step 5 “Back out from the truck” was added to the HTA.	Yes
M2	Reword 1.2 as 'lower bucket to carrying height'	6,10,11	Reworded.	Yes
M3	Add 1.3 'Keep the cutting edge of the bucket flat’ before 1.3 ‘Drive squarely toward stockpile’ + 1.4 ‘Lower bucket to digging level’ +	8	The additional step “keep the cutting edge of the bucket flat” was added.	Yes
M4	1.3 ‘Drive squarely toward stockpile’ + 1.4 ‘Slowly lower bucket to digging level’	1,4,7,8,9, 11	These two steps (originally 1.3 and 1.4) were renumbered as 1.1 and 1.2.	Yes
M5	Add 'Curl the bucket/ return to dig' before 1.1 ‘back away from the truck’	5,6,8,9, 10,11	This step was added to Step 5.	Yes

(continued on next page)

Table 7.4 continued.

Modification ID	Modification	Operator ID	Modification before verification	Final modification after verification with operators #12, 13, and 14
	Step 2			
M6	2.2 Lift bucket + 2.3 Curl bucket	All	Referring to M5, Step 2 included 2.1 “Keep the cutting edge of the bucket flat for penetration” and 2.2 “fill the bucket” consisting of four sub-operations. The original 2.1, 2.2 and 2.3 were renumbered as 2.2.1, 2.2.2 and 2.2.3, respectively. An additional step 2.2.4 “Apply crowd force” was added. This step requires verification to confirm the order of Steps 2.1.1 to 2.1.4.	The order for 2.2 “fill the bucket” is 2.2.1+2.2.2 > 2.2.2+2.2.3+2.2.4
M7	2.2 Lift bucket + 2.3 Curl bucket + Add “drive in (apply crowd force)”	2,3,7,8,9, 10,11		
	Step 3			
M9	3.1 Back away from pile + 3.2 Turn so the truck come into view	2,3,5,6,7, 8,9,10, 11	Require confirmation which one is the most accurate description.	M9 is the most accurate.
M10	3.1 Back away from pile + 3.2 Turn so the truck come into view + Add 'Raise boom'	7, 11		
M11	3.1 Back away from pile + 3.4 Raise bucket to the dumping height	4		

(continued on next page)

Table 7.4 continued.

Modification ID	Modification	Operator ID	Modification before verification	Final modification after verification with operators #12, 13, and 14
M12	3.3 Drive squarely toward truck bed + 3.4 Raise bucket to the dumping height	3,6,8,9, 10,11	Require confirmation which one is the most accurate description.	M12 is the most accurate. M13 depends on operator's habit
M13	3.3 Drive squarely toward truck bed + 3.4 Raise bucket to the dumping height (*** Use Boom Kickout)	10		
M14	Add "Lower the bucket in carrying position" while backing out	3,6,7	A new subgoal is created and added to Step 3.	Yes
M15	3.3 reword truck bed to truck box	1,4,7	Reworded	Yes
	Step 4			
M16	Expand Step4. Boom up & Bucket Dump	6,7,9	Require confirmation which one is the most accurate description.	Agree on M20
M17	Expand Step4. Boom up & Bucket Dump + driving forward	7		Yes

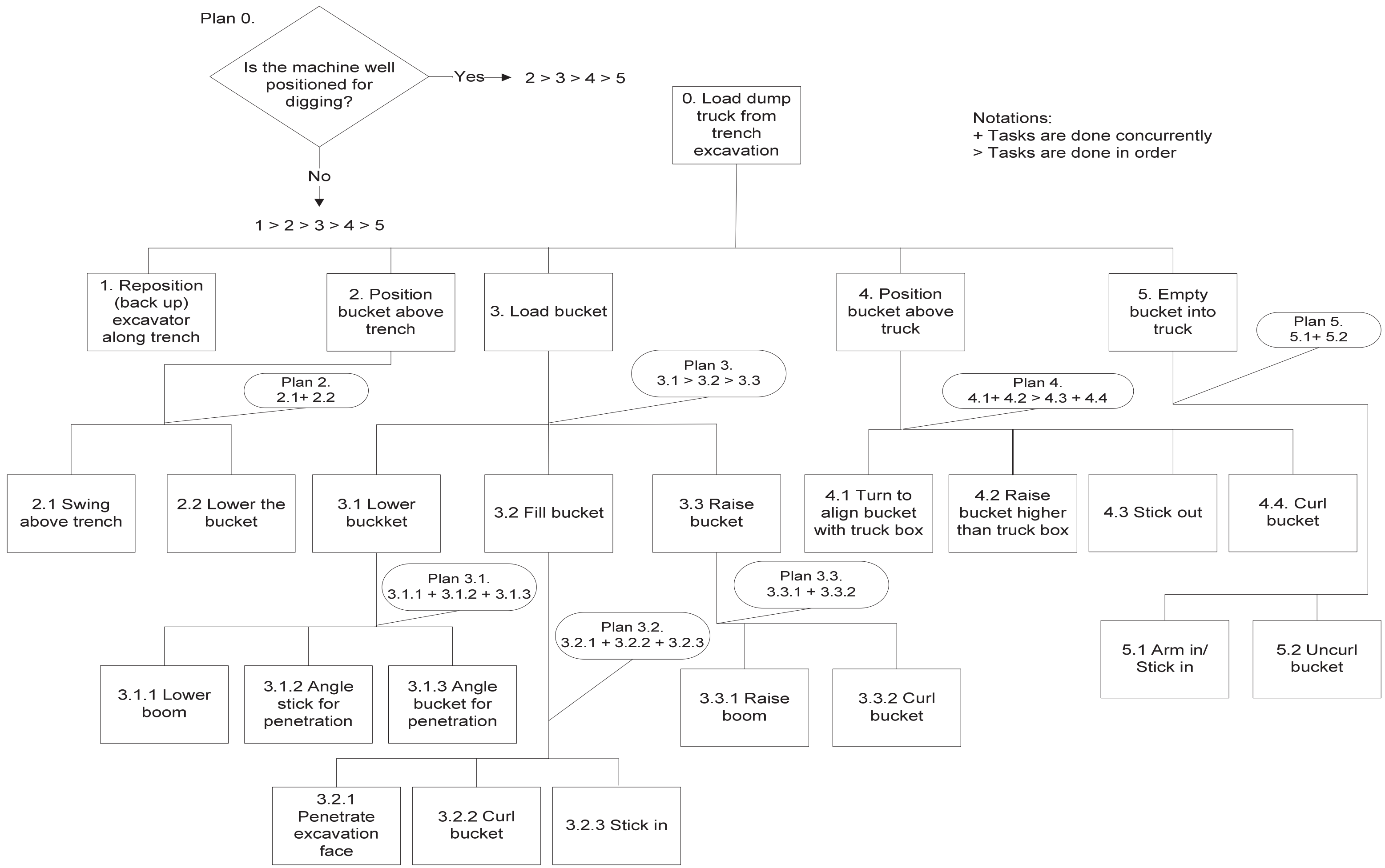


Figure 7.3. Finalized HTA for excavator digging a trench and loading a truck.

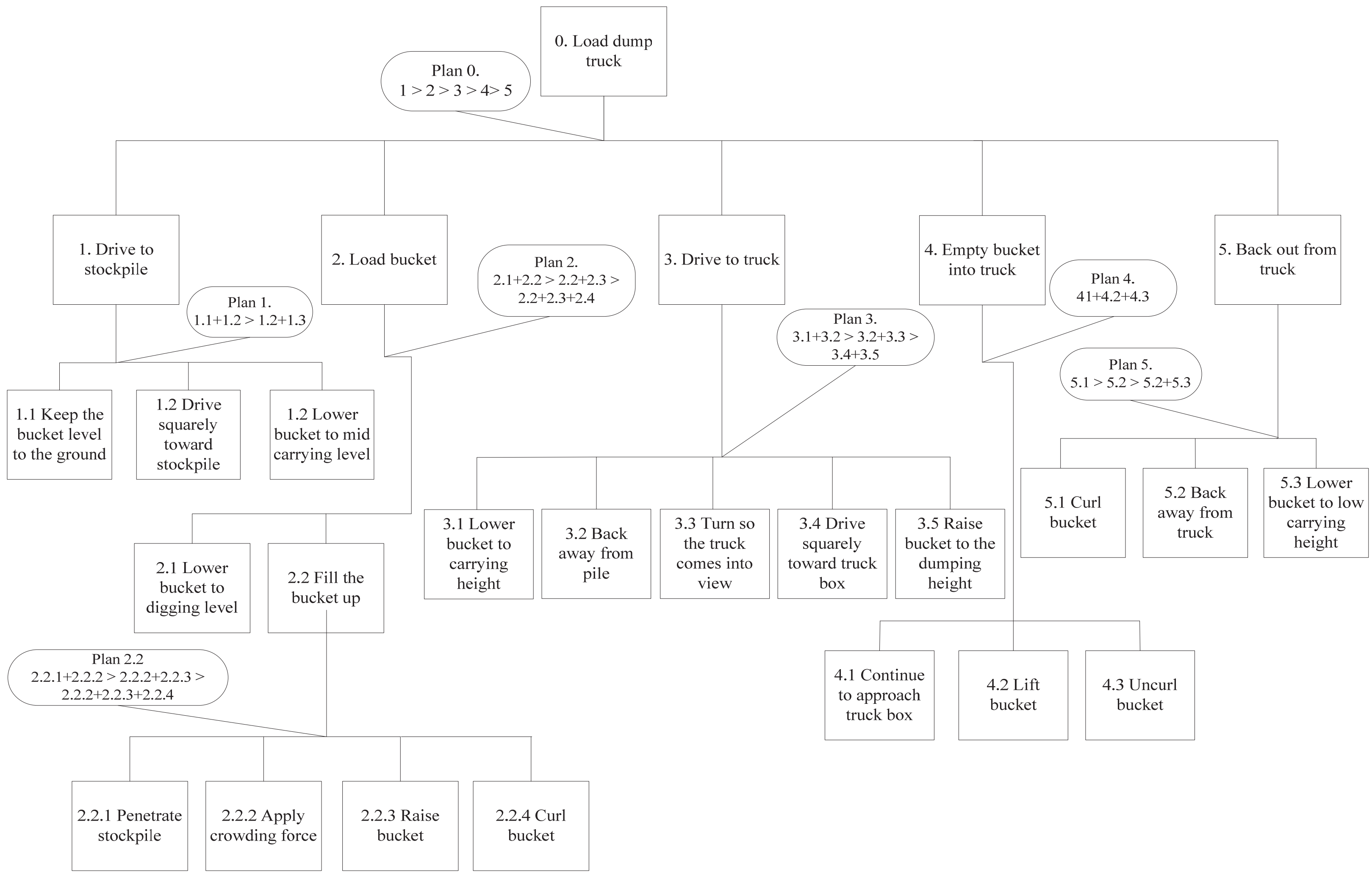


Figure 7.4. Finalized HTA for loader transferring soil from stockpile to truck.

Table 7.5. Responses Collected in the Post Questionnaires in Study 3

Questionnaires	Responses from 14 operators
1. Which machine is more difficult for truck loading task?	• 7 operators indicated that operating loader is more difficult. Their reasons included: operating loader requires a lot of turns (more sub-steps) and stops; timing is important from the stockpiles to truck to avoid collision with the truck; loader is more mobile requiring to look forward and backward, whereas excavator is more stationary and could swing with more clearance. • 3 operators said that operating the excavator is more difficult because of the simultaneous control of the two joysticks and requiring good timing when to raise the bucket. • 4 operators did not decide which one is more difficult.
2. What is the major difficulty you encountered with both the excavator simulator and the loader simulator? (in terms of the nature of the controls and tasks)	• 6 operators indicated that there was an issue of depth perception, where they found it is difficult to judge the distance between the truck and the machine. • 5 operators mentioned that the simulators have no tactile feedback, such as force feedback, feel of the machine, etc. • 3 operators indicated the limited peripheral vision, e.g., they cannot look at the shoulder's view.
3. What features of the simulator (both excavator and wheel loader) do you think were counterproductive to your learning during training? Why?	• In addition to the three difficulties mentioned in Question 2 --- lack of depth perception, absence of tactile feedback, and limited peripheral views, two operators indicated that the pedal is lighter and less sensitive compared to a real machine.

7.5. Discussion

HTA was employed in this study to illustrate the complexity of the truck loading tasks and distinguish the skills to be acquired for both excavator and loader. From the comparison of the HTAs depicted in Figures 7.3 and 7.4, a number of observations are made below:

First, the structures of subgoals (or subtasks) involved in each task as revealed by the HTAs suggest specific skills that need to be taught. For example, the proper alignment of the loader on its approaches is critical for loading the bucket from the piles or unloading the bucket to the truck; the proper angling of the excavator bucket is critical for efficient digging. Such insights from the HTAs allow research to evaluate the extent to which the tasks practiced in the training program should emphasize the component skills that need to be mastered.

Second, the HTAs, as expected, clarify the extent of similarity and differences between the two cases, in which the analysis suggests that skill at subtasks in common between tasks may transfer from one to the other. Although the two HTAs revealed that both truck loading tasks have the same goals – truck loading – and some of the same subgoals (e.g., empty bucket into truck, load bucket), the sub-operations are different. This may explain why little transfer or no transfer was found in Study 2 because the sub-operations level is most critical to learning since they differ across the loader and excavator. For example, through direct observation, it is noted that loading the bucket using both the excavator and the loader involve lowering the bucket, filling the bucket and raising the bucket. However, by comparing the goals and their subgoals of the two machines, HTAs provide a clearer picture to distinguish what are the dissimilarities and what makes them different from each other. In the HTA for the excavator in Figure 7.3, it was shown that the Goal 2 “Load bucket” is broken down into three subgoals: 2.1 “Lower the bucket”, 2.2 “Fill the bucket” and 2.3 “Raise the bucket” which are performed in a serial manner, whereas, the Goal 2 “Load bucket” in the loader is accomplished by two subgoals: 2.1 “Lower the bucket” and 2.2 “Fill the bucket”. By taking a closer investigation of their sub-operations of Subgoal 2.2, filling the bucket using the loader is accomplished by curling and raising the bucket at the same time (i.e. in a parallel (concurrent) process). In other words, the bucket is raising up while curling it during the filling process. It therefore explains the reason why there is no separate “Raise the bucket” sub-operation.

Another issue with the current HTAs is that the simultaneous movements of the controls are not captured in the HTA because of our focus on the bucket. For example, the excavator has higher degrees of freedom because the bucket location is controlled by both stick and boom, whereas the loader is only controlled by the boom. To control the bucket movement efficiently, an excavator operator needs to move the boom, stick and bucket concurrently. The functions of these components are not captured in the HTA because of our focus on the bucket. Extending the HTA to further levels of subgoals would begin to reveal these complexities.

One of the questions concerning use of HTA is “how to know when to stop an analysis”. As pointed out by Hoffman and Militello (2007, pp. 73), HTA can go into ever more detail to involve conditional dependencies (e.g. muscle movement). The ultimate stop rule, though, is just “stop when you have all the information you need to meet the purposes of the analysis” (Annett, 2003; 2004). In this chapter, the HTAs depicted in Figures 7.3 and 7.4 terminated with focus on the movements of the bucket, before involving the descriptions of the movement of controls. At that point, the levels of the HTAs were sufficient to reveal the similarity and dissimilarity of the component skills to be acquired for each task. The current analysis shows that having the same goals and subgoals does not guarantee facilitation of skill development and transfer, but the sub-operations level to achieve the goals and subgoals is most critical to learning and skill transfer. Also, the HTAs in the current analysis alone were not enough to capture the relative difficulty of the different tasks. The results in Study 2 showed that practicing on loader led to significantly higher overall mental workload comparing to the excavator. Seven operators indicated that operating the loader is more difficult than operating the excavator, whereas three operators thought the opposite and four were undecided. Simply comparing the number of subgoals or subgoal levels is not enough to derive which task is more difficult. For example, to control the bucket movement efficiently, an excavator operator needs to move the boom, stick, and bucket concurrently. For the loader, the right hand is used to control the boom and bucket, while the other hand is controlling the direction of the machine. The functions of these components are not captured in the depicted HTA. Extending the HTA to further levels of cognitive subgoals (e.g., Phipps et al., 2011) may help reveal these complexities, but this is out of the scope of this study and should be an objective of future research.

7.6. Conclusion

In this study, HTA was used to study skill transfer and found to serve as a useful tool for modeling the tasks in the form of goals and subgoals. It was able to reveal the complexity of tasks and suggest specific skills that need to be taught during training. By identifying elements that tasks have in common, the HTAs suggested where benefits of training may transfer. The HTAs revealed why no positive transfer was found in performing the truck loading task alternately with the excavator and loader. The lack of transfer was likely due to the differences between loader and excavator in terms of the controls, physical constraints, the goals and subgoals of the task. In addition, comparing the number of levels of subgoals did not reveal the level-of-difficulty differences between tasks. It is believed that mental workload measurement as well as performance measures on the tasks could provide indicators of the relative difficulty of the tasks.

CHAPTER 8. STUDY 4: VERBAL PROTOCOL ANALYSIS BY EXPERTS

A skilled construction operator is required to operate the machine to perform different tasks on the construction site. Some common machine types are equipped for switching the tool attachments so that a wider variety of tasks can be performed without bringing more machines to the site. Some wheel loaders, for example, allow detachment of the bucket for replacement with a fork attachment. Thus, the tasks for a wheel loader operator are not restricted to only bucket loading and dumping, but also include other carrying tasks such as loading and unloading fabricated materials with a fork. Different lessons and tasks with different tools (e.g. a bucket, a wide fork, and a narrow fork) are modeled and available in the John Deere training simulator used in this study. Studies have shown that different practice schedules for motor control tasks may differentially influence performance and learning (e.g. Lee & Simon, 2004; Schmidt & Lee, 2011). When training on multiple tasks is desired and becomes available, whether introducing intermixed practice trials within a machine during training facilitates transfer and retention is worth further investigation.

The training modules presented in the loader simulator constitute an array of easy to difficult tasks requiring basic to advanced skills. The obvious start is with controls familiarization, and it is evident that several later modules are based on the expectation of skills acquired from a previously learned module. However, no explicit instruction is given as to how to effectively utilize those modules provided in the simulator; it is only presumed plausible assumption that the order of the lessons listed in the simulator system menu will be followed because they appear to trend from easy to difficult. Reported research has been consistent in supporting the training difficulty principle, according to which, conditions that cause difficulty during learning facilitate later retention and transfer (e.g. Maxwell et al., 2001; Clawson et al., 2001). This practice implication raises another interesting question for the loader simulator: Given the different lessons embedded in the simulator system menu, in what order should the training modules be presented to the trainees in order to achieve the optimal training performance, eventually leading to optimal retention and transfer?

Chapters 5, 6 and 7 (Studies 1, 2, and 3) focused on skills transfer between an excavator and a loader. The second part of this research (Studies 4 and 5) is intended to determine whether there is a performance cost when switching between different types of training modules for a loader. Whereas the HTAs in Chapter 7 were conducted by direct observations by the researcher and then follow-up interviews with experienced operators, in this chapter a different method of initiating HTAs — verbal protocol — was used to decompose the experimental (training module) tasks. Although verbal protocol has been found the most direct elicitation tool in examining the on-going processes and intentions as and when learning happens, most studies using verbal protocol were interested in comparing the use of different methods for conducting verbal protocols (concurrent vs retrospective) (e.g., Banks, Stanton & Harvey, 2014; Ryan & Haslegrave, 2007a; 2007b), different verbal protocol instruction (classic vs explicit); Zhao, MacDonald, & Edwards, 2012), and novices vs experts (e.g., Hoffman et al., 2009), etc. When it comes to the analysis of verbal protocol, research has reported the procedure that the verbal protocols were transcribed and then segmented into identifiable units of speech to develop their own coding scheme (Banks et al., 2014) or taxonomy (Ryan & Haslegrave, 2007a). The next steps were to seek patterns and interpret patterns (Chi, 1997). However, most published research using think-aloud protocols have not presented the details of how coding and analysis were done (Gu, 2014). Indeed, no study has been reported using verbal protocols to derive HTAs. Therefore, the development of a systematic method of deriving HTA from verbal protocol is also one of the objectives of this study.

8.1. Objectives

In this study, experienced operators were tested on the loader simulator, and they were asked to use the ‘think aloud’ method to explain the what, how, and why of what they do during each module: bucket loading, filling a trench, moving a load with wide forks, and truck loading. The difficulty level of each task was classified by the experienced operators. A systematic method of how the HTAs were derived from think-aloud protocols was also developed in this study. The four HTAs generated from the verbal protocols and the difficulty level of each task classified by the expert operators were then used to bolster the hypotheses for Study 5. Also, Study 4 collected a couple of performance measures and opinions from experts to compare with those obtained from the novices in Study 5.

8.2. Method

8.2.1. Participants

Through contacts with the User Experience group at the Moline Technology Innovation Center of Deere & Company, a total of 8 machine evaluators from John Deere Dubuque Works, Dubuque, Iowa, experienced in the operation of numbers of construction equipment, including wheel loader, were invited to participate in this study. The participants were invited based on the availability of their work schedules and followed IRB Human Subject Protocol #1304013518 (Appendix F). Two operators from Study 3 (#s 4 and 10) also participated. The demographic information of the machine operators is summarized in Table 8.1.

Table 8.1. The Demographic Information of the Operators in Study 4

Operator #	Age	Loader experience (years)	Excavator experience (years)	No. of equipment known	Average years of experience
4*	24	1	0	1	1
10*	30	5	3	4	3
15	26	1	1	4	1
16	37	10	10	4+	10
17	40	20	20	4+	20
18	40	20	2	4	16
19	34	7	1	3	4
20	27	5	5	5	5

Note: * Operator participated in both Studies 3 and 4.

8.2.2. Experimental Setup

This study was conducted at the Virtual Reality Lab of John Deere Dubuque Works, in Dubuque, Iowa. The location and simulator were the same as described in the previous study. The simulator used in the study was John Deere's PC-based 4-Wheel Drive Loader Simulator, equipped with a 60-in. Samsung LED TV monitor, which simulates a John Deere 544K 4WD Loader.

8.2.3. Design

Verbal protocol analysis, a think-aloud method, for examining the on-going processes and intentions as and when learning happens (Gu, 2014) was used to decompose the selected training modules. Four training modules (as shown in Figure 8.1) were tested. The four modules all require driving on the job site and manipulating the tool attached to the front to complete the tasks.

The completion times of the four training modules are different in the simulator system. For example, there is a 7-minute limit for the truck loading task but no time limits for the other three. To control the training time, some of the criteria of the four modules were modified from the original modules. Detailed descriptions of each task are presented in the next section, and performance indicators that were recorded for each task are noted. The values for these indicators were recorded manually because the software did not maintain a thorough performance database for such use.

8.2.4. Experimental Task

Four training modules (Figure 8.1) from the John Deere wheel-loader simulator were selected for this study: bucket loading (B1), filling a trench (B2), moving a load

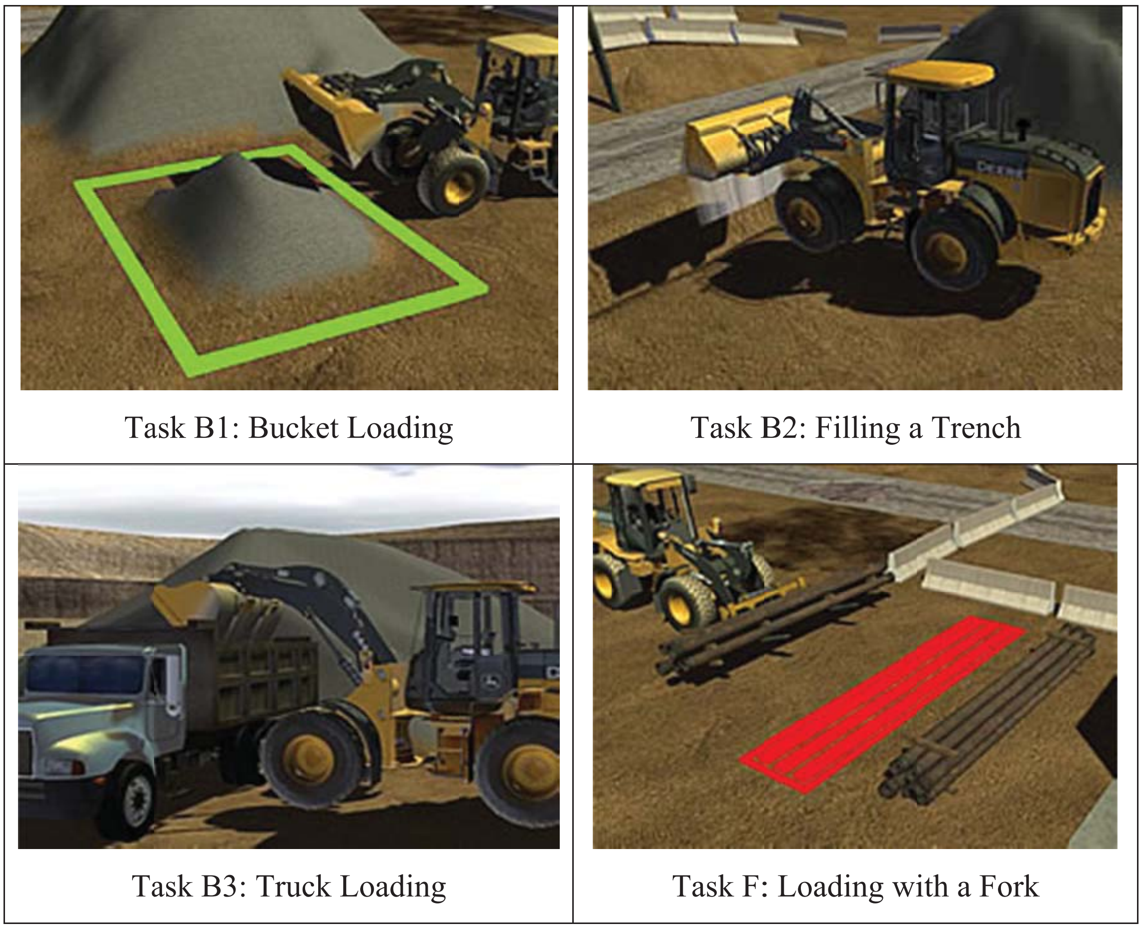


Figure 8.1. Screen shot from the four training modules.

with wide forks (F), and truck loading (B3). Tasks labeled with 'B' require the manipulation of the loader bucket, while 'F' requires the manipulation of a fork.

8.2.4.1. Simple Bucket Loading (B1)

This module teaches the trainee how to accurately approach an aggregate stockpile and position the boom height and bucket angle to achieve maximum bucket fill, involving four basic steps: 1) maneuver the 4WD Loader towards the aggregate stockpile, 2) adjust boom height and bucket angle to line up with the red bucket target near the aggregate stockpile, 3) start to fill their bucket with aggregate, 4) maneuver the 4WD Loader to the green highlighted dump area to the left of the aggregate stockpile and dump the bucket load into the dump area.

8.2.4.2. Filling a Trench (B2)

This task was modified from the Feather Bedding into a Trench module, which requires the 4WD Loader operator to be precise with their controls of the bucket for lightly dumping. Instead of doing a 'feathering' task, filling the trench with the full bucket of aggregate was the goal of the task. In this lesson, the participant was asked to fill a trench through four steps: 1) approach the aggregate pile to get a full bucket of aggregate, 2) approach the red 4WD Loader positioning target, 3) carefully start to dump aggregate into the trench, 4) back away from the trench and head towards the next red 4WD Loader positioning target. Such modification was due to the consideration that the novices to be tested in the next study may have found it difficult to perform the feathering task in the short period of practice and such modification also made the task similar to the B1 and B3 tasks which involve loading and dumping a full bucket.

8.2.4.3. Truck Loading (B3)

B3 was the task of greatest interest. This module teaches the trainee how to quickly and accurately load dump trucks. This is a common real-world application that requires the trainee to be fast, alert, efficient, and safe. Four steps are involved: 1) approach the aggregate pile, 2) get a full bucket of aggregate, 3) reserve the loader and approach the dump truck and 4) carefully dump the aggregate into the dump truck and avoid hitting the side of the truck.

8.2.4.4. Moving a Load With Wide Forks/Fork Lifting (F)

This module teaches the trainee how to properly transport a wide load using wide pallet forks, involving 3 steps: 1) pick up a wide heavy load of bundled 20' iron pipe, 2) maneuver through a jobsite while avoiding jersey barriers and safety hazards such as exposed rebar, high voltage lines, and utility poles, and 3) position the bundled pipe within the red target until it turns green and disappears.

8.2.5. Procedure

After signing the consent form, a preliminary questionnaire obtaining demographic information was administered before the first session began. Participants were informed of the study's aim and that the goal was to learn how they do the task by using a 'think aloud' method. The study was divided into three sessions and took approximately 45 minutes.

In the first session, the experts were randomly assigned to one of the experimental sequences of the four tasks: B1, B2, B3, and F. They were seated in the loader simulator to carry out the tasks. For each task, the operators were asked to try out the task once and then use the concurrent 'think aloud' method to verbalize how they were executing the tasks in their second attempt. A NASA TLX questionnaire was administered after each task performance. This procedure was followed consistently for all four tasks. The instruction for the concurrent 'think aloud' was adopted from Ericsson and Simon (1993) (see Table 2.2). The participants' verbal protocols were recorded throughout Session 1.

In Session 2, all operators were asked to execute the truck loading task 5 times by loading and dumping three buckets into the truck. The results obtained in Session 2 were intended to provide benchmarks of experienced loader operator performance for later comparison to the performance of the novices which were measured in Study 5.

Besides getting a benchmark of truck loading performance, it is also of interest how much attention the experienced operators give each type of displayed feedback. There were numerous feedback indicators on the screen while operating on the loader simulator as illustrated in Figure 8.2. The experienced operators were asked to indicate the approximated percentage (%) of time they spent on each type of feedback indicated in the figure (0% indicated that they did not refer to particular that feedback, the values put in each feedback should add up to 100%). This question was addressed in the post questionnaire (See Appendix H) which was administered in Session 3 of this experiment. The questionnaire consisted of questions regarding their perceptions of the difficulty of the experimental tasks and their opinions on operating the simulated loader.

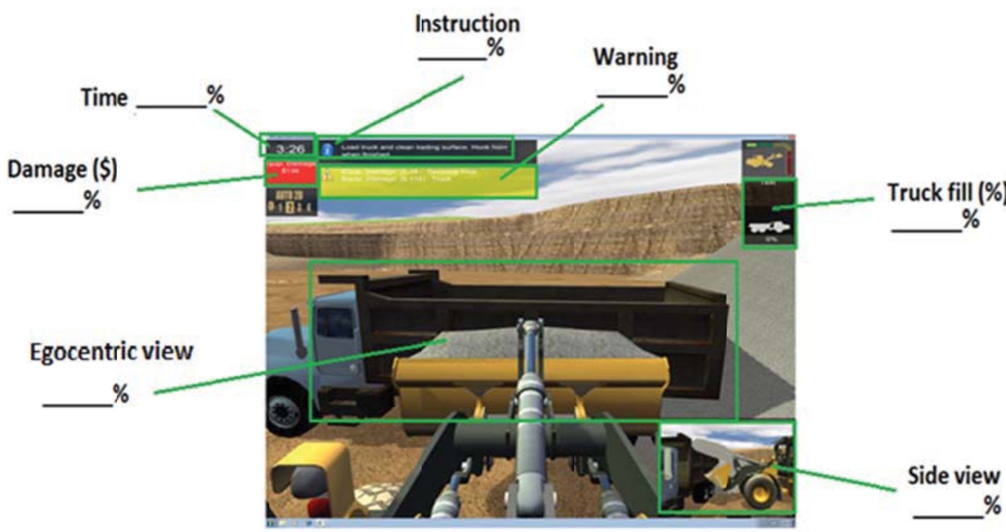


Figure 8.2. Feedback indicators shown on the screen of loader simulator.

8.3 Data Analysis: How to use Verbal Protocol Analysis to Develop HTA

Thematic analysis, a method for identifying, analyzing, and reporting patterns (themes) within qualitative data (Braun & Clarke, 2006) and a widely-used qualitative analytic method (e.g. Boyatzis, 1998; Braun & Clarke, 2006; Joffe, 2011) in qualitative psychology, was incorporated in the method devised for transforming verbal protocols to HTA diagrams. Six phases outlined by Braun and Clarke (2006) for conducting thematic analysis were adapted for developing HTA diagrams from the verbal protocols (see Table 8.2).

Table 8.2. Phases for Conducting Thematic Analysis as Outlined by Braun and Clarke (2006) and Modified for Developing HTA Diagrams From Verbal Protocols

Six phases outlined by Braun and Clarke (2006) for conducting thematic analysis	Nine phases proposed here to transform verbal protocols to HTA diagram
1) Familiarizing yourself with your data. 2) Generating initial codes. 3) Searching for themes. 4) Reviewing themes. 5) Defining and naming themes. 6) Producing the report	1) Transcription of verbal data & familiarizing yourself with your data. 2) Cleaning data 3) Identifying verbs of actions 4) Rank verbal reports by number of actions 5) Searching for themes 6) Reviewing themes and codes 7) Naming themes into goals and subgoals 8) Tracing the plans of the subgoals 9) Building the HTA diagram

A few terms used throughout the chapter are defined here to avoid confusion. *Verbal protocols* refers to all verbal data (utterance) collected from conducting the think-aloud method, whereas *verbal reports* refers to verbal data collected per experimental tasks. In other words, each operator will generate 4 verbal reports since they were given four tasks to complete. The procedures for how the verbal protocols are coded, cleaned, and analyzed are presented as follows and examples are offered to demonstrate each phase.

8.3.1. Phase 1: Transcription of Verbal Data and Familiarizing Yourself With Your Data

After verbal protocols are collected, the audio recordings need to be transcribed. To obtain inter-transcriber reliability of the transcripts, more than one transcriber should be involved during this process. Transcription also should be done independently for a validity check. Braun and Clarke (2006) emphasized that it is important that the researcher spend more time familiarizing oneself with the data, and also check the transcripts back against the original audio recordings for accuracy if the verbal data are transcribed by others.

In this study, all verbal protocols were transcribed by an undergraduate (a senior student) research assistant who had been trained on the four experimental tasks on the loader simulator for a total of 4 hours in two weeks to ensure the research assistant had gained sufficient knowledge of the training modules. The student assistant was instructed to listen and type out the scripts and record the time gap (in seconds) when there was silence in between words and phrases. Although the undergraduate research assistant did not report any difficulty in transcribing the scripts or show any concerns about ambiguity during the utterances, to ensure the transcripts were reliable, consistent, and accurate, a second individual—a doctoral student in Psychology— was invited to transcribe 4 of the verbal protocols independently. The purpose of having a second person here was to serve as a spot-check. The four selected verbal protocols (one from each experimental task) were either the one with most or least number of words spoken (utterance) by the operators. Table 8.3 shows the operators who gave the most and least utterances in the four experimental tasks. The goal in this step is to avoid selecting verbal protocols from the same operators for spot-checking. In this instance, the verbal protocol collected from operators #10, 15, 17 and 18 were chosen for spot-checking (see Appendix H). The results of spot-checking were that 98.56% (478 out of 485 spoken words) of the transcriptioning done by the graduate student was identical to the corresponding transcriptioning done by the student researcher, indicating the inter-transcriber reliability is very high. Furthermore, the differences in the transcriptions did not alter the meaning of the intended actions of the operators, e.g. “be sure no one is there” versus “make sure no one is there”, “need to watch out for…” versus “gotta watch out for…”, “cut off the throttle” versus “take off the throttle”. For the differences between the two transcription sets, the researcher listened to the recordings to find the most accurate transcription. Finally, all transcribed verbal protocol reports and the recordings were gone through and checked by the researcher to make sure the transcription was the same as the recordings.

8.3.2. Phase 2: Cleaning Data

Data should be cleaned by following three steps: 1) eliminate the verbal protocols that are irrelevant to the task and 2) fix the incomplete sentences, and 3) separate phrases by a period or comma.

Table 8.3. Operators With the Most and Least Utterance in the Four Experimental Tasks

	Simple bucket loading (B1)	Filling a Trench (B2)	Truck loading (B3)	Fork lifting (F)
Most Utterance	#17*	#10*	#17	#10
Least Utterance	#15	#15	#18*	#15*

Note: * denotes the operator whose audio recording was selected for spot-checking.

1. *To clean the data, it is important to keep in mind the goal of collecting verbal protocols in the experiment and to eliminate the verbal protocols that are irrelevant to the task.*

In the example presented here, the goal of collecting the verbal protocols is to understand how the operators perform the tasks on the simulated loader and ultimately to develop the HTAs from the verbal protocols. Even though the operators were given standard instructions for making their concurrent (to task performance) verbal protocol reports at the beginning of the task, it was unavoidable that the operators would mention something not directly descriptive of how they perform the task. For example, Operator #10 said "Now when I'm approaching the trench, it's a little more difficult here because I can't see where the bottom of the trench is." At another point, he said "Normally you would be able to see the edge of the trench and what I watch out for is that I don't want to get too close to the trench." Those comments about the operating environment conditions were removed and not included for the development of the HTA but they were kept in a separate table containing all comments from the operators (see Appendix H).

2. *There are two types of incomplete sentences: missing words or unclear words (e.g., pronouns). If a sentence is incomplete, the researcher will make the best guess to fill in the missing words; squared brackets [ ] are used to indicate the missing words. If a sentence contains a pronoun that does not refer to previous content,, a clarification is added and put in rounded brackets( ).*

For example, if the operator said “And going forward and trying to line up with…”, the sentence was modified as “And going forward and trying to line up with [the trench]…”. If the operator said “As I'm backing up, I raise the bucket just a little bit until I'm ready. So now I'm approaching ‘it’ and I'll start dumping the bucket…”, previous sentence did not infer what ‘it’ is, but the actions described preceding ‘it’ imply that the operator is approaching the dump area, thus the words ‘(dump area)’ were inserted after the word ‘it’ to the verbal report for clarification.

3. *Punctuations help to separate the task elements and help the reader to understand transitions where there maybe a few seconds needed to execute the action that the operator has just described. A comma is used to break down conditional sentences, a period (one dot) is used to indicate where the sentence is complete and there is*

1- to 3-second gap, and three dots (…) are used to indicate where there is more than 3 seconds of silence before the operator spoke again.

Three dots were used to indicate where there were 5 or 6 seconds of silence in this following example. “Put in forward. Scoop the bucket and then put it in reverse… (6 seconds pause) Turn … (5 seconds pause) Lower the bucket here in carry position.”

8.3.3. Phase 3: Identifying Verbs of Actions (Generating Initial Codes)

When sentence begins with the intending or considering verbs such as “make sure”, “want to”, “going to” , “trying to”, etc., only the verb that they are intended to do is coded.

According to Braun and Clarke (2006), codes were the identified feature of the data that appears interesting to the analyst, Boyatzis (1998, p.63) refer ‘unit of coding’ to “the most basic segment, or element, of the raw data or information that can be assessed in a meaningful way regarding the phenomenon”. Here, the *initial codes* refer to the verbs of the actions that are related to the tasks. For each task, the verbs appearing in the verbal protocols were highlighted as illustrated in Table 8.4. Two examples where sentences involve the words “make sure” help to illustrate this step: If the operator said “I want to make sure I ***load*** the pile”, the verb ***load*** will be coded. However, sometimes the verb to be used is not actually a verb in the utterance. If, for example, the operator said “I'm going to make sure my bucket is in position”, in this case, the operator needed to move the control in order to position the bucket. Thus, the word "position" is coded and highlighted.

8.3.4. Phase 4: Rank Verbal Reports by Number of Actions

After identifying of all the actions (or verbs) of each verbal report, the verbal reports are put in order according to the number of actions each operator has cited in that specific task. The one with the highest number of actions is ranked highest (and placed at the top of the ranking tabulation).

In this example, after the verbs were coded, the eight verbal reports were ranked in order following the number of actions being cited. The verbal report of Operator #10 with 17 verbs (the largest number of verbs) identified was ranked first, whereas the report of Operator #15 with 7 verbs (the smallest number of verbs) coded was ranked last, as illustrated in the Table 8.4. The ranked verbal reports of the other three tasks are shown in Appendix H.

8.3.5. Phase 5: Searching for Themes

When all the verbs (the actions) have been initially coded and ranked, the next step is to sort the different codes into potential themes starting from the top ranked verbal report. When analyzing the codes and considering how different codes may combine to form an overarching theme, visual representations such as tables, mind-maps could be used to help sort the different codes into themes (Braun & Clarke, 2006). In this stage, it is advised to keep the data as its original form in order not to lose the relationships between the actions, which will be very critical in the later steps to identify the sequences of the subgoals to achieve the goals for developing the plans of the HTA. For example, we do not want to remove the conjunction words 'when' and

Table 8.4. Eight Ranked Verbal Reports With Highlighted Actions for the Filling a Trench Task

Ranks	Operator ID	No. of Actions	Verbal Reports
1	10	17	Now, when I am going to the pile, I'm going to make sure my bucket is in position and boom is lowered to the ground. And when I go to the pile, I want to make sure it's square to the pile like this… so that I get a full bucket. So when I approach the pile the best I can, I want to make sure I load the pile… I'll keep the bucket low to the ground until I get to where I want to go. As I approach it then I'll square with it (trench), and I'll start dumping … and now I'm backing up and I don't want to hit the guy behind me. I'm lowering my boom and getting my bucket in the position for the next bucket…
2	19	16	So you keep your bucket flat on the ground with the cutting edge flat. Pull into the pile… and increase the throttle as you pull into the pile. As you start hitting the material, start raising the bucket and curling the bucket back. Raising your boom until it's full. So once your bucket is full, put it in reverse...Back up… Line it up with your tracks... put it in forward… Brakes. Raise the boom and dump. Put it in reverse…
3	20	15	Pull into the pile and you just load the bucket trying to stay close to the ground… Get a full bucket...Raise the bucket back up… Watch the guy behind you. Approach… Again, raise the bucket so you have full access to the trench… Hold your foot on the brake. Dump the bucket. So as you're backing out, turn your bucket to the carry position. Watch the guy behind you. Lower the bucket and approach the pile…
4	16	12	So I'm trying to get square with the pile… Watch the bucket come down to the ground and hits the pile… Okay, so I backed away from the pile… looking in the mirror watching for the people behind…Hold the bucket. Back away watching out behind again. Go to the red spot… Raising [the bucket], watching where I'm going to go to dump it next…
5	17	12	I'm going over to the pile and looking to see if my bucket is level with the ground … I'm backing up... now I'm going forward… and I'm looking at the spot where you want me to position for the trench…Moving up to the trench trying to judge if my bucket is over the trench. Stopping…raise and dump the bucket…backing up…

(continued on next page)

Table 8.4 continued.

6	4	9	Put in forward. Scoop the bucket and then put it in reverse…Turn …Lower the bucket here in carry position. And going forward and trying to line up with [the trench] …Put the bucket back to carry position. Back out…
7	18	9	Drive to the pile… Lower the boom...Fill my bucket. I'm going to back away and I'm going to center onto the grip. Approach the spot... Raise the bucket...I'm going to back away, fill another bucket…
8	15	7	Forward throttle to fill up my bucket… Reverse …Lower the boom when driving in forward …Brake in front of trench. Reverse…

'while' which indicate the timing when an action will be executed. The word 'and' is also important to indicate when two actions are performed concurrently.

In this example of analyzing the verbal protocols for the filling a trench task, the *codes* here refers to the verbs of action identified in Step 4 and the *potential themes* here refers to the possible scenarios during the task. The goal is to group the different codes into potential themes. Table 8.5 provides an illustration of how different codes were coded into potential themes from the verbal protocols. The themes generated for the other three tasks are shown in Appendix H. For example, in the first row of Table 8.1, when operators said "going to the pile", "pull into the pile", "drive to the pile" or "put in forward", all of these were put under the same theme — ***Approach the pile***. At this stage, the names of the themes need not be finalized until Step 7. The benefit of starting with analyzing the verbal report with the most verbs uttered is that such reports will have a higher chance of covering most themes, thus providing a better descriptive flow of the tasks. From the verbal report of Operator #10, seven scenarios were generated. When analyzing the verbal report with the second highest number of verbs,

Table 8.5. An Illustration of Sorting the Different Codes Into Potential Themes From the Verbal Protocols of Filling a Trench Task

Operator ID / *Theme*	***#10***	**#19**	**#20**	**#16**	**#17**	**#4**	**#18**	**#15**
A. Approach the pile	Now, when I am going to the pile		Pull into the pile	So I'm trying to get square with the pile…	I'm going over to the pile	Put in forward.	Drive to the pile.	
B. Adjust bucket	I'm going to make sure my bucket is in position and boom is lowered to the ground.	So you keep your bucket flat on the ground with the cutting edge flat.	trying to stay close to the ground…	Watch the bucket come down to the ground and hits the pile…	and looking to see if my bucket is level with the ground.		Lower the boom.	

(continued on next page)

Table 8.5 continued.

Operator ID / *Theme*	*#10*	#19	#20	#16	#17	#4	#18	#15
C. Load bucket	And when I go to the pile, I want to make sure it's square to the pile like this… so that I get a full bucket.	Pull into the pile… and increase the throttle as you pull into the pile.	and you just load the bucket.			Scoop the bucket	Fill my bucket	Forward throttle to fill up my bucket.
	So when I approach the pile the best I can, I want to make sure I load the pile…	As you start hitting the material, start raising the bucket and curling the bucket back.	Get a full bucket... Raise the bucket back up…					
		Raising your boom until it's full.						

(continued on next page)

Table 8.5 continued.

Operator ID / Theme	*#10*	#19	#20	#16	#17	#4	#18	#15
D. Lower bucket	I'll keep the bucket low to the ground until I get to where I want to go.			Hold the bucket.		Lower the bucket here in carry position.		Lower the boom
E. Back away				Back away watching out behind again				
F. Drive to trench	As I approach it	Line it up with your tracks... put it in forward…	Approach… Again, raise the bucket so you have full access to the trench…	Go to the red spot…	now I'm going forward… and I'm looking at the spot where you want me to position for the trench.	And going forward.	Approach the spot	when driving in forward

(continued on next page)

Table 8.5 continued.

Operator ID / *Theme*	***#10***	**#19**	**#20**	**#16**	**#17**	**#4**	**#18**	**#15**
G. Line up with trench	then I'll square with the [trench],				Moving up to the trench trying to judge if my bucket is over the trench	And trying to line up with [the trench] …		
H. Brake		Brakes.	Hold your foot on the brake.		Stopping…			Brake in front of trench
I. Raise bucket		Raise the boom		Raising [the bucket],	Raise		Raise the bucket.	
J. Dump bucket	and I'll start dumping	and dump.	Dump the bucket.	watching where I'm going to go to dump it next	and dump the bucket….			

(continued on next page)

Table 8.5 continued.

Operator ID / *Theme*	*#10*	#19	#20	#16	#17	#4	#18	#15
K. Back out for second bucket	and now I'm backing up and I don't want to hit the guy behind me.	Put it in reverse…	So as you're backing out, turn your bucket to the carry position.		backing up	Back out	I'm going to back away, fill another bucket	Reverse
	I'm lowering my boom and getting my bucket in the position for the next bucket		Watch the guy behind you.					
			Lower the bucket and approach the pile.					

two more themes (F & G) that were not included from the first operator were added to the table. Through this process, all possible scenarios were identified.

8.3.6. Phase 6: Reviewing Themes and Codes

Step 6 begins to consider the elements of HTAs by defining the themes and breaking those themes down into goal, subgoals and sub-operations of the HTA diagram. In Phase 6, some knowledge of the task from the researcher could help make the decision between goals and subgoals. The knowledge could be gained through direct observation of the task (here, it could be watching the demo videos on the simulators or studying the user manuals). It is important to keep in mind that the aim is to build a hierarchy consisting of goal, subgoals and sub-operation. In this stage, according to Braun and Clarke (2006), all the collated extracts for each theme are reviewed and re-considered whether they appear to form a coherent pattern. As indicated by Braun and Clarke, there are two purposes to re-read the entire data set. The first is to ascertain whether the themes fit to the data set. The second is to code more if there is any additional data within themes that has been missed in earlier coding stages.

So, in the current example, only 5 themes (A, C, F, J, & K) remained as the first level from the 11 potential themes generated from Step 5. The remaining six themes and the codes were then categorized accordingly into sub-operations of these five subgoals.

8.3.7. Phase 7: Naming Themes Into Goals and Subgoals

After decomposing the potential themes and codes into subgoals and subgoals, the next step is to give a generic name to each goal and subgoal.

For example, there are different ways of saying how the operators described their action when they were approaching the pile, such as “going to the pile”, “pull into the pile” and “drive to the pile”. In this case, “Drive to stockpile” was used as the name of the subgoal.

8.3.8. Phase 8: Tracing the Plans of the Subgoals

Plan is one of the important elements of HTAs, which offers the sequences of actions of each goal (or subgoal).To find the plans of the subgoals, you ought to understand the relations between goals and subgoals. To learn such relationship, the conjunction words “when”, “while”, “as”, “once”, etc. serve as an important identifier to indicate when the actions are being executed. There are four main types of plan introduced by Annett (2004): a ***simple sequence*** *of operations, a* ***conditional sequence*** *involving a decision, a* ***time-shared*** *procedure when two goals must be performed concurrently, and an* ***unordered*** *procedure where all subgoals must be performed but order is not critical. To trace the plan, the table generated in Step 5 is the starting point.*

For example, Operator #19 described how he loaded the bucket during a filling a trench task as “Pull into the pile…increase the throttle as you pull into the pile. As you start hitting the material, start raising the bucket and curling the bucket back… Raising your boom until it's full…” In such paragraph, three observations could be made:

1. “Pull into the pile” and “Increase the throttle” are concurrent motions.

2. "As you 'start' hitting the material" indicated a time-sharing procedure.
3. To get a full bucket, the operator would "increase the throttle", "raise the bucket" and "curl the bucket" at the same time.

8.3.9. Phase 9: Building the HTA Diagram

To build the HTA, the notation used is adopted from Annett et al. (2000). Similar procedure of developing preliminary HTAs is discussed in Section 7.2.1 in the previous chapter. The procedure involves labeling the goal and subgoals accordingly and using ">", "+" , " /", ":" to indicate whether those actions are performed in sequence, concurrently, either one, or unordered, respectively, in the corresponding plans.

An example HTA diagram for the Filling a Trench module developed from the verbal protocol is illustrated in Figure 8.3.

8.4. Results

This section mainly focuses on the HTAs of the four experimental tasks developed from the verbal protocols, the NASA TLX measures collected for the four experimental tasks and the post-questionnaires. Analysis and discussion of the benchmark performance collected on the simulated loader will be presented in the next chapter for the comparisons between experts and novices.

8.4.1. HTAs of Four Loader Tasks

How the HTA was developed for filling a trench (B1) task was illustrated step-by-step in the previous section. The HTAs of the remaining three tasks — [simple

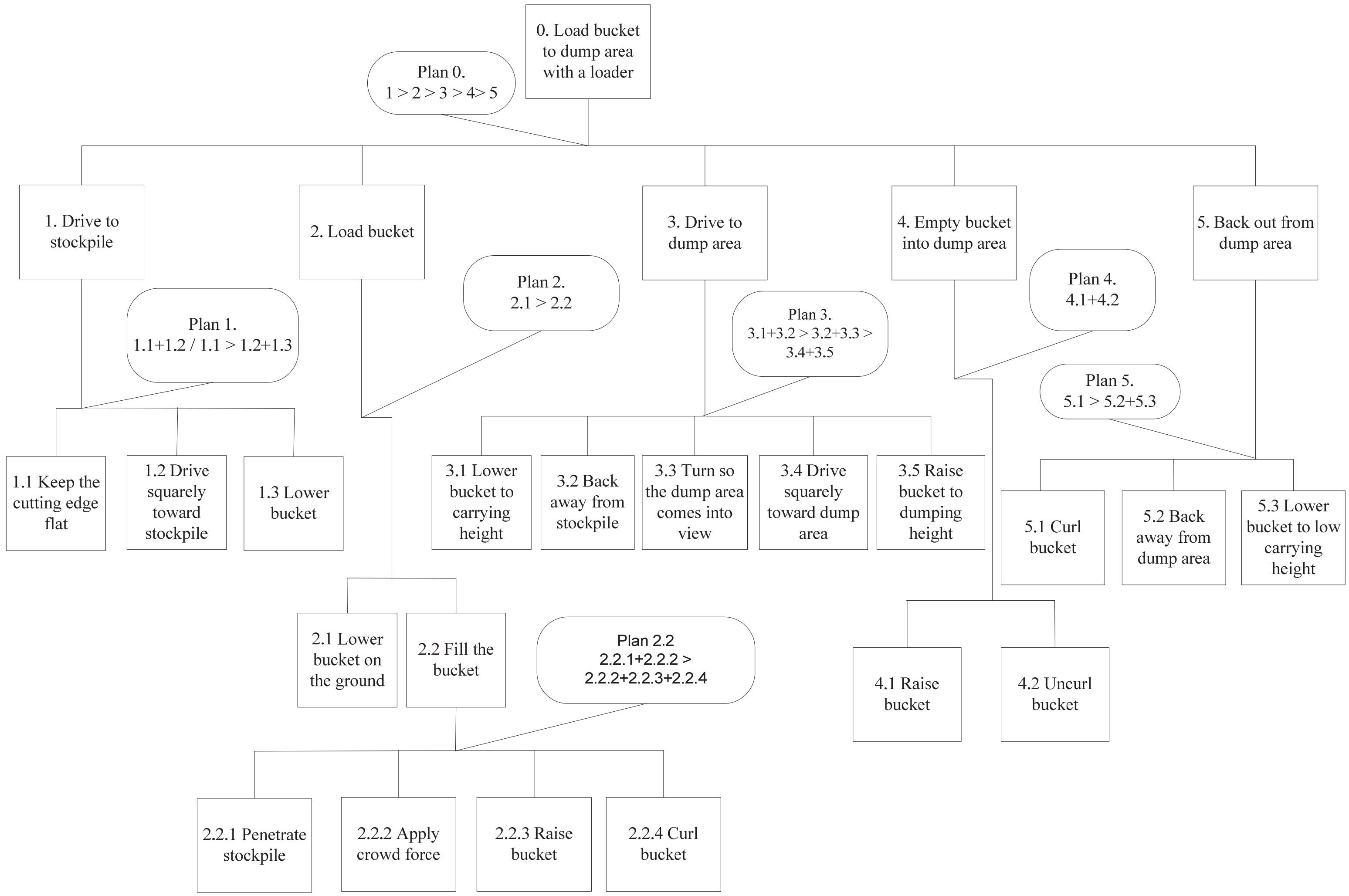


Figure 8.3. HTA for bucket loading module (B1).

bucket loading (B1), truck loading (B2), and fork lifting (F)] — were developed following the same step-by-step process as described in Section 0. The four HTA diagrams are illustrated in Figures 8.3-8.6. The tables containing the ranked verbal reports and initial themes for these three tasks can be found in Appendices L and M, respectively.

8.4.2. Workload Measures

The operators were asked to complete NASA-TLX questionnaires to obtain workload estimates after each time they verbalized how they were executing a specific task with the simulated loader. The six different subscales of the NASA-TLX were analyzed using a mixed design ANOVA with the mental measure as a within-subject factor and experimental task as a between-subjects factor. Similar to previous analysis of the workload measures in Studies 1 and 2, the scale for the Performance measure was reversed before the analysis, so that a higher score meant higher workload. The Huynh-Feldt correction for violations of sphericity was applied, but the results with Huynh-Feldt correction do not change the significance level of the results with sphericity assumed, so the results with sphericity assumed are reported here. The ANOVA results (Appendix H) showed a main effect of task, $F(1, 28) = 5.14$, $p = .006$, ${\eta_p}^2 = .355$, where fork lifting (F) task exhibited a significantly higher overall workload compared to the simple bucket loading (B1) task ($p < .001$, Appendix H). The main effect of measure was also significant, $F(5, 140) = 8.78$, $p < .001$, ${\eta_p}^2 = .239$, where participants rated all the tasks as requiring higher mental demand ($M = 5.25$ out of 10, $SD = .229$) and effort ($M = 4.69$, $SD = .237$), but lower temporal demand ($M = 3.52$,

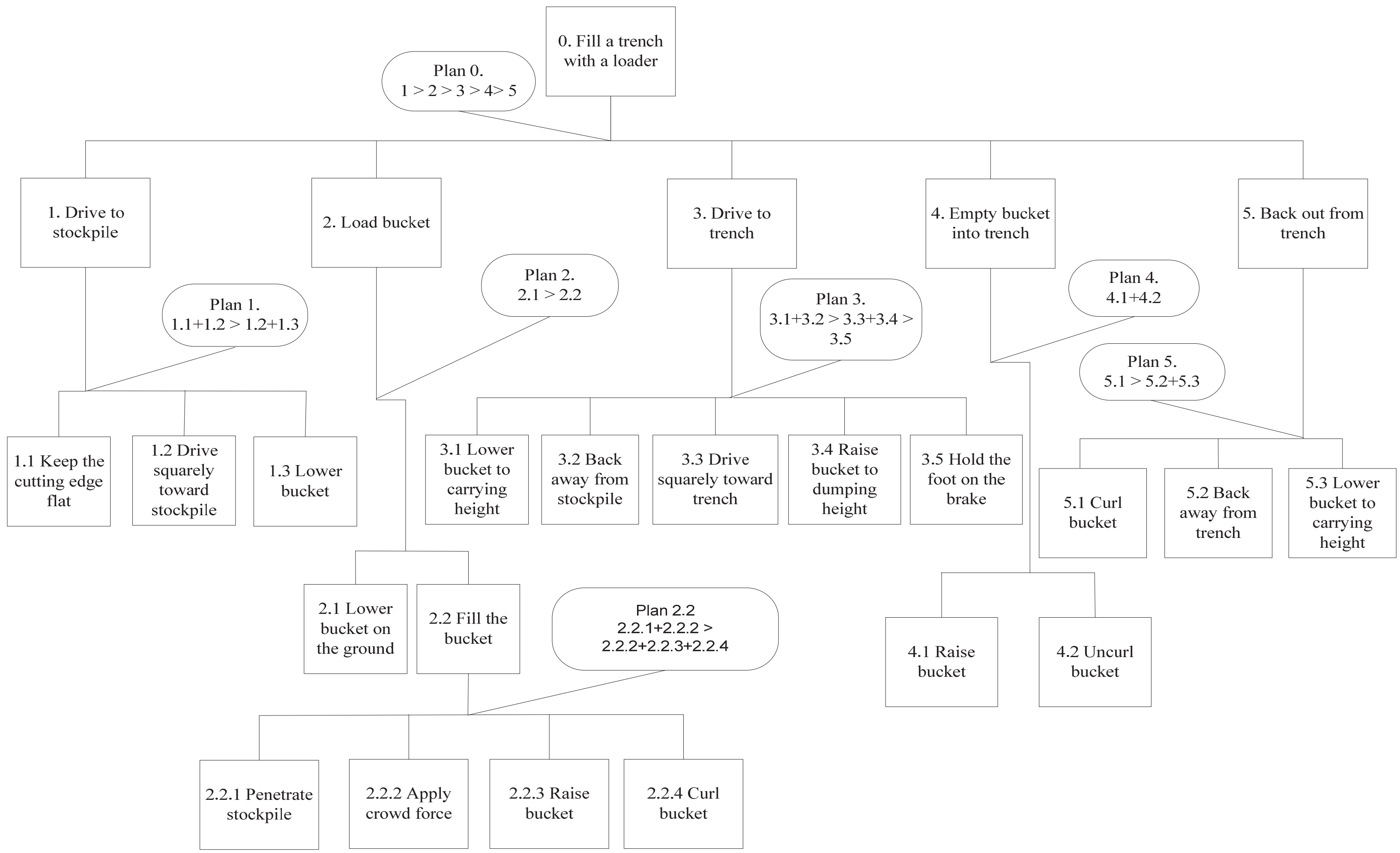


Figure 8.4. . HTA for filling a trench module (B2).

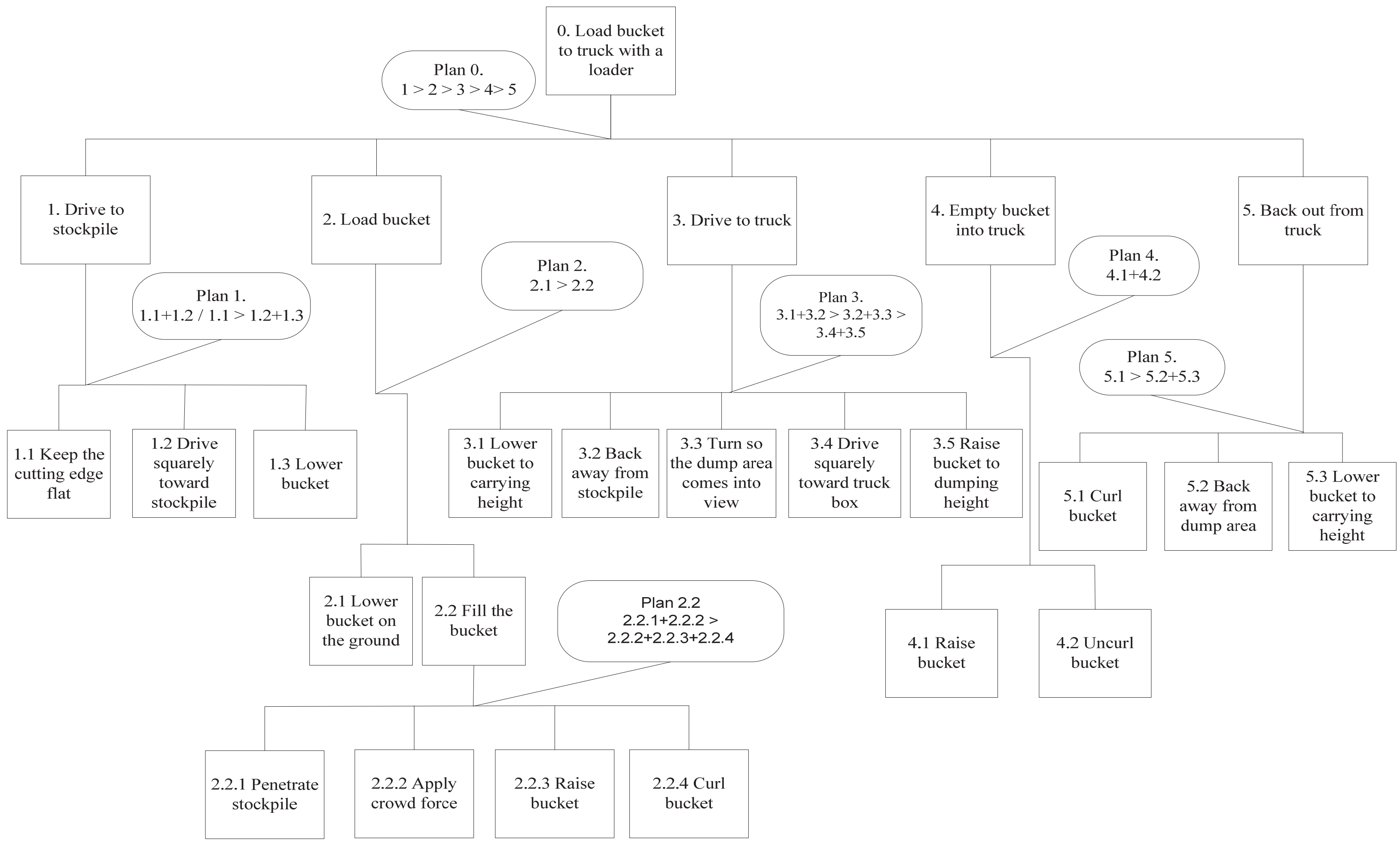


Figure 8.5. HTA for truck loading module (B3).

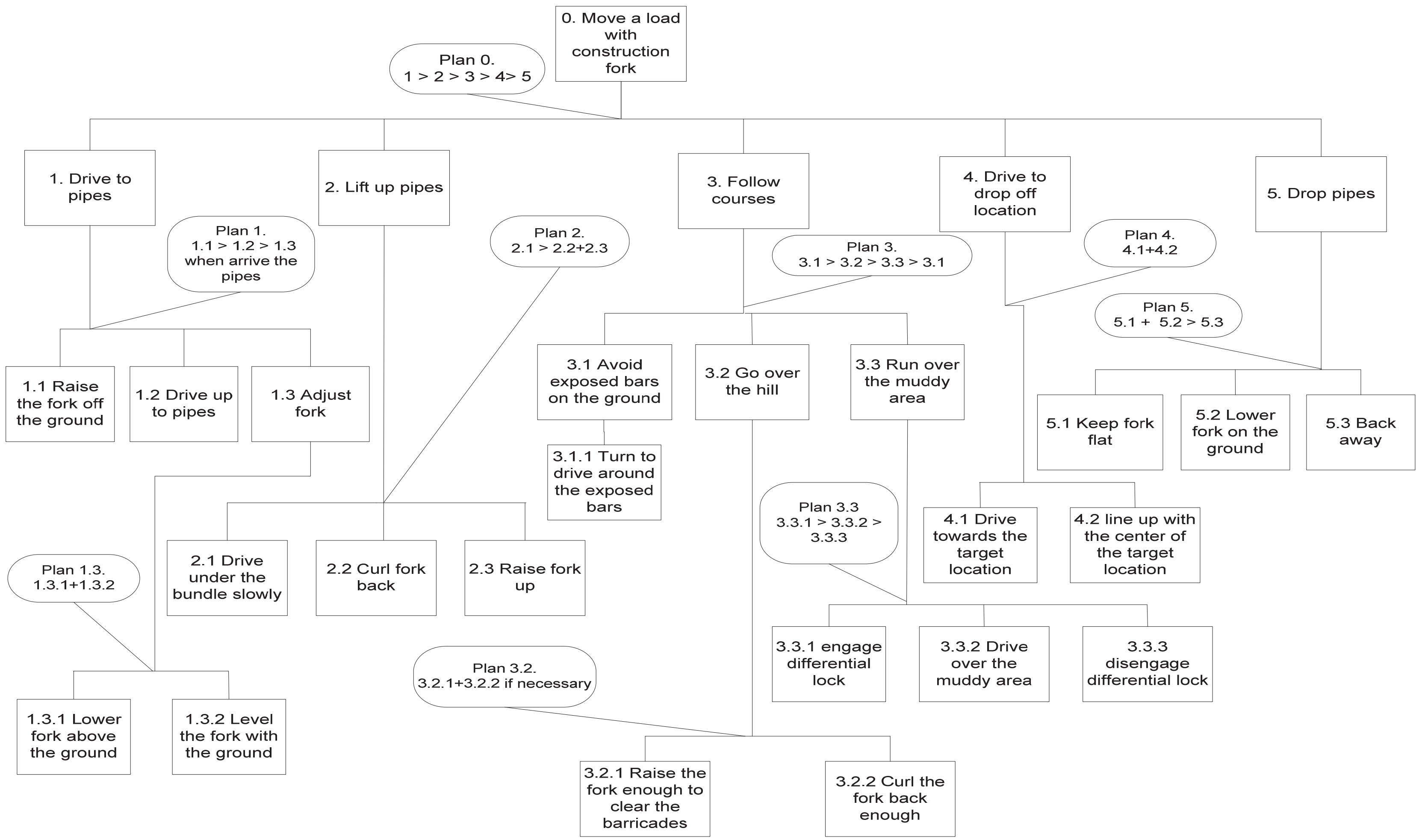


Figure 8.6. HTA for moving with a wide fork module (F).

SD = .340), physical demand (*M* = 3.53, *SD* = .284), less frustration (*M* = 3.61, *SD* = .361) and were very satisfied with their performance (*M* = 3.69, *SD* = .373). No interaction between measure and task was found.

8.4.3. Post Questionnaires

The operators' rankings of the difficulty of the experimental tasks and their opinions on operating the simulated loader were addressed in the post questionnaire which was administered in Session 3. The experienced operators were asked to rank the difficulty of the four tasks (B1, B2, B3 and F) from easiest to hardest (1: easiest to manipulate; 4: hardest to manipulate). The analysis of perceived difficulty by the operators from the four tasks was conducted with the Friedman test, a nonparametric statistical method of testing for differences between several related groups. There was a statistically significant difference in perceived difficulty depending on which training modules was performed, $\chi^2(3) = 21.750$, $p < 0.001$. The post hoc analysis with Wilcoxon Signed Rank tests (Appendix H) show that all six pairwise comparisons were significantly different, implying B1 (*M* = 1, all operators ranked it as the easiest) was rated significantly lower in difficulty compared to B2, B3 and F; B2 (*M* = 2.13, 7 out of 8 operators ranked it as the second easiest) was rated significantly lower in difficulty compared to B3 and F, and B3 (*M* = 3.00) was ranked significantly easier than F (*M* = 3.88, the most difficult). The distribution of the ranking of the four experimental tasks by 8 experienced operators is shown in Appendix H. In other words, among these four tasks, the Simple Bucket Loading (B1) module is the easiest, the Filling a Trench (B2) module is the second easiest, the Truck Loading (B3) module is the third easiest and the Fork Lifting (F) module is the most difficult. Other comments from the operators about the difficulty of the experimental tasks and their opinions on operating the simulated loader compared to a real loader are summarized in Table 8.6.

8.5. Discussion

8.5.1. Comparisons of B1, B2, B3 and F HTA Results

To answer in what sequence should the training modules be presented to the trainees in order to achieve the optimal training performance and eventually lead to optimal retention and transfer, four specific built-in training modules on the loader simulator were selected to investigate skill transfer and retention training issues. Four HTAs were built based on the verbal protocols collected from eight experienced operators using the think-aloud method to articulate how they perform the four tasks. By comparing the four HTAs, a few observations are made:

1. The HTA diagram of the Fork Lifting (F) module is significantly different from those of the three bucket loading tasks (B1, B2 and B3), where all three bucket loading tasks involved five similar subgoals—drive to stockpile, load bucket, drive to dump, empty bucket and back out—that the fork lifting task does not. Although fork lifting requires a similar operation to drive to pick up the pipes and to drop off the pipes, the module also involves maneuvering through a jobsite, avoiding exposed rebars on the ground, muddy area and jersey barriers, etc.

Table 8.6. Responses Collected via the Post Questionnaires in Study 4

Questionnaires	Responses from 8 operators
1. Did you find similarity across these four tasks: B1, B2, B3 and F?	• 4 operators indicated that B1 and B3 are similar in terms of controls. • 2 operators pointed out that all four tasks are all picking up the materials which are pretty similar. • 1 operator said 'B1 and B3 are almost identical mechanism' • 1 operator said 'B2 and B3 are similar, but the fork lifting task is different because it has a lot of obstacles.'

2. Is there any particularly difficult aspect of any of these tasks?	• Filling a trench
	  o 1 operator indicated that it is difficult to see the edge of the trench.
	• Truck loading
	  o 1 operator worried about hitting the truck and loading into a smaller target.
	• Fork lifting
	  o 3 operators said the course is very challenging, too tight for the barriers.
	  o 3 operators indicated that the task is more difficult because it requires a lot of attention, such as lining up, balancing, stabilizing the fork, following the course
	  o 2 operators said there are more things that they have to watch out for, e.g. how the fork is level, the edges of the piles, the elevation of the fork.

3. Are there any differences between the real machine and the simulator?	• 5 operators indicated that the controls, the basic functions, methods, and principles are the same. • 4 of them also mentioned that the simulator provide less feedback compared to real machine. • 3 operators indicated the simulator provides poor depth perception.

2. The plans of Goal 1. 'Drive to stockpile' in B1, B2 and B3 are slightly deviated from each other due to two reasons: the starting points of modules and the operator's habits. In the three training

modules, the loader is initially located differently relative to the stockpiles and the bucket height from the ground is slightly different. Some operators may adjust the bucket by keeping the cutting edge flat before starting to move forward the vehicle, but some may choose to adjust the bucket angle while driving squarely toward the stockpile at the same time. Either method would achieve the same goal.

3. The two HTAs of B1 and B3 have identical goals, subgoals and plan, except after filling the bucket, B1 requires the operator to drive toward and dump to a targeted dump area, whereas the operator drives and dumps to the truck box in B3. Indeed, four operators indicated that B1 and B3 are similar in terms of controls and 1 operatorid 'B1 and B3 are almost identical mechanism'.
4. The HTA of B2 (filling a trench) illustrated that the loader requires a complete stop (3.5 hold the foot on the brake) when reaching the trench (Goal 3) where the HTAs of B1 and B3 don't.

8.5.2. Experts' Verification Versus Verbal Protocol Analysis

Comparing the HTA diagrams of the truck loading task on the simulated loader in Study 4 in Chapter 7 (having experts to comment on and modify the preliminary HTA) and the one obtained in Chapter 8 (using verbal protocol analysis), all the elements (goal and subgoals) are identical, except an alternative sequence of subgoals was offered from the verbal protocol method because two operators preferred to adjust the bucket before driving towards the stockpile. HTA does allow the analyst to capture operators' different approaches to achieve the same goals, as illustrated as "1.1+1.2 / 1.1 > 1.2+1.3" in Plan 1. Indeed, with expert verification (the method described in Chapter 7), a more generic HTA was formed since the final three operators in the study could verify what most operators would do and could rule out exceptional cases where some operators may not follow norms.

In terms of the details that the operators provided using these two methods, the method of expert verification used in Study 7 may provide a more precise naming convention to each action. For example, instead of just 'lower bucket', the operators were able to point out that they lower the bucket to mid carrying level, digging height or low carrying height. These were specifics the operators failed to mention in their verbal protocols. However, the comments (see Appendix H) collected during the verbal protocols, but removed in Phase 2, did illustrate some instances where the simulator might not have performed exactly the same as the real machines, such as the need to always apply the brake to stop and the restricted view of the screen.

In terms of the time consumed in analyzing and developing the HTA using these two methods, both may take a similar amount of time to conduct the experiment with the experts, involving explaining the task, having participants try out the task and verifying the HTA (Chapter 7) or using the think-aloud method to verbalize what they do (Chapter 8). The similarity ends there, however, as transforming the verbal protocols required a fair amount of additional time and work for data analysis, i.e., transcribing the data and verifying the verbal data, cleaning the verbal report, extracting codes, searching for themes and developing the HTA structure. Chapter 8 did not involve experts to verify the final HTAs, indeed, it would be ideal to always go back to the experts for verification.

8.5.3. Difficulty Levels

In terms of difficulty levels of the four tasks, although all three bucket loading tasks shared a similar mechanism, the operators ranked B1 as the easiest, followed by B2 and B3 as the most difficult probably because positioning over the dumping target gets harder from B1 (a large green box), B2 (a narrow trench) to B3 (a high-sided truck box). The Fork Lifting module is ranked as the most difficult task which could be supported by the complexity shown in its HTA diagram, the workload index, as well as the subjective ratings by the operators. Indeed, some operators also indicated the fork lifting task was more difficult because this task has a very challenging course to drive through and requires a lot of attention to obstacles.

8.6. Conclusion

A different method of conducting HTAs — verbal protocol analysis — was used to deconstruct the four training modules on the loader simulator. A systematic method for how the HTAs can be derived from think-aloud protocols was also developed in this study. Four HTAs were successfully generated from the verbal protocols following the nine proposed steps. The primary downside to using verbal protocol analysis is the detailed and time-consuming nature of the process. The findings show that 1) the HTA of the Fork Lifting module is significantly different from those of the three bucket loading tasks 2) although all three bucket loading tasks shared a similar mechanism, the operators ranked B1 as the easiest, followed by B2 and B3, and fork lifting was ranked as the most difficult task.

CHAPTER 9. STUDY 5: SKILL TRANSFER AND RETENTION ON A MACHINE

Study 5 sought to verify whether an alternating practice sequence within the same machine, i.e., training with an alternative tool (a wide fork) and returning to the original learned tool (a bucket) on a loader simulator, yields better skill transfer and retention (after a one-week interval). The experiment investigated primarily whether the *specificity of training* principle, for which the conditions of practice should match the conditions of test to facilitate retention or transfer, and secondarily whether the *progressive difficulty training* principle, for which difficulty should impede the learning stage but facilitate retention, hold for training on one type of construction equipment. The four specific built-in modules (B1, B2, B3 and F, as descripted in Chapter 8) on the loader simulator were selected to investigate this question. The results of Study 4, which provided information on the similarity and dissimilarity of the four tasks and the difficulty level of each task, were used to bolster the hypotheses in this study.

9.1. Objectives

There are three major goals of Study 5:

1. To examine whether there are performance costs when training with an alternative tool (i.e., a fork) and returning to the original learned tool (i.e. a bucket).

2. To investigate whether the progressive difficulty principle, where introducing difficulty in training will impede progress in the learning stage (tests on the first session) but facilitate retention and transfer (a retention test after a week interval), holds for construction equipment training.
3. To compare the performance and attention focus between novices and experts.

9.2. Hypotheses

In Study 4, the HTA of the Fork Lifting (F) module was shown to be significantly different from those of the three bucket loading tasks (B1, B2 and B3) because the fork lifting task does not share the same subgoals as those in common between B1, B2 and B3. The task difficulty levels of the four tasks were ranked following this order from easiest to most difficult: B1< B2 < B3 < F. By comparing the HTAs and the task difficulty levels, the following hypotheses were formulated:

> *Hypothesis 5*: Thorndike and Woodworth's (1901) identical elements theory posited that transfer of learning depends on the proportion to which the learning task and the transfer task are similar. The procedural reinstatement principle (Lohse & Healy, 2012) also suggested that practicing a similar mental model (i.e., practicing bucket loading in B1: simple bucket loading and B2: filling a trench tasks) during the learning phase may facilitate subsequent retention and transfer in test phase where a similar mental model (B3: truck loading

task) is tested. It is anticipated that *groups provided practice with loader bucket manipulation during the training phase will perform better in the test phase (where a new task is given that also requires manipulation of bucket) and show better retention after a week interval.*

Hypothesis 6a: Schneider, Healy, and Bourne (2002) suggested that any manipulation of task difficulty during training may have facilitating effects during retention and transfer testing. *It is anticipated that groups provided practice with task presented in order of increasing difficulty will have better performance in both initial tests and retention tests.*

Hypothesis 6b: However, not all sources of difficulty are desirable. Some researchers argued that introducing difficulties during training is facilitative only when the training and retention tasks share task-relevant cognitive processes (McDaniel & Butler, 2011; McDaniel & Einstein, 2005). It is anticipated that *the group with training modules that share task-relevant cognitive processes, presented in order of increasing difficulty, will have better performance in both initial tests and retention tests.*

9.3. Method

9.3.1. Participants

Sixty undergraduate students (44 males and 16 females, distributed evenly across the three groups), ages 19–34 years (M = 20.1; SD = 2.3), participated for experimental credits toward an introductory psychology course requirement, according to Institutional Review Board (IRB) Human Subject Protocol #1110011339 (Appendix B). All were right-handed, physically capable of operating the simulator, and had no experience operating construction equipment.

9.3.2. Experimental Setup

The setup for the loader simulator was the same as presented in the Chapters 5 and 6 (Studies 1 and 2), i.e., John Deere's PC-based 4-Wheel Drive (4WD) Loader Simulator, which simulates a John Deere 544K 4WD Loader. Participants were presented with a virtual scene from the perspective of a person in the machine cabin and they controlled the virtual machine through the same interface mechanisms described in Studies 1 and 2 (Chapters 5 and 6).

9.3.3. Design

The experiment involved five sessions: 1) skill acquisition on the controls of the loader simulator; 2) Training on the first task; 3) Training on the second task; 4) Skill Transfer tests on a third task; 5) Retention test on the third task after a week interval. All sessions used training modules provided as part of the loader simulator software. The four training modules described in Study 4 were used in this study: bucket loading (B1), filling a trench (B2), moving a load with wide forks (F), and truck loading (B3). Four experimental groups (illustrated in Figure 9.1) were tested. The details of each task are described in Section 0.

The factors and levels studied are: two tests (initial, retention), five trials (1 to 5) within each test, and two practice order (start with B1 and start with B2) and two practice type (learning a different tool and learning one tool throughout the practice).

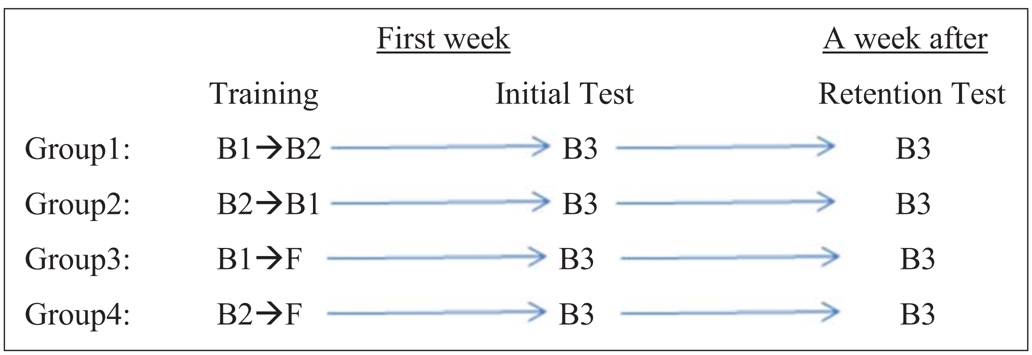


Figure 9.1. Four experimental groups in Study 4.

Tests and trials are within-subject factors, and practice order and practice type are between-subjects factors.

9.3.4. Performance Measures

The truck loading task (B3) is the task of most interest. Several performance measures obtained in B3 such as bucket fill (%), area/truck fill (%), damages ($), and execution time (minute:second) and warnings were recorded manually by reading from the performance indicators shown at the top two corners of the monitor display. In addition to obtaining performance measures, the subjective measures of workload were gathered by the end of each session using the NASA-TLX (Hart & Staveland, 1988). Mixed design ANOVAs were conducted to examine the effects of different practice schedules on the productivity of transfer task (B3) obtained in the initial test and in the retention test.

9.3.5. Procedure

Participants were informed of the study's aim and that they would learn the basic controls and functions of a wheel loader simulator, and carry out some related tasks through a series of sessions on the wheel loader simulator. A preliminary questionnaire obtaining demographic information was administered before the first session began. In Session 1, participants were given five minutes to study a three-page printed instruction handout. It describes the parts and basic functions of the excavator and the corresponding operation of the joystick and pedal controls. Participants would then be seated at the loader simulator and tested with the Controls Familiarization module once (30 trials).

In Session 2, two groups of participants (Groups 1 and 3) were asked to complete the B1 task (bucket loading), whereas the other two group (Groups 2 and 4) were asked to complete the B2 task (filling a trench). All participants started with studying the instructions in the form of a booklet for five minutes before they began the first trial of the task. Both tasks required the participants to do three bucket loads, and these three cycles were repeated over five trials. All participants were asked to rate the workload measures at the end of the session.

In Session 3, Group 1 was given the B2 training, and Group 2 was given the B1 training, with both tasks following the same procedures as in Session 2. Groups 3 and 4, however, were trained to move a load (a bundle of pipes) with the fork attachment (Task F). Similarly, all participants were given five minutes to study the instruction when new tasks were introduced. For Task F, participants were asked to perform 5 trials. All participants were asked to rate the workload measures at the end of the session.

Session 4 is the skill transfer test for the truck loading task on the simulated loader. The participants were given a five minutes instruction to study the task, followed by 5 trials of the truck loading task. In each trial, the participants were asked to load three buckets onto the truck. All participants also rated the workload measures by the end of the session, followed by a post questionnaire addressing some questions about the task difficulty and perceived attention focus.

Session 5 was the retention test for the truck loading task on the simulated loader after a one-week interval, in which all participants performed the same truck loading module as in Session 4. All participants were asked to rate the workload measures again at the end of the session.

This study took approximately 2 hours for the first part of the experiment (Sessions 1 through 4) and 30 minutes for the retention test.

9.4. Results

9.4.1. Initial Practice

After the introduction and training with the controls familiarization modules in Session 1, Groups 1 and 3 were asked to complete the B1 task, whereas Groups 2 and 4 were asked to complete the B2 task. Both tasks required the participants to do three bucket loads, and these three cycles were repeated over five trials. A repeated-measures ANOVA was used to test the effects of trial and experimental groups on the total execution time per module. The results showed a significant trial effect, $F(4, 224) = 118.42$, $p < 0.001$, $\eta_p^2 = .679$, indicating the execution time dropped from trial 1 (297.93 s) to trial 5 (154.75 s). but neither an experimental group main effect, $F(3, 56) = .027$, $p = .994$, $\eta_p^2 = .001$, nor interaction with trial, $F(12, 224) = .403$, $p =.961$, $\eta_p^2 = .021$, was significant. Thus, the participants took approximately the same amount of time when they were first trained on either B1 or B2.

In Session 3, Group 1 was given the B2 training, and Group 2 was given the B1 training, with both tasks following the same procedures as in Session 2. Groups 3 and 4, however, were asked to perform 5 trials on the F task. Similar analysis with ANOVA was conducted to test the effects of trial and experimental groups on the total execution time per module and similar results were obtained, where only main effect of trial was found, $F(4, 224) = 104.93$, $p < .001$, $\eta_p^2 = .652$, showing a continuous drop in the execution time from trial 1 (259.11 s) to trial 5 (145.38 s). No interaction with experimental group, $F(12, 224) = .65$, $p = .798$, $\eta_p^2 = .034$, nor main effect of experimental group, $F(3, 56) = .355$, $p = .785$, $\eta_p^2 = .019$, was found. These results also confirmed that the training time on a second task was consistent even though the tasks assigned in this session were not the same for all groups.

9.4.2. Performance Measures

To examine the effects of training with an alternative tool (a wide fork) and returning to the original learned tool (a bucket) on a loader simulator on skill transfer and retention, the performance measures on the truck loading task (B3), which was used as transfer and retention test, were analyzed. Several performance measures obtained in B3 included bucket fill (%), truck fill (%), damages ($), and execution time. The bucket fill and truck fill is highly correlated in both the transfer test, $r(900) = .606$. $p < .001$ and the retention test, $r(900) = .656$. $p < .001$, indicating that the aggregates that were picked up from the stockpiles were mostly transferred to the truck, with a low chance of spilling from the bucket. Also, only 2.78% (25 out of 900 bucket loads) and 1.89% (17 out of 900 bucket loads) of the total number of bucket loads (3 buckets x 5 trials x 60 subjects) recorded truck damages in the transfer test and retention test, respectively. In Session 1, the participants were reminded about safety concerns and they were told to do their best to avoid safety violations, including improper carry height, boom raised too high on incline, excessive steering with boom up, and tipping the machine. If they performed any unsafe acts during the module, they would not receive a score and they would need to do amake-up trial. Thus it is believed that the low damage rate and the high correlation of bucket fill and truck fill indicated that the participants did pay full attention when performing the truck loading task and did avoid unsafe acts even though they were on a simulator. Only truck fill percentage and the total execution time were used to calculate overall productivity per trials.

9.4.2.1. Productivity on Truck Loading Task

Productivity (m^3/hr) was calculated as the total volume (total truck fill percentage x 12 yards) transferred from the pile to the truck (converted into m^3), divided by the total execution time (converted to hr). Figure 9.2 illustrates the mean productivity on the loader across the 2 sessions (transfer and retention) for all groups.

A mixed-design repeated-measures ANOVA (Appendix I) was used to test the effects of session (skill transfer test, retention test) and trial (Trials 1 - 5) on productivity per trial on the loader, with method (four training sequences) as a between-subjects factor. The ANOVA showed a main effect of session, $F(1, 56) = 119.50$, $p < .001$, $\eta_p^2 = .681$, with productivity increasing significantly from 110.86 m^3/hr to 132.53 m^3/hr (a 19.54% increase) across the two sessions. Trial was a significant factor, $F(4, 224) = 150.827$, $p < .001$, $\eta_p^2 = .729$, showing an increase in productivity from trial 1 to trial 5 within a session. The result of the main effect of method, $F(3, 56) = 8.94$, $p < .001$, $\eta_p^2 = .324$, showed that Group1 (B1>B2>B3) obtained higher productivity than Groups 3 and 4, which involved training with a fork before returning to a bucket task. Group 2 (B2>B1>B3) also showed a higher productivity than Group 4 (B2>F>B3) (see Appendix I). The session × method

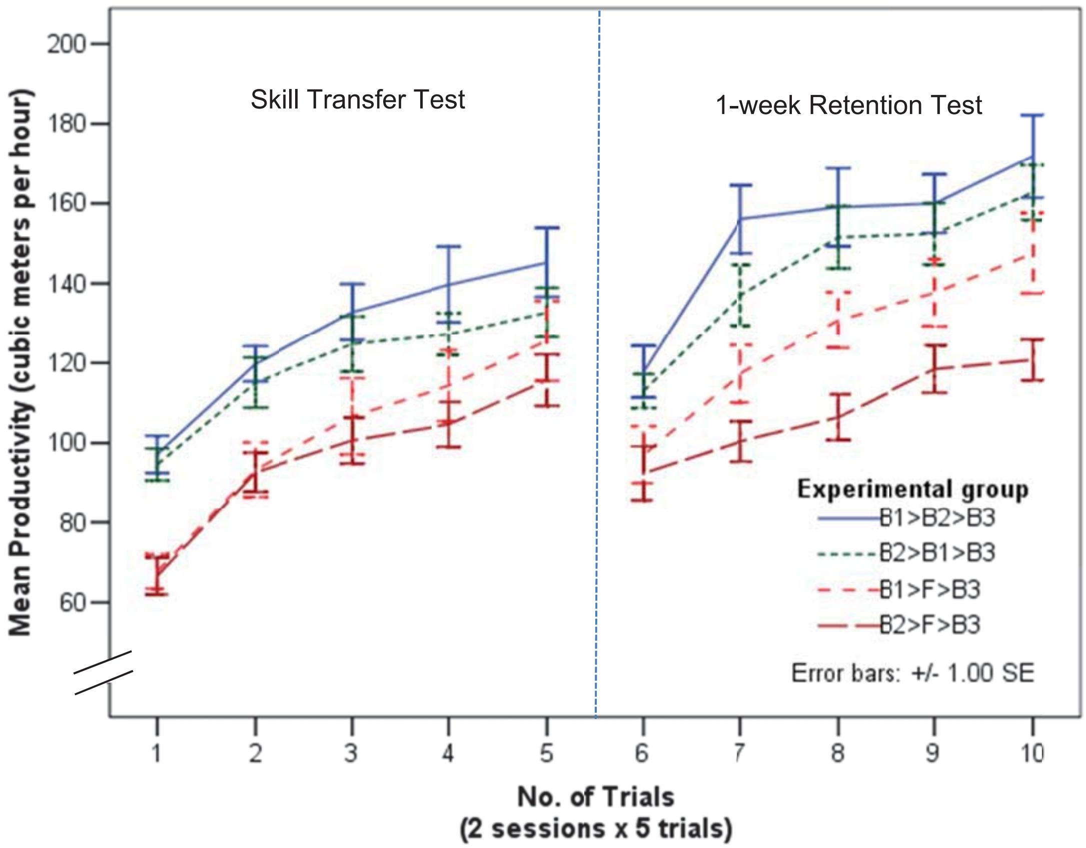


Figure 9.2. Productivity of truck loading task on the loader across the 2 sessions for the four experimental groups.

interaction (Figure 9.3) was significant, $F(3, 56) = 2.867$, $p < .05$, $\eta_p{}^2 = .133$, showing that Groups 1 and 2 practicing with bucket loading during the entire training phase showed a better transfer in both sessions (transfer test and retention test) compared to the groups that switched to the fork lifting task (Bonferroni pairwise comparisons tests, *p*s > 0.01). Group 3 showed a higher productivity in the retention test compared to Group 4 (Bonferroni pairwise comparisons tests, *p*s > 0.05).

9.4.2.1.1.<u>Effects on First and Last Trials in Each Session</u>

Further analysis was conducted to examine the effects on the productivity obtained in first and last trials on the transfer test (initial test on the bucket task after the training) and the retention test (a week after), with session (skill transfer test, retention test) and trial (first, last) on productivity as within-subject factors and method (four training sequences) as a between-subjects factor. The ANOVA (Appendix I) showed a main effect of session, $F(1, 56) = 98.85$, $p < .001$, $\eta_p{}^2 = .638$, with productivity better in the second session than in the first. Two other main effects were significant: trial, $F(1, 56) = 256.58$, $p < .001$, $\eta_p{}^2 = .821$, showing an increase in productivity from the first trial to last trial within a session, and method, $F(3, 56) = 8.94$, $p < .001$, $\eta_p{}^2 = .324$, showing that Group1 (B1>B2>B3) obtained higher productivity than Groups 3 and 4, the latter two groups involving training with a fork before returning to a bucket task. Group 2 (B2>B1>B3) also showed a higher productivity than Group 4 (B2>F>B3) (*ps* < .01). The session × trial × method interaction (Figure 9.4) was significant, $F(3, 56) = 2.806$, $p < .05$, $\eta_p{}^2 = .131$.

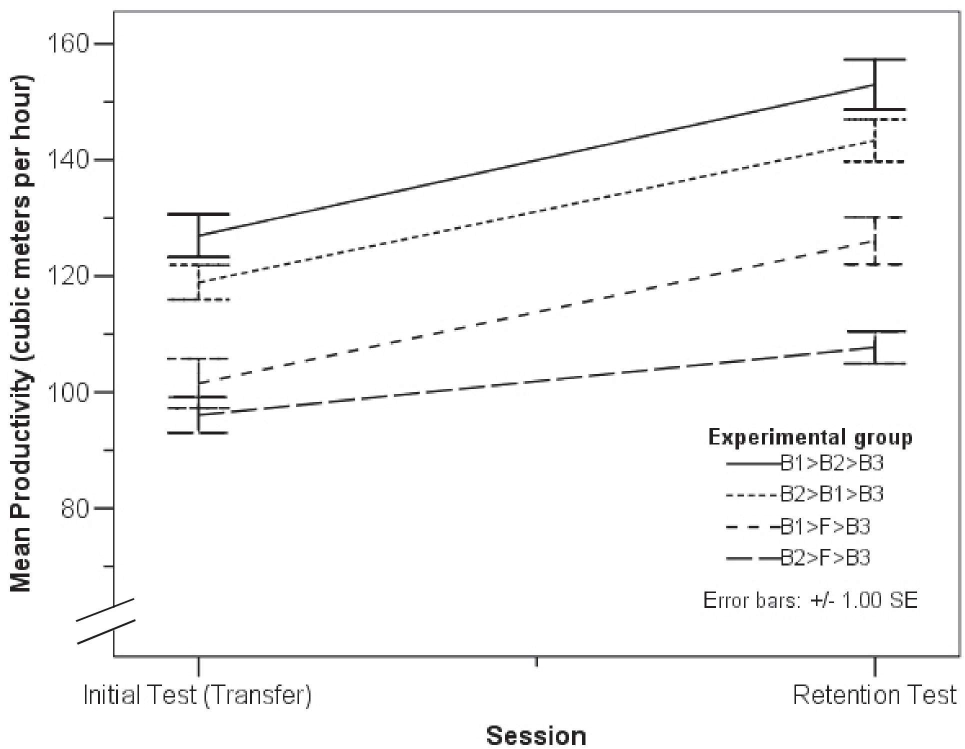


Figure 9.3. Two-way interaction plot of session × method on productivity of truck loading task.

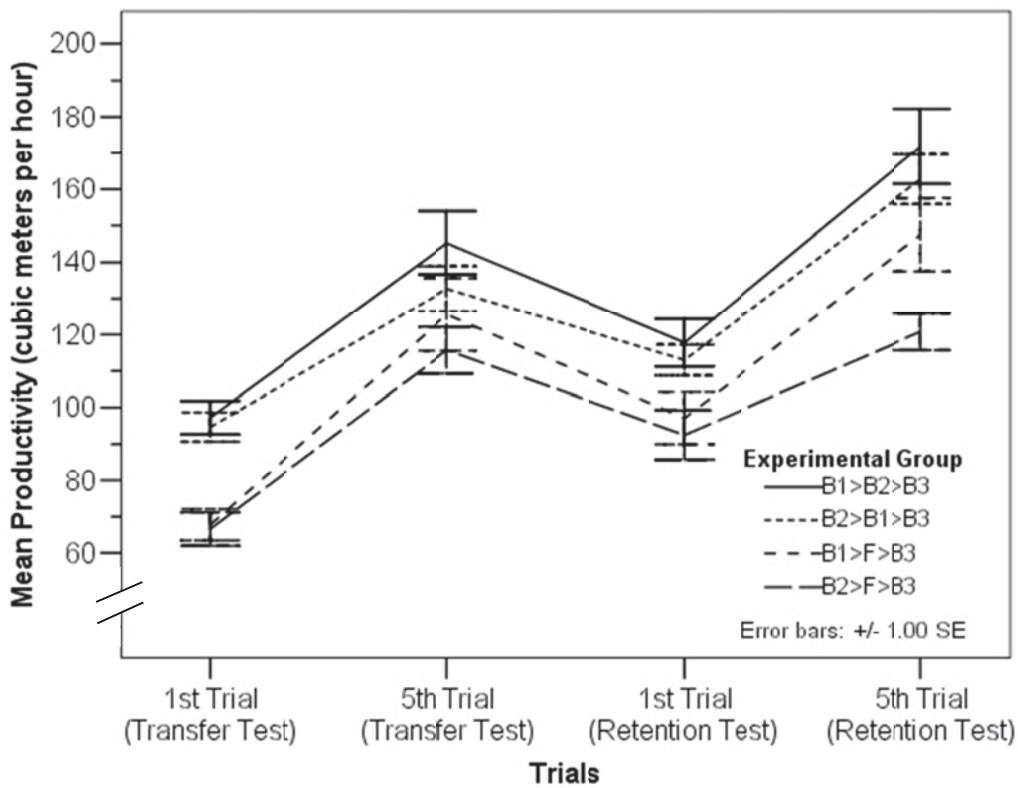


Figure 9.4. Three-way interaction plot of session × trial × method on productivity of truck loading task.

Bonferroni pairwise comparisons tests show that Groups 1 and 2 had a higher productivity in the first trial on the truck loading task compared to Groups 3 and 4 (*ps* $< .001$). No significant differences were found among the four groups in the last trial of the transfer test. When they returned to the truck loading after a-week interval, only Group 1 showed significantly higher productivity than Group 4 (*ps* $< .001$) in the first trial, and both Groups 1 and 2 obtained higher productivity than Group 4 in the last trial of the retention test, but did not differ significantly from Group 3.

9.4.2.2. Workload Measures

9.4.2.2.1. Transfer Test Versus Retention Test

A mixed design ANOVA with the mental measure (6 attributes) and session (transfer, retention) as within-subject factors and method (4 training sequence) as a between-subjects factor were conducted. The ANOVA (Appendix I) showed a main effect of session, $F(1, 56) = 47.00$, $p < .001$, $\eta_p^2 = .456$, the average workload measure dropped significantly from transfer test ($M = 4.47$) to retention test ($M = 3.71$). The main effect of measure was also significant, $F(5, 140) = 8.78$, $p < .001$, $\eta_p^2 = .239$, where participants rated the tasks as requiring higher mental demand ($M = 4.78$) and effort ($M = 4.69$), but lower temporal demand ($M = 4.10$), physical demand ($M = 3.84$), less frustration ($M = 3.93$), and the participants were satisfied with their performance ($M = 3.19$). No interaction between measure and task was found. The effect of training method was also found to be significant, $F(3, 56) = 48.91$, $p < .001$, $\eta_p^2 = .256$, showing that Group 1 had a lower average workload measure than the other three groups (*ps* < 0.001). The session × group interaction (Figure 9.5) was significant,

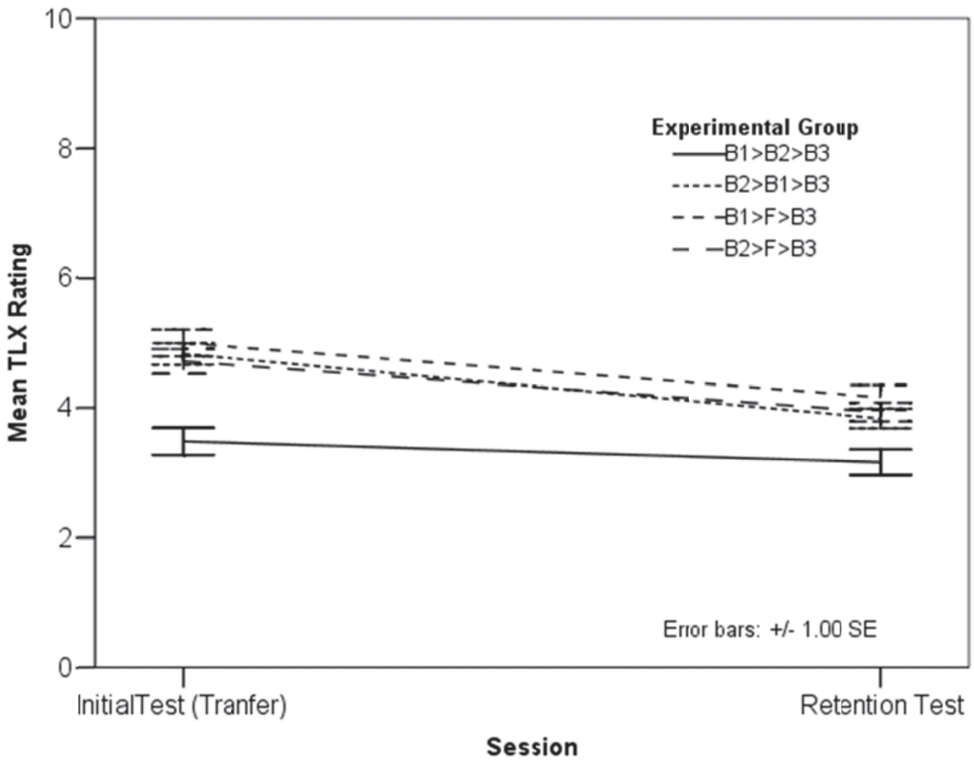


Figure 9.5. Two-way interaction plots of session × group on mean TLX rating across the two sessions.

reflecting that Group 1 rated the truck loading task as constituting a lower workload while the other three groups did show a drop in mean TLX rating from transfer test to retention test, indicating that the workload is reduced in the retention test.

9.4.2.2.2. Workload Measures of the Four Loader Tasks

Each participant was only trained with 3 out of 4 training modules during the experiment. Consequently, four separate repeated-measures ANOVAs with session and measure as within-subject factors were used to evaluate the effects of different training modules on the TLX ratings for each experimental group. All ANOVAs (see Appendix I) showed only a main effect of measure, indicating that participants rated the tasks as requiring higher mental demand and effort but lower temporal demand and physical demand, and less frustration, and that the participants were satisfied with their performance. No significant differences were found in the workload measure between B1, B2, B3 and F.

9.4.2.3. Task Difficulty Ranked by the Novices

At the end of Session 4, participants were asked to rank the difficulty of the three tasks that they were being trained during the experiment (depending which group they were assigned to) from easiest to hardest (1: easiest to manipulate; 3: hardest to manipulate). The analysis of perceived difficulty by the operators from the three tasks in each experiment group was conducted with the Friedman test. The post hoc analysis with Wilcoxon Signed Rank tests were used if Friedman test showed significant differences among the tasks. The distribution of the ranks of the loader tasks by each group is shown in Appendix I. The test results (see Appendix I) showed that the tasks were ranked differently by Group 1 (B1>B2>B3), $\chi^2(2) = 14.40$, $p < 0.001$, indicating

B1 is ranked significantly easier than B2 and B3, but no differences in difficulty levels between B2 and B3 were found. Similar results were found for group 2 (B2>B1>B3), $\chi^2(2) = 12.13$, $p < 0.005$, where B1 is ranked significantly easier than B2 and B3. The Friedman tests also showed significant difference in difficulty levels for Group 3, $\chi^2(2) = 28.13$, $p < 0.001$, and Group 4, $\chi^2(2) = 8.37$, $p < 0.05$. All participants in Group 3(B1>F>B3) ranked fork lifting task as most difficult, indicating that they are all consent that F is most difficult. The results of Wilcoxon Signed Rank tests also shown that B1 was ranked significantly easier compared to B3 and B3 was ranked significantly easier than F. In Group 4, only F is ranked more difficult than B2 and B3. No significant differences were found between B2 and B3, which such pattern was also found in Group 1 and Group 2.

In summary, it is consistent that B1 is the easiest task, followed by B2 and B3, whereas, F is the most difficult task. This ordering is in agreement with the results obtained by the experts, where the difficulty levels of tasks followed this order: B1 (easiest) < B2 < B3 < F (most difficult), although the experts' rankings did show a significant difference between B2 and B3.

9.5. Novices vs. Experts

In Study 4, all experienced operators were asked to execute the truck loading task 5 times on the loader simulator. The results were intended to provide benchmarks of experienced loader operator performance for comparison to the performance of the novices, which was measured in Study 5. A repeated-measures ANOVA was used to test the effects of group (Experts, 4 training groups) and trial (1 to 5) on productivity on the truck loading task. The performance obtained in retention test by the novices was used for comparison. The results showed a main effect of trial, $F(4, 252) = 59.523$, $p < .001$, $\eta_p^2 = .486$, with an increase in productivity from trial 1 to trial 5. Group is a significant factor, $F(4, 64) = 25.07$, $p < .001$, $\eta_p^2 = .614$, indicating that the productivity obtained by the experts was significantly higher than that for the novice groups. The trial × group interaction (Figure 9.6), $F(16, 252) = 2.26$, $p < .005$, $\eta_p^2 = .126$, illustrates that the benchmark obtained by experts is higher than that of the novices and exhibits a more steady performance throughout the trials, whereas the novices showed significant and continuous improvement from trials 1 to 5 (Bonferroni pairwise comparisons tests, *p*s > .001).

Besides getting a benchmark for truck loading performance, it is also of interest how much attention the experienced operators give each type of displayed feedback. Seven types of visual feedback were provided to the participants during each trial, including egocentric view, damages, time, instruction, warning, bucket fill percentage, and side view (see Figure 8.2). The ANOVA shows a main effect of feedback, $F(6, 378) = 96.66$, $p < .001$, $\eta_p^2 = .605$, with the majority of time focused on the egocentric view ($M = 43.93\%$) and side view ($M = 22.73\%$), followed by bucket fill % ($M = 22.59\%$). The feedback × group interaction (Figure 9.7), $F(24, 378) = 2.62$, $p < .001$, $\eta_p^2 = .143$, shows that experts primarily focused on the egocentric view ($M = 66.38\%$) and secondly relied on the side view ($M = 20.63\%$). The rest of the feedback only contributed 10% of the total time. Novices followed a similar trend with the majority of time focusing on the egocentric view ($M = 37.98\%$, but 1.75 times less compared to the experts) and secondly relied on the side view($M = 23.25\%$), except that the novices

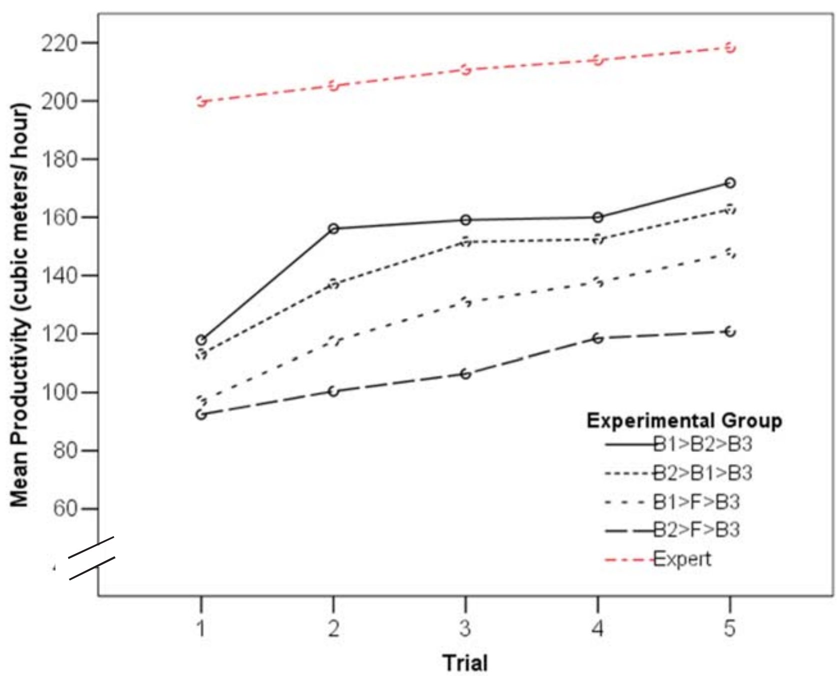


Figure 9.6. Two-way interaction plots of trial × group on productivity (red line indicates the benchmark).

also spend more than one-third of the time focusing on other feedback such as bucket fill(%), warning, time, etc. Group is not significant, $F(4, 64) = 1.985$, $p = .108$, $\eta_p^2 = .112$. The descriptive statistics showing the perceived percentage (%) of time spent on the visual feedback on the simulator screen of experts ($N = 8$) and novice ($N = 60$) are shown in Appendix I.

9.6. Discussion

This experiment was designed to investigate whether there were performance costs when training with an alternative tool and returning to the previously learned tool on the same machine. There was a significant main effect of method showing Group 1 (B1>B2>B3) obtaining higher productivity than Groups 3 (B1>F>B3) and 4 (B2>F>B3) and Group 2 (B2>B1>B3) also showing a higher productivity than Group 4. This outcome suggests that when groups are assigned to practice sequentially on two tasks involving the manipulation of the same tool, they perform better than a group that switches to a different tool in the new skill transfer test that also makes use of the original tool. These results supported Hypothesis 5 as the groups provided practice only with loader bucket manipulation during the training phase performed better in the test phase (skill transfer test on a new task) and showed better retention after a week interval. This outcome also supported the procedural reinstatement principle (Lohse & Healy, 2012), for which practicing a similar mental model (i.e., practicing bucket loading both in B1: simple bucket loading and in B2: filling a trench tasks) during the learning phase may facilitate subsequent retention and transfer in the test phase where a similar mental model (B3: truck loading task *with a bucket*) is tested.

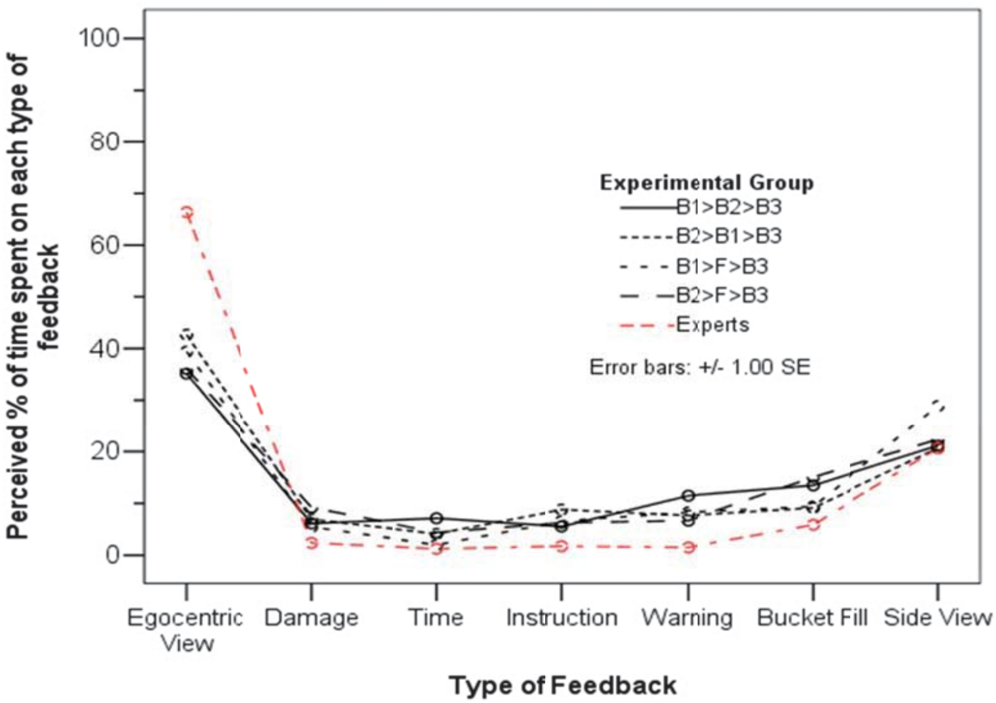


Figure 9.7. Two-way interaction plots of feedback × group on the perceived % of time spent on each type of feedback.

Group 4 (B2>F>B3) obtaining a significantly lower performance in both transfer and retention tests than Groups 1 and 2 could be explained by the identical elements theory (Thorndike & Woodworth, 1901), according to which transfer of learning depends on the proportion to which the learning task and the transfer task are similar. Previous development of the HTA for the Fork Lifting task yielded a lower similarity in comparison to those for the three bucket loading tasks. Among the three bucket loading tasks, B1, B2 and B3, the two HTAs of B1 and B3 have identical goals, subgoals, and plans, except after filling the bucket, B1 requires the operator to drive toward and dump to a targeted dump area, whereas the operator drives and dumps into the truck box in B3. Indeed, four operators indicated that B1 and B3 are similar in terms of controls and 1 operator also said 'B1 and B3 are almost identical mechanism'. This degree of similarity between B1 and B3 may offer an explanation why Group 3 showed higher productivity in the retention test than Group 4, even though both groups had practiced on the fork lifting. One critical thing to note is that in the last trial of the retention test, unlike Group 4, Group 3's performance was not significantly different from that of Groups 1 and 2. This may suggest that B1 was a critical task during the training phase to achieve a better performance in test (B3) involving a high similarity of controls and mechanisms (as seen in the HTA and the comments from the experienced operators). This finding also supports Speelman and Kirsner's (2001) finding that old skills continue to improve if task conditions are not altered, because the participants were performing in essence the same task as that in the immediate test.

In terms of task difficulty, from the previous study, the four tasks were ranked by the experts in this order: B1(easiest) < B2 < B3 < F(most difficult) while the novices' rankings of the tasks obtained in this study also indicated that B1 is the easiest and F is most difficult. In this study, Groups 1 (B1>B2), 3(B1>F), and 4(B2>F) were presented in order of increasing difficulty. Group 1, the only one for which the modules were presented completely (i.e., including the transfer task) in increasing difficulty order, showed a benefit in transfer and in retention, whereas for Groups 3 and 4, the difficulty introduced by the fork lifting task seemed to impede transfer and retention. This result is not consistent with *Hypothesis 6a*, that any manipulation of task difficulty during training may have facilitating effects during retention and transfer testing, as advocated by Healy, and Bourne (2002), but supports the claim of McDaniel and his colleagues that introducing difficulties during training is facilitative only when the training and retention tasks share task-relevant cognitive processes to the to-be-learned materials (McDaniel & Butler, 2011; McDaniel & Einstein, 2005). In this study, the benefit shown in Group 1, with practice on B1 first and then B2 and being tested on B3, may imply that having identical elements may be more important than introducing task difficulty, support *Hypothesis 6b*. Thus, it is suggested that when training perceptual-motor tasks, tasks being practiced during the learning phase should match the transfer task. Manipulation of task difficulty may play a role only if the tasks share task relevant cognitive processes and mental models.

In Figure 9.2, the observation is made that if the first trial of the retention test is excluded, the curves obtained by Group 1 (B1>B2>B3) and Group 2 both would show a fairly continuous curve from the 1st trial in the transfer test to the 5th trial in the retention test. Discounting the first trial of the retention test, performance effectively picks up where it had left off in the last trial of the previous session. A decrease in productivity on the first trial of the retention test may reflect a warm-up decrement associated with recollecting the old skills (Schmidt & Lee, 2011) or a “fast, transient dimension of adaptation”, as explained by Newell, Mayer-Kress, Hong, and Liu (2009). Alternatively speaking, the subjects were able to pick up the skills very quickly after a short warm-up period when they first returned to the task and were able to continue to improve throughout the session. However, this pattern may only be seen when learning was truly occurring during the training phase, since Groups 3 and 4, with the training of fork lifting, did not benefit when returned to a transfer task that involved manipulation of a previously learned tool. A similar result was obtained in an earlier study (So et al., 2013), which examined whether part-task training produces better learning and retention than whole-task training of a trench-and-load task performed on the hydraulic excavator simulator. The results from So et al.’s study (Figure 9.8) showed that the continuous projection of the performance improvement only occurred in part-task training, where the benefit of part-task training for better retention was found. Part-task training provided better learning during the training phase, which allowed the skills acquired from the part tasks to enable better performance in the retention test. Both studies tend to suggest that an effective training method which enables true skill acquisition during the learning phase allows participants to pick up the skills from where they left off at the end of the first session very quickly, after a quick warm-up period (1^{st} trial) when they return to the same task after an interval of a week or longer, and continue to improve their performance throughout the session.

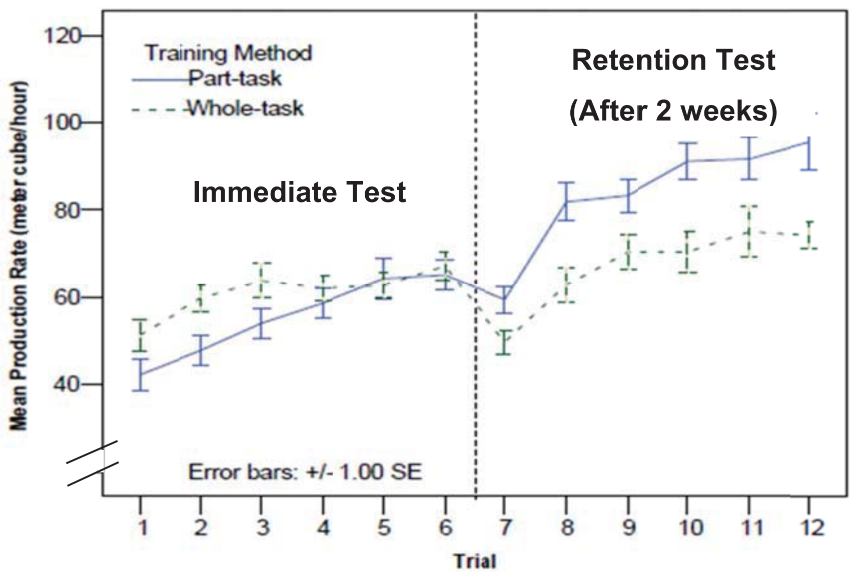


Figure 9.8. The performance on the simulated excavator of all trials in two sessions (initial test and 2-week retention test) in So et al.'s (2013) study.

9.7. Conclusion

This study (#5) investigated whether an alternating practice sequence with the same machine, i.e., training with an alternative tool and returning to the original learned tool, yields better skill transfer and retention. The results, obtained on a loader simulator, showed that when groups were assigned to practice on two different tasks involving the manipulation of buckets, they would perform better in the skill transfer test where a new task also involved the manipulation of the bucket. The finding of this study fully supported the identical elements theory and procedural reinstatement theory, but not the progressive difficulty principle. Indeed, this study suggested that when training perceptual-motor tasks, the elements trained in the tasks during the learning phase should match closely the transfer tasks. Manipulation of task difficulty may play a role only if the tasks during the learning phase share task-relevant cognitive processes and mental models.

CHAPTER 10. GENERAL DISCUSSION

10.1. Summary of Main Findings

This research consisted of two parts mainly focusing on skill acquisition, retention, and transfer between two machines (Part 1) and between tasks within a single machine (Part 2), as demonstrated on virtual reality-based training simulators. Two training principles—specificity and task difficulty—were explored. Whether introducing an alternative type of construction equipment or a different task to practice during training will have positive or negative effects on learning, retention and transfer was addressed. To understand skill development for the operations of construction equipment and to distinguish the skills to be acquired for each task or machine, interviews and verbal protocol analysis with expert operators were employed. TLX ratings were also gathered to measure the subjective cognitive load associated with each task.

Part 1 consists of three studies, where 2 experiments were designed to verify whether alternating practice sequence yields better skills transfer and retention for both simple response selection task and a complex task that involves multiple operations, based on the principle of specificity of training. The main finding of Studies 1 and 2 is that no cost or benefit was found from inserting practice on a simulated loader while also learning on a simulated excavator for both a simple task—controls familiarization task and a complex task—truck loading. Given a fixed amount of total training time, the two groups whose practice was intervened by the practice of a similar task with a loader continued to show improvement when returning to the excavator. Practice with the loader between the excavator sessions did not alter the excavator learning, as performance picked up at the level of the prior excavator session and continued to improve.

To better understand what caused the loader group to improve less when returning to the excavator compared to the control group, HTA was used to reveal common elements that define the essential similarities at various levels in the overall task structures. HTAs revealed why no positive transfer was found in performing the truck loading task alternately with the excavator and loader. The lack of transfer was likely due to the differences between loader and excavator in terms of the controls, physical constraints, the goals, and subgoals of the task. However, there is a limitation using HTA, whereby simply comparing the number of levels of subgoals did not reveal the level-of-difficulty differences between tasks. Thus, using TLX ratings which maybe more sensitive to capture the workload measurement of different tasks and measuring the actual performance on the tasks may serve as better indicators to evaluate the relative task difficulty.

Part 2, containing Studies 4 and 5, focused on training multiple tasks within a machine. In Study 4, a different method of conducting HTAs—verbal protocol—was used to decompose the four training modules on the loader simulators. A systematic method for how the HTAs can be derived from think-aloud protocols was also developed in this study. The four HTAs were successfully generated from the verbal protocols following the nine proposed steps. The findings show that 1) the HTA of the Fork Lifting module is significantly different from those of the three bucket loading tasks and 2) although all three bucket loading tasks shared a similar mechanism, the operators ranked B1 as the easiest, followed by B2 and B3, due to the reduced size and accessibility of the area to which the operators had to attend, and fork lifting was ranked as the most difficult task.

Study 5 was an experiment conducted with student participants to verify whether an alternating practice sequence with the same machine, i.e., training with an alternative tool (a wide fork) and returning to the original learned tool (a bucket) on a loader simulator, yields better skill transfer and retention. Four experimental groups were tested. Two groups were given practice on tasks involving bucket loading, whereas the other two groups were at first given a bucket loading task to practice and then switched to practice a fork lifting task in the next session before they returned to test on a new task which involved bucket loading. The results showed that the groups who were assigned to practice on two tasks involving the manipulation of buckets would perform better in the skill transfer test which involved the manipulation of the bucket. These results supported the specificity of training principle but did not conform with the progressive difficulty training principle. It is suggested that when training perceptual-motor tasks, tasks being practiced during learning phase should match the transfer task. Manipulation of task difficulty may play a role only if the tasks share task-relevant cognitive processes and mental models.

10.2. Research Questions

10.2.1. How much does training on one machine transfer (positively or negatively) to other machines?

In this project, two machines were used to examine the issue of transfer. In Studies 1 and 2, participants were given practice on a simple task (controls familiarization, which requires prompt operation of a correct control action in response to a visual command) or a complex task (truck loading, which requires actual operation of the machinery to complete a particular task) on the excavator and moved to the loader for the same task and returned to the excavator. Both studies showed no performance cost on either task attributable to inserting practice on a loader while also learning on an excavator. The group whose practice on the excavator was interrupted by the practice on the loader continued to show improvement on the excavator, with performance picking up where it had left off. Hence, neither positive nor negative transfer was found when practicing on excavator but being interrupted by practice on loader. In other words, training on one machine did not transfer positively nor negatively to training on the other machine.

10.2.2. Does insertion of training on various machines facilitate (or inhibit) learning and retention on a previously practiced machine?

There are two types of retention being studied and measured in this research: 1) the retention on the previously learned machine right after insertion of performance on an alternative machine, and 2) retention on the same task after 1-week interval. This question refers to the first type. Study 1 examined the retention of one group on the controls familiarization task on the simulated excavator after practicing on the same task on the simulated loader in three sessions (E>L>E), and their performances were compared against the control group who practiced on the excavator throughout three session (E>E>E). The results did not show effects of having received the intervening training on the loader. The lack of significant difference from the group who practiced on the same machine (i.e., excavator) for all three sessions suggests that switching from one machine to another does not degrade the original performance. In Study 2, the number of sessions was increased, from three to five, to examine the possible influence when participants continue to switch between the machines twice and the simple controls familiarization task was replaced with a complex truck loading task. Similar results were obtained in both retention tests where no performance costs were found on the excavator after practicing on a loader. Also, practicing on the excavator between the two loader sessions for the loader group (E>L>E>L>E) did not negatively impact their returning performance on the loader, indeed showing continuous improvement in performance on the loader throughout the session. This result implies having practice on an alternative machine does not inhibit learning and retention of another previously learned machine.

In both studies, the participants were able to resume their skills where they left off and continue to improve the performance throughout the session when they returned to the previously learned machine. It is possible that insertion of training on various machines may have indeed facilitated learning and retention when the participants returned to the previously practiced machine. One thing that could be confirmed from these two studies is that insertion of training on more than one machine did not degrade their learning and retention on the previously learned machine.

10.2.3. When should an alternative machine be introduced in the training if skills on multiple machines are required of an operator?

In Studies 1 and 2, participants were exposed to a different machine after 20 to 25 minutes practice on the initial learned machine. Although the participants would not obtain proficiency of that machine during that training period before switching to another machine, they were able to recover their skills from where they had left off in the previous training session when they returned to the previously learned machine. In Study 2, besides the two experimental groups (control and loader groups), an additional group, was added to address the question of how the duration of insertion of practice on an alternative machine matters to the performance on the previous learned machine. This group was given practice on the loader for two consecutive sessions before switching back to the excavator (E>L>L>E>E). The results showed that there was no negative transfer due to a longer practice on the loader. Alternatively, the skills learned previously on the excavator simulator were retained even after the participants learned and practiced on the loader simulator for two consecutive training sessions. Such findings suggest that the timing of when an alternative machine should be introduced may not be critical and do not require the operator to fully master one machine before a new machine is introduced.

10.2.4. What is contributing to positive or negative transfer when switching between machines?

Both Studies 1 and 2 did not show positive or negative transfer when switching between machines. In particular, the lack of transfer with the truck loading task in Study 2 was likely due to the differences between the loader and excavator in terms of the controls, physical constraints, the goals and subgoals of the tasks, as examined in Study 3, where HTA was used to study skill transfer and found to serve as a useful tool for modeling the tasks in the form of goals and subgoals. Thus, by identifying elements that tasks have in common, the HTAs could suggest where benefits of training may transfer.

10.2.5. Is there positive or negative transfer due to switching tasks within a machine?

The two types of switching tasks were examined in Study 5: 1) switching tasks that did not share task-relevant cognitive processes and mental models (fork lifting vs bucket loading) and 2) switching tasks that do share task relevant cognitive processes and mental models (tasks that involve the manipulation of buckets, such as truck loading and filling a trench). The results showed that the groups provided practice only with loader bucket manipulation during the training phase performed better in the test phase (skill transfer test on a new task) and showed better retention after a week interval than the groups that switched to fork lifting. This outcome also supported the procedural reinstatement principle (Lohse & Healy, 2012), for which practicing a similar mental model (i.e., practicing bucket loading both in B1: simple bucket loading and in B2: filling a trench tasks) during the learning phase may facilitate subsequent retention and transfer in the test phase where a similar mental model (B3: truck loading task) is tested. It is suggested that when training perceptual-motor tasks, the elements trained in the tasks during the learning phase should match closely the transfer tasks, in order to obtain positive transfer.

10.2.6. Can the complex perceptual-motor operator skills acquired during simulator training be retained for at least a week over which there is no interaction with the simulator or related equipment?

Yes. Study 5 examined whether the participants could retain the skills of performing the truck loading task on the simulated loader after a week. The results showed that the participants were able to pick up the skills very quickly after a short warm-up period when they returned to the task after week and were able to continue to improve throughout the session. A similar result was obtained in an earlier study (So et al., 2013), which examined whether part-task training produces better learning than whole-task training of a trench-and-load task performed on the hydraulic excavator simulator and whether the skills could be retained after 2 weeks. Both studies tend to suggest that an effective training method which enables skill acquisition during the learning phase allows participants to pick up the skills from where they have left off at the end of the first session very quickly when they return to the same task after an interval of a week or two, and continue to improve their performance throughout the session.

10.3. Practical Implications

The present project has attempted to provide better understanding of skill development for the operation of construction equipment and how the trainees may better spend their practice time for (a) single machine and (b) multiple machines training. The practical implications of the findings of this research are:

1. If a trainer wants to maximize learning to operate a machine during a finite time period, practice should be devoted to that machine, whereas if the trainer wants to provide experience with two machines, this can be done without the practice on one machine having a negative effect on the learning of the other.
2. To the extent that operation of alternate equipment types is found to be dissimilar, concurrent practice on one equipment type should not set back learning on the other. This finding especially inspires consideration of concurrent simulator-based training rather than the practice of learning to operate only one machine at a time.
3. When training different tasks on the same machine, the elements trained in the tasks during the learning phase should match closely the transfer tasks. Manipulation of task difficulty may play a role only if the tasks during the learning phase share task relevant cognitive processes and mental models.
4. Similarity in the overall goals of the tasks, e.g., truck loading, is less important than similarities among the subgoals that comprise the tasks as performed on the respective equipment types. Detailed task analyses should reveal common elements that define the essential similarities at various levels in the overall task structures. HTA could serve as a useful tool for modeling the tasks in the form of goals and subgoals to study skill transfer. By identifying

elements that tasks have in common, the HTAs suggest where benefits of training may transfer.

5. Both the method of having experts to comment on and modify a preliminary HTA and the method of verbal protocol analysis successfully deconstruct the training module and develop the HTA diagrams. However, the primary downside to using verbal protocol analysis is the detailed and time-consuming nature of the process, in which transforming the verbal protocols requires a substantial amount of additional time and work for transcribing and verifying the verbal data, cleaning the verbal report, extracting codes, searching for themes and developing the HTA structure. Also, verbal protocol analysis may not involve experts to verify the final HTAs. Indeed, it would be ideal to always go back to the experts for verification to provide a more generic HTA to rule out exceptional cases where some operators may not follow norms.

10.4. Limitations and Future Direction

10.4.1. Variety of Machines

In this current research, only two pieces of construction equipment were studied. However, since skilled operators of heavy construction equipment may require skills at operating even more machine types, goals for training may facilitate that reality. Due to the recent availability of multiple different simulators in today's construction training schools, it is necessary to investigate the most effective way to train the beginning operators with several machines within a fixed amount of training time. It remains to be seen whether adding a third or even a fourth machine type to concurrent training will affect learning and retention in some way not revealed in this research.

10.4.2. Training Modules

This research only tested a few built-in training modules on the two simulators. The first half of this research focused on Controls Familiarization modules and Truck Loading modules which are available on both excavator and loader simulators. The second half tested 4 training modules on the loader simulator, with a focus on investigating how training with an alternative tool attached to the front of the loader affected the performance when returning to the original learned tool. Future studies could look into the skill transfer of training tasks sharing similar goals of other built-in training modules on the simulators, for example, digging heavy boulders from a bench vs. digging light material from a bench (on the excavator simulator). These tasks share the same bucket movement elements but manipulate dramatically different sizes of aggregates. Another possibility is transporting a tall load with a fork (e.g., a portable toilet) versus transporting a flat load (e.g., a bundle of pipes) on the loader simulator, which share the same fork lifting elements, but presenting different lifting and movement challenges. Finally, transfer between bench climbing/descending (on the excavator simulator) versus driving on a jobsite (on the loader simulator) may be investigated, as both require safely maneuvering the machine through jobsite obstacles.

10.4.3. Training Time

In this research, each participant was given 3-5 training sessions on the same day (except for a one-week interval retention test in Study 5). Each session took only about 20-25 minutes. This research did not focus on how daily practice on one machine/task and switching to an alternative machine/task would affect their performance when returning to a previously learned machine/task. It would be of interest how an increase in the practice time on simulators, e.g., practice on the excavator for a couple of hours and switching to a loader on the next day for a couple of hours and returning to the excavator the following day, would affect their returning performance on the excavator.

10.4.4. Transfer to Real Machine

For simulator training to be effective, one of the most important questions that must be answered is how skills learned from the simulators are transferred to other tasks, devices, situations and, ultimately, to real machines. Research studying skill transfer from simulator to real machine is still rare in the literature. The rise and continuing improvement of immersive simulators offers a range of tools to supplement training and great opportunity to explore training issues more generally.

10.4.5. Research-Friendly Simulators

The two training simulators of construction equipment employed in this study are modeled after specific models of real machines, with different training modules intended to develop skills in basic machine controls, proper operator technique, and safe job site operation. Since the simulators are not designed primarily for research purposes, in this research, many data were recorded manually, e.g., the loader simulator did not display the summary report after each trial in the Controls Familiarization module.

Among the three common training principles on skills retention and transfer: specificity of training (Identical Elements Theory), variability of practice, and task difficulty (Progressive Difficulty Training), only two training principles were assessed. The current simulators do not capture the degree of movement and how much forces applied on the controls by the trainees. A modification on the programming of the simulator, may make the assessment of greater variations of practice on these two machines feasible, e.g., to improve the bucket fill and truck fill per cycle by a correct movement of the controls to obtain optimal bucket alignment.

10.5. Conclusion

Construction equipment operation is used as an example of a complex perceptual-motor skill that is unique compared to cognitive or simple motor tasks that have been widely studied. There are indications from this study that its unique complexity negates simplistic application of fundamental principles of skill acquisition. This research has contributed to our understanding of how established skill acquisition principles govern the learning of such complex tasks by addressing some of the training issues on how to facilitate transfer and retention through different practice schedules that are based on the understanding of two common principles: Specificity of Training and Task Difficulty. The findings especially inspire consideration of concurrent simulator-based training rather than the practice of learning to operate only one machine at a time. When training different tasks on the same machine, the elements trained in the tasks during the learning phase should match closely the transfer tasks. The tasks trained in the learning phase should also share task relevant cognitive processes and mental models if manipulation of task difficulty is considered.

To study skill transfer, HTAs can be used for modeling the tasks in the form of goals and subgoals. By identifying elements that tasks have in common, the HTAs suggest where benefits of training may transfer. Both the method of having experts to comment on and modify a preliminary HTA and the method of verbal protocol analysis could be used to deconstruct the task and develop the HTA diagrams. A nine phase, systematic method for how the HTAs can be derived from think-aloud protocols was also developed in this study. However, the primary downside to using verbal protocol analysis is the detailed and time-consuming nature of the process. The implications of the findings advance our ability to predict the outcomes from implementing a particular practice schedule, especially when training on multiple machine types is in view. The findings are expected to generalize to heavy equipment training in related domains, such as forestry and mining, which require shoveling, drilling, loading, hauling, dozing, excavation, etc. The results may also generalize to domains requiring instrument handling skills, such as surgery, domains requiring operation of robotic arms in explosive area or even orbital space vessel external operations.

LIST OF REFERENCES

LIST OF REFERENCES

Advisory Group for Aerospace Research and Development. (1980). *Fidelity of simulation for pilot training (Tech. Rep. No. AGARD-AR-159)*. Neuilly sur Seine, France: North Atlantic Treaty Organization.

Ahlberg, G., Heikkinen, T., Iselius, L., Leijonmarck, C. E., Rutqvist, J., & Arvidsson, D. (2002). Does training in a virtual reality simulator improve surgical performance? *Surgical Endoscopy and Other Interventional Techniques, 16*, 126–129.

Allport, A., Styles, E., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Annett, J. (2003). Hierarchical Task Analysis. In E. Hollnagel, (Ed.), *Handbook of Cognitive Task Design* (pp. 17–35), Mahwah, NJ: Lawrence Erlbaum Associates.

Annett, J. (2004). Hierarchical Task Analysis. In: D. Diaper and N.A. Stanton (Eds.), *The Handbook of Task Analysis for Human-Computer Interaction* (pp. 67–82). Mahwah, NJ: Lawrence Erlbaum Associates.

Annett, J., Cunningham, D., & Mathias-Jones, P. (2000). A method for measuring team skills. *Ergonomics, 43*, 1076–1094.

Annett, J., & Duncan, K. D. (1967). Task analysis and training design. *Occupational Psychology*, *41*, 211-221.

Annett, J., Duncan, K. D., Stammers, R. B., and Gray, M. J. (1971). *Task analysis*. London: Her Majesty's Stationary Office.

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, *21*, 627–635.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612–637.

Bernardin, J.B., & Mason, A. H. (2011). Bimanual coordination affects motor task switching, *Experimental Brain Research*, *215*, 257–267.

Banks, V. A., Stanton, N. A., & Harvey, C. (2014). What the drivers do and do not tell you: using verbal protocol analysis to investigate driver behaviour in emergency situations. *Ergonomics, 57*, 332–342.

Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 23-44). Hillsdale, NJ: Erlbaum.

Beauchamp, M. K., Harvey, R. H., & Beauchamp, P. H. (2012). An integrated biofeedback and psychological skills training program for Canada's Olympic short-track speedskating team. *Journal of Clinical Sport Psychology, 6*, 67–84.

Bhalerao, B. N. (2009). Influence of hands-on exploration and classroom orientation on learning for developing basic motor control skills for equipment operators. *MS Thesis, Purdue University*, West Lafayette, Indiana, USA.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development.* Thousand Oaks, CA: Sage.

Boyle, L. N., & Lee, J. D. (2010). Using driving simulators to assess driving safety. *Accident Analysis & Prevention, 42,* 785–787.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience and School.* Washington, DC: National Academy Press.

Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3,* 77–101.

Briggs, G. E., & Naylor, J. C. (1962). The relative efficiency of several training methods as a function of transfer task complexity. *Journal of Experimental Psychology, 64,* 505–512.

Burdick, K. J. (1977). Effects of massed and distributed practice on the learning and retention of a novel gross motor skill. Master's Thesis, Western Illinois University.

Bürki-Cohen, J., Soja, N. N., & Longridge, T. (1998). Simulator platform motion—The need revisited. *International Journal of Aviation Psychology, 8*, 293–317.

Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Ed.), *Advances in Industrial Ergonomics and Safety* (pp. 481–485). Taylor & Francis.

Caro, P. W. (1976). *Some Factors Influencing Transfer of Simulator Training (Tech. Rep. No. HumRRO-PP-1-76)*. Alexandria, VA: Human Resources Research Organization.

Chamberland, C., & Tremblay, S. (2011). Task switching and serial memory: Looking into the nature of switches and tasks, *Acta Psychologica, 136*, 137–147.

Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences, 6*, 271–315.

Clawson, D. M., Healy, A. F., Ericsson, K. A., & Bourne Jr, L. E. (2001). Retention and transfer of morse code reception skill by novices: Part-whole training. *Journal of Experimental Psychology: Applied, 7*, 129-142.

Cornelissen, M., Salmon, P. M., McClure, R., & Stanton, N. A. (2013). Using cognitive work analysis and the strategies analysis diagram to understand variability in road user behaviour at intersections. *Ergonomics, 56*, 764-780.

Cree, V. E., & Macaulay, C. (2000). *Transfer of learning in professional and vocational education*. Psychology Press.

De Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds*.), Control of Cognitive Processes: Attention and Performance XVIII* (pp. 357–376). Cambridge, MA: MIT Press.

Donchin, O., & Cardoso De Oliveira, S (2004). Electrophysiological approaches to bimanual coordination in primates. In S.P. Swinnen, & J. Duysens (Eds.), *Neuro-behavioral Determinants of Interlimb Coordination: A Multidisciplinary Approach* (pp. 131–153). Boston, MA: Kluwer Academic.

Dopico, D., Luaces, A., & González, M. (2010). A soil model for a hydraulic simulator excavator based on real-time multibody dynamics. In *Proceedings of the 5th Asian Conference on Multibody Dynamics 2010* (pp. 23–26).

Duffy, V. G., Ng, P. P., & Ramakrishnan, A. (2004). Impact of a simulated accident in virtual training on decision-making performance. *International Journal of Industrial Ergonomics, 34*, 335–348.

Dunston, P. S., Proctor, R. W., & Wang, X. (2014). Challenges in evaluating skill transfer from construction equipment simulators. *Theoretical Issues in Ergonomics Science, 15*, 354–375.

Ericsson, A. (2003). Valid and Non-Reactive Verbalization of Thoughts During Performance of Tasks Towards a Solution to the Central Problems of Introspection as a Source of Scientific Data. *Journal of Consciousness Studies, 10*, 1–18.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol Analysis: Verbal Reports as Data*, revised edition. Cambridge, MA:MIT Press.

Fitts, P. M., (1964). Perceptual Motor Skill Learning. In A. W. Melton (Ed.), *Categories of Human Learning* (pp. 243–285). Academic, NY.

Gilbert, S.J., Shallice, T. (2002) Task switching: a PDP model. *Cognitive Psychology, 44*, 297–337.

Grantcharov, T. P., Kristiansen, V. B., Bendix, J., Bardram, L., Rosenberg, J., & Funch-Jensen, P. (2004). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery, 91*, 146–150.

Gu, Y. (2014). To code or not to code: Dilemmas in analyzing think-aloud protocols in learning strategies research. *System, 43*, 74–81.

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In *Proceedings of the SIGCHI conference on human factors in computing systems, CHI'06* (pp. 1253–1262). New York: ACM.

Gundry, A. J. (1976). Man and motion cues. Paper presented at the Third Flight Symposium on Theory and Practice in Flight Simulation, London, UK.

Hagman, J. D. (1980). *Effects of training schedule and equipment variety on retention and transfer of maintenance skill (Report No. 1309*). Alexandria, VA: U.S. Army Research Institute.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, No. 9, pp. 904–908). Sage Publications.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of emprical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland Press

Healy, A. F. (2007). Transfer: Specificity and Generality. In H. L.Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of Memory: Concepts* (pp. 271–275). Oxford University Press, NY.

Healy, A. F., Schneider, V. I., & Bourne, L. E., Jr. (2012). Basic research on training principles. In A. F. Healy & L. E. Bourne, Jr. (Eds), *Training cognition: Optimizing Efficiency, Durability, and Generalizability,* New York: Psychology Press.

Healy, A. F., Kole, J. A., & Bourne Jr, L. E. (2014). Training principles to advance expertise. *Frontiers in Psychology, 5*, 1-4.

Healy, A. F., Wohldmann, E. L., & Bourne Jr, L. E. (2005). The Procedural Reinstatement Principle: Studies on Training, Retention, and Transfer. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 59–71). Washington, DC: American Psychological Association.

Henry, F. M. (1958). Specificity vs. generality in learning motor skills. *Proceeding of the College Physical education Association*, *61*, 126-128.

Hildreth, J. C., & Heggestad, E. (2010). Effect of simulation training methods on operator anxiety and skill development. In K. Makanae, N. Yabuki, & K. Kashiyama, (Eds.), *Proceedings of the 10th International Conference on Construction Applications of Virtual Reality (CONVR 2010)* (pp. 251–259). Sendai, Japan: CONVR2010 Organizing Committee.

Hildreth, J. C., & Stec, M. (2009). Effectiveness of simulation-based operator training. In X. Wang & N. Gu (Eds.), *Proceedings of the 9th International Conference on Construction Applications of Virtual Reality (CONVR 2009)* (pp. 333–342). Sydney, Australia: University of Sydney Press.

Hoffman, K. A., Aitken, L. M., & Duffield, C. (2009). A comparison of novice and expert nurses' cue collection during clinical decision-making: Verbal protocol analysis. *International Journal of Nursing Studies, 46,* 1335–1344.

Hoffman, R. R., & Militello, L.G. (2008). *Perspectives on Cognitive Task Analysis*. Boca Raton, FL: CRC Press/Taylor and Francis.

Hsieh, S. (2012). Two decades of research on task switching: What more can we ask? *Chinese Journal of Psychology, 54,* 67–93.

Ivry, R., Diedrichsen, J., Spencer, R., Hazeltine, E., & Semjen, A. (2004). A cognitive neuroscience perspective on bimanual coordination and interference. In S. Swinnen & J. Duysens (Eds.), *Neuro-behavioral determinants of interlimb coordination* (pp. 259–295). Boston, MA: Kluwer Academic Publishing.

Jacobs, J. W., Prince, C., Hays, R. T., & Salas, E. (1990).A meta-analysis of the flight simulator training research (Tech. Rep. No. TR-89-006). Orlando, FL: Naval Training Systems Center.

Jersild, A. (1927). Mental set and shift. *Archives of Psychology* (Whole No. 89).

Jiang, B., & Kuehn, P. (2001). Transfer in the Academic Language Development of Post-secondary ESL Students. *Bilingual Research Journal, 25*, 417–436.

Joffe, H. (2011). Thematic analysis. In D. Harper & A. R. Thompson (Eds.), *Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners (*pp. 209–223). Chichester, UK: John Wiley & Sons.

Kamezaki, M., Iwata, H., & Sugano, S. (2008). Development of an operation skill-training simulator for double-front work machine. In *Advanced Intelligent Mechatronics, 2008. AIM 2008. IEEE/ASME International Conference* (pp. 170-175). IEEE.

Kamezaki, M., Iwata, H., & Sugano, S. (2009a). Primitive static states for intelligent operated-work machines. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference* (pp. 1334–1339). IEEE.

Kamezaki, M., Iwata, H., & Sugano, S. (2009b). Work State Identification using Primitive Static States–Implementation to Demolition Work in Double-Front Work Machines. In *26th International Symposium on Automation and Robotics in Construction (ISARC 2009)* (pp. 24–27), International Association for Automation and Robotics in Construction.

Kecskes, I., & Papp, T. (2000). *Foreign Language and Mother Tongue*. Hillsdale, NJ: Erlbaum.

Kole, J. A., Healy, A. F., Fierman, D. M., & Bourne, L. E. (2010). Contextual memory and skill transfer in category search. *Memory & Cognition, 38*, 67–82.

Koonce, J. M., & Bramble Jr, W. J. (1998). Personal computer-based flight training devices. *The International Journal of Aviation Psychology*, *8*, 277–292.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15,* 126–147.

Kuusela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *American Journal of Psychology*, *113*, 387–404.

Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor skill acquisition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*, 730–746.

Lee, S. W., Park, J., Kim, A. R., & Seong, P. H. (2012). Measuring situation awareness of operation teams in NPPs using a verbal protocol analysis. *Annals of Nuclear Energy, 43,* 167–175.

Lee, T. D., & Simon, D. (2004). Contextual interference. In A. M. Williams & N. J. Hodges (Eds.), *Skill acquisition in sport: Research, theory and practice* (pp. 29–44). London, UK: Routledge.

Lehmann, K. S., Ritz, J. P., Maass, H., Çakmak, H. K., Kuehnapfel, U. G., Germer, C. T., & Buhr, H. J. (2005). A prospective randomized study to test the transfer of basic psychomotor skills from virtual reality to physical reality in a comparable training setting. *Annals of surgery, 241*, 442–449.

Lewis, C. H. (1982). *Using the "Thinking Aloud" Method In Cognitive Interface Design (Technical report RC-9265)*. IBM.

Lin, F., Ye, L., Duffy, V. G., & Su, C. J. (2002). Developing virtual environments for industrial training. *Information Sciences*, *140*, 153–170.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G.D. and Gordon, R.D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review. 108*, 393–434.

Lohse, K. R., & Healy, A. F. (2012). Exploring the contributions of declarative and procedural information to training: a test of the procedural reinstatement principle. *Journal of Applied Research in Memory and Cognition*, *1*, 65–72.

Lopez-Santamaria, B. N. (2011). Evaluating skill acquisition and transfer in heavy machinery operation using a simulation-based approach. *MS Thesis, Purdue University*, West Lafayette, Indiana, USA.

Lyall, B., & Wickens, C. D. (2005). Mixed fleet flying between two commercial aircraft types: An empirical evaluation of the role of negative transfer. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 45–48). Santa Monica, CA: HFES.

MacLeod, W. B., Butler, D. L., & Syer, K. D. (1996, April). Beyond achievement data: Assessing changes in metacognition and strategic learning. In *Annual Meeting of the American Educational Research Association,* New York.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science*, *9*, 241–289.

Martin, E. A. (1985).The influence of tactual seat-motion cues on training and performance in a roll-axis compensatory tracking task setting (Doctoral dissertation). Columbus, OH: Ohio State University.

Mason, A.H. (2008). Coordination and control of bimanual prehension: effects of perturbing object location. *Experimental Brain Research, 188*, 125–139.

Maxwell, J. P., Masters, R. S. W., Kerr, E., & Weedon, E. (2001). The implicit benefit of learning without errors. *The Quarterly Journal of Experimental Psychology: Section A, 54*, 1049–1068.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1423–1442.

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In A. S. Benjamin (Ed.), *Successful Remembering and Successful Forgetting: A festschrift in honor of Robert A. Bjork* (pp. 175–199). New York: Psychology Press.

McDaniel, M. A., & Einstein, G. O. (2005). Material appropriate difficulty: a framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed*.), Experimental Cognitive Psychology and Its Applications* (pp. 73–85). Washington, DC: American Psychological Association.

Messerli, D. (2005). *1001 Great Stories, Volume 1*. Køben-havn: Green Integer.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134–140.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Nieuwenhuis, S., & Monsell. S. (2002). Residual costs in task switching: Testing the failure-to-engage hypothesis. *Psychonomic Bulletin & Review, 9.1*, 86-92.

Newell, K. M., Mayer-Kress, G., Hong, S. L., & Liu, Y.-T. (2009). Adaptation and learning: Characteristic time scales of performance dynamics. *Human Movement Science, 28*, 655–687.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the power law of practice. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Norman, D.A. (1981). Categorization of Action Slips. *Psychological Review*, 88 (1), 1-15.

Ober, G. J., (2010). *Operating techniques for the tractor loader backhoe* (Rev. ed.). Northridge, CA: Equipment Training Resources.

Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning.* New York: Cambridge University Press.

Ormrod, J. E. (2004). *Human learning*. (4th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Ormerod, T. C., Richardson, J., & Shepherd, A. (1998). Enhancing the usability of a task analysis method: A notation and environment for requirements specification. *Ergonomics, 41*, 1642–1663.

Orlansky, J., & String, J. (1977). *Cost-effectiveness of Flight Simulators for Military Training: vol. I. Use and Effectiveness of Flight Simulators* (Paper No. P-1275). Arlington, VA: Institute for Defense Analyses.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review, 56*,132–143.

Patrick, J. (1992*). Training: Research and practice.* London: Academic Press.

Pavlov, I. P. (1935/1955). *Selected Works.* Moscow: Foreign Languages Publishing House.

Pellegrino, J. W., Doane, S. M., Fischer, S. C., & Alderton, D. (1991). Stimulus complexity effects in visual comparisons: the effects of practice and learning context. Journal of experimental psychology. *Human Perception and Performance, 17*, 781–791.

Pfeiffer, M. G., & Horey, J. D. (1987). *Training Effectiveness of Aviation Motion Simulation: A Review and Analysis of the Literature (Special Rep. No. 87-007).* Orlando, FL: Naval Training Systems Center.

Phipps, D., Meakin, G. H., Beatty, P. C. W., Nsoedo, C., and Parker, D. (2008). Human factors in anaesthetic practice: insights from a task analysis. *British Journal of Anaesthesia, 100,* 333–343.

Phipps, D. L., Meakin, G. H., and Beatty, P. C. W. (2011). Extending hierarchical task analysis to identify cognitive demands and information design requirements. *Applied Ergonomics, 42*, 741–748.

Pool, D. M., Mulder, M., Van Paassen, M. M., & Van der Vaart, J. C. (2008). Effects of peripheral visual and physical motion cues in roll-axis tracking tasks. *Journal of Guidance, Control, and Dynamics, 31,* 1608–1622.

Rogers, R. D., & Monsell, S. (1995). The costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O., (2001). Acquisition of Intellectual and Perceptual Motor Skills. *Annual Review of Psychology, 52*, 453–470.

Proctor, R. W., Dunston, P. S., So, J. C. Y., Lopez-Santamaria, B. N., Yamaguchi, M., & Wang, X. (2013). Specificity of transfer in basic and applied perceptual-motor tasks. *American Journal of Psychology, 4,* 401-415.

Proctor, R. W., Dunston, P. S., So, J. C.Y., & Wang, X. (2012). Task Analysis for Improving Training of Construction Equipment Operators. In *Construction Research Congress 2012: Construction Challenges in a Flat World* (pp. 169–178). ASCE.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103*, 734–760.

Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1404–1419.

Ryan, B., & Haslegrave, C. M. (2007a). Use of concurrent and retrospective verbal protocols to investigate workers' thoughts during a manual-handling task, *Applied Ergonomics, 38,* 177–190.

Ryan, B., & Haslegrave, C. M. (2007b). Developing a verbal protocol method for collecting and analysing reports of workers' thoughts during manual handling tasks. *Applied Ergonomics*, *38*, 805–819.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207–217.

Schmidt, R. A. (1991). *Motor Learning and Performance: from Principles to Practice.* Champaign, IL: Human Kinetics Books.

Schmidt, R. A., & Lee, T. D. (2011). *Motor Control and Learning: A Behavioral Emphasis.* 5th ed. Champaign, IL: Human Kinetics.

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

Schneider, V. I., Healy, A. F., & Bourne, L. E., Jr. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language, 46*, 419–440.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery, 236*, 458–464.

Shea, J. B. & Morgan, R. L. (1979). Contextual Interference Effects on the Acquisition, Retention and Transfer of a Motor Skill. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 179–187.

Shepherd, A. (2001). *Hierarchical Task Analysis*. London, UK: Taylor & Francis.

Sigman, M., & Dehaene, S. (2006). Dynamics of the central bottleneck: dual-task and task uncertainty. *PLoS Biology, 4,* 1227–1238.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Smode, A. R, Beam, J. C, & Dunlap, J. W. (1959). Motor habit interference: A resume of the literature and the development of principles for its minimization in training. Contract Nonr 2515(00) with the Office of Naval Research and Dunlap and Associates, January.

Speelman, C. P., & Kirsner, K. (2001). Predicting transfer from training performance. *Acta Psychologica, 108*, 247–281.

Sutherland, L. M., Middleton, P. F., Anthony, A., Hamdorf, J., Cregan, P., Scott, D., & Maddern, G. J. (2006). Surgical simulation: a systematic review. *Annals of Surgery*, *243*, 291-300.

So, J.C.Y., Proctor, R.W., Dunston, P.S. & Wang, X., (2013), Better Retention of Skill Operating a Simulated Hydraulic Excavator after Part-Task than after Whole-Task Training, *Human Factors, 55*, 449–460.

Sohn, M. H., & Carlson, R. A. (2000). Effects of repetition and foreknowledge in task-set reconfiguration. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 26*, 1445–1460.

Soini, T. (1999). Preconditions for active transfer in Learning Processes. Doctoral Thesis, University of Helsinki, Department, Helsinki.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89*, 669–679.

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics, 37,* 55–79.

Stinchcombe, A., & Gagnon, S. (2013). Aging and driving in a complex world: exploring age differences in attentional demand while driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 17,* 125–133.

Su, X., Dunston, P. S., Proctor, R. W., & Wang, X. (2013). Influence of training schedule on development of perceptual-motor control skills for construction equipment operators in a virtual training system. *Automation in Construction, 35,* 439–447.

Swinnen, S. P., & Wenderoth, N. (2004). Two hands, one brain: cognitive neuroscience of bimanual skill. *Trends in Cognitive Sciences, 8*, 18–25.

Tan, S. S., & Sarker, S. K. (2011). Simulation in surgery: A review. *Scottish Medical Journal, 56*, 104–109.

Taylor, H. L., Lintern, G., Hulin, C. L., Talleur, D., Emanuel, T., & Philips, S. (1997). *Transfer of training effectiveness of personal computer-based aviation training devices.* Washington, DC: U.S Department of Transportation.

Taylor, H. L., Lintern, G., & Koonce, J. M. (1993). Quasi-transfer as a predictor of transfer from simulator to airplane. *The Journal of General Psychology, 120,* 257–276.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247–261.

Tichon, J., & Diver, P. (2012). Interactive Simulator Training in Civil Construction: Evaluation from the Trainer's Perspective. *Journal of Interactive Learning Research, 23*, 143–163.

Torkington, J., Smith, S. G. T., Rees, B. I., & Darzi, A. (2001). Skill transfer from virtual reality to a real laparoscopic task. *Surgical Endoscopy, 15*, 1076–1079.

Torres-Rodriguez, H. I., Parra-Vega, V., & Ruiz-Sanchez, F. J. (2004). Dynamic haptic training system for the operation of an excavator. In *Electrical and Electronics Engineering, 2004.(ICEEE). 1st International Conference* (pp. 350–355). IEEE.

Torres-Rodriguez, H. I., Parra-Vega, V., & Ruiz-Sanchez, F. J. (2005). Integration of force-position control and haptic interface facilities for a virtual excavator simulator. In *Advanced Robotics, 2005. ICAR'05 Proceedings, 12th International Conference* (pp. 761–768). IEEE.

Travlos, A. K. (2010). Specificity and variability of practice, and contextual interference in acquisition and transfer of an underhand volleyball serve. *Perceptual and Motor Skills, 110*, 298–312.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Valverde, H. H. (1973). A review of flight simulator transfer of training studies. *Human Factors, 15*, 510–523.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47,* 513–539.

Visser, T., Tichon, J., & Diver, P. (2012). Reducing the dangers of operator distraction through simulation training. In *SimTecT 2012: Asia-Pacific Simulation and Training Conference and Exhibition*. Simulation Australia.

Wang, X., & Dunston, P. S. (2005). Heavy equipment operator training via virtual modeling technologies. In *Proceeding of Construction Research Congress 2005* (pp. 618–622). Reston, VA.: ASCE.

Waszak F, Hommel B, Allport A (2003). Task-switching and long-term priming: role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46,* 361–413.

Wek, S. R. & Husak, W. S. (1989). Distributed and massed practice effects on motor performance and learning of autistic children. *Perceptual and Motor Skills, 69*, 107–113.

Wickelgren, W.A. (1974). Single-trace fragility theory of memory dynamics. *Memory & Cognition, 2*, 775–780.

Wickens, C. D., Hutchins, S., Carolan T., & Cumming, J. (2011). Investigating the impact of training on transfer: A meta-analysis. In *Proceedings of the 2011 Conference of the Human Factors and Ergonomics Society.* Santa Monica, CA: Human Factors.

Williams, A. M., & Ward, P. (2003). Perceptual expertise: Develop-ment in sport. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 219–247). Champaign, IL: Human Kinetics.

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science, 18*, 133–134.

Wohldmann, E. L., Healy, A. F., & Bourne, L. E., Jr. (2008). Global inhibition and midcourse corrections in speeded aiming. *Memory & Cognition, 36*, 1228–1235.

Zhao, T., McDonald, S., & Edwards, H. M. (2014). The impact of two different think-aloud instructions in a usability test: a case of just following orders? *Behaviour & Information Technology, 33*, 163–183.

APPENDICES

Appendix A: NASA-TLX

1. Please place an "X" along each scale at the point that best indicates your experience.

   a) **Mental Demand**: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, search, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

   b) **Physical Demand**: How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

   c) **Temporal Demand**: How much time pressure did you feel due to the rate or pace at which the task occurred? Was the pace slow or leisurely or rapid and frantic?

   d) **Performance**: How successful do you think you were in accomplishing the goals of the task? How satisfied were you with your performance in accomplishing these goals?

   e) **Effort**: How hard did you have to work (mentally and physically) to accomplish your level of performance?

   f) **Frustration**: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during your task?

Appendix B: Consent Form (IRB Protocol #1110011339) for Studies1, 2 and 5

Research Project Number 1110011339

APPROVED
OCT 04 2012
EXPIRES 10-3-13
PURDUE UNIVERSITY
INSTITUTIONAL REVIEW BOARD

CONSENT FORM
Sona Systems Number XXX
Use of Simulated Environments for Training Operation of Construction Equipment

Dr. Robert W. Proctor
(Principal Investigator)
Purdue University, Psychological Sciences
Dr. Phillip S. Dunston
(Co-PI)
Purdue University, Civil Engineering

Purpose of Research The purpose of this research is to examine skill acquisition and skill development through the use of a Virtual Training System (VTS). This research aims at analyzing the complexity of equipment operation and distinguishing the skills to be acquired for operating different construction machines.

Specific Procedures to be used The session has 3 phases. In Phase 1, after providing demographic information, you will be introduced to the basic components and controls of a specific kind of construction equipment, ending with hands-on practice. In the Phase 2, you will perform a number of practice trials of specified tasks. In Phase 3, you will perform a specific test task and will be asked to complete a post-experience questionnaire.

Duration of Participation The range of estimated duration of participation is 1 – 3 hours.

Risks to the Individual You will be exposed to minimal risk, which is no greater than every day activities. The Virtual Reality simulator functions basically like a computer game you will play while seated in a chair. The display setup is a 30" desktop computer monitor.

Benefits to the Individual or Others You may gain insight into use of simulators for training and how training tasks are designed for developing particular skills. The knowledge gained from the study will help us understand applicable basic cognition principles and develop programs for training real-world skills.

Compensation You will earn up to 6 course credits (1 credit per 30 minutes).

Confidentiality Breach of confidentiality is a risk related to the research, but safeguards are in place to minimize the risk. Information collected for this study that identifies you will remain confidential and will be disclosed only with your permission or as required by law. The data will be stored on computers in CIVL B141, behind locked doors, only accessible to the researchers. Informed consent forms will be destroyed after three years following the IRB's closure of this protocol. Raw data will be kept on lab computers indefinitely for possible subsequent reanalysis. However, these data files will not include your name or any other personal information. The project's research records may be reviewed by the National Science Foundation and by departments at Purdue University responsible for regulatory and research oversight.

Voluntary Nature of Participation Your participation in this study is voluntary; you may withdraw at any time without consequence. Participation or non-participation will not affect your status in the university. The investigator may withdraw you from this research if circumstances warrant doing so.

Contact Information If you have any questions about this research project, you can contact Dr. Phillip S. Dunston at (765) 494-0640, or E-mail: dunston@purdue.edu or Dr. Robert W. Proctor at (765) 494-0784, or Email: proctor@psych.purdue.edu. If you have concerns about the treatment of research participants, you can contact Institutional Review Board at Purdue University, Ernest C. Young Hall, 10th floor- room 1032, 155 S. Grant Street, West Lafayette, IN 47907-2114. The phone number for the Committee's secretary is (765) 494-5942. The email address is irb@purdue.edu.

I HAVE HAD THE OPPORTUNITY TO READ THIS CONSENT FORM, ASK QUESTIONS ABOUT THE RESEARCH PROJECT AND AM PREPARED TO PARTICIPATE IN THIS PROJECT.

____________________________ __________________
Participant's Signature Date

____
Participant's Name

____________________________ __________________
Researcher's Signature Date

Appendix C: Preliminary Questionnaires for Studies 1, 2, and 5

Participant No:_______ Group:_____ Date:_________ Time:_________

Preliminary Questionnaires

1. Name: ____

2. Age: _____________ years

3. Gender: Male Female

4. Ethnic Background:

Hispanic or Latino	American Indian or Alaska Native	Asian	Black or African American	White	Native Hawaiian or other Pacific Islander	More than One Race	Other	Do not wish to provide

5. Handedness: Left Right

6. Have you taken part in similar experiments or performed tasks in a Virtual Reality environment before?

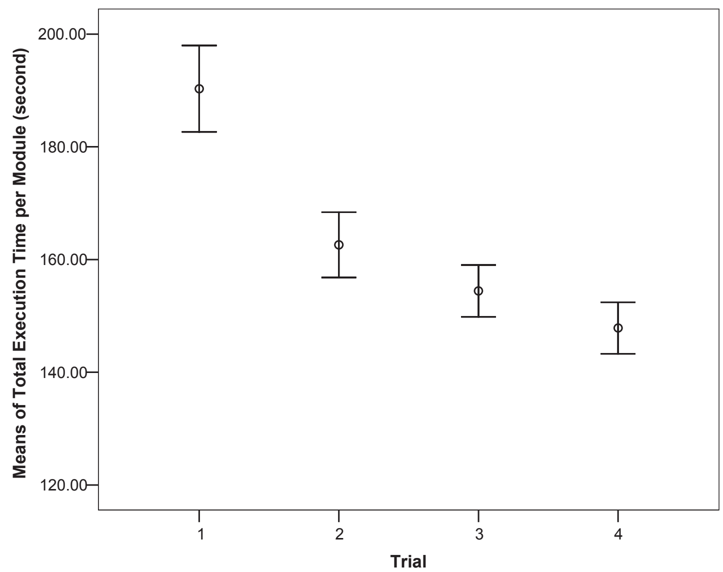
☐ Many times ☐ Couple of times ☐ Once ☐ Never

Appendix D: Study 1

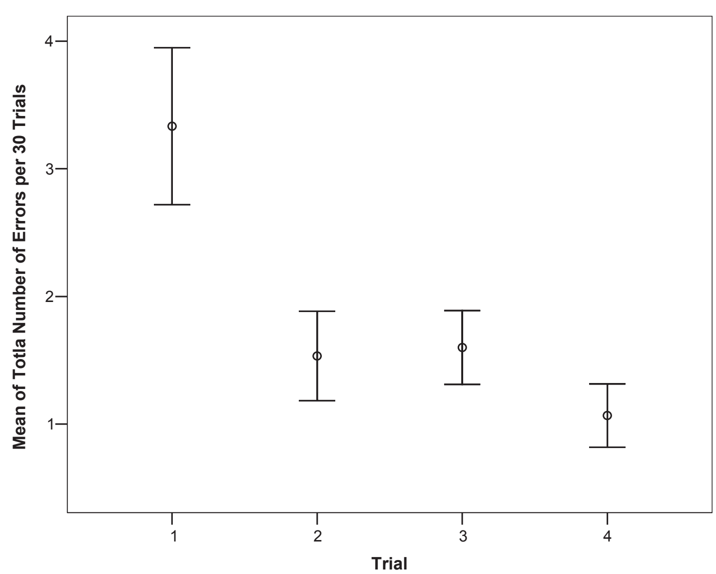
The mixed-design repeated-measures ANOVA comparing the performance measures on simulated excavator across three sessions among control group.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Execution Time						
Session	33264.022	2	716632.011	75.547	.000	.844
Error(Session)	6164.363	28	220.156			
Trial	16646.137	3	5548.712	68.555	.000	.830
Error(Trial)	3399.398	42	80.938			
Session x Trial	18151.875	6	3025.312	39.430	.000	.738
Error (Session x Trial)	6445.072	84				
Number of Errors						
Session	80.478	2	40.239	6.506	.005	.317
Error(Session)	173.189	28	6.185			
Trial	57.644	3	19.215	5.007	.005	.263
Error(Trial)	161.189	42	3.838			
Session x Trial	40.456	6	6.743	1.932	.085	.121
Error (Session x Trial)	293.211	84	3.491			

Data Analysis on the execution time and no. of errors on Controls Familiarization module on loader simulator.



Results: This figure shows the mean execution time (per 30 trials) in Control Familiarization module on the loader simulators. The results of ANOVA show that trial is a significant factors, $F(3, 42) = 31.087$, $p < .001$, $\eta_p^2 = .689$, with a significant drop of execution time was found from trail 1 to trial 2 (Bonferroni pairwise comparisons tests, *ps* > 0.001).



Results: This figure shows the mean no. of errors (per 30 trials) in Control Familiarization module on the loader simulators. Similar result pattern was found for the number of error. The ANOVA show that trial is a significant factors, $F(3, 42) = 7.039$, $p = .001$, $\eta_p^2 = .335$, with a significant drop of execution time was found from trail 1 to trial 2 (Bonferroni pairwise comparisons tests, *ps* > 0.001).

The mixed-design repeated-measures ANOVA comparing the workload measures of three experimental groups in two sessions.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	162.830	2	81.415	14.228	0.000	0.258
Session × Group	74.611	4	18.653	3.260	0.016	0.137
Error(Session)	469.220	82	5.722			
Measure	634.804	5	126.961	21.593	0.000	0.345
Measure × Group	46.797	10	4.680	0.796	0.633	0.037
Error(Measure)	1205.343	205	5.880			
Session × Measure	64.331	10	6.433	3.651	0.000	0.082
Session × Measure × Group	41.234	20	2.062	1.170	0.277	0.054
Error(Sessions × Measure)	722.472	410	1.762			
Between-Subject Factor						
Group	7.284	2	3.642	.184	.833	.009
Error (Group)	811.788	41	19.800			

The mixed-design repeated-measures ANOVA comparing the productivity on simulated excavator between Group 1 vs Group 2.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	14574.165	2	7287.083	91.379	.000	.706
Session x Group	1745.741	2	872.870	10.946	.000	.224
Error(Session)	6060.648	76	79.745			
Trial	1417.937	1	1417.937	65.667	.000	.633
Trial x Group	108.865	1	108.865	5.042	.031	.117
Error(Trial)	820.522	38	21.593			
Session x Trial	416.653	2	208.326	9.262	.000	.196
Session x Trial x Group	43.935	2	21.967	.977	.381	.025
Error (Session x Trial)	1709.369	76	22.492			
Between-Subject Factor						
Group	1512.177	1	1512.177	3.474	.070	.084
Error (Group)	16540.841	38	435.285			

The mixed-design repeated-measures ANOVA comparing the productivity on simulated excavator between Group 1 vs Group 3.

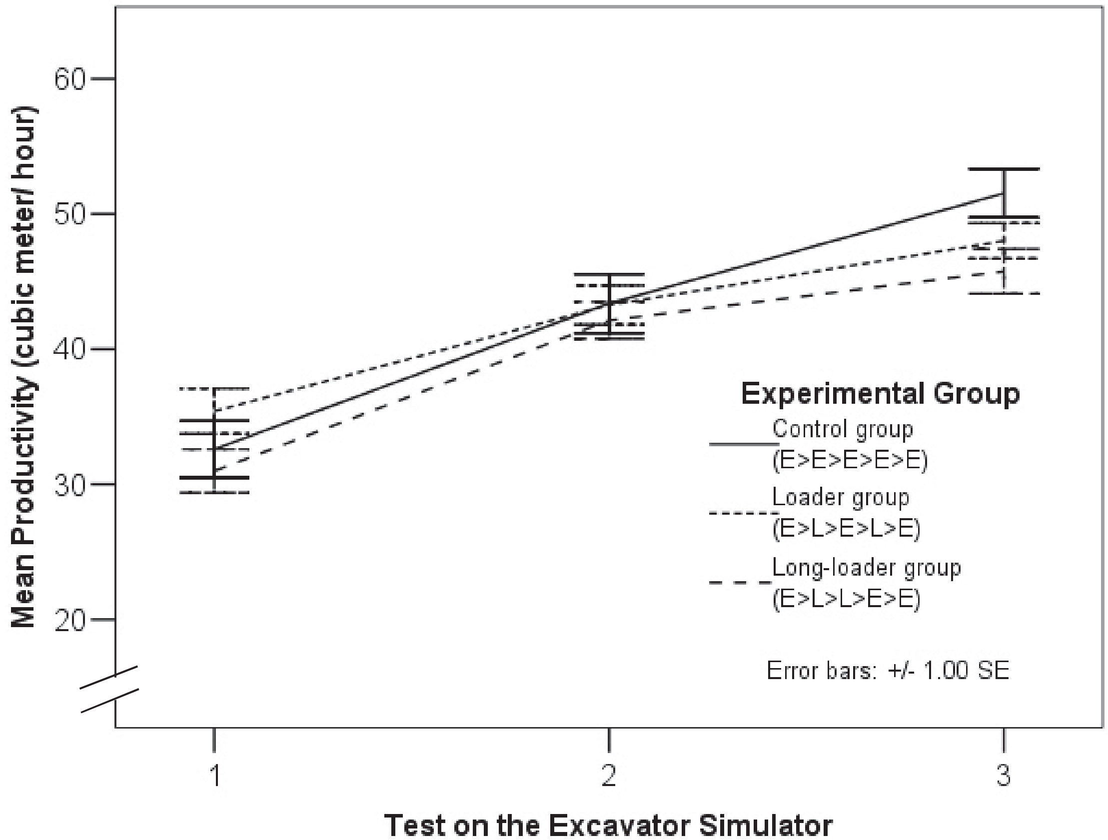
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	18046.020	2	9023.010	106.698	.000	.737
Session x Group	1460.794	2	730.397	8.637	.000	.185
Error(Session)	6426.994	76	84.566			
Trial	674.683	1	674.683	25.163	.000	.398
Trial x Group	22.709	1	22.709	.847	.363	.022
Error(Trial)	1018.880	38	26.813			
Session x Trial	417.563	2	208.781	5.200	.008	.120
Session x Trial x Group	26.805	2	13.402	.334	.717	.009
Error(Session x Trial)	3051.227	76	40.148			
Between-Subject Factor						
Group	4328.728	1	4328.728	9.610	.004	.202
Error(Group)	17116.631	38	450.438			

Appendix E: Study 2

The mixed-design repeated-measures ANOVA for the truck loading task performance on the simulated excavator for the first time, second time and third time by all three groups.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	14050.356	2	7025.178	108.416	.000	.655
Session x Group	402.535	4	100.634	1.553	.192	.052
Error(Session)	7387.034	114	64.799			
Trial	1669.393	1	1669.393	64.524	.000	.531
Trial x Group	185.616	2	92.808	3.587	.034	.112
Error(Trial)	1474.724	57	25.872			
Session x Trial	547.042	2	273.521	8.880	.000	.135
Session x Trial x Group	35.627	4	8.907	.289	.885	.010
Error(Session x Trial)	3511.503	114	30.803			
Between-Subject Factor						
Group	412.806	2	206.403	.470	.628	.016
Error(Group)	25052.122	57	439.511			

Two-way interaction plots of session x group on productivity of the first three tests on excavator simulator.



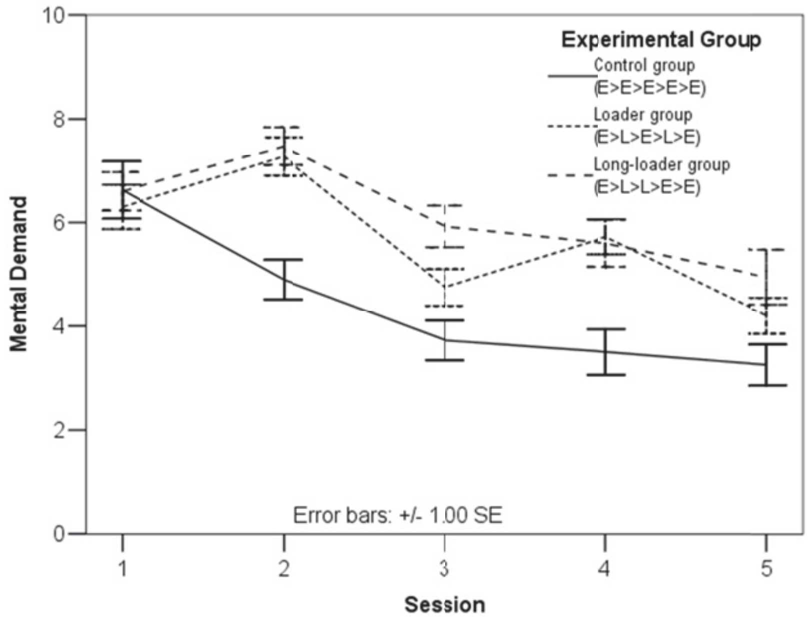
The mixed-design repeated-measures ANOVA comparing the productivity on simulated loader between Group 2 vs Group 3.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	39119.944	1	39119.944	109.089	.000	.742
Session x Group	70.331	1	70.331	.196	.660	.005
Error(Session)	13627.052	38	358.607			
Trial	12173.643	1	12173.643	68.124	.000	.642
Trial x Group	.747	1	.747	.004	.949	.000
Error(Trial)	6790.548	38	178.699			
Session x Trial	84.303	1	84.303	.498	.485	.013
Session x Trial x Group	105.370	1	105.370	.622	.435	.016
Error(Session x Trial)	6436.775	38	169.389			
Between-Subject Factor						
Group	182.076	1	182.076	.102	.751	.003
Error(Group)	67672.369	38	1780.852			

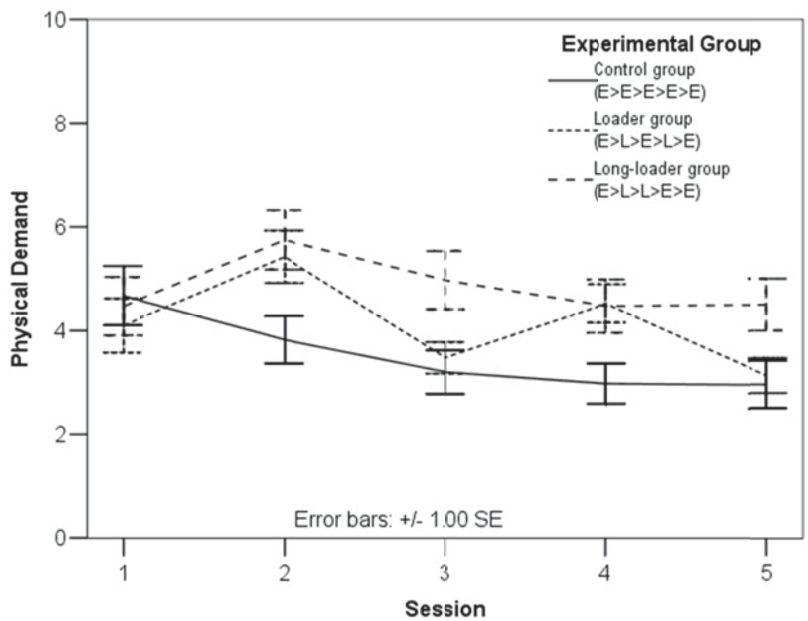
The mixed-design repeated-measures ANOVA for the subjective ratings across all 5 sessions.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	986.594	4	246.648	48.078	0.000	0.458
Session x Group	339.524	8	42.441	8.273	0.000	0.225
Error(Session)	1,169.682	228	5.130			
Measure (TLX)	1,152.646	5	230.529	30.582	0.000	0.349
Measure x Group	50.688	10	5.069	0.672	0.750	0.023
Error (Measure)	2,148.349	285	7.538			
Session x Measure	83.359	20	4.168	3.105	0.000	0.052
Session x Measure x Group	37.552	40	0.939	0.699	0.922	0.024
Error (Session x Measure)	1,530.488	1,140	1.343			
Between-Subject Factor						
Group	597.369	2	298.684	9.481	0.000	.250
Error (Group)	1795.781	57	31.505			

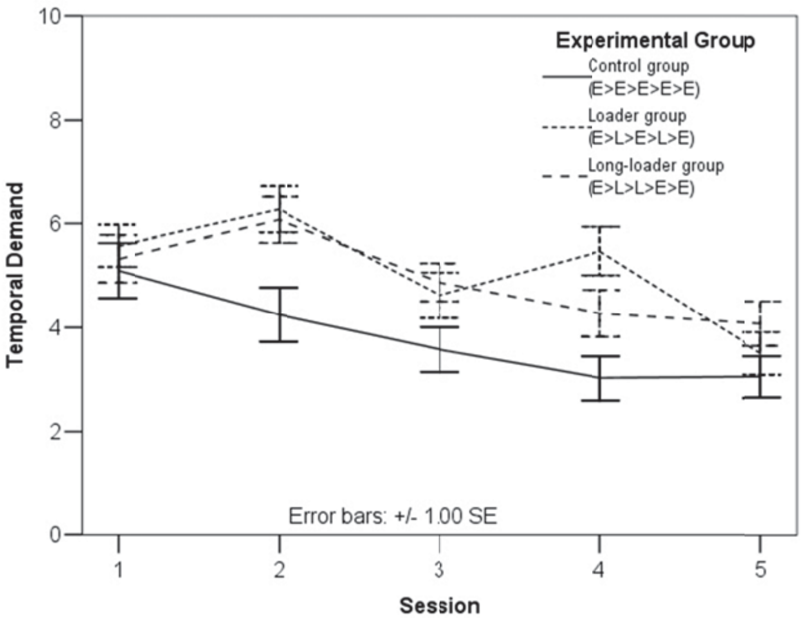
Two-way interaction plots of session x group of the six subscale of NASA-TLX across the five sessions.



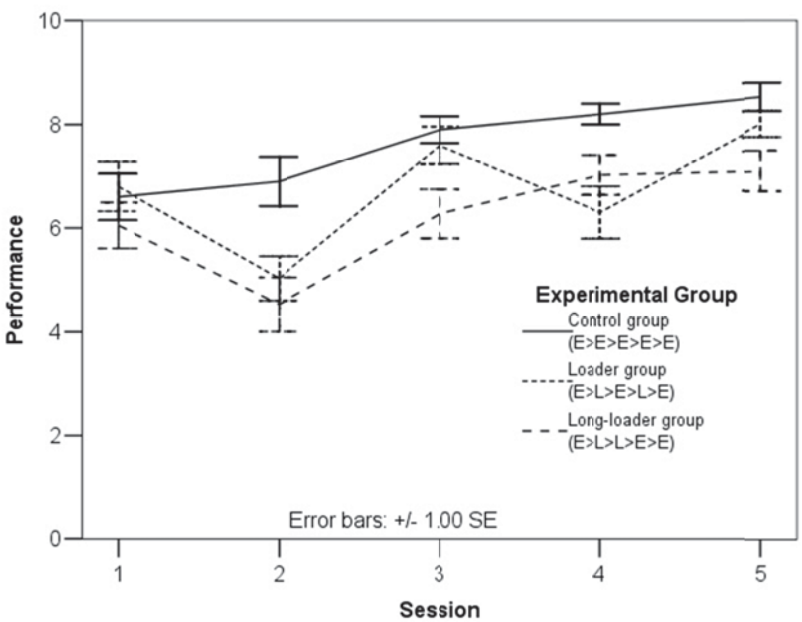
(a) Mental Demand



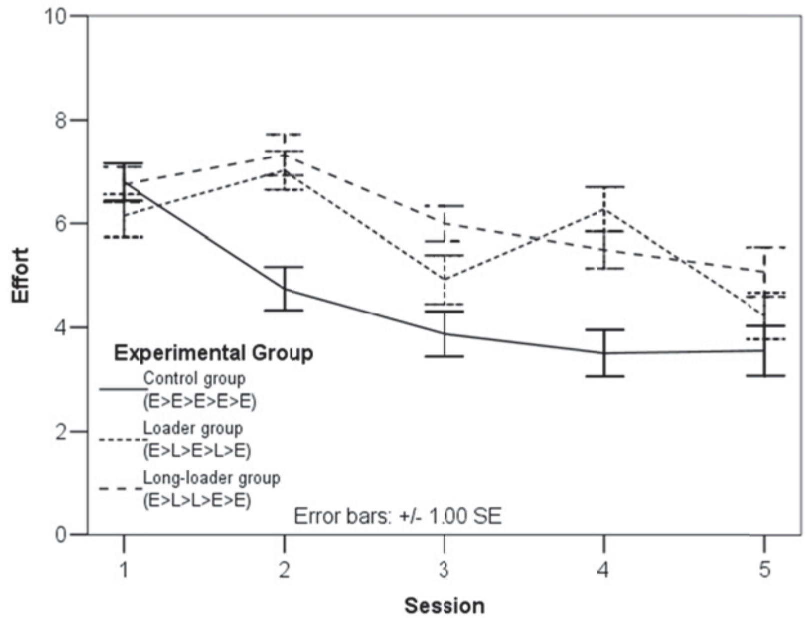
(b) Physical Demand



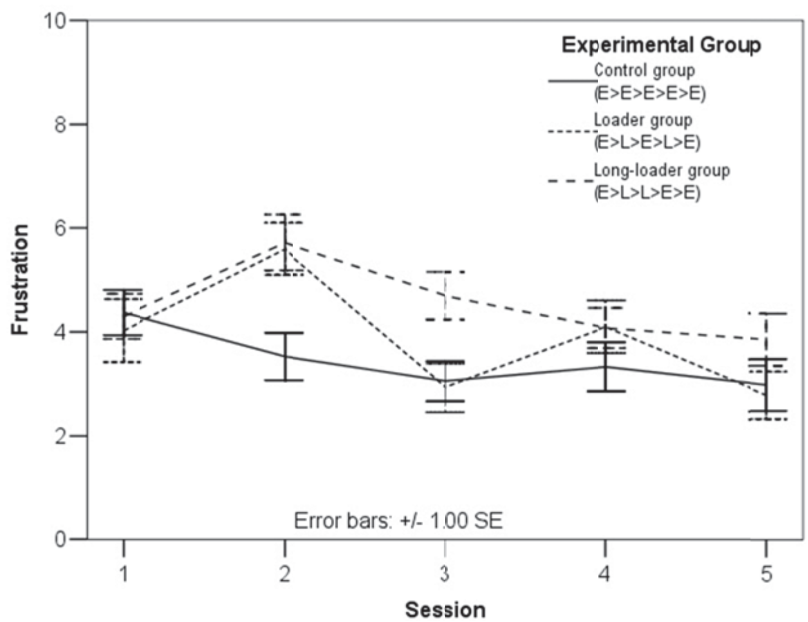
(c) Temporal Demand



(d) Performance (original data, the higher the rating, more satisfied with their performance)



(e) Effort



(f) Frustration

Appendix F: Consent Form (IRB Protocol #1304013518) for Studies 3 and 4 Conducted at John Deere Sites

Research Project Number ____________

For IRB Office Use Only

APPROVED
MAY 28 2013
EXPIRES 5-27-14
PURDUE UNIVERSITY INSTITUTIONAL REVIEW BOARD

CONSENT FORM
Simulated Construction Equipment Training

Dr. Phillip S. Dunston
(Principal Investigator)
Purdue University, Psychological Sciences
Dr. Robert W. Proctor
(Co-PI)
Purdue University, Civil Engineering

Purpose of Research The purpose of this research is to examine skill acquisition and skill development through the use of a Virtual Training System (VTS). This research aims at analyzing the complexity of equipment operation and distinguishing the skills to be acquired for operating different construction machines.

Specific Procedures to be used The study is divided into two parts. Part 1 consists of three (3) phases. In Phase 1, after providing demographic information, you will be introduced to the basic components and controls of a specific kind of construction equipment simulator. In the Phase 2, you will perform a number of practice trials of specified tasks on the simulator. In Phase 3, you will perform another specific task and will be asked to complete a post-experience questionnaire followed by an interview. Your verbal explanation of your task approach and interview responses will be recorded during the session.

Duration of Participation Your participation will take up to 2 hours.

Risks to the Individual You will be exposed to minimal risk, which is no greater than every day activities. The Virtual Reality simulator functions basically like a computer game you will operate while seated in a chair.

Benefits to the Individual or Others You may gain insight into use of simulators for training and how training tasks are designed for developing particular skills. The knowledge gained from the study may help us understand applicable basic perceptual-motor principles and develop programs for training real-world skills.

Compensation You will not be compensated, other than time allowed during your office hours.

Confidentiality Breach of confidentiality is a risk related to the research, but safeguards are in place to minimize the risk. Information collected for this study that identifies you will remain confidential and will be disclosed only with your permission or as required by law. The data will be stored on computers at Purdue University behind locked doors, only accessible to the researchers. Informed consent forms will be destroyed after three years following the IRB's closure of this protocol. Raw data will be kept on lab computers indefinitely for possible subsequent reanalysis. However, these data files will not include your name or any other personal information. The project's research records may be reviewed by departments at Purdue University responsible for regulatory and research oversight.

Voluntary Nature of Participation Your participation in this study is voluntary; you may withdraw at any time without consequence. Participation or non-participation will not affect your status in the company.
Contact Information If you have any questions about this research project, you can contact Dr. Phillip S. Dunston by phone at (765) 494-0640 or by e-mail at dunston@purdue.edu or Dr. Robert W. Proctor by phone at (765) 494-0784 or by email at proctor@psych.purdue.edu. If you have concerns about the treatment of research participants, you can contact Institutional Review Board at Purdue University, Ernest C. Young Hall, 10th floor- room 1032, 155 S. Grant Street, West Lafayette, IN 47907-2114. The phone number for the Committee's secretary is (765) 494-5942. The email address is irb@purdue.edu.

I HAVE HAD THE OPPORTUNITY TO READ THIS CONSENT FORM, ASK QUESTIONS ABOUT THE RESEARCH PROJECT AND AM PREPARED TO PARTICIPATE IN THIS PROJECT.

____________________ Participant's Signature

____________________ Date

____________________ Participant's Name

____________________ Researcher's Signature

____________________ Date

Appendix G: Preliminary Questionnaire for Studies 3 and 4

Participant No:_______ Date:_________ Time:_________

Preliminary Questionnaire

1. Age: _____________ years

2. Gender: Male Female

3. Handedness: Left Right

4. Experience profile

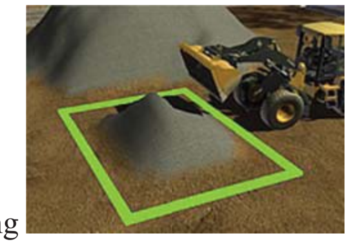
Which types of equipment do you know how to operate?	Calendar year you learned to operate the equipment?	Years of experience operating the equipment?

Appendix H: Study 4

Post-test Questionnaire

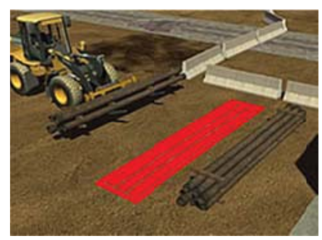
1. Rank the difficulty of the four tasks from easiest to hardest (1: easiest to manipulate; 4: hardest to manipulate).

__________ Task A: Simple Bucket Loading



__________ Task B: Filling a Trench

__________ Task C: Loading with a Fork



__________ Task D: Truck Loading

2. Why do you rank in this order? Is there any particularly difficult aspect of any of these tasks ?
   Task A: Simple Bucket Loading
   ____
   Task B: Filling a Trench
   ____

Task C: Loading with a Fork

____

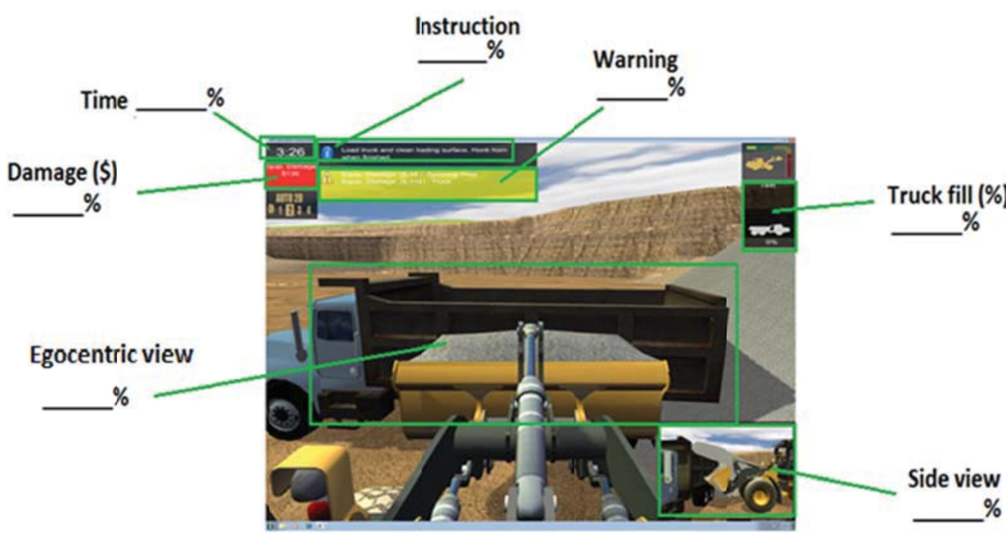
Task D: Truck Loading

____

3. Did you find any similarity across these four tasks, in terms of controls and techniques?

4. Are there any differences between the real machine and the simulator?

5. There were numerous feedback indicators on the screen while you are operating the machine. Please indicate the approximated percentage (%) of time you spent on each type of feedback indicated in the figure below. (Put 0% wherever you did not refer to particular that feedback.)



Spot-checking for verbal data transcription

	Transcription by Undergraduate Research Assistant	Transcription by a graduate student
Simple Bucket loading (B1)/ Operator #17/ 139 words	So the first thing I do is release the parking brake, shift into gear, then start driving [it]. So I got to line up with the existing bucket… You can hear that feedback, but I don't know what the cutting noise, angled where the bucket is at… This is kind of awkward because this is where I would want to look around the side of the machine to make sure no one is there. This gives me the depth perception so I can see with the sight… So as I'm backing out, I'm trying to reposition my bucket so I don't have to move it as far when I start to go forward into the pile… So as I back out, raise it up, ~~take~~ cut off the throttle and shift into gear… The sight really helps with the depth perception…	So the first thing I do is release the parking brake, shift into gear, then start driving it. So now I gotta line up with the existing bucket… You can hear that feedback but I don't know what cutting noise, angled where the bucket is at. This is kind of awkward because this is where I'd want to look around the side of the machine to make ~~be~~ sure no one's there. This gives me the depth perception so I can see with the sight. So as I'm backing out I'm trying to reposition my bucket so that I don't have to move as far when I start to go forward into the pile… So as I back out, raise it up, cut off the throttle and shift the gear. Again, that site view really helps with the depth perception…
Filling a trench (B2)/ Operator #10/ 287 words	Now this simulator is a little different because it doesn't show me... it keeps starting in rear view on the bottom corner, which is a little different than what I'm used to because it doesn't show me the bucket but I can see that there is a man behind me that I ~~need to~~ gotta watch out for when backing out… Um, Now when I'm approaching the trench, um, it's a little more difficult here because I can't see where the bottom of the trench is… Okay, it's showing me... Normally you would be able to see the edge of the trench…and what I watch out for is that I don't want to get to close to the trench. I don't want to cave the trench in, so…and now I'm backing up a little behind me and I don't want to hit the guy behind me. I'm lowering my boom and getting my bucket in the position for the next bucket… Now, again when I look at going to the pile I'm going to make sure my bucket is in position and boom is lowered to the ground. And when I go to the pile, I want to make sure it's square to the pile like this… so that I get a full range, full bucket. I do not want to go into the pile like this because then I would just get that corner of the bucket loaded. So when I approach the pile the best I can, I want to make sure I load the pile… I'll keep the bucket low to the ground until I get to where I want to go. As I approach it then I'll square with it, and I'll start dumping…	Now this simulator is a little different because it doesn't show me... it keeps starting in rear view on the bottom corner, which is a little different than what I'm used to because it doesn't show me the bucket but I can see that there is a man behind me that I gotta watch out for when backing out… Um, Now when I'm approaching the trench, um, it's a little more difficult here because I can't see where the bottom of the trench is… Okay, it's showing me... Normally you would be able to see the edge of the trench…and what I watch out for is that I don't want to get to close to the trench. I don't want to cave the trench in, so…and now I'm backing up a little behind me and I don't want to hit the guy behind me. I'm lowering my boom and getting my bucket in the position for the next bucket… Now, again when I look at going to the pile I'm going to make sure my bucket is in position and boom is lowered to the ground. And when I go to the pile, I want to make sure it's square to the pile like this… so that I get a full range, full bucket. I do not want to go into the pile like this because then I would just get that corner of the bucket loaded. So when I approach the pile the best I can, I want to make sure I load the pile… I'll keep the bucket low to the ground until I get to where I want to go. As I approach it then I'll square with ~~that~~ it, and I'll start dumping…

	Transcription by Undergraduate Research Assistant	Transcription by a graduate student
Truck loading (B3)/ Operator #18/ 37 words	So I'm driving into the pile… I lower the bucket. And as I'm pushing into the pile, I'm raising the bucket at the same time… Reverse, turn a little… Forward. Raise the bucket. Reach high and dump.	So I'm driving to the pile…I lower the bucket and as I'm pushing it to the pile I'm raising the bucket at the same time… Reverse, turn a little… Forward. Raise the bucket. Reach high and dump.
Fork lifting (F)/ Operator #15/ 22 words	Ok so I'm just picking this up and driving to the... [So,] I release the brake. Drive forward … Reverse… Over the barrier…	Ok so I'm just picking this up and driving to the… So, I release the brake. Drive forward… Reverse… Over the barrier.

The comments removed from the transcribed verbal data.

Simple bucket loading (B1)	Operator #17 • This is kind of awkward because this is where I would want to look around the side of the machine to make sure no one is there. • The sight really helps with the depth perception.
Filling a trench (B2)	Operator # 10 • Now this simulator is a little different because it doesn't show me... it keeps starting in rear view on the bottom corner, which is a little different than what I'm used to. • It doesn't show me the bucket. • Now when I'm approaching the trench, it's a little more difficult here because I can't see where the bottom of the trench is. • Normally you would be able to see the edge of the trench and what I watch out for is that I don't want to get to close to the trench. Operator #15 • I noticed the brakes don't quite stop it. It should stop completely. Operator #19 • Brakes don't work.. It still won't stop in neutral.. see this brake doesn't work.
Truck loading (B3)	No comments found.
Fork lifting (F)	Operator #16 • It's hard to see where the spot was.. I need to put my boom down. Operator #17 • It's kind of hard to see the barricades.

Tables generated after Phases 3 and 4 with 8 ranked verbal reports with highlighted verbs for a) bucket loading (B1), b) moving a load with wide forks (F), and c) truck loading (B3).

a) Simple bucket loading (B1)

Ranks	Operator ID	No. of verbs	Verbal Reports
1	4	15	Raise the bucket to carry position. Setting the blade toward the ground. Drive forward to the pile… Curl the bucket and raise the boom to load the bucket… Reserve and turn toward the dump area… back to forward… line up and perpendicular to the dump area… Raise the bucket. Dump the bucket. Curl bucket and reserve and turn toward the pile again…
2	10	13	So I go into the pile and what I'm doing now is I'm looking at the bucket. Square on the ground. And so then I take the brake off, and I'm curling the bucket back as I go into the pile. As I'm backing up, I raise the bucket just a little bit until I'm ready. So now I'm approaching it (dump area) and I'll start dumping the bucket… And now as I'm backing up, I'm lowering the boom and getting my bucket positioned at the same time.
3	19	13	Roll the bucket out so the cutting edge is flat. Line the machine up with the bucket… Then I drive over to the pile, so I load the bucket, go forward, accelerate into the pile, raise the boom and curl the bucket back until your bucket is full…Reverse… Line up with the green box. Go forward, raise the boom and dump it in the green box.
4	20	13	Make sure there's no one around before you drive… As you're coming into the pile, keep the bucket close to the ground. Creep up to the pile… Load my bucket… So we'll lower the bucket, creep up to the pile, curl it up trying to get a full bucket… Back out with the bucket in paddle position… Swing around… And dump the bucket… Back out, lower the bucket …
5	16	12	Lower the boom back down onto the ground then approach the pile…Drive in then release the throttle, slowly raising and curling the bucket …Pull back and away…I'm going to line up with the green box, drive towards it raising the boom and dump the bucket. Turn and reverse while you're turning.
6	18	12	Keep my bucket flat and now I'm going to pull up, turn the parking brake off, then pull up to fill the bucket… As I'm booming up, I'm going to fill my bucket so its full. Back away… I'm going to bring the bucket back to level. What I just did was I raised my boom and dumped my bucket out, then returned to dig. Go back into this…
7	17	11	So the first thing you are going to do is release the parking brake, shift into gear, then start driving it. Then I got to line up with the existing bucket… So as I'm backing out, I'm trying to reposition my bucket so I don't have to move it as far when I start to go forward… So as I back out, raise it up, take off the throttle and shift into gear…
8	15	6	Push the pile, curl up with the bucket … Get close, raise the bucket… Reverse… bucket dump…

b) Truck loading (B3)

Ranks	Operator ID	No. of verbs	Verbal Reports
1	10	17	Now when I approach the gravel pile, I want to make sure the bucket is level with the ground. And as soon as I approach the pile, I level the bucket so it's on the ground. And I want to make sure the bucket is square with the pile so that I don't just load one corner of the bucket… As I approach the pile and get a load in the bucket, I slowly lift up and curl the bucket back at the same time until I get a full bucket. Once I get a full bucket, I start backing up and turning… Now as I'm backing up, I'm slowly raising the bucket to prepare for loading the truck… Now I see my bucket is square with the truck…
2	16	15	So basically I drive towards the pile. Start loading a bucket close to the truck, lower it down bucket to the ground… Stay on the throttle until the bucket is full. Back it away, turning until I'm lined up with the truck. Move forward, raising the boom the whole time… Start dumping the bucket once you get close to the truck… Once the bucket is empty, return to dig and back away from the truck. Lower the boom and change directions.
3	17	15	So you put your bucket on the ground and then lift up a little bit and curl at the same time to get a full bucket… And then as you back out, you want to start turning right away so that when you pull up to the truck you are straight… And then when I start going forward, I start raising my boom at the same time… As I start to dump, I also lift the boom and dump at the same time… Once my bucket is dumped all the way, then I stop raising the boom, then I start backing up…
4	4	12	So, lift the bucket a little off the ground and approach the pile… As you get closer, lower it back down, right off the ground. Curl and lift the loader at the same time… Put it in reverse and back out away from the pile and the truck at the same time. Put it in forward, raise the bucket so I can dump it in the truck… Dump then raise [bucket]…
5	20	12	Approach the pile. Bucket low to the ground. Raise the boom, curl and fill the bucket…Then I'm going to back out. Square up with the truck… Raise the boom as you approach the truck…As you get up to the truck, you're going to want to slow down… Dump the bucket… Lower the boom back down…
6	15	9	Driving forward into the pile with my boom raised. [Apply] full throttle… So you got a full bucket then go in reverse… Drive forward and raise the boom… Apply the brakes and bucket dump…
7	19	9	Drive into the pile, raise the boom and curl the bucket back until the bucket is full. Reverse… Line up with the truck. Then you approach the truck. Raise the boom… Raise the boom and dump it into the truck…
8	18	9	So I'm driving to the pile…I lower the bucket and as I'm pushing it to the pile I'm raising the bucket at the same time… Reverse, turn a little… Forward. Raise the bucket. Reach high and dump.

c) Moving a load with a wide fork (F)

Ranks	Operator ID	No. of verbs	Verbal Reports
1	10	31	Ok so from a stopped position, I raise the boom off the ground slightly and tilt the forks so that they are parallel to the ground. That way so when I'm running forward I don't hit anything. Before I start, I put the brake on and I put it in gear… Try to approach the pipes in the center… As I approach them, I lower the forks down trying to center them on the plate, making sure that the corks are level… Once the forks start going through the pipes, I slowly inch forward so the pipe is far enough back so it will tilt… Once it is underneath the pipes, I will wrap the forks back and lift the boom up… and when I know they are on securely I will continue going backwards and go where I need to go… Once I get to where I need to go... When I set them down, I make sure that I am square with the bumpers. I lower the forks down to the ground… Raise the fork back up… Slowly back in… So when I'm done, I set the forks back on the ground and put the parking brake back on.
2	4	20	Raise the fork a little up the ground. Turn and approach… Approach the pipe. Parallel to the ground… Put the fork down, so it slides underneath the pipe… Keep it as straight as possible. Slowly lift up and curl back… So that they are off the ground by maybe 6 inches. I'm going to reverse, then do a 4-point turn and turn around … [Watch] barricades… Raising my boom as I'm going down the hill… And then lower to carry position… Muddy situation: Stuck in mud so I apply the differential lock… Try and line up with the [target location]… Dump the power fork so it's on the ground…
3	16	15	I'm trying to watch and see where the ends of the pipes are so that they don't hit the barriers. Keep myself lined up in the center of the path. Approach the hill, adjust my throttle and check the end of my pipes to make sure it's not going to run into anything… My tires spin so I just hit the diff. lock button. Hold it… To see where the spot was.. I need to put my boom down… I'm trying to get the pipe lined up with the center… Go over the top of it (the spot)…
4	20	15	So approach the pile, [get] forks down level. I got to straighten it out…Curl up the pipe, lift it up to the carrying position… I'm just watching the barrier… So it's pretty narrow so I'm definitely going to want to hold it up above… Make sure it stays curled back enough going up the hill so it doesn't slide out… Go over the muddy area…So you want to square up [with the spot]… Keep your forks flat, lower the pipes, and back away…
5	18	12	I'm lifting up and making sure the boom stays level… So now, I'm going to go left. I'm going to turn around… So I'm going to back up now. I can see the fence there… Now I have to go high enough to clear the things there… I'm going to lower the boom to get into position for it … I'll put it into second gear… drop over there by the red arrow…
6	17	11	Ok so I'm going to put it in gear… I want to make sure my forks are level… So I'm going to curl the bucket a little bit back towards me so the pipes do not fall off… I'm going to raise it enough to try and clear the barricade… So I'm just trying to follow the course and when I see it move, I try to slow down… So that's mud… Use the differential lock to move over it.
7	19	11	So you approach the pile and get your forks level and low to the ground… Drive under it slowly and get it, curl the bucket back and raise the boom… Raise the thing above so it doesn't hit any objects… Keeping them low to the ground… Avoid this barrier. This is the exposed part… Drive through the muddy area…
8	15	6	Ok so I'm just picking this (the fork) up and driving to the [pipe]… I release the brake. Drive forward… Reverse… [Drive] over the barrier…

Tables generated after Phase 5: Searching for themes from the verbal protocols for a) simple bucket loading, b) truck loading and c) fork lifting.

a) Simple bucket loading (B1)

Operator ID / *Theme*	#4	#10	#19	#16	#20	#18	#17	#15
A. Adjust bucket	Raise the bucket to carry position.	So I go into the pile and what I'm doing now is I'm looking at the bucket.	Roll the bucket out so the cutting edge is flat.	Lower the boom back down onto the ground		Keep my bucket flat		
	Setting the blade toward the ground.	Square on the ground.	Line the machine up with the bucket…					
B. Approach pile	Drive forward to the pile…	And so then I take the brake off,	Then I drive over to the pile,	then approach the pile…	Make sure there's no one around before you drive…	and now I'm going to pull up, turn the parking brake off,	So the first thing you are going to do is release the parking brake, shift into gear, then start driving	
					As you're coming into the pile, keep the bucket close to the ground.		Then I got to line up the bucket…	

(continued on next page)

Table continued.

Operator ID / *Theme*	#4	#10	#19	#16	#20	#18	#17	#15
C. Load bucket	Curl the bucket and raise the boom to load the bucket…	and I'm curling the bucket back as I go into the pile.	so I load the bucket, go forward, accelerate into the pile, raise the boom and curl the bucket back until your bucket is full…	Drive in then release the throttle, slowly raising and curling the bucket …	Creep up to the pile… Load my bucket… So we'll lower the bucket, creep up to the pile, curl it up trying to get a full bucket…	then pull up to fill the bucket… As I'm booming up, I'm going to fill my bucket so its full.		Push the pile, curl up with the bucket …
D. back out from pile	Reserve and turn toward the dump area…	As I'm backing up, I raise the bucket just a little bit until I'm ready.	Reverse… Line up with the green box.	Pull back and away…	Back out with the bucket in paddle position…	Back away… I'm going to bring the bucket back to level.	So as I'm backing out, I'm trying to reposition my bucket so I don't have to move it as far when I start to go forward…	
E. Towards the dump area	back to forward… line up and perpendicular to the dump area…	So now I'm approaching it (dump area)	Go forward,	I'm going to line up with the green box, drive towards it				Get close,
F. Raise bucket	Raise the bucket.		raise the boom	raising the boom				raise the bucket…

(continued on next page)

Table continued.

Operator ID / *Theme*	#4	#10	#19	#16	#20	#18	#17	#15
G. Dump bucket	Dump the bucket.	and I'll start dumping the bucket…	and dump it in the green box.	and dump the bucket.	Swing around… And dump the bucket…	What I just did was I raised my boom and dumped my bucket out,		bucket dump…
H. Back out from	Curl bucket and reserve and turn toward the pile again…	And now as I'm backing up, I'm lowering the boom and getting my bucket positioned at the same time.		Turn and reverse while you're turning.	Back out, lower the bucket …	then returned to dig. Go back into this…	So as I back out, raise it up, take off the throttle and shift into gear…	Reverse…

b) Truck loading (B3)

Operator ID / *Theme*	#10	#16	#17	#4	#20	#15	#18	#19
A. Approach the pile	Now when I approach the gravel pile, I want to make sure the bucket is level with the ground.	So basically I drive towards the pile.		So, lift the bucket a little off the ground and approach the pile	Approach the pile.		So I'm driving to the pile…	Drive into the pile,

(continued on next page)

Table continued.

Operator ID / *Theme*	#10	#16	#17	#4	#20	#15	#18	#19
B. Adjust bucket	And as soon as I approach the pile, I level the bucket so it's on the ground. And I want to make sure the bucket is square with the pile so that I don't just load one corner of the bucket…		So you put your bucket on the ground	… As you get closer, lower it back down, right off the ground.	Bucket low to the ground.			
C. Load bucket	As I approach the pile and get a load in the bucket, I slowly lift up and curl the bucket back at the same time until I get a full bucket.	Start loading a bucket close to the truck, lower it down bucket to the ground… Stay on the throttle until the bucket is full.	and then lift up a little bit and curl at the same time to get a full bucket…	Curl and lift the loader at the same time…	Raise the boom, curl and fill the bucket…	Driving forward into the pile with my boom raised.	I lower the bucket and as I'm pushing it to the pile I'm raising the bucket at the same time…	raise the boom and curl the bucket back until the bucket is full. …
						[Apply] full throttle…		

(continued on next page)

Table continued.

Operator ID / *Theme*	#10	#16	#17	#4	#20	#15	#18	#19
D. back out from pile until it lined up with the truck	Once I get a full bucket, I start backing up and turning… Now as I'm backing up, I'm slowly raising the bucket to prepare for loading the truck…	Back it away, turning until I'm lined up with the truck..	And then as you back out, you want to start turning right away so that when you pull up to the truck you are straight…	Put it in reverse and back out away from the pile and the truck at the same time.	Then I'm going to back out. Square up with the truck…	So you got a full bucket then go in reverse…	Reverse, turn a little…	Reverse… Line up with the truck.
E. Toward the truck	Now I see my bucket is square with the truck…	Move forward, raising the boom the whole time…	And then when I start going forward, I start raising my boom at the same time…	Put it in forward, raise the bucket so I can dump it in the truck…	Raise the boom as you approach the truck…	Drive forward and raise the boom…	Forward. Raise the bucket.	Then you approach the truck. Raise the boom…
G. Dump bucket		Start dumping the bucket once you get close to the truck… Once the bucket is empty, return to dig and back away from the truck. Lower the boom and change directions	As I start to dump, I also lift the boom and dump at the same time…	Dump then raise [bucket]…	As you get up to the truck, you're going to want to slow down… Dump the bucket…	Apply the brakes and bucket dump…	Reach high and dump.	Raise the boom and dump it into the truck
H. Back out from truck			Once my bucket is dumped all the way, then I stop raising the boom, then I start backing up…		Lower the boom back down…			

c) Moving a load with wide fork (F)

Operator ID / Theme	#10	#4	#16	#20	#18	#17	#19	#15
A. Drive to the pile	Ok so from a stopped position, I raise the boom off the ground slightly and tilt the forks so that they are parallel to the ground.	Raise the fork a little up the ground.	I'm trying to watch and see where the end of the pipes are so that they don't hit the barriers	So approach the pile,		Ok so I'm going to put it in gear…	So you approach the pile	Ok so I'm just picking this (the fork) up and driving to the [pipe]…
	That way so when I'm running forward I don't hit anything.	Turn and approach… Approach the pipe.	Keep myself lined up in the center of the path.					Drive forward…
B. Adjust bucket	Before I start, I put the brake on and I put it in gear…	Parallel to the ground…		[get] forks down level.	I'm lifting up and making sure the boom stays level…	I want to make sure my forks are level….	and get your forks level and low to the ground…	
	Try to approach the pipes in the center…	Put the fork down, so it slides underneath the pipe…						
		Keep it as straight as possible.						
C. Pick up pipes	As I approach them, I lower the forks down trying to center them on the plate, making sure that the corks are level…	Slowly lift up and curl back…		I got to straighten it out…		So I'm going to curl the bucket a little bit back towards me so the pipes do not fall off…	Drive under it slowly and get it, curl the bucket back and raise the boom…	

(continued on the next page)

Table continued.

Operator ID / *Theme*	#10	#4	#16	#20	#18	#17	#19	#15
C. Pick up pipes	Once the forks start going through the pipes, I slowly inch forward so the pipe is far enough back so it will tilt…	So that they are off the ground by maybe 6 inches.		Curl up the pipe, lift it up to the carrying position…				
	Once it is underneath the pipes, I will wrap the forks back and lift the boom up…							
D. back out from pile until it lined up with the truck		I'm going to reverse, then do a 4-point turn and turn around …			So now, I'm going to go left.			Reverse…
					I'm going to turn around… So I'm going to back up now.			
E. Over the hill	and when I know they are on securely I will continue going backwards and go where I need to go…	[Watch] barricades…	Approach the hill, adjust my throttle and check the end	I'm just watching the barrier…	I can see the fence there…	I'm going to raise it enough to try and clear the barricade…	Raise the thing above so it doesn't hit any objects…	I release the brake.
		Raising my boom as I'm going down the hill…	of my pipes to make sure it's not going to run into anything…	So it's pretty narrow so I'm definitely going to want to hold it up above…	Now I have to go high enough to clear the things there…	So I'm just trying to follow the course and when I see it move, I try to slow down…	Keeping them low to the ground…	[Drive] over the barrier…
	Once I get to where I need to go...	to carry And then lower position…		Make sure it stays curled back enough going up the hill so it doesn't slide out…			Avoid this barrier.	

(continued on next page)

Table continued.

Operator ID / *Theme*	#10	#4	#16	#20	#18	#17	#19	#15
F. muddy area		Muddy situation: stuck in mud so I apply the differential lock…	My tires spin so I just hit the diff. lock button.	Go over the muddy area…		So that's mud…	This is the exposed part…	
			Hold it…			Use the differential lock to move over it	Drive through the muddy area…	
G. Line up with the targeted location	When I set them down, I make sure that I am square with the bumpers.	Try and line up with the [target location]…	To see where the spot was... I need to put my boom down…	So you want to square up [with the spot]…	I'm going to lower the boom to get into position for it …			
			I'm trying to get the pipe lined up with the center…		I'll put it into second gear…			
			Go over the top of it (the spot)…					
H. Drop the pipes.	I lower the forks down to the ground…	Dump the power fork so it's on the ground…		Keep your forks flat, lower the pipes,	Drop over there by the red arrow…			
	Raise the fork back up…							
I. Back out from the pipes	Slowly back in…			and back away…				
	So when I'm done, I set the forks back on the ground and put the parking brake back on.							

The mixed-design repeated-measures ANOVA comparing the workload measures in four experimental tasks.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Measure	87.25	5	17.450	8.783162	0.000	0.239
Measure * Task	43.188	15	2.879	1.44918	0.133	0.134
Error(Measure)	278.146	140	1.987			
Between-Subjects						
Task	129.890	3	43.297	5.134651	0.006	0.355
Error(Task)	236.104	28	8.432			

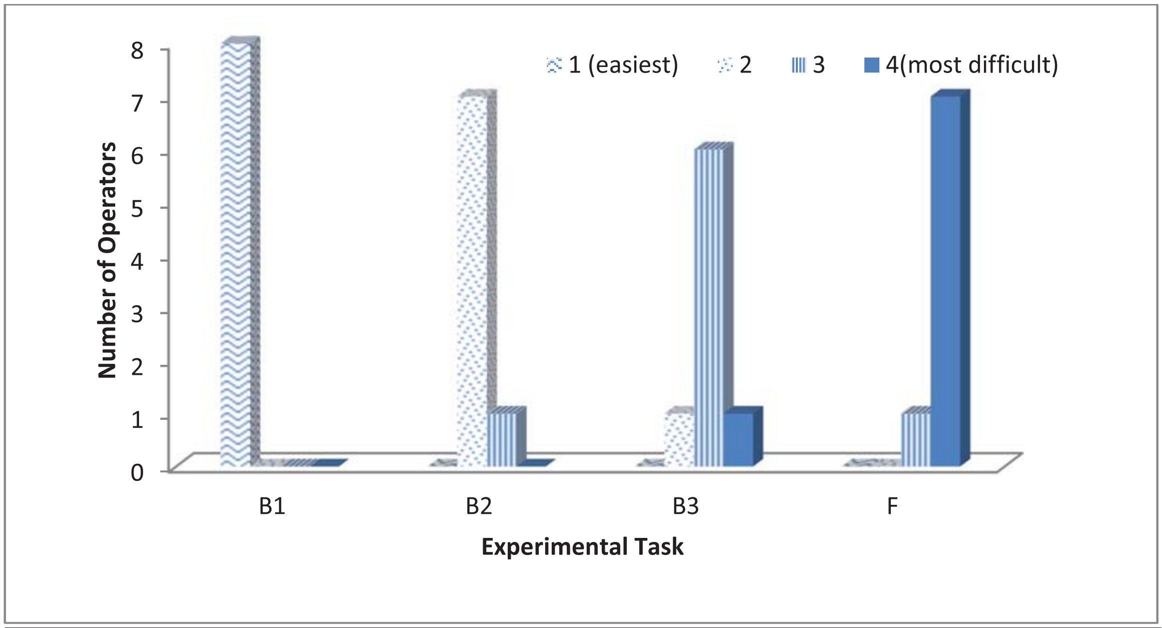
Results of Bonferroni multiple comparisons performed on average workload measure with the main factor of experiment task.

Task (I)	Task (II)	Mean Difference (I-II)	Std. Error	Sig.
B1 Simple bucket loading	B2 Filling a trench	-0.938	0.593	0.750
	B3 Truck loading	-1.063	0.593	0.503
	F Fork lifting	-2.313	0.593	0.003
B2 Filling a trench	B3 Truck loading	-0.125	0.593	1.000
	F Fork lifting	-1.375	0.593	0.167
B3 Truck loading	F Fork lifting	-1.250	0.593	0.264

Results of Wilcoxon Signed Rank tests on difficulty ranks for B1, B2, B3 and F.

	B2 - B1	B3 - B1	F - B1	B3 - B2	F - B2	F - B3
Z	-2.714	-2.636	-2.714	-2.111	-2.640	-2.111
Asymp. Sig. (2-tailed)	0.007	0.008	0.007	0.035	0.008	0.035

Distribution of the difficulty levels of the four loader tasks ranked by 8 operators.



Appendix I: Study 5

The mixed-design repeated-measures ANOVA comparing the productivity on simulated loader.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-subject Factors						
Session	70436.636	1	70436.636	119.496	0.000	0.681
Session * Method	5070.390	3	1690.130	2.867	0.045	0.133
Error(Session)	33009.052	56	589.447			
Trial	156471.451	4	39117.863	150.827	0.000	0.729
Trial * Method	3993.806	12	332.817	1.283	0.230	0.064
Error(trial)	58095.833	224	259.356			
Session * Trial	199.416	4	49.854	0.199	0.939	0.004
Session * Trial * Method	4159.602	12	346.633	1.385	0.174	0.069
Error(Session*Trial)	56068.312	224	250.305			
Between-Subjects Factor						
Method	131548.769	3	43849.590	8.936	0.000	0.324
Error	274788.700	56	4906.941			

Results of Bonferroni multiple comparisons performed on productivity on the loader with the main factor of training method.

Method (I)	Method (II)	Mean Difference (I-II)	Std. Error	Sig.
B1>B2>B3	B2>B1>B3	8.828	8.089	1.000
	B1>F>B3	26.146	8.089	0.012
	B2>F>B3	38.069	8.089	0.000
B2>B1>B3	B1>F>B3	17.318	8.089	0.220
	B2>F>B3	29.242	8.089	0.004
B1>F>B3	B2>F>B3	11.923	8.089	0.876

The mixed-design repeated-measures ANOVA comparing the productivity on simulated loader for first and last trials across four training method.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-Subject Factors						
Session	29614.197	1	29614.197	98.847	0.000	0.638
Session * Method	948.278	3	316.093	1.055	0.376	0.053
Error(Session)	16777.470	56	299.598			
Trial	132259.215	1	132259.215	256.579	0.000	0.821
Trial * Method	2186.974	3	728.991	1.414	0.248	0.070
Error(Trial)	28866.391	56	515.471			
Session * Trial	102.735	1	102.735	0.364	0.548	0.006
Session * Trial * Method	2372.785	3	790.928	2.806	0.048	0.131
Error(Session*Trial)	15783.824	56	281.854			
Between-Subjects Factor						
Method	43088.978	3	14362.993	8.020	0.000	0.301
Error	100286.961	56	1790.839			

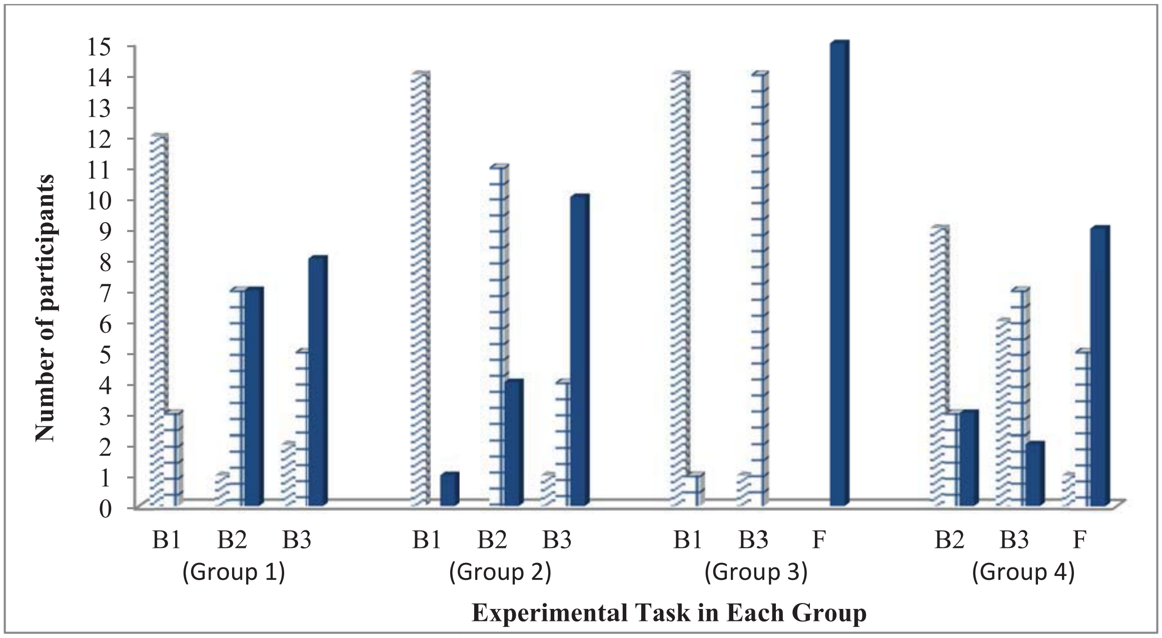
The mixed-design repeated-measures ANOVA comparing the workload measure on truck loading task.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Within-Subject Factors						
Session	228.094	1	228.094	47.001	0.000	0.456
Session * Method	20.800	3	6.933	1.429	0.244	0.071
Error(Session)	271.767	56	4.853			
Measure	170.561	5	34.112	10.247	0.000	0.155
Measure * Method	54.871	15	3.658	1.099	0.357	0.056
Error(Measure)	932.126	280	3.329			
Session * Measure	52.021	5	10.404	5.709	0.000	0.093
Session * Measure * Method	40.016	15	2.668	1.464	0.118	0.073
Error(Session*Measure)	510.270	280	1.822			
Between-Subjects Factor						
Method	146.738	3	48.913	6.439	0.001	0.256
Error	425.375	56	7.596			

The mixed-design repeated-measures ANOVA comparing the TLX ratings on simulated loader for all three session across four training method.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Group 1: B1>B2>B3						
Within-Subjects Factors						
Session	6.724074	2	3.362037	1.154885	0.329652	0.076205
Error(session)	81.51204	28	2.911144			
Measure	111.0741	5	22.21481	3.082998	0.014253	0.180472
Error(measure)	504.3912	70	7.205589			
session * measure	17.79815	10	1.779815	1.041629	0.412004	0.06925
Error(session*measure)	239.2157	140	1.708684			
Group 2: B2>B1>B3						
Within-Subjects Factors						
Session	4.212963	2	2.106481	1.219372	0.310626	0.08012
Error(session)	48.37037	28	1.727513			
Measure	297.038	5	59.40759	15.81043	2.14E-10	0.530366
Error(measure)	263.0245	70	3.757493			
session * measure	19.22037	10	1.922037	1.347472	0.21124	0.087798
Error(session*measure)	199.6963	140	1.426402			
Group 3: B1>F>B3						
Within-Subjects Factors						
Session	4.22963	2	2.114815	0.351178	0.706914	0.02447
Error(session)	168.6176	28	6.022057			
Measure	218.2157	5	43.64315	8.528459	2.4E-06	0.378564
Error(measure)	358.2148	70	5.117354			
session * measure	22.54815	10	2.254815	1.280947	0.246775	0.083826
Error(session*measure)	246.438	140	1.760271			
Group 4: B2>F>B3						
Within-Subjects Factors						
Session	4.505556	2	2.252778	0.482759	0.62212	0.033333
Error(session)	130.6611	28	4.666468			
Measure	240.3111	5	48.06222	17.44495	3.52E-11	0.554777
Error(measure)	192.8556	70	2.755079			
session * measure	20.08333	10	2.008333	1.271768	0.252026	0.083276
Error(session*measure)	221.0833	140	1.579167			

Distribution of the difficulty levels of the loader tasks ranked by different group.



Data Analysis on the difficulty levels of the loader tasks ranked by different experimental group.

a) Results of Friedman tests on the ranks of the three tasks given to each experimental group.

	Group			
	1 (B1>B2>B3)	2 (B2>B1>B3)	3 (B1>F>B3)	4 (B2>F>B3)
N	15	15	15	15
Chi-Square	14.400	17.733	28.133	8.373
Df	2	2	2	2
Asymp. Sig.	0.001	0.000	0.000	0.015

b) Results of Wilcoxon signed-rank tests on difficulty levels ranked by each experimental group

	Comparison	Z	Asymp. Sig. (2-tailed)
Group 1	B2 - B1	-3.218	0.001
	B3 - B1	-2.982	0.003
	B3 - B2	-0.030	0.976
Group 2	B1 - B2	-3.220	0.001
	B3 - B2	-1.291	0.197
	B3 - B1	-2.981	0.003
Group 3	F - B1	-3.771	0.000
	B3 - B1	-3.357	0.001
	B3 – F	-3.771	0.000
Group 4	F - B2	-2.280	0.023
	B3 - B2	-0.361	0.718
	B3 – F	-2.448	0.014

Descriptive Statistics showing the perceived percentage (%) of time spent on the visual feedback on the simulator screen of experts and novice.

	Expert (N=8)		Novice (N=60)	
Feedback	Mean	SD	Mean	SD
Egocentric View	66.38	19.40	37.98	18.61
Damage	2.38	2.33	6.93	6.11
Time	1.25	2.31	4.42	6.63
Instruction	1.75	2.31	6.80	6.63
Warning	1.50	1.85	8.52	6.67
Bucket Fill (%)	5.88	6.01	12.10	11.26
Side View	20.63	13.48	23.25	12.87

VITA

VITA

Chung Yin (Joey) So was born in Hong Kong, China. She received her Bachelor's degree in Manufacturing Systems Engineering at City University of Hong Kong in 2007. During her undergraduate study, Joey participated in a student exchange program at Carnegie Mellon University, Pittsburgh, United States in 2005 fall and a co-op training program as a research assistant at University of Greenwich, United Kingdom in 2006 summer. She was awarded the Faculty Medal in 2007 (renamed as 'College Medal' in 2008), which is the most prestigious award by the College of Science and Engineering of City University of Hong Kong, in recognition of both academic excellence and demonstration of well-round personal qualities. Upon graduation, she completed a Master of Philosophy Degree in Manufacturing Engineering and Engineering Management from City University of Hong Kong. Her Master's thesis was titled 'The effects of different design factors on comprehension performance and subjective evaluation for reading Chinese on LED display', supervised by Dr. Alan H. S. Chan. She received Sir Edward Youde Memorial Fellowships 2008/09, Student Member with Honors Award granted by the national Human Factors and Ergonomics (HFES) in 2009, and the 'Outstanding Project Award 2010 (master level)' from the Hong Kong Ergonomics Society.

After graduation from City University of Hong Kong, she enrolled in the Industrial Engineering PhD program at Purdue University, where she has been advised by Drs. Robert Proctor and Phillip Dunston. Later in 2012, she decided to complete an additional Master's degree in Cognitive Psychology, which she received in December 2013. During her time at Purdue, Joey was elected as the graduate student representative for the Engineering Advisory Council of Purdue University between 2012 and 2014. Joey served as the Vice President of the Industrial Engineering Graduate Student Organization between 2012 and 2013, the President of the HFES Purdue Student Chapter between 2011 and 2012 and was re-elected for 2012-2013 academic year. Joey was recognized as an outstanding graduate teaching assistant, receiving the 2011-2012 Magoon Award for Excellence in Teaching from the Engineering Professional Education at Purdue. She also completed two summer internships with John Deere, as Ergonomics Intern and Product Engineering Intern, in Moline, IL, in 2012 and 2013, respectively.

PUBLICATIONS

PUBLICATIONS

Journal Articles

Proctor, R. W., Dunston, P. S., **So, J. C. Y.**, Lopez-Santamaria, B. N., Yamaguchi, M., & Wang, X. (2013). Specificity of Transfer in Basic and Applied Perceptual-Motor tasks. *American Journal of Psychology*, 126, 401-415.

***So, J.C.Y.**, Proctor, R.W., Dunston, P.S. and Wang, X., (2013), Better Retention of Skill Operating a Simulated Hydraulic Excavator After Part-Task Than After Whole-Task Training, *Human Factors,* 55, 449 - 460.

So, J.C.Y. and Chan, A.H.S., (2013) Effects of Display Method, Text Display Rate and Observation Angle on Comprehension Performance and Subjective Preferences for reading Chinese on an LED Display, *Displays*, 34, 371-379.

So, J.C.Y. and Chan, A.H.S., (2013). Effects of Display Method, Number of Message Lines, and Text Colour on Chinese Comprehension and Subjective Preferences for an LED Display, *Journal of the Society for Information Display*, 21, 181–191.

Chan, A.H.S. and **So, J.C.Y.** (2009). Task Factor Usability Ratings for Different Age Groups Writing Chinese. *Ergonomics*, 52, 1372–1385.

So, J.C.Y. and Chan, A.H.S. (2008). Display Factors on Dynamic Text Display. *Engineering Letters*, 16, 368-371.

<u>Conference Presentations/ Proceedings</u>

*__So, J.C.Y.__, Proctor, R.W. and Dunston, P.S., (2014). Training on Perceptual-Motor Tasks Using Simulated Construction Equipment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2014*, pp. 2360-2364.

*__So, J.C.Y.__, Proctor, R.W. and Dunston, P.S., (2012). Impact of Interrupting Simulated Hydraulic Excavator Training with Simulated Loader Training. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2012*, pp. 2482-2486.

*__So, J.C.Y.__, Proctor, R.W., and Dunston, P.S., (2012). Transfer of Learning across Two Earthmoving Equipment Simulators. *Applied Human Factors and Ergonomics 2012 (AHFE International Conference 2012)*, 21-25 July, San Francisco, California, USA.

So, J.C.Y., Proctor, R.W., Dunston, P.S. and Wang, X., (2012). Acquisition And Retention Of Skill At Operating A Simulated Hydraulic Excavator Using Part-And Whole-Task Training. *Midwest Cognitive Science Conference 2012*, 7 May, Bloomington, Indiana, USA.

Proctor, R. W., Dunston, P. S., **So, J.C.Y.**, Wang, X., (2012). Task Analysis for Improving Training of Construction Equipment Operators, Construction Challenges in a Flat World. *Proceedings of 2012 Construction Research Congress*, Hubo Cai, Amr Kandil, Makarand Hastak, Phillip S. Dunston (editors), American Society of Civil Engineers, 169-178.*

Wang, X., Dunston, P.S., Proctor, R. W., Hou, L. and **So, J.C.Y.** (2011). Reflections On Using A Game Engine To Develop A Virtual Training System For Construction Excavator Operators Pages, *Proceedings of the 28th ISARC*, Seoul, Korea, pp. 631-636.

Chan, A.H.S., **So, J.C.Y.** and Tsang, N.H., (2011). Developing Optimum Interface Design for On-screen Chinese Proofreading Tasks. *Proceedings of HCI International*, 9-14 July 2011, Orlando, Florida, USA (CD ROM).

So, J.C.Y. and Chan, A.H.S. (2009). Design Factors of LED Display for Improving Message Comprehension. *Proceedings of the 17th World Congress on Ergonomics* (IEA 2009), 9-14 August 2009, Beijing, China (CD ROM).

So, J.C.Y. and Chan, A.H.S. (2009). Influence of Highlighting Validity on Dynamic Text Comprehension Performance. *Proceedings of International MultiConference of Engineers and Computer Scientists 2009* (IMECS 2009), 18-20 March 2009, Hong Kong, China, pp. 1970-1973. **(Best Student Paper Award of 2009 IAENG International Conference on Industrial Engineering)**

So, J.C.Y. and Chan, A.H.S. (2008). Variation of error rates for different task factors in Chinese writing. *Proceedings of Applied Human Factors and Ergonomics 2008* (AHFE International Conference 2008), Las Vegas, USA, 14-17 July 2008 (CD ROM).

So, J.C.Y. and Chan, A.H.S. (2008). Effects of Design Factors on Comprehension Performance with Dynamic Text Display: A Review. *Proceedings of International MultiConference of Engineers and Computer Scientists 2008* (IMECS 2008), 19-21 March 2008, Hong Kong, China, pp. 1843-1845. **(Best Student Paper Award of 2008 IAENG International Conference on Industrial Engineering – Human Factors and Ergonomics)**

Book Chapters

So, J.C.Y. and Chan, A.H.S. (2009). Validity of Highlighting on Text Comprehension. *IAENG Transactions on Engineering Technologies* (published by America Institute of Physics), 3, pp. 217-224.

So, J.C.Y. and Chan, A.H.S. (2009). Display Factors and Subjective Evaluation of Dynamic Text Display. *IAENG Transactions on Engineering Technologies* (published by America Institute of Physics), 1, pp. 123-131.

*Note: This work contains some of the research in this dissertation.