An Approach to Identify Effective Learning Outcomes for a Training Program

Yoon Suk Lee

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Industrial and Systems Engineering

Advisory Committee:

Dr. Tonya L. Smith-Jackson, Chair
Dr. Brian M. Kleiner, Member
Dr. Maury A. Nussbaum, Member

Dec 14, 2007
Blacksburg, Virginia

Keywords: Training system, Revised Bloom's Taxonomy, Learning outcomes

An approach to Identify Effective Learning Outcomes for a Training Program

Yoon Suk Lee

ABSTRACT

Low back disorders (LBDs) are one of the most commonly occurring injuries in industry. To attempt to reduce these work-related injuries, billions of dollars are being budgeted for formal training in the U.S. However, the outcomes of this training are below a satisfactory level. The majority of organizations utilize the Four-level Evaluation Model to evaluate their training program. However, previous studies have pointed out some limitations regarding this evaluation model. Moreover, most organizations collect only trainee reaction, the first level of the Four-level Evaluation Model, to determine the effectiveness of their training program. Many studies reveal that trainee reaction is an invalid indicator to determine the effectiveness of a training program, and further suggest multi-dimensional categorization within each level of the Four-level Evaluation.

Therefore, in this study, the Revised Bloom's Taxonomy was employed to enable multidimensional categorization of learning outcomes in a lifting and lowering training program. The learning outcomes of interest in such a training program relate to procedural knowledge and the cognitive process involved are categorized as remembering, understanding, applying, and evaluating the contents of the training program. Two research questions were asked. What types of learning outcomes are most predictive of training performance? How do the learning outcomes predict training performance compared to affective and utility type reactions?

The ability of different types of learning outcomes to predict training performance was tested by multiple regression analyses. The results revealed that apply-procedural learning outcomes and the interaction variable of understand-procedural and apply-procedural learning outcomes were the most predictive of training performance. Further, these learning outcomes were more predictive of training performance than the trainee reactions (affective and utility type reactions) to explain training performance. The results of this study yielded a set of recommendations that may be useful in designing and evaluating lifting and lowering training programs. Moreover, this study examined the Revised Bloom's Taxonomy as a novel method of considering the multidimensional nature of learning and provided a potential application of the Revised Bloom's Taxonomy in the training discipline.

ACKNOWLEDGEMENTS

Throughout this journey to achieve a small accomplishment, there were so many people who have devoted their time and effort to emotionally support me and to share their knowledge when I was burdened with so many problems that I could not solve by myself.

I would like to show my gratitude to my advisory committee: Dr. Smith-Jackson, Dr. Kleiner, and Dr. Nussbaum. Thanks for all of you for your questions and comments that made me learn how to think as a researcher. I would like to especially thank my advisor, Dr. Smith-Jackson, for her guidance and patience. Without your unflagging support, this work would not have been possible.

I would also like to show my deep appreciation to Don Taylor, Joel Dighans, and Stephen Jones for providing the opportunity for me to work on a project that was funded by the Department of Labor. I would like to especially extend my appreciation to Don who originally made me think about this topic and who always listened and provided constructive feedback to me. Thank you.

To all KISE and fellow HFEE students: I have learned how important peers can be. Your support and collegiality has been a source of encouragement throughout this study. I would like to extend my sincere appreciation to Sunwook Kim for always listening to me and sharing thoughts with me. Your ceaseless encouragement made me overcome all sorts of problems that challenged me. You have showed me what a true mentor should be. Special thanks to Harry Yoo, Gyu Hyun Kwon, and Kichol Lee for always listening to my complaints and arguments. Thanks to Jason Meyer and Liana Algarin for devoting your valuable time to proofreading my document.

Finally, there is no such word to describe how thankful I am to my parents in Korea. Your forever love has been and will be the most valuable source of strength for me as I continue through my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

According to recent statistics from the Bureau of Labor Statistics (BLS, 2005b), a total of 375,540 musculoskeletal disorders (MSDs) were reported in 2005. Roughly 24% of these cases were attributable to workers' motions or positions. On average, workers with occupational injuries lost more than $10,000 in lifetime earnings (Reville, Bhattacharya, & Sager, 2001; Weil, 2001). Employers also spend a substantial amount of money due to lost productivity, hiring, training of replacements, and administrative supervisory activities (Reville, Bhattacharya, & Sager, 2001). Thus, work-related injuries and fatalities clearly have significant impacts on workers, their families, and their employers.

Organizations have implemented different types of interventions (e.g., training) to control work-related injuries, especially occupational back pain. Training is defined as "the systematic acquisition of skills, rules, concepts, or attitudes that result in improved performance in another environment" (Goldstein & Ford, 2002, p. 1). In 2005, $51.1 billion was budgeted for formal training (Dolezalek, 2005). However, many studies noted that trainees utilize less than 10% of what they learn during training (Baldwin & Ford, 1988; Fitzpatrick, 2001; Georgenson, 1982; Saks & Belcourt, 2006; Wexley & Baldwin, 1986). Likewise, training programs to reduce low back disorders (LBDs) showed little success (Gagnon, 2003). Some reasons for this are a lack of control conditions, inappropriate measurement techniques (Fabio, 1995), and inadequate training methods (St-Vincent, Tellier, & Lortte, 1989).

1.1. Problem statement

Training evaluation has become an important issue for training researchers and practitioners (Alliger *et al.*, 1997). Training evaluation is both costly and labor intensive (Salas

& Cannon-Bowers, 2001), and evaluation criteria must be psychometrically sound (Alliger *et al.*, 1997). One of the most popular frameworks of evaluation criteria is Kirkpatrick's typology (Kirkpatrick, 1976). This consists of four evaluation criteria, known as the Four-level Evaluation Model. However, some studies have revealed a number of conceptual flaws in this framework (Alliger & Janak, 1989; Snyder, Raben, & Farr, 1980).

Kraiger *et al.* (1993) questioned whether the Four-level Evaluation Model differentiates between skills and facts, since the model measures them with the same assessment tools. This is problematic, giving that these elements are substantively different; skills represent the "how" of knowledge, whereas facts reflect the "what" of knowledge. Knowing a fact does not always mean that the person has the skills to demonstrate it. Therefore, different types of assessments are required to measure different types of knowledge. Although many studies have modified the Four-level Evaluation Model (Brinkerhoff, 1987; Hamblin, 1974; Kaufman & Keller, 1994), revised models are still questionable due to incomplete implementation and a lack of empirical testing (Holton III, 1996).

Training evaluation becomes more critical when considering how organizations conduct their evaluation. A survey completed by the American Society for Training and Development found that 75% of organizations collected trainee reactions (level 1 of the Four-level Evaluation Model) to evaluate their training program (Sugrue, 2003). This indicates that a majority of the organizations believe that trainee reactions are valid and reliable indicators for determining the effectiveness of a training program. However, researchers have empirically demonstrated the weak correlation between trainee reactions and other evaluation criteria, suggesting that trainee reactions should not be utilized as their only indicator for training evaluation (Alliger & Janak, 1989; Alliger *et al.*, 1997; Arthur *et al.*, 2003; Noe & Schmitt, 1986b).

Despite these researchers' conclusions, there are several reasons that organizations continue to solely use trainee reactions. First, it is easy to implement and collect data, since it is a self-reported measure. Second, it is cheaper to plan and implement than levels 2, 3, and 4 from the Four-level Evaluation Model. From an organizational perspective, the optimal solution is to evaluate their training program effectively with less cost and input. However, an organizational decision regarding training solely based on potentially invalid measures (trainee reactions) could eventually lead to an increase in expenses. It is thus important to develop inexpensive and easily applicable evaluation criteria that are pertinent to training effectiveness.

The majority of the previous studies regarding training evaluation examined the relationships among different criteria of the Four-level Evaluation Model. The correlations and the predictive validity among the criteria were tested to analyze the effectiveness of a training system and to further develop a sound and meaningful one. As a result, developing more detailed criteria to measure different aspects of the effectiveness of a training system became a paramount issue. More suggestions and recommendations were made to highlight the importance of more sophisticated assessments in order to diagnose different types of learning outcomes (Salas & Cannon-Bowers, 2001). Kraiger *et al.* (1993) suggested that more detailed and thorough taxonomies of learning outcomes should be applied in the field of training. Analogously, researchers from educational and instructional design underlined that extensive taxonomies are considered more appropriate to modern learning theories, reflecting a multifaceted nature of learning (Bloom, 1956; Gagne, 1984; Krathwohl, Bloom, & Masia, 1964).

1.2. Background

This study was part of a larger project to design a state-of-the art safety training system for a commercial delivery company. This safety program reflects the organization's goal to foster the most effective safety practices in all phases of their employees' work. Therefore, it covers every aspect of the safety issues related to delivery work, such as job preparation, package handling, slips and falls, powered equipment, driving, and hazard identification.

The delivery company organized the training modules into three different phases: Prequalification, Workshop, and Post-workshop. The Virginia Tech research team focused on integrating their design into the existing framework. For example, Learning Stations and Kinetic Learning Modules (KLM) were included as sub-phases of the Workshop. In addition, just-in-time (JIT) feedback, peer-assisted learning (PAL), peer evaluation (buddy system), and scenario based learning were employed and designed by the Virginia Tech research team. Figure 1 explains the different phases of the training program in more detail.

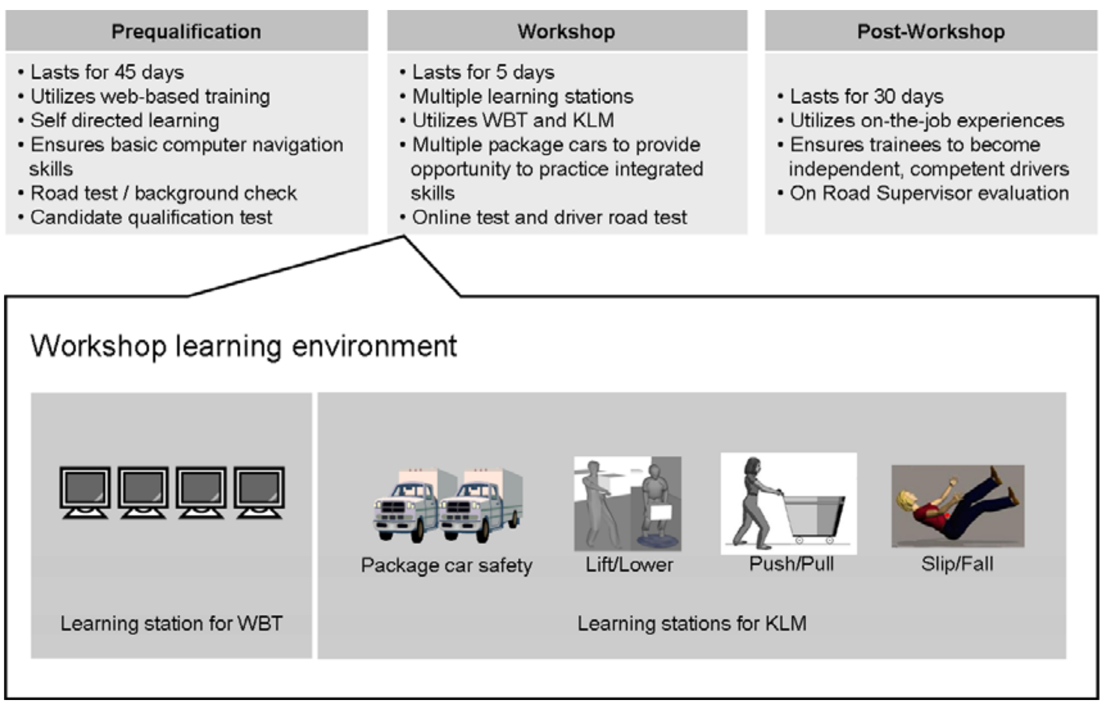


Figure 1. Overview of the training program

1.2.1. Prequalification

The Prequalification phase is the beginning of the training system. In it, trainees were required to learn the provided web-based training (WBT) materials in 45 days. It was designed for the trainees as self-directed learning to allow them to complete the WBT during the required time. Only the key concepts were introduced in this phase in order to provide basic ideas of what the trainees would be learning in the subsequent phases. Frequent assessments along with just-in-time (JIT) feedback were employed to enable self-checks to help the trainees enhance their level of understanding. At the end of this phase, there was a Candidate Qualification Test to ensure that the trainees had learned the materials well enough to proceed to the next training phase.

1.2.2. Workshop

Participants who passed the Candidate Qualification Test in the Prequalification phase were qualified for the Workshop, which consisted of two subsections: Learning Stations using web-based training (WBT) and Kinetic Learning Modules (KLM). WBT was utilized during the Learning Stations to introduce more detailed concepts with various scenarios as examples, along with more complicated assessments. The scenarios and assessments reflected more of the actual work context. The second subsection, the KLM, was designed to allow trainees to have hands-on experience of previous materials that they have learned. For example, general concepts were introduced in the Prequalification phase, and these concepts were elaborated during the Learning Stations. Then in KLM, trainees applied these concepts to their demonstrations. The KLM phase provided opportunities to synthesize declarative and procedural knowledge and apply it to specific situations.

During the KLM, peer assisted learning (PAL) was employed by providing performance feedback. The first advantage of PAL is that it addresses and overcomes trainees' various ability levels (Kroeger & Kouche, 2006). Furthermore, it strengthens learning through social interactions among peers (Brown, 2005). Therefore, trainees were paired and demonstrations took place one person at a time while the trainee's counterpart observed and completed a checklist (Appendix E) to evaluate his/her peer. The demonstrations were recorded by a digital camcorder, and each pair watched the recorded video to discuss each other's performance based on their checklist evaluations.

Using this training program, this study utilized a laboratory environment to determine the predictive validities of training performance in lifting and lowering. Therefore, the training contents that were not related to lifting and lowering will be excluded hereafter.

1.3. Research purpose

The goals of this study were to apply multidimensional categorization to expand the second level learning of the Four-level Evaluation Model, and to determine the predictive validity of different types of learning outcomes to training performance in a lifting and lowering training program. To achieve these research purposes, the Revised Bloom's Taxonomy was utilized to categorize learning outcomes and the two following research questions were asked.

1.4. Research questions

1. What types of learning outcomes are most predictive of training performance?

2. How do the learning outcomes predict training performance compared to affective and utility type reactions?

1.5. Hypotheses

1. For the first research question, it was hypothesized that learning outcomes that were classified as apply-procedural would be more predictive of training performance in lifting and lowering. This hypothesis was tested by classifying learning outcomes with the Revised Bloom's Taxonomy, obtaining assessment scores for each learning outcome, and comparing total scores of each learning outcome to the training performance scores.

2. For the second research question, it was hypothesized that the learning outcomes would more strongly correlate with training performance than trainee reactions (affective type, utility type). This hypothesis was tested by collecting affective reactions and utility reactions after the training, and comparing each variable to the training performance scores that were obtained in research question #1.

CHAPTER 2. LITERATURE REVIEW

2.1. Low back disorder

The National Institute for Occupational Safety and Health (NIOSH) has considered MSDs as the leading research priority and disease prevention effort in the United States (NIOSH, 1996). Many researchers have indicated that low back disorders (LBDs) are among the most costly and common MSDs (Marras, 2000; Snook, 1978). According to the BLS, 270,890 cases related to back injuries or illnesses were reported in 2005, causing 21.9% of all non-fatal occupational injuries and illnesses (BLS, 2005a). In an organizational perspective, back disabilities account for approximately one hundred million lost work days each year. Furthermore, Webster and Snook (1994) noted that 16% to 19% of all worker compensation claims were due to LBDs. The estimated overall annual costs for LBDs has been as high as $100 billion (Bigos *et al.*, 1986), resulting, making it one of society's most expansive health disorders (Marras, 2000; Punnett *et al.*, 2005).

Inappropriate manual material handling (MMH) has been a primary factor in the development of the occupational LBDs (Bigos *et al.*, 1986). Lifting and MMH account for 50% to 75% of all back injuries (Bigos *et al.*, 1986), which are considered the most frequent health problems for industrial workers (Svensson & Andersson, 1983). The critical reviews investigated by Bernard (1997) and Hoogendoorn *et al.* (1999) have also drawn strong associations between lifting and forceful movements and LBD risk.

Another rationale for focusing on lifting and lowering is that the training program utilized in this study includes some unique design features. The 2005 Industry Report indicated that 70% of formal training is conducted in a classroom environment with live instructors despite a large emphasis on technology-based learning such as e-learning, mobile learning and on-demand learning, (Dolezalek, 2005). However, the training program that will be tested in this research differs from traditional training or most currently available training programs. As described in the background section, this training program includes WBT with interactive assessments along with JIT feedback, and most importantly, KLM with peer-assisted learning. This type of training environment is expected to be adequate for testing various types of learning outcomes (e.g., skills and abilities) whereas a traditional classroom environment is more adequate for testing knowledge.

2.2. Trainee reactions (Level 1)

Trainee reaction, the first level of the Four-level Evaluation Model, is a measure of trainee's feelings about a training program (Tan, Hall, & Boyce, 2003). It is the most common criterion used in the industry to evaluate training programs (Bassi, Benson, & Scott, 1996; Saari *et al.*, 1988). However, for over twenty years, researchers have consistently questioned the over-usage of trainee reaction as a sole predictor for determining the effectiveness of a training program (Ruona *et al.*, 2002; Tan, Hall, & Boyce, 2003). Noe and Schmitt (1986a) have suggested that trainee satisfaction has no significant relationship with learning. Similarly, Warr and Bunce (1995b) found no significant correlation between reported enjoyment of training, usefulness of training, and learning scores. However, there is a strong correlation between learning score and job performance. One effort to overcome this problem was to distinguish trainee reactions into affective and utility reactions (Alliger *et al.*, 1997). Other studies have further emphasized the importance of understanding the multidimensional nature of trainee reactions (Morgan & Casper, 2000; Tan, Hall, & Boyce, 2003). In their meta-analyses, Alliger *et al.* (1997) showed that stronger correlations were found between utility reactions and learning or job performance than between affective reaction measures and learning or job performance.

A similar conclusion was made by Ruona *et al.*(2002), who searched for the relationship between utility reactions and the predictors of learning transfer. They suggested utility reactions for learning outcome evaluations, rather than traditional affective reactions.

2.3. Learning (Level 2)

Learning, the level 2 of the Four-level Evaluation Model, is defined as "principles, facts, and techniques understood and absorbed by the trainees" (Alliger & Janak, 1989). As mentioned in the problem statement, many researchers have argued the need for more detailed taxonomies for learning outcomes; these would better present the multifaceted nature of learning. To fulfill these needs, many researchers in the field of education, industrial organizational psychology, and training have developed a number of frameworks.

Krathwohl *et al.* (1964) expanded the taxonomy developed by Bloom (1956) to incorporate affect-oriented objectives along with cognitively-based learning outcomes. Later, Gagne (1984) categorized learning outcomes into intellectual skills (procedural knowledge), verbal information (declarative knowledge), cognitive strategies (executive control process), motor skills, and attitudes. This classification strengthened instructional objectives, enabling them to examine not only the behavioral domain but also various cognitive, skill-oriented, and attitudinal learning outcomes (Gagne, 1984). Based on these taxonomies, Kraiger *et al.* (1993) proposed categories of learning outcomes that involved cognitive, skill-based, and affective; they then explained each category with recommendations for evaluation measurements.

Alliger *et al.* (1997) in their meta-analysis classified learning (level 2) into three categories: immediate knowledge, knowledge retention, and behavior/skill demonstration. The first category, immediate knowledge, utilizes multiple choice test responses, open-ended questions, and listing of facts to evaluate trainees' knowledge. The second category, knowledge retention, is similar to immediate knowledge but administered at a later point of time rather than just after training. The last category, behavior/skill demonstration, comprises any indicators of behavioral proficiency that are measured during the training.

All frameworks have emphasized the multi-dimension of learning outcomes. In this study, an alternative categorization of learning outcomes will be provided by utilizing the Revised Bloom's Taxonomy from education and instructional design. The latter sections will offer more detailed description and rationales for employing the taxonomy.

2.4. Skill acquisition

Skill is "an ability that allows a goal to be achieved within some domain with increasing likelihood as a result of practice" (Rosenbaum & Carlson, 2001, p. 454). In this regard, skill acquisition can be considered as attaining capabilities by practice, which helps increase the possibility of goal achievement. Many researchers have broken down skill acquisition into various phases for a better understanding. For example, Snoddy (1926) introduced two stages, adaptation and facilitation, to explain the learning curve. In the adaptation stage, the improvement depends on more than repetition while, the facilitation stage depends on repetition alone.

Fitts (1964) subsequently suggested three stages of skill acquisition. In the first stage, or cognitive stage, a considerable amount of cognitive activity is required to encode the skill into a form that will make it possible for learners to generate the desired behavior. The primary goal for the learners in this stage is to know what is to be done rather than to improve their motor patterns. Therefore, improvements can be observed through verbal mediation. The second stage of skill acquisition, the associative stage, involves a gradual improvement on performance. Errors can be detected and eliminated by practice so that the performance level becomes more consistent. Further, verbal mediation reduces and the learner focuses more on "how to do" rather than "what to do." In Fitts' last phase, the autonomous stage, learners can automatically perform the desired skill. The performance requires less mental workload, and secondary tasks can be performed with less interference.

Later, Anderson (1982) offered a broader explanation of Fitts' three stages, but used a different lexicon. In his construction, the declarative stage, the knowledge compilation, and the procedural stage respectively correspond to Fitts' three stages of skill acquisition. However, while utilizing similar underlying concepts, Anderson emphasized the knowledge type in each stage. For example, he notes that the declarative stage mostly utilizes verbal mediation, and the procedural stage focuses on performance itself. The knowledge compilation is a conversion phase from the declarative to procedural stage.

In the lifting and lowering training program, factual and conceptual information is provided to support trainees' understanding of knowledge and skills. However, the training program gradually emphasizes the trainees' performance more, since the ultimate goal focuses more on the level of performance. The first hypothesis of this study, that learning outcomes regarding application will be more predictive of training performance, can be supported by Fitts and Anderson's three stages of skill acquirement. It can be expected that predictors that involve the actual application of performance will more accurately anticipate training performance than will those that rely upon factual or conceptual knowledge. Additionally, it is likely that multi-dimension of knowledge and cognitive abilities are involved throughout skill acquisition. This supports the multi-dimension of learning outcomes related to the lifting and lowering training program, which is reviewed in the following paragraphs.

2.5. Multi-dimension of learning outcomes

Training is defined as "the systematic acquisition of skills, rules, concepts, or attitudes that result in improved performance in another environment" (Goldstein & Ford, 2002, p. 1). It is a specialized form of education that focuses on developing or improving performance. Education, on the other hand, focuses on whether learners have gained knowledge and skill, but does not evaluate actual performance or application in a different environment (Rekus, 1999). That is, simply understanding and knowing a specific subject does not warrant an effective training system. An effective training system must ensure that learners are able to perform the expected outcomes in such way that they have demonstrably learned during the training. Researchers have become more aware of the significance of the multidimensional nature of learning (Bloom, 1956; Gagne, 1984; Krathwohl, Bloom, & Masia, 1964) and have tried to apply those theories in designing and evaluating training systems (Kraiger, Ford, & Salas, 1993).

2.5.1. The Revised Bloom's Taxonomy

The Revised Bloom's Taxonomy guides the process of classifying learning objectives, as well as defining and selecting the clearer assessments that are more relevant to the learning objectives (Airasian & Miranda, 2002). This taxonomy framework consists of two dimensions: knowledge and cognitive process. The knowledge dimension involves four general types of knowledge: factual, conceptual, procedural, and meta-cognitive. The cognitive process dimension involves six categories: remembering, understanding, applying, analyzing, evaluating, and creating (Anderson & Krathwohl, 2001). Each category and subcategory are described in Table 1 and 2.

Table 1. The major types and subtypes of the knowledge dimension [adopted from (Anderson & Krathwohl, 2001, p. 29)]

A. Factual knowledge: the basic elements students must know to be acquainted with a discipline or solve problems in it
A1. Knowledge of terminology A2. Knowledge of specific details and elements
B. Conceptual knowledge: the interrelationships among the basic elements within a larger structure that enable them to function together
B1. Knowledge of classifications and categories B2. Knowledge of principles and generalizations B3. Knowledge of theories, models, and structures
C. Procedural knowledge: how to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods
C1. Knowledge of subject-specific skills and algorithms C2. Knowledge of subject-specific techniques and methods C3. Knowledge of criteria for determining when to use appropriate procedures
D. Metacognitive knowledge: knowledge of cognition in general as well as awareness and knowledge of one's own cognition
D1. Strategic knowledge D2. Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge D3. Self-knowledge

Table 2. The six categories of the cognitive process dimension and related cognitive processes [Adopted from (Anderson & Krathwohl, 2001, pp. 67-68)]

1. Remember: Retrieve relevant knowledge from long-term memory
1.1 Recognizing 1.2 Recalling
2. Understand: Construct meaning from instructional messages, including oral, written, and graphic communication
2.1 Interpreting 2.2 Exemplifying 2.3 Classifying 2.4 Summarizing 2.5 Inferring 2.6 Comparing 2.7 Explaining
3. Apply: Carry out or use a procedure in a given situation
3.1 Executing 3.2 Implementing
4. Analyze: Break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose
4.1 Differentiating 4.2 Organizing 4.3 Attributing
5. Evaluate: Make judgments based on criteria and standards
5.1 Checking 5.2 Critiquing
6. Create: Put elements together to form a coherent or functional whole; reorganize elements into a new pattern or structure
6.1 Generating 6.2 Planning 6.3 Producing

The two dimensions are now tabulated and form the Taxonomy Table which is illustrated in Table 3.

Table 3. The Taxonomy Table [adopted from (Anderson & Krathwohl, 2001, p. 28)]

The Knowledge Dimension	**The Cognitive Process Dimension**					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual Knowledge						
B. Conceptual Knowledge						
C. Procedural Knowledge						
D. Metacognitive Knowledge						

The rows and columns in the Taxonomy Table respectively correspond to each category of the knowledge dimension and the cognitive process dimension. Thus, a learning objective can be classified by recognizing the verb or verb phrase and the noun or noun phrase in the learning objective. The verb is then examined according to the cognitive process dimension. Similarly, the noun is examined according to the knowledge dimension. For example, if a learning objective is defined as "the trainees will learn how to apply the proper procedure of using a carry aid," the first step is to identify the verb and the noun in the learning objective (step 1). Then the verb will be examined across the categories in the cognitive process dimension, and the noun will be examined through the categories in the knowledge dimension. In this example, it is the verb "apply" corresponds to "apply" from the cognitive process dimension. Likewise, the noun phrase "proper procedure of using a carry aid" can be categorized as procedural knowledge (step 2). Accordingly, the learning objective fits into one of the cells from the Revised Bloom's Taxonomy (step 3). Figure 2 illustrates the classification process.

Step 0 Instructional Objective:
The trainees will learn how to apply the proper procedure of using a carry aid.

Step 1 Locating the verb or verb phrase and the noun or noun phrase

Verb: apply

Noun: proper procedure of using a carry aid

Step 2 Examining according to the cognitive process dimension and knowledge dimension

Verb: apply → Cognitive process dimension
1. Remember
2. Understand
3. Apply
4. Analyze
5. Evaluate
6. Create

Noun: proper usage of the Power Zone → Knowledge dimension
A. Factual knowledge
B. Conceptual knowledge
C. Procedural knowledge
D. Metacognitive knowledge

Step 3 Placing the learning objective to the Taxonomy Table

The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual Knowledge						
B. Conceptual Knowledge						
C. Procedural Knowledge			●			
D. Metacognitive Knowledge						

Figure 2. Classification process of learning objectives

The original Bloom's Taxonomy was developed in 1956 when the behaviorist learning theory was predominant. However, in subsequent years, the development of the educational and psychological disciplines have lead to advancement of other theories of learning, including Constructivism (Underhill, 2006), Metacognition (Foster *et al.*, 2003), and Self-regulated Learning (Boekaerts, 1997). Accordingly, it has been necessary to revise Bloom's Taxonomy to accommodate more recent developments. The most significant change was the expansion from one dimension to two dimensions. By adding another dimension, the revised taxonomy adequately represents the dual nature of the knowledge category; i.e., verb and noun components (Krathwohl, 2002). In addition, a fourth category, Metacognitive Knowledge, was included in the knowledge dimension.

Metacognitive Knowledge is defined as "knowledge about cognition in general as well as awareness of and knowledge about one's own cognition" (Pintrich, 2002, p. 219). The significance of metacognitive knowledge to a learning strategy is that it enables students to learn and perform better (Amer, 2006). More importantly, metacognitive knowledge is related to the learning transfer, which is the ability to apply the previously gained knowledge in another context or situation (Bransford & Cocking, 1999).

The Revised Bloom's Taxonomy also had significant changes in its cognitive process dimension. Knowledge, comprehension, and synthesis were renamed to remember, understand, and create respectively. Categories with noun forms were changed to verb forms to accommodate how those are used in instructional objectives; thus, application, analysis, and evaluation were changed to apply, analyze, and evaluate. In addition, the order of Synthesis/Create and Evaluation/Evaluate was interchanged because the former requires a more complex process. Anderson & Krathwohl (2001) noted that create is a more complex process than evaluate, since creation involves induction while evaluation involves deduction.

Table 4. The original Taxonomy

1.0 Knowledge
- **1.10 Knowledge of specifics**
  - 1.11 Knowledge of terminology
  - 1.12 Knowledge of specific facts
- **1.20 Knowledge of ways and means of dealing with specifics**
  - 1.21 Knowledge of conventions
  - 1.22 Knowledge of trends and sequences
  - 1.23 Knowledge of classifications and categories
  - 1.24 Knowledge of criteria
  - 1.25 Knowledge of methodology
- **1.30 Knowledge of universals and abstractions in a field**
  - 1.31. Knowledge of principles and generalizations
  - 1.32 Knowledge of theories and structures

2.0 Comprehension
- 2.1 Translation
- 2.2 Interpretation
- 2.3 Extrapolation

3.0 Application

4.0 Analysis
- 4.1 Analysis of elements
- 4.2 Analysis of relationships
- 4.3 Analysis of organizational principles

5.0 Synthesis
- 5.1 Production of a unique communication
- 5.2 Production of a plan, or proposed set of operations
- 5.3 Derivation of a set of abstract relations

6.0 Evaluation
- 6.1 Evaluation in terms of internal evidence
- 6.2 Judgments in terms of external criteria

2.5.2. Rationale for applying the Revised Bloom's Taxonomy

The revised taxonomy addresses evolving theories in education and psychology. Thus, it reflects a better understanding of how people learn, which is explained as the multifaceted nature of learning outcomes. The knowledge and cognitive process dimensions yield twenty four combinations of learning outcomes, which is seemingly detailed enough to classify learning outcomes in the field of training. The taxonomy also satisfies the suggestions made by other researchers in the training domain that stressed to the importance of applying more detailed and sophisticated taxonomies of learning outcomes in the field of training.

The Revised Bloom's Taxonomy is designed to provide clear learning objectives, support designing assessments, and align correspondence among the objectives and assessments (Anderson & Krathwohl, 2001). Clear learning objectives are assured by the classification process. For example, a vague learning objective, such as having more than one concept, cannot be easily placed into one of the cells in the Taxonomy Table, because the learning objective potentially fits into more than one cell. This classification difficulty impels training designers to reconsider the learning objective and to divide the combined concepts into different learning objectives.

Having a clear learning objective can support the development of valid assessments. The type of assessment differs by the type of knowledge and type of cognitive skill required. For example, to assess the factual knowledge and remember categories, the trainees should be required to memorize factual knowledge. Comparably, procedural knowledge and apply categories can be assessed by examining whether trainees can use proper procedures in their application of knowledge.

Lastly, aligning the learning objectives and assessments increases the validity of the training evaluation. If misalignment occurs between instruction and assessments, even well designed instructions will not affect the trainees' performance on the assessments. Likely, if misalignment occurs between learning objectives and assessments, the assessment results will not indicate the achievement of the learning objectives. For these reasons, the Revised Bloom's Taxonomy can be utilized as training evaluation criteria, which potentially supports training design.

2.6. An instructional system

A systems approach to training was presented by Goldstein (Goldstein, 1992) who offered an instructional system as shown in the figure below. The model consists of four phases, which are needs assessment, training and development, evaluation, and training goal.

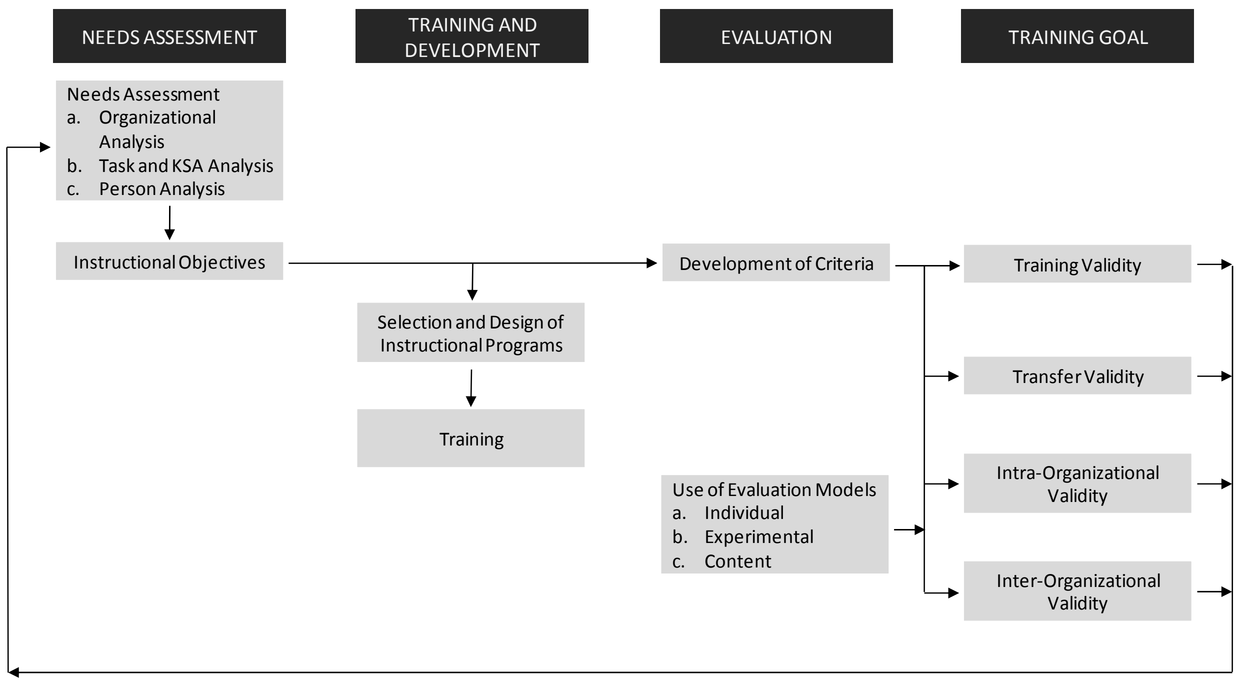


Figure 3. An instructional system [adopted from (Goldstein, 1992, p. 21)]

The first phase of Goldstein's instructional system, needs assessment, provides the information necessary to design and evaluate a training program. Organizational analysis focuses on the macro level of analysis, which involves organizational goals, training climate, and environmental constraints. The primary objective of this phase is to determine whether training is an appropriate solution for resolving problems caused by human resources. The second phase of the needs assessment (i.e., task and knowledge, skills, and abilities (KSAs) analysis) is somewhat significant in terms of developing the detailed goals of a training program. For a given job, tasks are identified and KSAs are determined in this phase so as to support the trainees to achieve the selected task. The decisions made in this phase play a major role in guiding the next phase, which is training and development. Person analysis involves determination of the individual's knowledge or skill. As a result of the needs assessment, specific instructional objectives are yielded. These instructional objectives are the cornerstones for designing training contents, materials, and assessments.

In the training and development phase, the primary goal is to blend instructional learning principles with the effective selection of media in order to determine the type of training program that will yield the best outcome. Therefore, it is important to consider how learning objectives can effectively represent the results from job task and KSA analysis on the trainees.

The evaluation phase mainly focuses on two procedures, which establish measures of success (criteria) and determine the changes that occurred during and after the training (transfer process). The criteria must outline the required skills that the trainee must achieve, the specific conditions under which the trainee is to perform, and the lowest acceptable boundaries. The criteria must be developed to evaluate the trainees' performance immediately after the training program and during on-the-job performance. Therefore, several different measures must be designed and collected for a valid evaluation. In his "Training in Organization," Goldstein mentioned Kirkpatrick's evaluation model as an example that consists of different levels of evaluation measures.

The second focus of the evaluation phase is to determine how to evaluate the training program. This depends on the type of outcome or information that is needed as a result of the evaluation. Different types of data can be collected with different experimental designs. The results in the evaluation phase will ultimately lead to the training goals.

Finally, four potential goals are listed in the last column of the model (Figure 3). Each item questions different aspects of training evaluation. Training validity checks whether trainees learned during the training, and transfer validity investigates the on-the-job performance. Intra-organizational validity examines the internal consistency of the training program, whereas inter-organizational validity tests the generalizability of the training program for other organizations. Further, the model emphasizes the flow of continuous feedback from the evaluation to modify the training program, which is represented by an arrow that connects the training goals to the needs assessment.

2.7. Transfer of training

The transfer of training is defined as "the degree to which trainees effectively apply the knowledge, skills, and attitudes gained in training context to the job" (Baldwin & Ford, 1988, p. 63). It is not surprising that transfer of training is one of the major concerns in training research. The transfer of training can be considered as the purpose of training and an organization's expectations for its employees who go through training. As mentioned in the introduction, the transfer problem in organizational training is increasing. Learned behavior must be generalized to the work context and maintained over a period of time on the job (Baldwin & Ford, 1988). Therefore, generalization and maintenance can be the conditions for transfer of training. Figure 4 illustrates the framework for the transfer process.

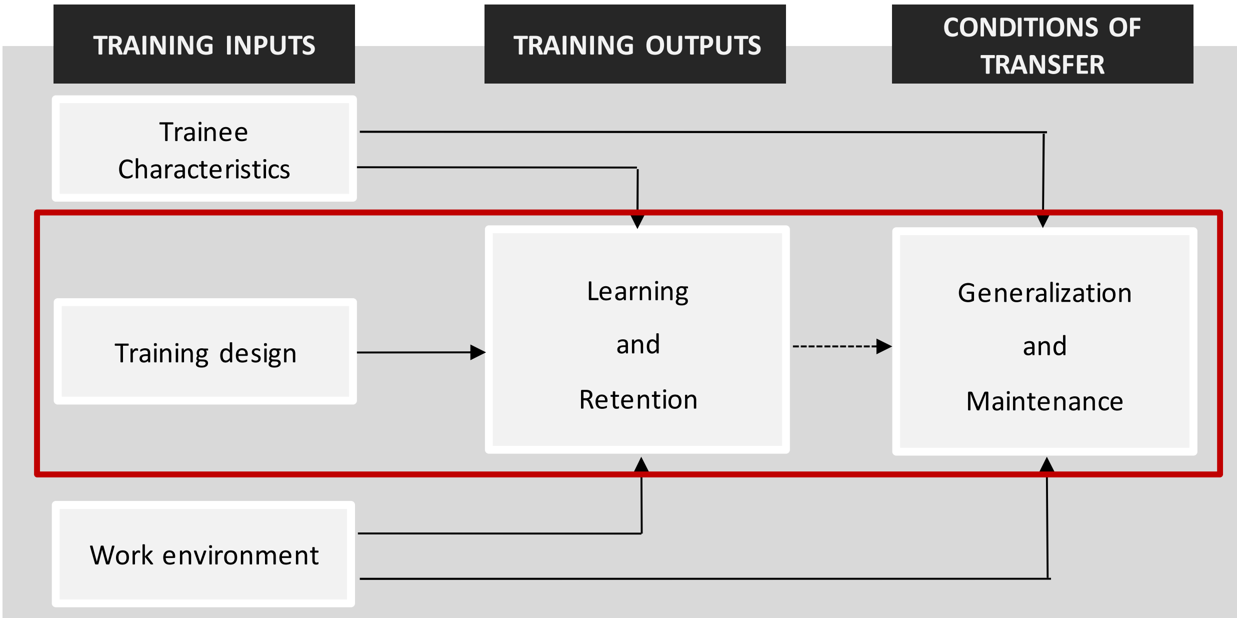


Figure 4. A Model of the Transfer Process [Adopted from (Baldwin & Ford, 1988, p. 65)]

The transfer process model is comprised of training input factors, training output factors, and conditions of transfer. Training outputs can be defined as the amount of original learning that has occurred with the training and retention of that learned material after the training program (Baldwin & Ford, 1988). Training inputs include trainee characteristics, training design, and work-environment characteristics. Many empirical studies have looked into the relationships of each entity and have found that all of these factors have relatively significant correlation to the training outputs as well as to the transfer of training. Since this study focuses on finding the predictive validities of different types of learning outcomes to training performance, the following paragraphs will discuss the relationship of training design, training outputs, and transfer of training.

Working backward within the Transfer Process Model (Figure 4), training transfer is directly correlated with the training outputs of learning and retention. Training performance, which is a similar to training outputs but more specialized to performance, is defined as “the ability to perform a newly acquired skill at the end of training, prior to transfer, and is measured through observable demonstration that a trainee can implement the knowledge acquired in training” (Alvarez, Salas, & Garofano, 2004, p. 397). It is important to note that many researchers support the idea that training performance has positive influences on transfer performance (Alliger *et al.*, 1997; Kirkpatrick, 1976; Kraiger, 2002). Tannenbaum *et al.* (1991) also revealed that trainees whose performance outcomes were acceptable during the training tend to have a positive relationship with post-training motivation. Successful learners are expected to perform better and are more motivated to transfer what they have learned to their work (Holton III, 1996). Training design, which is one of the training inputs, includes principles of learning, sequencing, and training content. Additionally, Baldwin and Ford (1988) in their review paper on the transfer of training addressed the importance of applying instructional theories in training design as future research directions.

The following diagram (Figure 5) illustrates how the transfer process model is applied in this study.

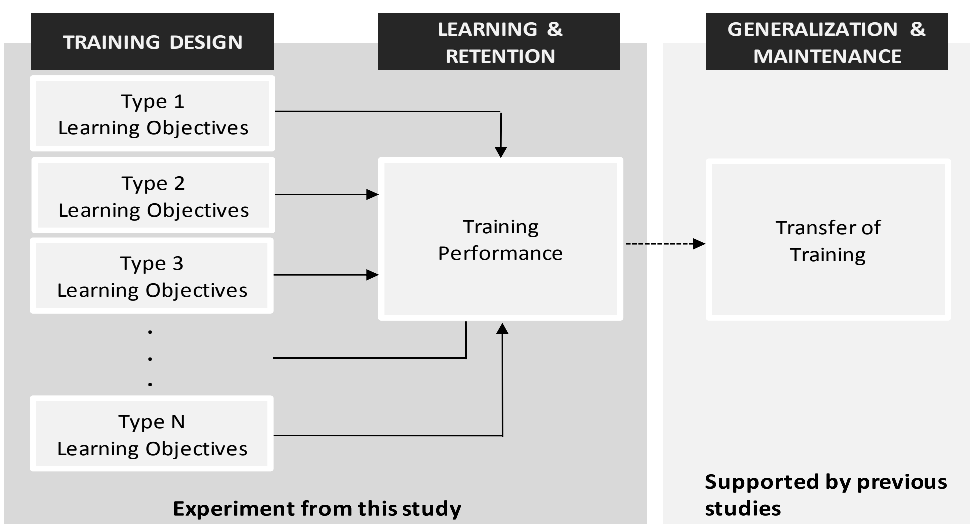


Figure 5. A Transfer Process of Types of Learning Outcomes

Different types of learning objectives, which form the training content of the training design, will have a relationship with the training performance in training outputs. The Revised Bloom's Taxonomy was applied to categorize the learning objectives and was further tested to predict the training performance. As with existing studies on the transfer of training, this study considers the connection between learning objectives and transfer; specifically, it asks what types of learning objectives in a lifting and lowering training program have the strongest correlation to the transfer of training. Although this study limits its boundary to finding the strongest correlation between different types of learning outcomes and training performance, the transfer of training can be assumed because various studies support the positive correlation of training performance and transfer of training.

2.8. Research purpose

The goals of this study were to apply the multidimensional categorization scheme to expand the second level learning of the Four-level Evaluation Model and to determine the predictive validity of different types of learning outcomes to training performance. To achieve the first goal, the Revised Bloom's Taxonomy was reviewed and employed to classify the learning outcomes. The classified learning outcomes were then utilized to predict the training performance with respect to affective and utility type reactions.

The Revised Bloom's Taxonomy not only provides a systematic classification of learning outcomes but may also offer some advantages in training design and evaluation. Having knowledge of a type of learning outcome that best explains the training performance may support and provide guidelines in various aspects of training design and evaluation. Details will be explained in the discussion section based on the theories and models that were introduced in the literature review section.

CHAPTER 3. METHODS

A laboratory-based experiment was conducted to train the participants to safely lift and lower packages and to obtain an assessment of their learning outcomes and their training performance scores. As illustrated in Figure 6, the assessment scores were classified into four different types of learning outcomes, and the performance scores were computed by using training performance checklist scores and calculating the NIOSH lifting index. The four learning outcomes were then tested for their ability to predict the checklist scores and the NIOSH lifting index (Hypothesis 1). Then, the most predictive learning outcomes and two types of reactions to training (affective and utility type reaction), which were collected after the training were compared to determine the variable that best explains the training performance scores (Hypothesis 2). Multiple regression analyses and correlation analyses were utilized to examine the ability of combinations of variables to predict training performance. The following diagram illustrates the overall study design.

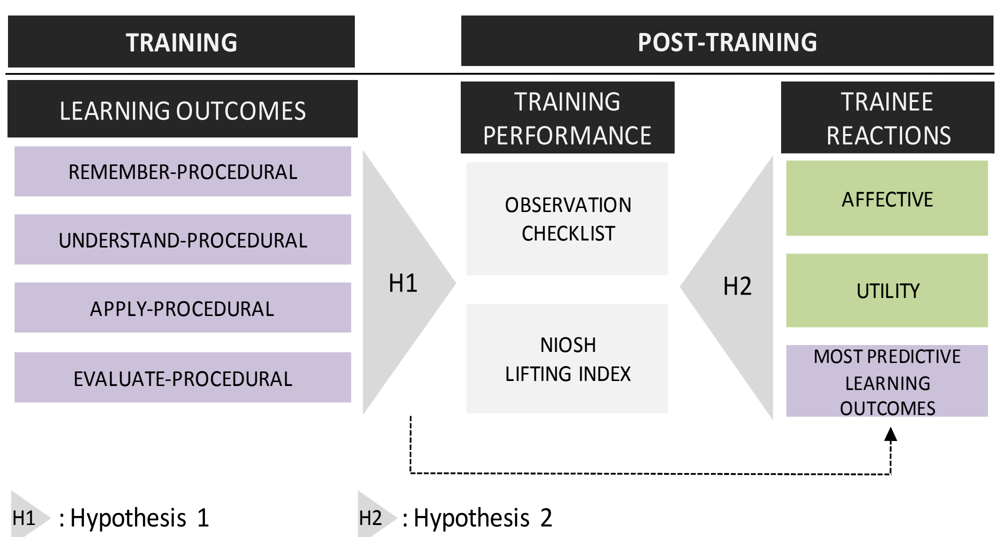


Figure 6. Overall study design

3.1. Learning outcomes

An organization should have expectations and goals for the trainees, and a training system must satisfy these expectations and goals to be successful. The organization's expectations and goals are implied by the learning objectives, which will later be described in terms of their status as independent and dependent variables in this study. The following sections will act as a guide through the process of generating the independent and dependent variables.

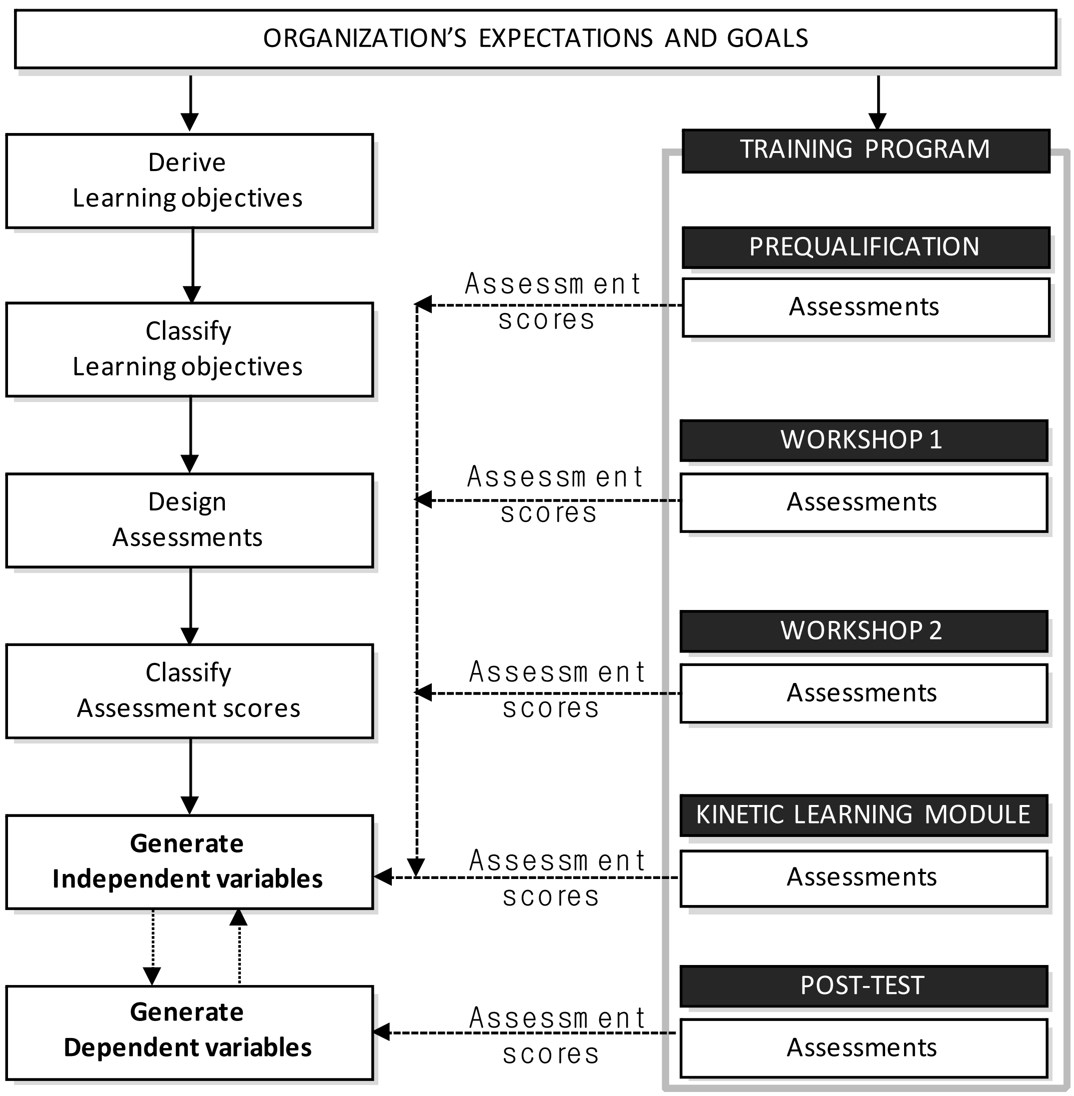


Figure 7. Generating independent and dependent variables

3.1.1. Learning objectives

Learning objectives were intended to explicitly indicate what the trainees should learn and accomplish. Table 5 shows the learning objectives that were used to train lifting and lowering.

Table 5. List of learning objectives

PREQUALIFICATION
P1. Evaluate proper usage of the Power Zone
P2. Evaluate proper usage of the End-Range Motion
P3. Understand proper footing
P4. Evaluate proper method of lifting
P5. Understand testing objects for weight and shifting contents
P6. Understand the concept of working within the Power Zone
P7. Understand the concept of bending at the knees (keeping the natural curve of the back)
P8. Understand the concept of moving feet, stepping or pivoting (without twisting)
WORKSHOP
W1. Understand how force should be applied to the back
W2. Remember to keep the load close to the body
W3. Understand how to distribute the load in package handling
W4. Remember to use smooth and controlled motions
W5. Evaluate proper package handling by maintaining the natural curve of the back
W6. Remember to use leg muscles when handling packages
W7. Remember to smoothly and cautiously handle packages (do not jerk)
W8. Understand the cause of how package handling affects the back anatomy
W9. Evaluate (determine) End-range of Motion from an incorrect lifting posture
W10. Evaluate (determine) improper package handling
W11. Remember how to avoid End-Range Motions
W12. Remember to test object's weight and for shifting contents
W13. Remember to avoid grasping package by its straps or bands
W14. Remember to move feet, step or pivot to prevent twisting while handling packages.
KINETIC LEARNING MODULE
K1. Apply proper usage of the Power Zone
K2. Apply proper method of avoiding End-Range Motion
K3. Apply proper foot position
K4. Apply how to keep the natural curve of the back
K5. Apply proper method of bending at the knees
K6. Apply how to test object for weight and shifting contents before handling a package
K7. Apply proper method of getting a firm grip and grasp
K8. Apply smooth and steady motion
K9. Apply how to avoid twisting by moving feet
K10. Apply the concept of properly distributing the load

* P=Prequalification, W= Workshop, K= Kinetic Learning Module

3.1.2. Classification of learning objectives using the Revised Bloom's Taxonomy

The learning objectives were then classified into different types of categories according to the Revised Bloom's Taxonomy. All of the learning objectives that were listed in Table 5 were classified along two dimensions of the taxonomy and were distributed among the taxonomy's categories as shown in Table 6.

Table 6. Classification of the learning objectives

The Knowledge Dimension	**The Cognitive Process Dimension**					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual						
B. Conceptual						
C. Procedural	L5, W7, W9, W10, W14, W15, W16, W17 (8)	P4, P6, P7, P8, P9, W4, W6, W11 (8)	K1, K2, K3, K4, K5, K6, K7, K8, K9, K10 (10)		P1, P2, P5, W8, W12, W13 (6)	
D. Metacognitive						

*( ): total number of learning objectives

3.1.3. Designing assessments

The assessments were designed by the Virginia Tech research team. Each learning objective had its own assessment to ensure that the trainees had mastered the learning objectives. A total of seven types of assessments were used: (1) multiple choice questions, (2) multiple response questions, (3) drag and drop (image positioning), (4) true / false, (5) short answer, (6) ordering, and (7) observation using checklist. For example, the assessments and feedback for learning objective P2, which states that trainees should evaluate proper usage of End-Range Motion, is shown in Figure 8.

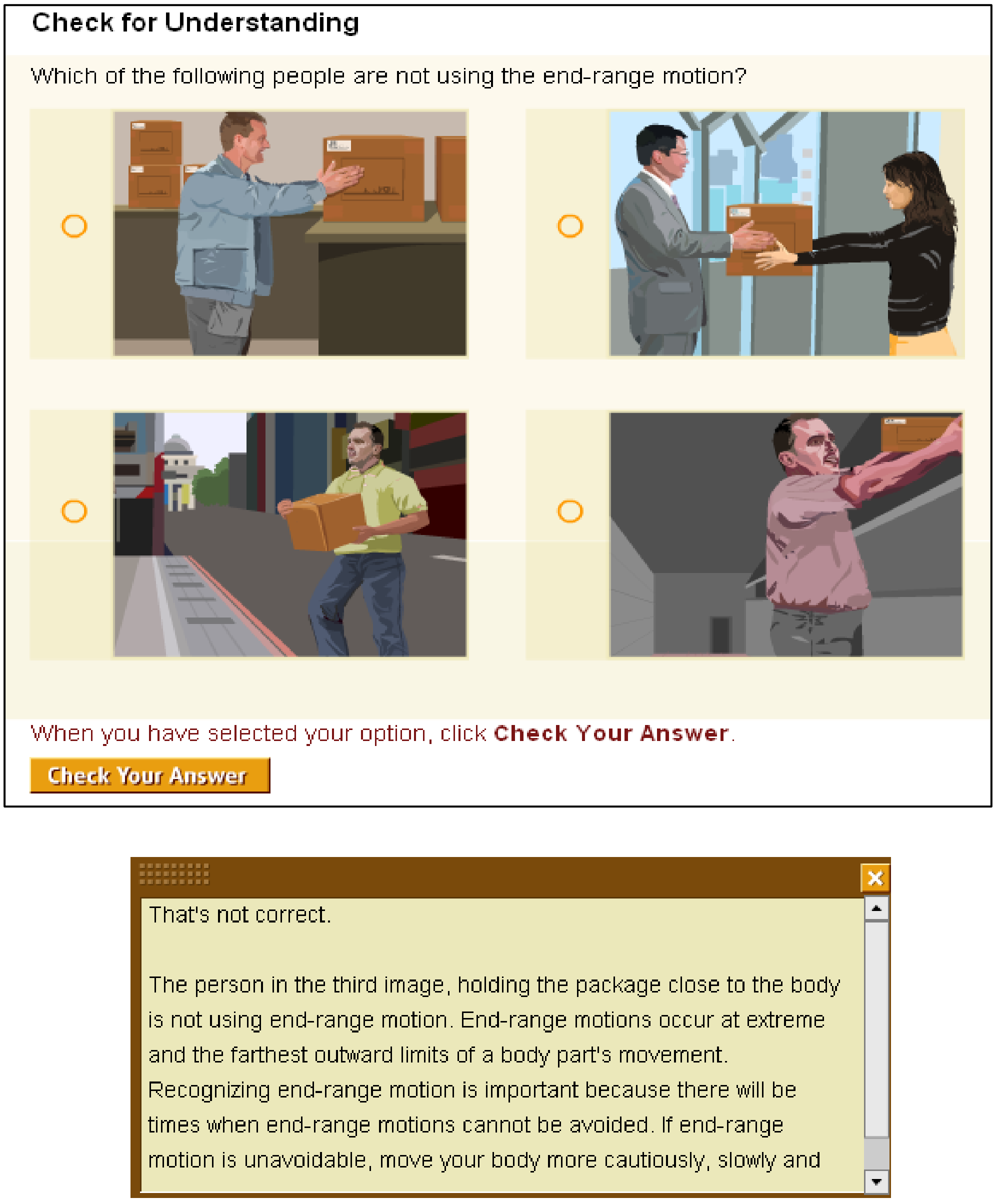


Figure 8. Assessment and feedback example

The question for this image asked "Which of the following people are not using the End-Range Motion?" The trainee had to evaluate whether the people in the images are using End-Range Motion and to select an image of a person who is not using End-Range Motion. Table 7 lists the assessment type associated with each learning objective.

Table 7. Learning objectives and type of assessment

The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual						
B. Conceptual						
C. Procedural	L5(*D*) L7(*D*), L9(*A*),L10(*D*), L14(*E*),L15(*D*), L16(*D*),L17(*E*) (Total :8)	P4(*C*), P6(*A*), P7(*A*), P8(*A*), P9(*A*), L4(*A*), L6(*E*), L11(*B*) (Total: 8)	K1(*G*),K2(*G*), K3(*G*),K4(*G*), K5(*G*),K6(*G*), K7(*G*),K8(*G*), K9(*G*),K10(*G*) (Total: 10)		P1(*A*),P2(*A*), P5(*A*),L8(*F*), L12(*B*),13(*B*) (Total: 6)	
D. Metacognitive						
Assessment A: Multiple choice questions *Assessment B*: Multiple response questions *Assessment C*: Drag and drop (positioning image) *Assessment D*: True / False questions			*Assessment E*: Short answer *Assessment F:* Ordering *Assessment G*: Observation with checklist			

*( ): total number of learning objectives

3.1.4. Gathering and classifying assessment outcomes

Assessment scores were gathered and classified according to the Revised Bloom's Taxonomy after the completion of the training program. The classification of the assessment scores was based on the classification of the learning objectives because the assessments were designed to measure the intended knowledge or skills that were expected from each learning objective.

3.1.5. Independent (predictor) variables

The independent variables were the sum of the assessment scores categorized under each cognitive process in the taxonomy table (Table 7). Therefore, there were a total of four independent variables:

- IV1: Sum of all assessment scores which were categorized as *Remember* and *Procedural* (remember-procedural)
- IV2: Sum of all assessment scores which were categorized as *Understand* and *Procedural* (understand-procedural).

- IV3: Sum of all assessment scores which were categorized as *Apply* and *Procedural* (apply-procedural).
- IV4: Sum of all assessment scores which were categorized as *Evaluate* and *Procedural* (evaluate-procedural).

Table 8. Independent variables

The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual						
B. Conceptual						
C. Procedural	L5, L7, L9, L10, L14, L15, L16, L17 (IV1)	P4, P6, P7, P8, P9, L4, L6, L11 (IV2)	K1, K2, K3, K4, K5, K6, K7, K8, K9, K10 (IV3)		P1, P2, P5, L8, L12, L13 (IV4)	
D. Metacognitive						

3.1.6. Dependent (response) variables

Training performance checklist scores

Training performance was evaluated based on the observation of participants' demonstrations, and scores were then calculated based on the checklist (Table 9) results. The checklist consists of ten items, which were mainly the key concepts that were emphasized during the training program. To increase the sensitivity of the checklist, each item was divided into six different phases. The column differentiated three phases of the task: lifting, carrying, and lowering. The first row within each item was to evaluate the participants while they carried the package to the designated location. Likewise, the second row was to evaluate the participants while they carried the package back to the original place. The gray areas in the checklist were phases where the items could not be utilized due to the nature of the task. For example, participants did not have to test the object for weight and shifting contents (#6 from the checklist) while they were carrying or lowering a package.

Table 9. Checklist for training performance assessment

No.	Does the participant do the following?	Lifting		Carrying		Lowering	
1.	Power Zone	Y	N	Y	N	Y	N
		Y	N	Y	N	Y	N
2.	End-Range Motion	Y	N	Y	N	Y	N
		Y	N	Y	N	Y	N
3.	Proper foot position	Y	N	-	-	Y	N
		Y	N	-	-	Y	N
4.	Keep the natural curve of the back	Y	N	Y	N	Y	N
		Y	N	Y	N	Y	N
5.	Bend at the knees	Y	N	-	-	Y	N
		Y	N	-	-	Y	N
6.	Test object for weight and shifting contents	Y	N	-	-	-	-
		Y	N	-	-	-	-
7.	Get a firm grip and grasp opposite corners	Y	N	Y	N	Y	N
		Y	N	Y	N	Y	N
8.	Lift / lower with a smooth, steady motion; don't jerk	Y	N	-	-	Y	N
		Y	N	-	-	Y	N
9.	Move your feet. Step or pivot; don't twist	Y	N	Y	N	Y	N
		Y	N	Y	N	Y	N
10.	Properly distributing the load	Y	N	Y	N	Y	N
		Y	N	Y	N	Y	N

Some of the checklist items had specific criteria as described in the following table.

Table 10. Checklist item criteria

No.	Item	Criteria
1.	Power Zone	The Power Zone is from mid-thigh to the arm pits.
2.	End-Range Motion	Holding package away from their body with arms extended and elbows locked.
3.	Proper foot position	Feet should be approximately shoulder-width apart, with one foot slightly ahead of the other.
4.	Keep the natural curve of the back	-
5.	Bend at the knees	Bending knees instead of bending waist when lifting / lowering.
6.	Test object for weight and shifting contents	Tilting or sliding objects to test the weight or shifting contents
7.	Get a firm grip and grasp opposite corners	Getting firm grip and grasping opposite corners
8.	Lift / lower with a smooth, steady motion; don't jerk	-
9.	Move your feet. Step or pivot; don't twist	-
10.	Properly distributing the load	Using the environment or objects to distribute the load while lifting / lowering

<u>NIOSH lifting index</u>

The NIOSH lifting index was used as a second method to evaluate training performance. The lifting index is used for the design and evaluation of lifting and lowering tasks and is computed by examining the ratio of actual weight divided by recommended weight limit (RWL). The RWL was computed by using the NIOSH lifting equation, which included six multipliers. The six multipliers were: (1) Horizontal multiplier, (2) vertical multiplier, (3) distance multiplier, (4) asymmetric multiplier, (5) coupling multiplier, and (6) frequency multiplier. To calculate each multiplier, the participants' demonstrations for the training performance assessment phase were recorded and then analyzed later.

3.2. Trainee reactions

Training reactions were collected by sending e-mail attachments with reaction questionnaires to the participants after their participation in the training program. There were a total of 19 questions, which were categorized into two types, affective and utility type questions. The questionnaire items were selected from previous studies, which examined trainee reactions (Morgan & Casper, 2000; Ruona *et al.*, 2002; Tan, Hall, & Boyce, 2003).

Table 11. Affective type reaction questionnaire [adopted from (Morgan & Casper, 2000, pp. 309-311)]

No.	Questions
1	How satisfied are you with quality of this training course overall?
2	How satisfied are you with the quality of course materials?
3	How satisfied are you with exam coverage and importance of material tested?
4	How satisfied are you with feedback you received as result of course testing?
5	How satisfied are you with match of course objectives with your idea of what would be taught?
6	How satisfied are you with the relevance of the course content to your job?
7	How satisfied are you with course's emphasis on most important information?
8	How satisfied are you with the extent to which the course prepared you to perform current job tasks more effectively?
9	How satisfied are you with the length of training course?
10	How satisfied are you with the pace of the course material presented?

As shown in table 11, there were ten affective type questions. Affective type questions mainly asked about participants' satisfaction with the training program. Five point Likert scales were used to indicate the level of agreement (5=strongly agree to 1=strongly disagree). There was a total of nine utility type questions, which asked about the usefulness of the training program (Table 12) and five point Likert scales were used.

Table 12. Utility type reaction questionnaire [adopted from (Ruona *et al.*, 2002, p. 222; Tan, Hall, & Boyce, 2003, p. 403)

No.	**Questions**
11	I have an overall good feeling about how the training program was carried out
12	The training program allowed me to develop specific skills that I can use on the job
13	The training program was, overall, very effective
14	The training program is very useful
15	My time was well spent
16	The session objectives were met
17	I would recommend this program to others who have the opportunity
18	I learned something I can apply immediately to my work
19	The course provided me with new ways of thinking about my job

Participants' responses were collected, and the sum of the responses on the Likert scale for each type of question was calculated for use in the analyses.

3.3. Participants

The participants for this study were students recruited by the SONA system from the Psychology Department at Virginia Tech. A total of 22 participants (11 females and 11 males) were recruited. The mean age of the participants was 19.77 years old (*SD* = 1.27 years). The mean height and weight were 173.99 cm (*SD* = 11.02) and 164.91 kg (*SD* = 45.03) respectively. The ethnicity of the participants included eighteen European-Americans, one Asian-American, two Hispanic-Americans, and one Native American.

The participants were limited to individuals who had (1) no working experience in a delivery or service industry related to lifting and lowering tasks, (2) no prior experience with training or lectures related to lifting and lowering, (3) basic computer proficiency, (4) no visual or hearing disabilities, and (5) no past injuries experience related to lifting and lowering. Two participants reported having work experience in the service industry (one year and 2 months, and 2 months, respectively). However, their work experience was not related to lifting and lowering. Six participants reported having training experiences. One was baseball training, four were weight training, and the other was emergency medical training (EMT). However, none of their training experience was related to lifting and lowering training. All of the participants had basic computer proficiency. Although two participants reported having visual disability because they wore corrective lenses, that condition did not make them ineligible. One participant reported a past injury. For this participant, the researcher thoroughly explained about the physical tasks involved in this study, and emphasized safety issues to be the first priority. The researcher also reminded participants that they were allowed to withdraw from the study at any time without any penalty. No participants withdrew the study, and all completed the entire study without any problems.

All of the participants received two extra credits in their coursework for their participation in the experiment. To collect participants' reaction to the training program, reaction questionnaires were sent via e-mail. There were two reasons for not collecting the reaction questionnaires at the end of the training program. First, the reaction questionnaires were submitted to the Virginia Tech Institutional Review Board (IRB) after the initiation of the experiment. Second, it would have extended the training session beyond the promised time for completion. For these reasons, responses to the reaction questionnaires were collected a few days after the end of the training program. A total of eighteen participants out of twenty-two in the original sample responded to the reaction questionnaires.

3.4. Procedure

This experiment was conducted in the Virginia Tech Safety Engineering Laboratory. The average time required to complete the experiment was about 90 minutes. The experiment included three major sections: (1) introduction, (2) training, and (3) training performance assessment. The overall experimental procedure is depicted on Figure 9 with major sections, subsections, and their estimated time for completion is noted.

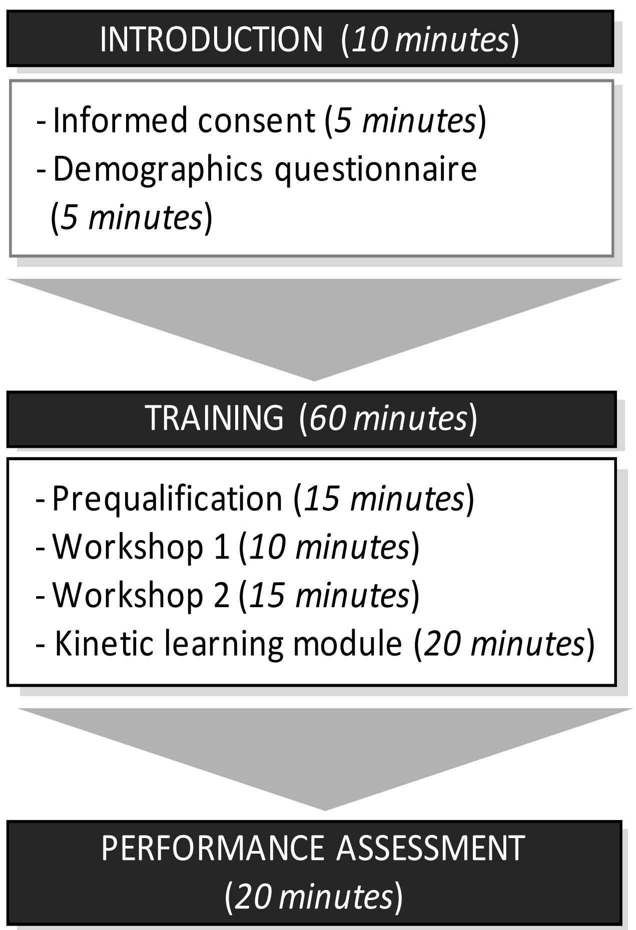


Figure 9. Experimental procedure

3.4.1. Introduction

Prior to training, all participants were required to read the informed consent form (Appendix H) and listen to an explanation provided by the researcher. Those who agreed to participate in this study signed the form and completed the demographics questionnaire (Appendix A). The demographics questionnaire included general background information, such as gender, ethnicity, age, height, and weight along with past work or training experience related to lifting and lowering. Participants' years of work experiences related to lifting and lowering packages were checked to reduce the bias that previous experience would have on the study results. Ree *et al.* (1995), in their longitudinal study, concluded that experience had a strong causal relationship to training performance. Participants were asked about their prior knowledge of lifting and lowering from lectures or training. There were additional inquiries about the presence of visual and hearing disabilities as well as computer proficiency. This study excluded participants who had visual and hearing disabilities and required participants to have a minimum level of computer proficiency, such as the ability to use a mouse to navigate through the WBT.

3.4.2. Training

The training program included four different phases; Prequalification, Workshop 1, Workshop 2, and Kinetic Learning Module (KLM). As mentioned in the background section of the introduction, the training contents and materials were originally designed for a commercial delivery company. For the Prequalification phase, verbal and written instructions as well as the Prequalification self-check score sheet (Appendix B) were provided to the participants. The score sheet was utilized as the method to collect the participants' assessment scores. Specifically in the Prequalification phase, there were a total of eight assessment questions. Since the WBT utilized just-in-time feedback, the participants could determine whether their answers were correct or not. Therefore, after the feedback, the participants were instructed to mark on the Prequalification self-check score sheet whether they got the question correct.

After completing the Prequalification phase, Workshop 1 was introduced to the participants. Workshop 1 was mainly a summary of materials introduced in the Prequalification phase, and there was no assessment during this phase. Workshop 2 included more detailed concepts and complex scenario-based examples. Verbal and written instructions as well as the Workshop 2 Self-check Answer Sheet (Appendix C) were provided to the participants. There were fourteen assessment questions, and the participants were instructed to write the answers on the answer sheet.

After completing Workshop 2, the researcher introduced the Kinetic Learning Module (KLM), the last phase of the training program. The KLM utilized the WBT as an introduction, but the main focus was after the WBT, when the participants demonstrated simple lifting and lowering tasks. Peer-assisted learning (PAL) was utilized during hands-on activities and participants alternatively observed each other's demonstrations to complete the KLM Peer Review Checklist (Appendix E). The checklist included 10 items, which were mainly the key concepts from the training material. Considering that the participants were not experts, images to guide their decisions were included for some items. During the activities, the researcher recorded participants' demonstrations for peer evaluation. After the demonstrations, the participants watched the recorded video of their demonstrations and provided feedback to each other. The following table shows the overall procedure of the KLM.

Table 13. Overall KLM procedure

	PARTICIPANT A	PARTICIPANT B	RESEARCHER
WBT	Introduction to the KLM	Introduction to the KLM	Introduce the KLM Handout the KLM Peer Review Checklist
HANDS-ON ACTIVITY	Perform KLM task Observe and complete the checklist	Observe and complete the checklist Perform KLM task	Record participants' demonstrations
FEEDBACK	Watch their demonstrations		Play the video and guide the discussion

For the KLM task, three packages with numbers labeled from 1 through 3 were prepared. Two packages (1 and 3) were on the shelf and the other package (2) was on the ground as illustrated in the following figure.

Figure 10. KLM laboratory environment

Participants were instructed to carry the packages to the designated location, which was about 16ft (4.87m) away from the shelf, and then, were asked to carry them back to their original places. The following table reproduces the written instructions about the KLM task scenario that were provided to the participants.

Table 14. KLM task scenario

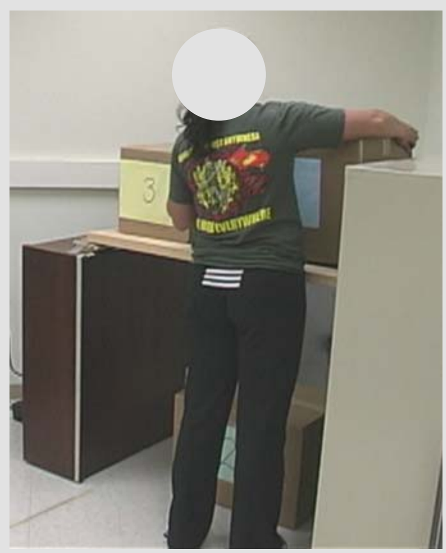
Task scenario #1 1. Get packages that are labeled as “1”, “2”, and “3”. 2. Place them at the designated location “C”. 3. Re-place each package at its original place.

The following diagram illustrates a participant performing the KLM task scenario.

1. INITIAL STATE

Packages 1 and 3 are on the shelf, which are 3.22 ft (98.5 cm) above the ground, and package 2 is on the ground.

2. LIFTING



Lift the package one by one

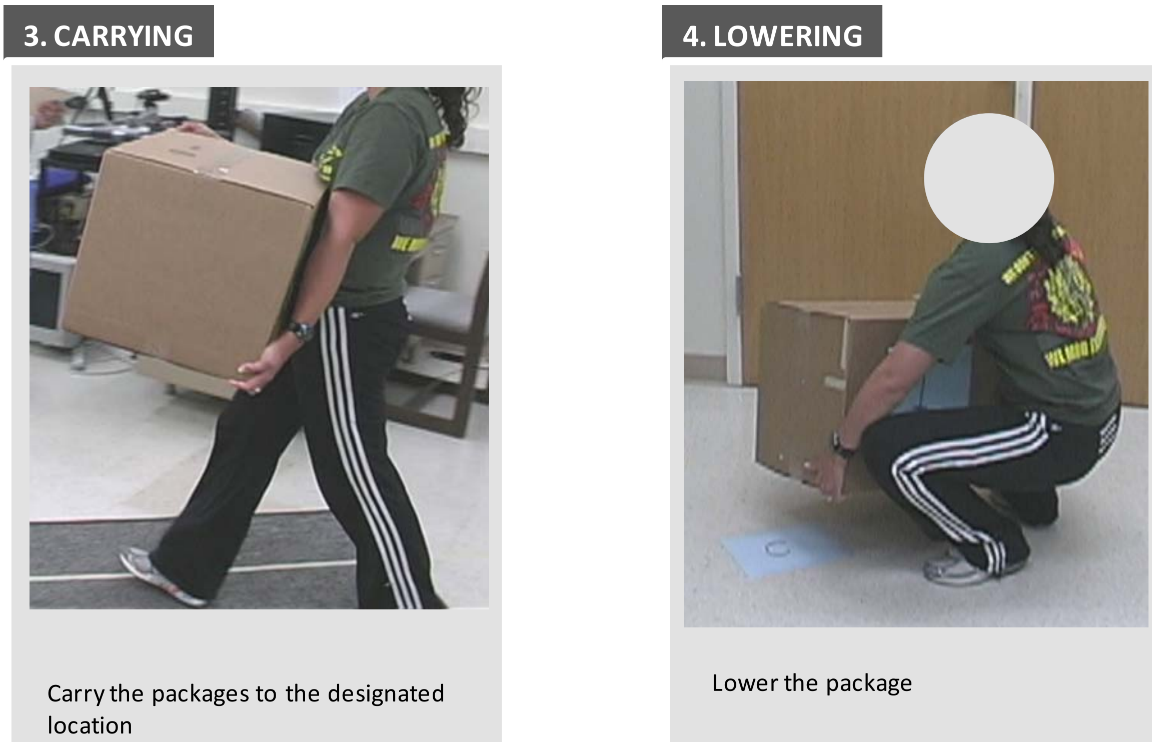


Figure 11. KLM demonstration

After the demonstrations, the researcher showed the recorded video to the participants. The participants sat in front of a television and provided feedback about each other's performance based on the checklist that they had completed while observing their peers. Figure 12 illustrates the peer-enabled learning environment. Open discussions were encouraged by the researcher. However, if both peers did not engage in discussion, the researcher asked the following questions from the discussion guideline to provoke discussion.

Table 15. Discussion guideline

Time for Feedback!
1. Do you think you used good methods to perform the activity?
2. Did your peer use good methods to perform the activity?
3. Point out places you think your peer could improve his / her demonstration.

Figure 12. Peer evaluation environment

3.4.3. Training performance assessment

After the training, participants were introduced to the training performance assessment phase. For training performance assessment, four digital cameras were used to record participants lifting and lowering tasks. Specifically, one camera was fixed at the original location, two cameras were fixed at the two other destination locations, and one camera was carried by the researcher. The following figure illustrates the schematic view of the training performance environment.

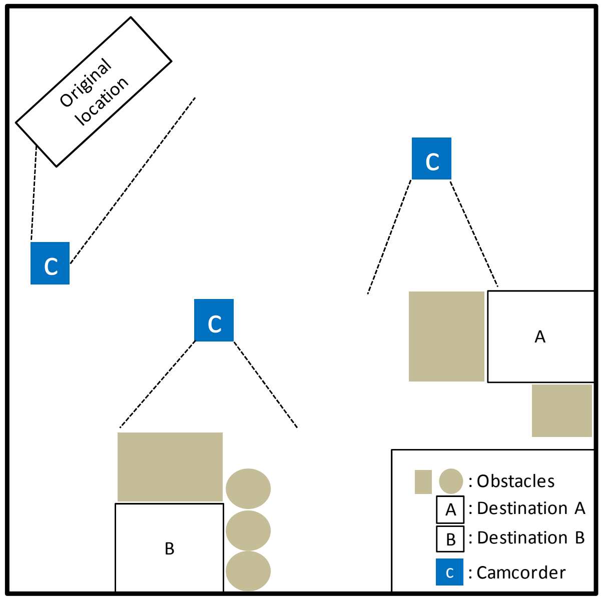


Figure 13. Schematic view of the training performance laboratory environment

The task for the performance assessment slightly differed from the KLM task scenario by implementing several risk factors. First, there were obstacles at the designated locations (Figure 14) that increased the possibility of not using the Power Zone and (or) not avoiding the End Range Motion.

(a) (b)

Figure 14. (a) Destination A, (b) Destination B

Second, white-tape attached on the ground represented a narrow path, causing participants to walk cautiously while carrying the package. Last, a wooden box was placed in the middle of the pathway to represent stairs. Therefore, the participants had to step up onto the wooden box and step down from it while carrying a package.

Figure 15. Training performance laboratory setup

The following table shows the written instructions for the training performance tasks.

Table 16. Instructions for training performance

Participant 1: 1. Lift package 1, which is on the shelf and carry it to the designated location “A”. 2. Lift the same package from “A”, and carry it back to its original location and place it on the ground underneath the shelf. **Participant 2:** 1. Lift package 1, which is on the ground and carry it to the designated location “B”. 2. Lift the same package from “B”, and carry it back to its original location and place it on the shelf.

Participants were required to perform slightly different tasks to prevent the likelihood of the first participant influencing the next participant's performance. Additionally, during the first participant's demonstration, the second participant was not allowed to watch the demonstration. The following diagram illustrates the performance assessment task for both tasks. The first participant was asked to lift package #1, which was on the shelf, and carry it to the designated location A, then carry it back to the original location and place it on the ground under the shelf. The second participant was required to lift package #1, which was on the ground, and carry it to designated location B, then carry it back to the original location and place it on the shelf. These demonstrations are illustrated on the following diagram.

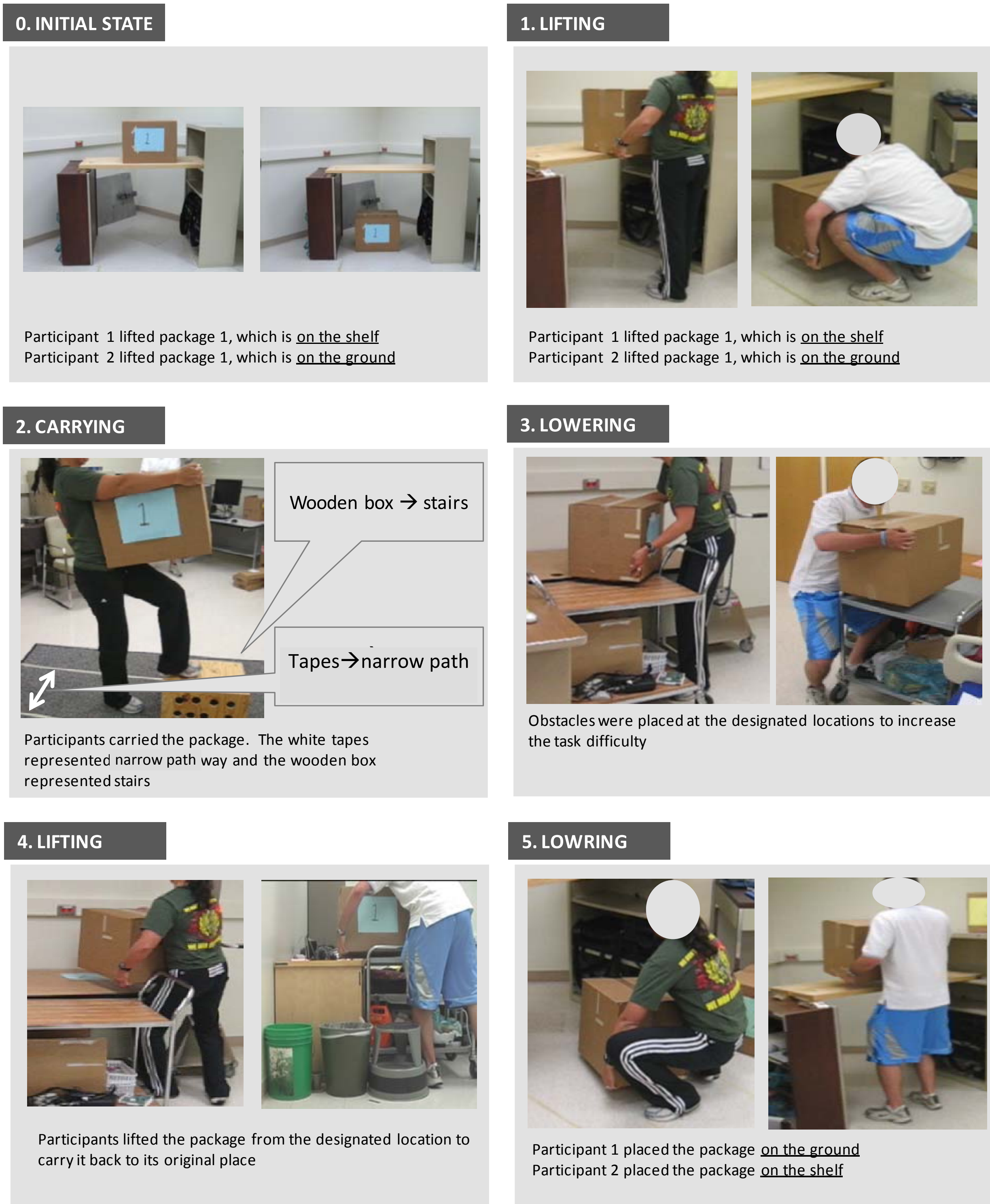


Figure 16. Training performance demonstration

3.5. Equipment and Apparatus

3.5.1. Laboratory set-up

The study was conducted in the Safety Engineering Laboratory in the Human Factors Engineering and Ergonomics Center at Virginia Tech. Since two participants were paired to be trained and assessed throughout the experiment, two laptops were utilized with all the training materials pre-installed. There were four video camcorders to record participants' demonstrations during the experiment. For the KLM task and the training performance assessment, three packages were prepared.

Table 17. Package information

Package #	Weight (lb)	Dimensions (inch)
Package 1	33 (15kg)	20*15*15
Package 2	25 (11.3kg)	20*15*15
Package 3	33 (15kg)	12*12*12

Packages 1 and 3 weighed the same amount, but differed in dimensions. A maximum weight of 32 lb (14.5 kg) is considered to be safe for more than 90% of female workers (Waters *et al.*, 1993).

3.5.2. WBT computer training software

In this study, the participants were trained using a beta version of a proprietary software, which was developed for a major delivery company. These WBT programs were utilized during the Prequalification and Workshop phases of the training program. Since this software included instructional materials on topics other than lifting and lowering, the participants were instructed to skip the topics that were not related to lifting and lowering, with the guidance of the researcher.

3.5.3. ImageJ (version 1.38x)

ImageJ is a Java based image processing program that was developed at the National Institutes of Health. This software was used in this study to measure the horizontal multiplier to calculate the NIOSH lifting index. The advantage of ImageJ is that pixels/cm can be calculated if there is any known distance in an image. As shown in Figure 17, the width of the package was calculated in pixels (390.74 pixels) by drawing a line using one of the functions that was provided by the program (#1 in the image). By inserting the actual package width (39cm) into the "known distance" dialogue box, the pixels/cm scale was calculated (10.019 pixels/cm). Next, again using the line drawing function, the horizontal distance was measured in pixels (446.31 pixels, #2 in the image). Since the pixels/cm is known, the actual horizontal distance can be calculated.

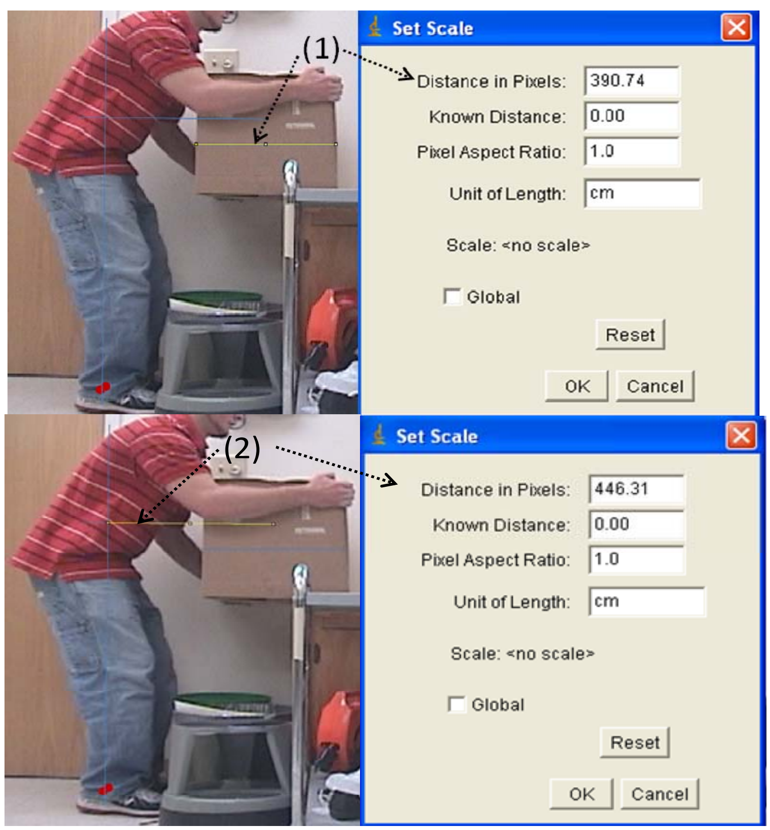


Figure 17. Measuring horizontal distance using ImageJ

3.6. Data analysis

A statistics analysis program, JMP 7.0, was used for the data analyses. Forward stepwise regression, with .25 of probability to enter and .10 of probability to leave the model, was utilized in the regression analyses.

To test hypothesis one, four independent variables and two dependent variables were recorded. The independent variables were collected by using the self-check score sheets that were provided to the participants during the training program. The participants were instructed to mark on the score sheet whether they chose a correct answer for each assessment question. One point was given for correct answers and no point was given for incorrect answers. The assessment scores were later classified into four different categories (remember-procedural, understand-procedural, apply-procedural, and evaluate-procedural) and the sum of assessment scores was calculated for each category.

Measurement of the dependent variables was collected by observing the recorded demonstrations. First, for the training performance assessment checklist scores, the researcher observed participants' demonstrations and completed the checklist based on the criteria that was developed for each checklist item. Then, total numbers of "yes" responses were considered as the total checklist score. To calculate the NIOSH lifting index, which was the second dependent variable, the recorded videos of participants' demonstrations were analyzed. The video frames of lifting and lowering postures were selected and were saved as image files. The horizontal distances were calculated using ImageJ. The vertical and distance multipliers were calculated based on the type of task in which the participant performed. Asymmetric multipliers were calculated by observing the recorded performances. Coupling multipliers were all set as fair, because the packages did not have handles but were all regular sized and in good condition.

Frequency multipliers were calculated based on the numbers of lifts per minute and the duration of the task.

Then, the independent variables were tested for their ability to predict the dependent variables using multiple regressions. Additionally, interaction variables were included in the regression model to consider possible interaction effects among the independent variables. The multiple linear regression equation is shown below:

$$Y_{ij} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_0 X_{3i} + \beta_0 X_{4i} + \beta_0 X_{12i} + \beta_0 X_{13i} + \beta_0 X_{14i} + \beta_0 X_{23i} + \beta_0 X_{24i} + \beta_0 X_{34i} + \varepsilon_i$$

Where:

i = participant number (i = 1,2,…,22)

j = training performance assessment method (j = 1, 2)

Y_{i1} = Training performance assessment checklist score.

Y_{i2} = Training performance assessment NIOSH lifting index.

X_{1i} = Sum of all assessment scores that were categorized as remember and procedural.

X_{2i} = Sum of all assessment scores that were categorized as understand and procedural.

X_{3i} = Sum of all assessment scores that were categorized as apply and procedural.

X_{4i} = Sum of all assessment scores that were categorized as evaluate and procedural.

X_{12i} = Interaction variable of remember-procedural and understand-procedural

X_{13i} = Interaction variable of remember-procedural and apply-procedural

X_{14i} = Interaction variable of remember-procedural and evaluate-procedural

X_{23i} = Interaction variable of understand-procedural and apply-procedural

X_{24i} = Interaction variable of understand-procedural and evaluate-procedural

X_{34i} = Interaction variable of apply-procedural and evaluate-procedural

ε_i = error term

To test the second hypothesis, the independent variables that were identified during the analysis of the first research question were included along with the variables for the affective and utility type reaction questionnaires in a second regression equation to predict training performance. The sum for the affective and utility type reaction questionnaires were calculated from the survey results.

CHAPTER 4. RESULTS

As stated in chapter 1, the goals of this study were to answer the following research questions:

1. What types of learning outcomes are most predictive of training performance?

2. How do the learning outcomes predict training performance compared to affective and utility type reactions?

4.1. Learning outcomes (research question 1)

Multiple linear regression analyses were conducted to develop a model to predict training performance by different types of learning outcomes (remember-procedural, understand-procedural, apply-procedural, and evaluate-procedural) that were classified by the Revised Bloom's Taxonomy. The following sections will describe the results regarding the checklist scores followed by the results regarding the NIOSH lifting index. Basic descriptive statistics of the total assessment scores for each learning outcome are presented in the following table.

Table 18. Summary statistics of independent & dependent variables

Variables	Description	*MAX*	*MIN*	*M*	*SD*
Remember-procedural	Sum of all assessment scores that were categorized as remember-procedural.	8	6	7.50 (8)	.74
Understand-procedural	Sum of all assessment scores that were categorized as understand-procedural.	8	4	6.27 (8)	1.32
Apply-procedural	Sum of all assessment scores that were categorized as apply-procedural.	50	31	44.32 (50)	4.68
Evaluate-procedural	Sum of all assessment scores that were categorized as evaluate-procedural.	5	2	3.73 (6)	.77
Checklist score	Total checklist scores from observation	49	34	44.36 (50)	3.77
NIOSH LI	The lifting index calculated by NIOSH Lifting Equation	3.4	.83	1.59	.54

n = 22 *( ): total possible points

4.1.1.Training performance checklist score

The predictors were first regressed onto the training performance scores that were collected by observation using the checklist. The checklist scores had a range from 34 to 49 with a mean of 44.36 (SD = 3.77). The assumptions for conducting regression analysis were met. The normality of the residuals were assured by the Shapiro-Wilk W test ($W = .94$, $p = .27$). The linearity of the model was checked according to the residual by predicted plot as illustrated in the following figure.

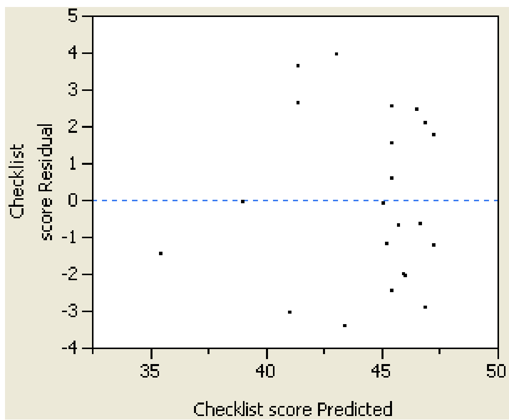


Figure 18. Residual by predicted plot

Additionally, variance inflation factors (VIF) were calculated to test the multicollinearity among independent variables. The VIF values ranged from 1.02 to 1.12, indicating no multicolinearity threat.

In step 1 of the regression analysis, the learning outcomes that were classified as apply-procedural (IV#3) were entered into the model. The model was significant ($p \leq .001$) with adjusted R^2 of .57, indicating that almost 60% of the variance in the training performance checklist scores was explained by apply-procedural.

Table 19. Results of regression analyses for training performance (checklist score)

Variable	*β*	R^2	Adj. R^2	*F-ratio*	*Cp*	*p*
Step 1		.59	.57	28.93***	1.84	2
Apply-procedural	.62***					
Step 2		.72	.66	11.11***	1.43	5
Apply-procedural	.55***					
Understand*Evaluate	1.62					

***$p \leq .001$, $n = 22$

Step 2 of the regression analysis added the interaction variable of understand-procedural and evaluate-procedural (understand*evaluate) to the model. However, this interaction variable was not a significant predictor of training performance checklist score. The step 2 model remained significant ($p \leq .001$) with the adjusted R^2 of .66 and apply-procedural remained as a significant predictor.

As a result, the step 1 regression model was chosen to predict the training performance checklist score based on the comparison between *Cp* and *p* value (Mallows, 2000). The regression equation is as shown below.

$$Y_i = 16.91 + 0.62\ X_{3i}$$

Where:

i = participant number

X_{3i} = sum of all assessment scores that were categorized as apply-procedural

4.1.2. NIOSH lifting index

The lifting index was calculated using the NIOSH Lifting Equation as an alternative method to evaluate the training performance. The lifting index varied from .83 to 3.4 with a mean of 1.59 (SD = .54). The normality was assured by the Shpairo-Wilk W test (W = .95, p

= .28). The linearity assumption was met by checking the residual by predicted plot as shown in Figure 19.

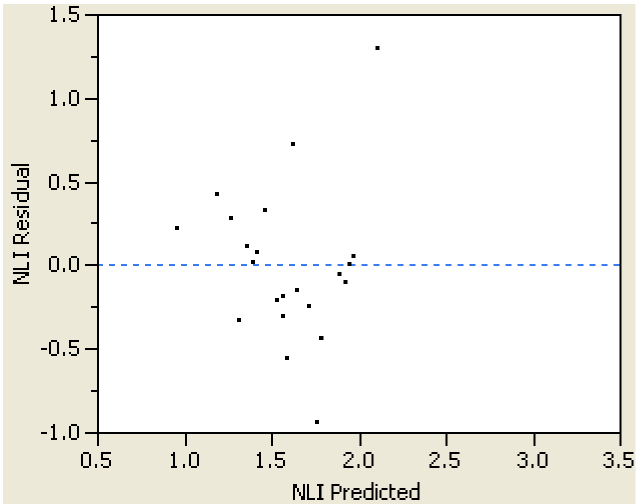


Figure 19. Residual by predicted plot

The regression analysis revealed that interaction variable of understand-procedural and apply-procedural (understand*apply) was the most predictive of the NIOSH lifting index. The model was significant with adjusted R^2 of .36, which means that about 36% of the variance in training performance measured by the NIOSH lifting index can be explained by the interaction between understand-procedural and apply-procedural.

Table 20. Results of regression analysis for training performance (NIOSH lifting index)

Variable	***β***	R^2	*Adj.* R^2	***F-ratio***	***Cp***	***p***
Step 1		.45	.36	4.96**	8.09	4
Understand*apply	.06**					
Step 2		.56	.46	5.46**	5.7	5
Understand*apply	.08**					
Remember-procedural	.26					

** $p \leq .05$, $n = 22$

Step 2 of the regression analysis added remember-procedural (IV#1) to the model, but the predictor was not statistically significant to predict the NIOSH lifting index. The adjusted R^2 increased to .46, which means that 46% of the variance can be explained by the model. The step 1 regression model was chosen to predict the training performance measured by the NIOSH lifting index, because remember-procedural (IV#1) was not a significant variable to predict the NIOSH lifting index. The regression equation is as shown below.

$$Y = 6.07 - 0.19 * X_{2i} - 0.07 * X_{3i} + 0.06 X_{23i}$$

Where:

i = participant number

X_2 = Sum of all assessment scores that were categorized as understand and procedural

X_3 = Sum of all assessment scores that were categorized as apply and procedural

X_{23} = Interaction effects of understand*apply

Based on the results, the variable apply-procedural and the interaction variable understand-procedural and apply-procedural (understand*apply) were the significant predictors of training performance. Therefore, these two predictors were compared with the affective and utility type reactions for their ability to predict the training performance (research question 2).

4.2. Learning outcomes and reaction questionnaires (research question 2)

Multiple linear regression analyses were conducted to develop a model to predict training performance based on reaction questionnaires and the learning outcomes (apply-procedural and understand*apply) previously found to be predictive of training performance in the analyses done for research question #1. The following table illustrates the descriptive statistics for the two types of reaction questionnaires.

Table 21. Summary statistics of reaction questionnaires

	Alpha	***MAX***	***MIN***	***M***	***SD***
Affective type reaction	.82	50.00	32.00	41.11	4.97
Utility type reaction	.92	44.00	23.00	35.44	6.23

n = 18

Overall, satisfaction level (affective type reaction) was slightly higher than the usefulness of the training program (utility type reaction). To check the internal consistency of the reaction questionnaires, Cronbach's alpha was calculated for the affective type questionnaire and the utility type questionnaire. The reliability coefficients were .82 and .92, respectively. These values for the reliability coefficient can be concluded as very reliable (Nunnaly, 1978).

4.2.1.Training performance checklist score

The regression analysis results (Table 22) indicated that apply-procedural was the most significant variable to explain the training performance checklist score. The model was significant with an adjusted R^2 of .53, which means that about 53% of the variance can be explained by the model.

Table 22. Results of regression analysis for training performance (checklist score)

Variable	***β***	R^2	*Adj.* R^2	***F-ratio***	***Cp***	***p***
Step 1		.56	.53	20.07***	-.23	2
Apply-procedural	.60***					

*** $p \leq .001$, $n = 18$

The regression equation is as shown below.

$Y_i = 16.97 + 0.60\ X_{3i}$

Where:

i = participant number

X_{3i} =Sum of all assessment scores categorized as apply-procedural

4.2.2. NIOSH lifting index

The regression analysis results (Table 23) indicated that understand-procedural was the most significant variable to explain the NIOSH lifting index. Step 1 of the regression analysis added understand-procedural variable to the model. The model was significant with an adjusted R^2 of .18, meaning that 18% of the variance can be explained by the model.

Table 23. Results of regression analysis for training performance (checklist score)

Variable	*β*	R^2	Adj. R^2	F ratio	Cp	p
Step 1		.22	.18	4.61*	5.89	2
Understand-procedural	-.24*					
Step 2		.42	.30	3.38*	5.66	4
Understand-procedural	-.27*					
Understand*apply	.064					

*$p \leq .05$, $n = 18$

Step 2 of the regression analysis added the interaction variable of understand-procedural and apply-procedural to the model. The model remained significant with adjusted R^2 of .30, yet the interaction variable (understand*apply) was not significant to explain the NIOSH lifting index. The regression equation is as follows.

$Y_i = 3.18 - 0.24\ X_{2i}$

Where:

i = participant number

X_{2i} = Sum of all assessment scores categorized as understand-procedural

Based on the results, the learning outcomes were more predictive of training performance than the affective and utility type reaction questionnaires. This supports the second hypothesis and answers the second research question.

4.2.3. Correlation analysis

To further explore the relationships among the predictors and training performances, Pearson's product-moment correlation analysis was conducted. First, the assessment scores of learning objectives that were classified as apply-procedural were included. This predictor was identified to be the most significant predictor from the first research question and, based on the Four-level Evaluation Model, it is categorized as a learning (level 2) criterion. Second, affective and utility type reactions were included. Affective and utility type reactions were the most widely utilized evaluation measurements, which are categorized in reactions (level 1) from the Four-level Evaluation Model. Lastly, training performance scores, which are checklist scores, and the NIOSH lifting index were included. The following table provides the summary of the correlation analysis.

Table 24. Summary of correlation analysis

Pearson Product-Moment Correlation Coefficient (*r*)					
	Apply-Procedural	Affective Type Reaction	Utility Type Reaction	Checklist Score	NIOSH lifting index
Apply-Procedural	------	.17	.18	.77**	-.32
Affective Type Reaction		------	.73**	.13	-.12
Utility Type Reaction			------	.05	-.13
Checklist Score				------	-.47*
NIOSH lifting index					------

** =$p < .001$, and * = $p < .05$, $n = 18$

The results indicated few significant relationships among the predictor variables and the variables that measured the training performance. Within reactions (level 1), affective type reactions and utility type reactions had a strong correlation [r (16) = .73, $p < .001$] as depicted in the following figure.

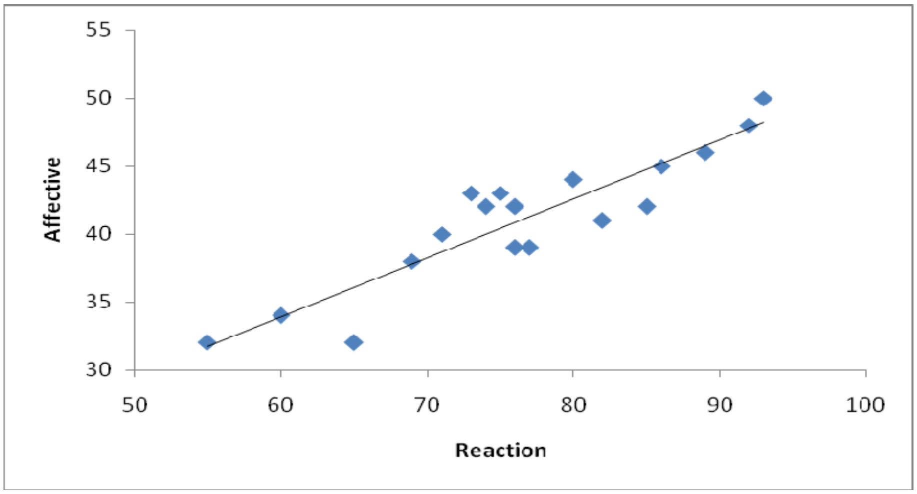


Figure 20. Scatter plot of affective and utility type reactions

However, neither of these variables showed significant correlations with training performance checklist scores or the NIOSH lifting index. The assessment scores categorized as apply-procedural was significantly correlated with the checklist scores [r (16) = .77, $p < .001$] but did not show any significant correlation with the NIOSH lifting index (Figure 21). As illustrated on Figure 22, the NIOSH lifting index was significantly correlated with the checklist scores [r (16) = -.47, $p < .05$].

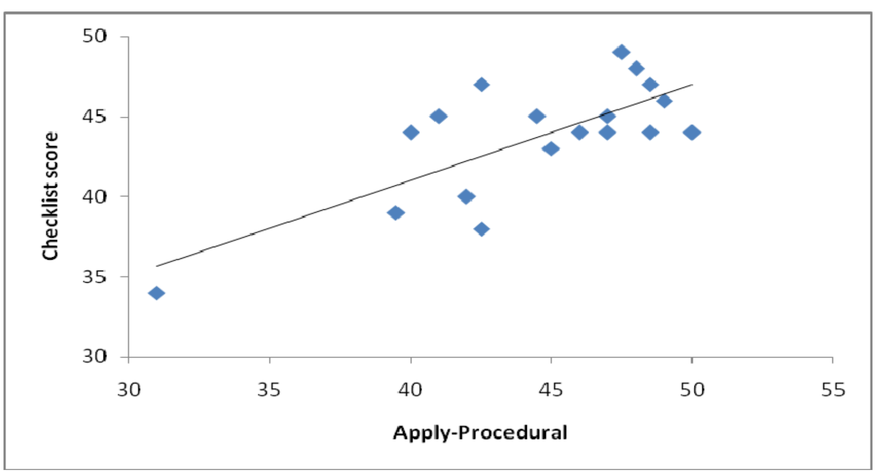


Figure 21. Scatter plot of checklist score and apply-procedural

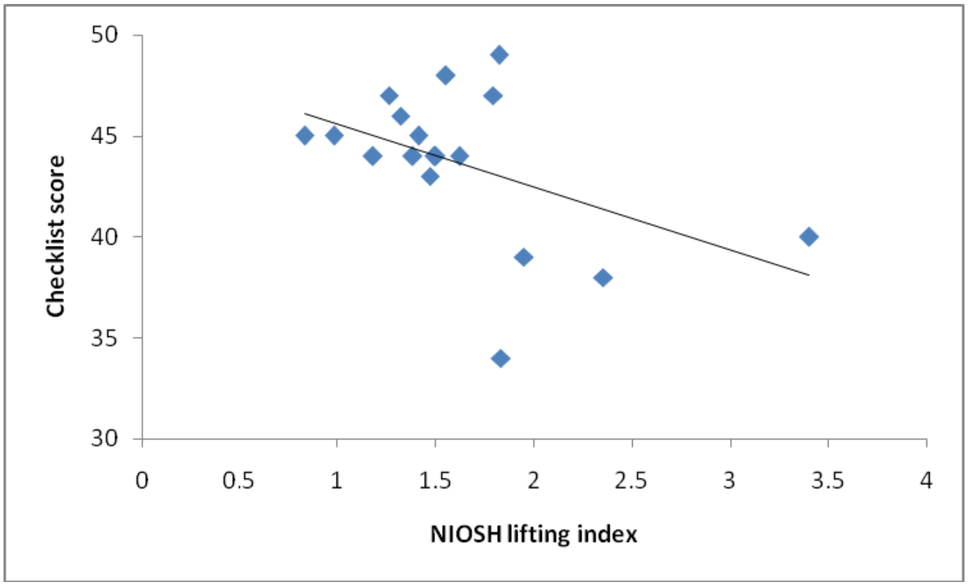


Figure 22. Scatter plot of checklist score and the NIOSH lifting index

CHAPTER 5. DISCUSSION

The goals of this study were to expand the learning, the second level of the Four-level Evaluation Model, using multidimensional categorization, and to determine the predictive validity of different types of learning outcomes on training performance in a lifting and lowering training program. To achieve these research purposes, the Revised Bloom's Taxonomy was employed to categorize learning outcomes, and two research questions were asked:

1. What types of learning outcomes are most predictive of training performance?

2. How do the learning outcomes predict training performance compared to affective and utility type reactions?

The discussion is organized into four sections: (1) interpretations of the results; (2) implications of the predictive learning outcome variables; (3) advantages of the Revised Bloom's Taxonomy to expand learning criteria; and (4) potential application of the Revised Bloom's Taxonomy to training programs.

5.1. Interpretation of the results

5.1.1. Learning outcomes (research question 1)

The first research question was examined by examining the relationship between predictor variables and training performance scores using multiple regression analyses. It was hypothesized that apply-procedural learning outcomes would be more predictive of training performance compared to other types of learning outcomes. The hypothesis was partially supported by the results such that apply-procedural was found to be the most significant predictor to explain training performance.

Training performance checklist scores

The assessment scores categorized as apply-procedural had a strong positive correlation with the training performance checklist scores. That is, participants who achieved higher scores during the KLM phase of the training also achieved higher scores for the training performance. This result can be explained by associating learning outcomes with their corresponding stages in the three stages of skill acquisition that were suggested by Anderson (1982) (Figure 20).

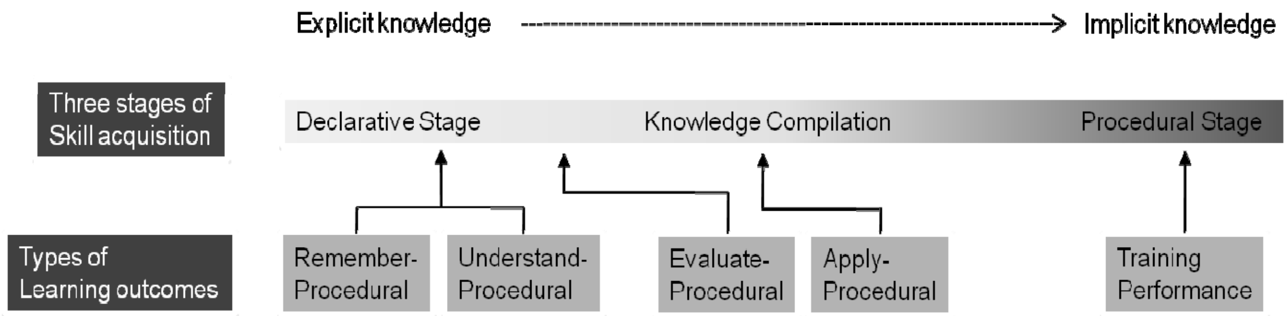


Figure 23. The three stages of skill acquisition and the learning outcomes

First, remember-procedural and understand-procedural learning outcomes are associated with the declarative stage in the three stages of skill acquisition. The assessments associated with remember-procedural and understand-procedural learning objectives appeared to reflect whether the participants remembered or understood specific knowledge. These assessments thus measured the level of knowledge or skill acquired during the declarative stage.

Second, evaluate-procedural learning outcomes are located between declarative and knowledge compilation stages. The assessments associated with evaluate-procedural learning objectives examined the level of trainees' judgment on the properness of lifting and lowering postures. To make correct judgments, different concepts related to proper lifting and lowering postures must be effectively synthesized, yet do not require actual demonstrations of proper lifting and lowering. That is, evaluate-procedural learning outcomes require a more enhanced level of explicit knowledge than is required for remember-procedural and understand-procedural. However, assessments regarding evaluate-procedural do not necessarily measure the actual demonstration of proper lifting and lowering. This result was supported by conducting post-hoc correlation analyses. A positive trend was found between evaluate-procedural and apply-procedural (p = .10), whereas, nonsignificant correlations were found between evaluate-procedural and remember-procedural (p = .71) and between evaluate-procedural and understand-procedural (p = .89).

Third, apply-procedural learning outcomes are associated with knowledge compilation stage. The assessments for apply-procedural learning objectives corresponded to participants' lifting and lowering performances, thus reflecting the level of acquired skills by practicing during the knowledge compilation stage.

Last, training performance was evaluated by measuring the level of lifting and lowering skills after the completion of the training program. That is, the participants were instructed to perform lifting and lowering tasks that were designed to be more difficult and complex than those in the training program so as to represent realistic work scenarios. Therefore, training performance assessment can be considered as an assessment of the procedural stage.

In summary, based on the three stages of skill acquisition, the assessments that are more closely associated with the procedural stage are more predictive of training performance. Again, this result was supported by the post-hoc correlation analyses, which revealed a significant correlation between apply-procedural and training performance ($p < .001$) and a positive trend between evaluate-procedural and training performance (p = .06). These results empirically supported the concept that declarative knowledge was not a valid predictor of behavior (Haccoun & Saks, 1998; Kraiger, Ford, & Salas, 1993).

NIOSH lifting index (understand-apply)

The interaction variable of understand-procedural and apply-procedural (understand*apply) was the most predictive variable of the NIOSH lifting index. A speculative interpretation for such result is that the level of understanding the training materials may have moderated the participants' ability to perform in a safe manner in the lifting and lowering phases during the training performance tasks. The training content involved realistic scenarios of the injuries that result from the improper handling of packages, and further elaborated the reasons for such injuries by explaining the basic anatomy of the human body in relation to lifting and lowering postures. This effort may have influenced the participants' risk awareness which may have further created incentive for participants to perform tasks safely. However, considering the difficulty of replicating interaction effects in the real world, the understand*apply variable will be excluded from the recommendations, which are described in the following sections.

It is interesting to note that the understand*apply variable was more predictive of the NIOSH lifting index than the apply-procedural variable. A possible explanation for this result is that the lifting indices were computed only at the lifting and the lowering phase of the training performance demonstration, whereas, the checklist assessed the overall demonstration. The checklist items are considered to be more contextually related to the lifting and lowering training program, since the checklist items were developed based on the instructional objectives of the training program. On the contrary, the NIOSH lifting guide does not represent all learning objectives that were emphasized throughout the training program. Representative examples of the training program's learning objectives are testing the object for weight and shifting contents, getting a firm grip and grasping opposite corners, lifting with a smooth and steady motion, and properly distributing the load. To improve training effectiveness at the organizational level, training outcomes should be linked to organizational outcomes (Kraiger, McLinden, & Casper, 2004). Therefore, the NIOSH lifting guide may serve as a valid indicator to measure risk level involved in general lifting and lowering tasks but may not reflect all of the learning outcomes that were emphasized in the lifting and lowering training program.

5.1.2. Learning outcomes and reaction questionnaires (research question 2)

The second research question was to determine how do the learning outcomes identified from research question one predict training performance compared to affective and utility type reactions. It was hypothesized that the learning outcomes would more strongly correlate with training performance than trainee reactions (affective and utility type). This hypothesis was supported by the results of the regression analyses.

Learning outcomes vs. trainee reactions

Learning outcomes categorized as apply-procedural were more predictive of training performance than affective and utility type trainee reactions. This result supports the findings of previous studies that questioned the validity of trainee reaction as an indicator of training effectiveness (Dixon, 1990; Holton III, 1996). These studies revealed weak correlations between reactions and learning (Noe & Schmitt, 1986a; Warr & Bunce, 1995a). Understand-procedural learning outcomes were a better predictive variable to explain the NIOSH lifting index than affective and utility type trainee reactions. Thus, participants who understood more of the training contents achieved lower NIOSH lifting index for their training performances. The level of understanding may have increased the level of risk awareness involved in lifting and lowering tasks and further lead to a safe performance.

5.1.3. Correlations among predictors

To further explore the relationships among the predictors and training performances, Pearson's product-moment correlation analyses were conducted.

Correlations involving apply-procedural

The results indicated a significant and strong correlation between apply-procedural and the training performance checklist scores. Additionally, no correlation was found between apply-procedural and the NIOSH lifting index (p = .2). The negative correlation between apply-procedural and the NIOSH lifting index relates to better performance being represented as a lower number on the NIOSH lifting index. That is, participants who achieved high scores during the KLM phase of the training were likely to have a low level of risk associated with their demonstrations.

Correlations involving affective and utility type reactions

Both affective and utility type reactions were not significantly correlated with the checklist scores and the NIOSH lifting index. This result differs from that in previous studies. Alliger and colleagues (1997), in their meta-analyses, revealed a modest but significant correlation between utility judgments and immediate learning. The difference in findings can be explained by examining the construct of a utility type reaction. A utility type reaction reflects the degree to which the trainees can apply the training content to actual work (Morgan & Casper, 2000). Participants' responses regarding utility reactions may be an underestimate based on the participants being college students. However, during the experiment, this problem was minimized by providing additional explanation of this issue to the participants. Although this conclusion does not fully support the hypothesis, it is important to note that a novel method of classifying the learning (level 2) evaluation criteria predicted training performance better than the current widely used criterion, trainee reaction. Lastly, the affective type reaction and utility type reaction both exhibited significant correlation with each other. This result supports a previous correlational study that examined the correlations within the reaction criterion (Alliger *et al.*, 1997).

Correlations involving the NIOSH lifting index

A trend toward a negative correlation was identified between the checklist scores and the NIOSH lifting index, which could indicate that two different measurements had a certain level of consistency in assessing training performance. As explained previously, the dimensions of the checklist items examined whether participants completed specific actions, which were derived from the learning objectives. The NIOSH lifting guide reflected more of the risk levels involved in the lifting or lowering activities. Therefore, it can be assumed that the training performance checklist items were verified as a valid assessment to examine risk levels involved in the lifting and lowering demonstrations. Interestingly, the NIOSH lifting index was negatively correlated with all of the other variables. That is, participants who achieved a higher score for apply-procedural assessment questions demonstrated a lower level of risk for injury in their performance of lifting and lowering a package. Likewise, participants who were satisfied with the training program and who felt that the training program was useful were likely to exhibit lower level of risk during their performance assessments.

In conclusion, the assessment scores of apply-procedural predicted training performance better than did trainee reactions. Considering that many organizations only utilize trainee reactions to predict different evaluation criteria, the findings from this study become more meaningful to industry.

5.2. Implications of the predictive learning outcome

The outcomes of this study may yield contributions that will be helpful to training practitioners and human resource managers in designing lifting and lowering training programs.

5.2.1. Guidelines to set minimum acceptable criteria

The learning objectives that were classified as apply-procedural would be the most important factor in setting minimum acceptable criteria that trainees must achieve in a lifting and lowering training program. The minimum acceptable criteria are commonly determined based on the policies in an organization and there are not many guidelines that explain how to set this minimum acceptable level. Especially when the training program concerns critical safety issues, setting the minimum acceptable criteria becomes important. This study demonstrated that application-related assessment scores were the most predictive of training performance scores. That is, participants who earned higher scores in the Kinetic Learning Module (KLM) phase also gained higher scores in training performance assessment. This finding could guide training designers to focus on application-related assessment scores to predict training performance. Likewise, minimum acceptable criteria could be set proportionate to the predictive powers of learning outcomes on training performance scores.

5.2.2. Predictions of transfer of training

The apply-procedural learning outcomes were found to be the best predictors of training performance in the lifting and lowering training program. This finding is meaningful in light of previous studies that found a positive correlation between training performance and transfer of training. Many researchers in the training discipline suggest that training performance has positive influences on transfer of training (Alliger *et al.*, 1997; Kirkpatrick, 1976; Kraiger, 2002). Kozlowski *et al.* (2001) empirically tested the relationships among multidimensional training outcomes and performance adaptability and concluded that training performance was the most predictive of performance adaptability. Further, Tannenbaum *et al.* (1991) found that performance during training was associated with post-training motivation. Based on the expectancy theory, more successful trainees would be more motivated to transfer their knowledge and skills to the work settings then less successful trainees (Vroom, 1964). Thus, training performance could be accepted as a potentially valid predictor of transfer of training.

5.2.3. Supports instructional design (training content design)

In a lifting and lowering training program, apply-procedural learning outcomes should be emphasized to ensure better training performance. For example, trainees should be more exposed to hands-on activities rather than memorizing concepts related to lifting and lowering. Therefore, training designers should assign higher priority to apply-procedural learning outcomes when scheduling a training program, so that sufficient time is ensured for the trainees to fully practice and acquire the expected skills.

5.3. Advantages of the Revised Bloom's Taxonomy to expand learning criteria

This study applied the Revised Bloom's Taxonomy to create subcategories in the level 2 learning of the Four-level Evaluation Model. Specifically, the four subcategories were remember-procedural, understand-procedural, apply-procedural, and evaluate-procedural. Compared to the original Four-level Evaluation Model, which does not have subcategories within the learning criterion, there are benefits to applying the Revised Bloom's Taxonomy to categorize learning outcomes as described below.

5.3.1. Considered multi-dimension of learning

The model expansion incorporates the multidimensional nature of learning. The significance of considering learning as multidimensional can be explained as follows. The multi-dimensional scheme enables a more detailed prediction of different measures in different levels of evaluation criteria. It is important to understand the relationships among different levels of evaluation criteria, because with that information, factors that may improve the effectiveness of a training program can be more easily understood and manipulated (Warr, Allan, & Birdi, 1999). Additionally, many studies have stressed the importance of the multidimensional aspect of learning (Bloom, 1956; Gagne, 1984; Krathwohl, Bloom, & Masia, 1964). However, there are only a few studies that have classified learning outcomes into subcategories. The results of this study have empirically revealed the advantages of multidimensional categorization by identifying the most predictive learning outcomes (apply-procedural) of training performance. Although the study scope is limited to the prediction among different learning outcomes of training performance, further research can examine the relationships among types of learning outcomes and different measures in different levels of evaluation criteria.

5.3.2. Theory-based classification

Theory-based classification may support understanding and interpreting various results involved in a training program. For example, there can be various reasons for a poor performance score. However, limiting this problem to a consideration of learning, a good explanation can be provided by the cognitive process dimension of the Revised Bloom's Taxonomy. Trainees have different cognitive abilities, and these differences may influence the task performance (Motowidlo, Walter, & Schmit, 1997), as well as the ability of on-the-job performance (Hunter & Schmidt, 1996). Recognizing what type of cognitive ability is deficient, by examining the classified assessment scores that were categorized by the cognitive process dimension, may provide potential strategies to improve the likelihood of successful knowledge and skill transfer. This relates to the learning management system (LMS), which many organizations have implemented over the past few years (Watson & Ahmed, 2004). One of the key features of LMS is to monitor trainees' level of learning. More importantly, the information must be useful for the training managers to support their decision makings for more effective training based on an understanding of why learning was not evident for some trainees. One example can be a customized training program, which focuses more on the learning outcomes that requires a specific cognitive ability that a trainee appears to lack. Likewise, this idea can be applied to scheduling retraining for disqualified trainees.

5.4. Potential application of the Revised Bloom's Taxonomy to training programs

This study demonstrated the usage of the Revised Bloom's Taxonomy to a lifting and lowering training program. However, the taxonomy has potential to be applied to other safety training programs, since the classification of learning outcomes does not depend on training contents but is general enough to be applied in various types of safety training programs. Likewise, the majority of the training programs follow the procedures from a generic framework to develop, design, and evaluate a training system. An instructional system introduced by Goldstein (1992) is a widely utilized model in the training domain. Therefore, the following sections will elaborate on how the Revised Bloom's Taxonomy can be integrated into an instructional system model to further support its potential application in designing and evaluating training programs.

The model has four phases: needs assessment, training and development, evaluation, and training goal (Figure 24).

(1) Needs assessment

- Instructional objectives are derived as a result of conducting organizational analysis, task and KSA (knowledge, skill, and attitudes) analysis, and person analysis in the needs assessment phase.
- Criteria are developed based on the instructional objectives

(2) Training and development

- Classification of instructional objectives:

  The Revised Bloom's Taxonomy is utilized to systematically classify the instructional objectives. One of the advantages of using the taxonomy is that it enables training designers to specify the objectives clearly so that planning evaluation methods become easier (Peggy, 2006) .
- Designing assessments:

  Assessments are designed to measure trainees' knowledge, skills, or attitudes during the training. In this phase, the taxonomy supports designing valid assessments that reflect the learning objectives (Anderson & Krathwohl, 2001). Later, the classified assessment scores will be utilized as a basis for deriving the advantages of using the Revised Bloom's Taxonomy, such as enabling detailed prediction among evaluation criteria and considering trainees' levels of cognitive ability. Therefore, the validity of the assessments must be ensured by proper alignment between assessments and corresponding learning objectives. Additionally, researchers in training disciplines have noted the importance of assessing during training (Ghodsian, Bjork, & Benjamin, 1997) and finding more rigorous and diagnostic assessments (Salas & Cannon-Bowers, 2001), which increases the importance of carefully designing assessments.

- Selection and design of instructional programs:

  A type of training program that yields the best result is determined in this phase. The Revised Bloom's Taxonomy provides clear instructional objectives (what is to be learned) and the classification guides what types of cognitive skills are required to yield the expected learning outcomes. This enables the training designers to recognize how learning should occur during training (e.g., remember, understand, apply, or evaluate) and further supports their decision in selecting a type of training program.

- Collection and classification of assessment scores:

  The assessment scores are collected during training. The classified assessment scores will be analyzed to evaluate the training program and further enhance learning, level 2 of the Four-level Evaluation Model.

(3) Evaluation

- Training performance evaluation:

  Training performance evaluation examines trainees' knowledge, skills, and attitudes right after their completion of the training program. Therefore, the assessments that reflect the overall goals of a training program should be measured in this phase.

- On-the-job evaluation:

  On-the-job evaluation is the assessment of the actual application of knowledge, skills, and attitudes in the work environment. Different types of measures that reflect the transfer of training must be designed and collected for a valid evaluation.

(4) Training goal

- Further analysis can be conducted to examine the validity of various methods for achieving training goals.

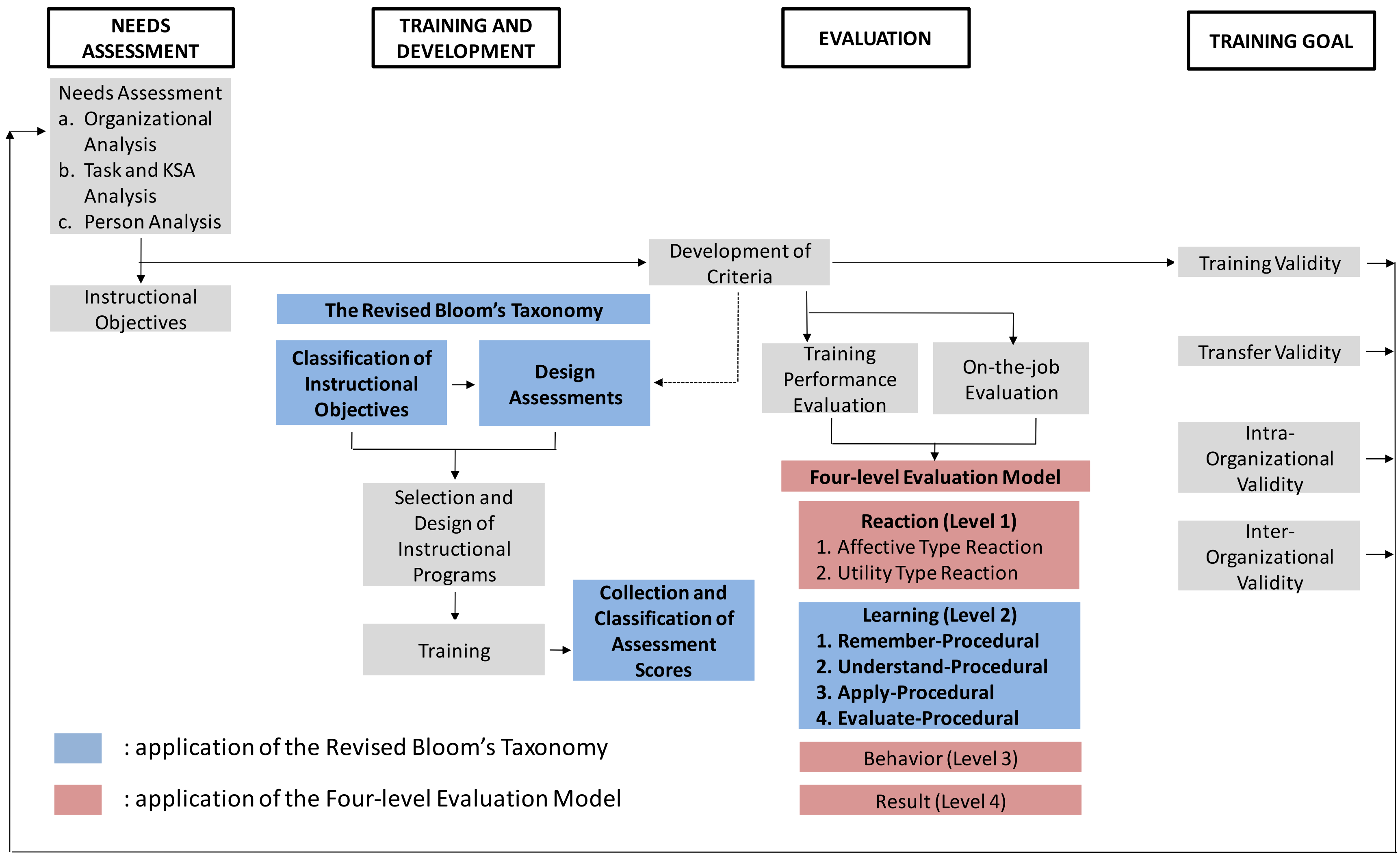


Figure 24. Expansion of the instructional system model

CHAPTER 6. CONCLUSION

6.1. Summary

The Revised Bloom's Taxonomy was applied to enable multidimensional categorization of learning outcomes and to determine which of those components were the most predictive of training performance in a lifting and lowering training program. The results identified that apply-procedural learning outcomes were the most predictive of training performance. Additionally, apply-procedural learning outcomes were more predictive of training performance than trainee reactions, which are the most widely utilized measurement in industry to evaluate training programs. The methods and findings of this study were used to develop a list of recommendations to support designing a lifting and lowering training program.

6.2. Recommendations

6.2.1. Apply-procedural, the most predictive learning outcome

Setting minimum acceptable criteria

When setting minimum acceptable criteria in a lifting and lowering training program, higher priority should be assigned to apply-procedural learning outcomes. Practical application examples can be:

- Setting higher standards for the assessments that are classified as apply-procedural.
- Considering apply-procedural assessment scores to support differentiating between qualified and disqualified trainees, so as to plan retraining.

Predicting transfer of training

It is assumed that transfer of training can be predicted by apply-procedural assessment scores. This information may assist human resource managers to effectively select and monitor trainees. Practical application examples can be:

- Utilizing apply-procedural assessments for trainee selection (prior to training).
- Utilizing apply-procedural assessments to identify the likelihood of successful transfer (during training).

Supporting training design

In lifting and lowering training programs, apply-procedural learning outcomes should be emphasized to ensure better training performance. Practical application examples can be:

- Providing hands-on activities, such as the Kinetic Learning Module in the training program.
- Ensuring sufficient training time for hands-on activities when planning training contents.
- Emphasizing and reviewing application related topics more frequently.

6.2.2. Utilizing the Revised Bloom's Taxonomy

Theory-based classification

The cognitive process dimension supports understanding and interpreting training outcomes that are due to trainees' different levels of cognitive ability. Since cognitive ability influences trainees' levels of performance, it should be considered as an important design factor. Practical application examples can be:

- Considering cognitive process dimension to design a customized training program that supports trainees' different levels of cognitive ability.

Potential application to design and evaluation of training programs

The Revised Bloom's Taxonomy has potentials to be applied to other safety training programs. Practical application examples can be:

- Applying the Revised Bloom's Taxonomy for systematic categorization of learning outcomes to expand level 2 learning of the Four-level Evaluation Model.
- Having clear learning objectives.
- Determining effective instructions.
- Designing valid assessments that measure whether learning objectives have been achieved.

6.3. Limitations and future research

There are several limitations involved in this study. First, the training environment was laboratory based, rather than at an actual training site. Studies have indicated that learning environment can significantly that influence the trainees' motivation to transfer training to their work (Seyler *et al.*, 1998). Moreover, three digital camcorders were installed in the laboratory and one mobile digital camera was carried by the researcher to record trainees' performances. This equipment may have physically hindered the participants' natural performance or potentially created a Hawthorne effect.

Second, the participant pool was mainly psychology undergraduate students at Virginia Tech. Considering that the training program was originally designed for generation X and Y learners, the sample population corresponded to the actual population. However, their motivation towards this training program may have differed from actual workers.

Third, the later learning outcomes could have been influenced by the prior learning outcomes in the training program. For example, application of proper lifting and lowering skills may have required the participants to remember the proper lifting and lowering concepts. An observational study was conducted in this study to focus on the realistic aspect of the training program. However, to overcome this problem, a true experimental design should be conducted. Different groups of participants should be randomly assigned to different types of learning objectives in the training program, so that the interaction among different types of learning outcomes can be minimized.

Fourth, the study infers transfer of training by support from previous researches that examined the relationship between training performance and transfer of training. However, more empirical studies should be conducted to assure that the inference is correct and to identify more detailed relationships among different evaluation criteria to further determine the factors that ensure the transfer of training.

Lastly, this study can be considered as an example of applying the Revised Bloom's Taxonomy in designing and evaluating a training program. Although the study results demonstrated a usage of the taxonomy in the field of training by integrating it into the instructional system model, more implementation should be conducted to further verify its benefits in training design and evaluation.

REFERENCE

Airasian, W. P., & Miranda, H. (2002). The Role of Assessment in the Revised Taxonomy. *Theory Into Practice, 41*(4), 249-254.

Alliger, M. G., & Janak, A. E. (1989). Kirkpatrick's Levels of Training Criteria: Thirty Years Later. *Personnel Psychology, 42*(2), 331-342.

Alliger, M. G., Tannenbaum, I. S., Bennett, W., Traver, H., & Shotland, A. (1997). A Meta-Analysis of the Relations Among Training Criteria. *Personnel Psychology, 50*(2), 341-358.

Alvarez, K., Salas, E., & Garofano, M. C. (2004). An Integrated Model of Training Evaluation and Effectiveness. *Human Resource Development Review, 3*(4), 385-416.

Amer, A. (2006). Reflections on Bloom's Revised Taxonomy. *Electronic Journal of Research in Educational Psychology, 4*(1), 213-230.

Anderson, L. W., & Krathwohl, R. D. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Anderson, R. J. (1982). Acquisition of Cognitive Skill. *Psychological Reivew, 89*(4), 369-406.

Arthur, W., Bennett, W., Edens, S. P., & Bell, T. S. (2003). Effectiveness of Training in Organizations: A Meta-Analysis of Design and Evaluation Features. *Journal of Applied Psychology, 88*(2), 234-245.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of Training:A Review and Directions for Future Research. *Personnel Psychology, 41*(1), 63-105.

Bassi, J. L., Benson, G., & Scott, C. (1996). The Top Ten Trends. *Training and Development, 50*(1), 28-42.

Bernard, B. P. (1997). Musculoskeletal disorders and workplace factors: a critical review of epidemiologic evidence for work-related mus culoskeletal disorders of the neck, upper extremity, and low back. US Department of Health and Human Services (DHHS) publication no. 97-141 (Cincinnati: National Institute for Occupational Safety and Health).

Bigos, S. J., Spengler, D. M., Martin, N. A., Zeh, J., Fisher, L., MNachemson, A., & Wang, M. H. (1986). Back Injuries in Industry: A Restrospective Study II: Injury Factors. *Spine, 11*(3), 246-251.

Bloom, B. (1956). *Taxonomy of Educational Objectives: The Cognitive Domain*. New York: Donald Mckay.

BLS. (2005a). Nonfatal Occupational injuries and Illnesses Requiring Days Away from Work, 2005. *Retrieved from http://www.bls.gov/news.release/pdf/osh2.pdf*.

BLS. (2005b). Number of nonfatal occupational injuries and illnesses involving days away from work by selected worker and case characteristics and musculoskeletal disorders. *Retrieved from http://www.bls.gov/iif/oshwc/osh/case/ostb1655.pdf*.

Boekaerts, M. (1997). Self-Regulated Learning: A New Concept Embraced by Researchers, Policy Makers, Educators, Teachers, and Students. *Learning and Instruction, 7*(2), 161-186.

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

Brinkerhoff, R. O. (1987). *Achieving Results from Training*. San Francisco: Jossey-Bass.

Brown, M. (2005). Learning Spaces. *Educating the Net Generation*, Retrived August 20, 2006, from www.educase.edu/educatingthenetgen/.

Dixon, M. N. (1990). The Relationship Between Trainee Responses on Participation Reaction Form and Postest Scores. *Human Resource Development Querterly, 1*, 129-137.

Dolezalek, H. (2005). 2005 Industry Report. *Training, 42*(12), 14-28.

Fabio, P. R. (1995). Efficacy of Comprehensive Rehabilitation Programs and Back School for Patients with Low Back Pain: A Meta-analysis. *Psycial Therapy, 75*(10), 865-878.

Fitts, P. M. (1964). *Perceptual-motor Skills Learning*. New York: Academic Press.

Fitzpatrick, R. (2001). The Strange Case of the Transfer of Training Estimate. *Industrial-Organizational Psychologist, 39*(2), 18-19.

Foster, G., Sawicki, E., Schaeffer, H., & Zelinski, V. (2003). *I Think, Therefore I Learn*. Ontario, Canada: Pembroke Publishers.

Gagne, M. R. (1984). Learning Outcomes and Their Effects: Useful Categories of Human Performance. *American Psychologist, 39*(4), 377-385.

Gagnon, M. (2003). The Efficacy of Training for Three Manual Handling Strategies Based on the Observation of Expert and Novice Workers. *Clinical Biomechanics, 18*(7), 601-611.

Georgenson, L. D. (1982). The Problem of Transfer Calls for Partnership. *Training and Development Journal, 36*(10), 75-78.

Ghodsian, D., Bjork, R., & Benjamin, A. (1997). *Evaluating Training During Training: Obstacles and Opportunities*. Washington DC: Americal Psychology Association.

Goldstein, L. I. (1992). *Training in Organizations: Needs Assessment, Development, and Evaluation* (3rd ed.): Brooks/Cole Publishing Company.

Goldstein, L. I., & Ford, J. K. (2002). *Training in Organizations: Needs Assessment, Development, and Evaluation* (4th ed.). Belmont, CA.: Wadsworth Publishing.

Haccoun, R. R., & Saks, M. A. (1998). Training in the 21st Century: Some Lessons from the Last One. *Canaian Psychology, 39*(1-2), 33-51.

Hamblin, A. C. (1974). *Evaluation and COntrol of Training*. New York: McGraw-Hill.

Holton III, F. E. (1996). The Flawed Four-Level Evaluation Model. *Human Resource Development Querterly, 7*(1), 5-21.

Hoogendoorn, W. E., Van Poppel, M. N., Bongers, P. M., Koes, B. W., & Bouter, L. M. (1999). Physical Load During Work and Leisure Time as Risk Factors for Back Pain. *Scandinavian Journal of Work and Environmental Health, 25*(5), 387-403.

Hunter, E. J., & Schmidt, L. F. (1996). Intelligence and Job Performance: Economic and Social Implications. *Psychology, Public Policy, and Law, 2*(3-4), 447-472.

Kaufman, R., & Keller, J. M. (1994). Levels of Evaluation: Beyond Kirkpatrick. *Human Resource Development Querterly, 5*, 371-380.

Kirkpatrick, D. L. (1976). *Evaluating of Training*. New York: McGraw-Hill.

Kozlowski, W. J. S., Gully, M. S., Brown, G. K., Salas, E., Smith, M. E., & Nason, R. E. (2001). Effects of Training Goals and Goal Orientation Traits on Multidimensional Training Outcomes and Performance Adaptability. *Organizational Behavior and Human Decision Processes, 85*(1), 1-31.

Kraiger, K. (2002). *Decision-Based Evaluation*. San Francisco, CA: Jossey-Bass.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of Cognitive, Skill-Based, and Affective Theories of Learning Outcomes to New Methods of Training Evaluation. *Journal of Applied Psychology, 78*(2), 311-328.

Kraiger, K., McLinden, D., & Casper, J. W. (2004). Collaborative Planning for Training Impact. *Human Resource Management, 43*(4), 337-351.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. White Plains, NY: Longman.

Krathwohl, R. D. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice, 41*(4), 212-218.

Kroeger, D. S., & Kouche, B. (2006). Using Peer-Assisted Learning Strategies to Increase Response to Intervention in Inclusive Middle Math Settings. *Teaching Exceptional Children, 38*(5), 6-13.

Mallows, C. L. (2000). Some Comments on Cp. *Technometrics, 42*(1), 87-94.

Marras, W. S. (2000). Occupational Low Back Disorder Causation and Control. *Ergonomics, 43*(7), 880-902.

Morgan, B. R., & Casper, J. W. (2000). Examining the Factor Structure of Participant Reactions to Training: A Multidimensional Approach. *Human Resource Development Querterly, 11*(3), 301-317.

Motowidlo, J. S., Walter, B. C., & Schmit, J. M. (1997). A Theory of Individual Differences in Task and Contextual Performance. *Human Performance, 10*(2), 71-83.

NIOSH. (1996). National Occupational Research Agenda. U.S. Department of Health and Human Services, Public Health Service, Center for Disease Control and Prevention, National Institute for Occupational Safety and Health (NIOSH), April 1996.

Noe, A. R., & Schmitt, N. (1986a). The Influence of Trainee Attitudes on Training Effectiveness: Test of a Model. *Personnel Psychology, 39*(3), 497-523.

Noe, R. A., & Schmitt, N. (1986b). The Influence of Trainee Attitudes on Training Effectivenss: Test of a Model. *Personnel Psychology, 39*(3), 497-523.

Nunnaly, J. (1978). *Psychometric Theory*. New York: McGraw-Hill.

Peggy, D. (2006). New Blooms in Established Fields: Four Domains of Learning and Doing. *Roeper Review, 28*(2), 70-78.

Pintrich, R. P. (2002). The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing. *Theory Into Practice, 41*(4), 219-227.

Punnett, L., Pruss-Ustun, A., Nelson, I. D., Fingerhut, A. M., Leigh, J., Tak, S., & Phillips, S. (2005). Estimating the global burden of low back pain attributable to combined occupational exposures (Vol. 48, pp. 459-469).

Ree, J. M., Carretta, R. T., & Teachout, S. M. (1995). Role of ability and prior job Knowledge in Complex Training Performance. *Journal of Applied Psychology, 80*(6), 721-730.

Rekus, J. F. (1999). Is Your Safety Training Program Effective? *Occupational Hazards, 8*, 38-40.

Reville, R. T., Bhattacharya, J., & Sager, L. R. (2001). New Methods and Data Sources for Measuring Economic Consequences of Workplace Injuries. *American Journal of Industrial Medicine, 40*, 452-463.

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

Ruona, E. A. W., Leimbach, M., Holton III, F. E., & Bates, R. (2002). The Relationship Between Learner Utility Reactions and Predicted Learning Transfer Among Trainees. *International Journal of Training and Development, 6*(4), 218-228.

Saari, M. L., Johnson, R. T., Mclaughlin, D. S., & Zimmerle, M. D. (1988). A Survey of Management Training and Education Practice In U.S. Companies. *Personnel Psychology, 41*(4), 731-743.

Saks, M. A., & Belcourt, M. (2006). An Investigation of Training Activities and Tansfer of Training Organizations. *Human Resource Management, 45*(4), 629-648.

Salas, E., & Cannon-Bowers, A. J. (2001). The Science of Training: A Decade of Progress. *Annual Review of Psychology, 52*, 471-499.

Schulte, A. P. (2005). Characterizing the Burden of Occupational Injury and Disease. *Journal of Occupational and Environmental Medicine, 47*(6), 607-622.

Seyler, D. L., Holton Iii, E. F., Bates, R. A., Burnett, M. F., & Carvalho, M. A. (1998). Factors Affecting Motivation to Transfer Training. *International Journal of Training and Development, 2*(1), 16.

Snoddy, S. G. (1926). Learning and Stability: A Psychophysiological Analysis of a Case of Motor Learning with Clinincal Applications. *Journal of Applied Psychology, 10*(1), 1-36.

Snook, S. H. (1978). The Design of Manual Handling Tasks. *Ergonomics, 21*(12), 963-985.

Snyder, A. R., Raben, S. C., & Farr, L. J. (1980). A Model for the Systemic Evaluation of Human Resource Development Programs. *Academy of Management Review, 5*(3), 431-444.

St-Vincent, M., Tellier, C., & Lortte, M. (1989). Training in Handling: an Evaluative Study. *Ergonomics, 32*(2), 191-210.

Sugrue, B. (2003). *State of the Industry: ASTD's Annual Review of U.S. and International Trends in Workplace Learning and performance*. Alexandria, VA: ASTD.

Svensson, H. O., & Andersson, G. B. (1983). Low-Back Pain in 40 to 47 Year Old Men: Work History and Work Environment Factors. *Spine, 8*(3), 272-276.

Tan, A. J., Hall, J. R., & Boyce, C. (2003). The Role of Employee Reactions in Predicting Training Effectiveness. *Human Resource Development Querterly, 14*(4), 397-411.

Tannenbaum, I. S., Mathieu, E. J., Salas, E., & Cannon-Bowers, A. J. (1991). Meeting Trainees' Expectations: The Influence of Training Fulfillment on the Development of Commitment, Self-Efficacy, and Motivation. *Journal of Applied Psychology, 76*(6), 759-769.

Underhill, F., Anthony. (2006). Theories of Learning and Their Implications for On-Line Assessment. *Turkish Online Journal of Distance Education, 7*(15), 165-174.

Vroom, V. (1964). *Work and Motivation*. New York: Wiley.

Warr, P., Allan, C., & Birdi, K. (1999). Predicting Three Levels of Training Outcomes. *Journal of Occupational and Organizational Psychology, 72*(3), 351-375.

Warr, P., & Bunce, D. (1995a). Trainee Characteristics and Outcomes of Open Learning. *Personnel Psychology, 48*, 347-375.

Warr, P., & Bunce, D. (1995b). Trainee Characteristics and the Outcomes of Open Learning. *Personnel Psychology, 48*(2), 347-375.

Waters, R. T., Putz-Anderson, V., Garg, A., & Fine, J. L. (1993). Revised NIOSH Equation for the Design and Evaluation of Manual Lifting Tasks. *Ergonomics, 37*(7), 749-776.

Watson, J., & Ahmed, K. P. (2004). Learning in the Age of Global Information: Development of Generic Architecture for an Advanced Learning Management System. *Campus-wide Information Systems, 21*(1), 4-21.

Webster, B. S., & Snook, S. H. (1994). The Cost of 1989 Workers' Compensation Low Back Pain Claims. *Spine, 19*(10), 1111-1115.

Weil, D. (2001). Valuing the Economic Consequences of Work Injury and Illness: A Comparison of Methods and Findings. *American Journal of Industrial Medicine, 40*, 418-437.

Wexley, N. K., & Baldwin, T. T. (1986). Posttraining Strategies for Facilitating Positive Transfer: An Empirical Exploration. *Academy of Management Journal, 29*(3), 503-520.

APPENDIX A: Demographics Questionnaire

DEMOGRAPHICS QUESTIONNAIRE

The purpose of this questionnaire is for you to provide some basic background information about yourself and your experience. Please complete the following demographics questionnaire.

BACKGROUND INFORMATION

Participant number: _____________

1. Gender: _____ Female _____ Male

2. Ethnicity: _____ African-American (Black)

_____ Asian-American

_____ European-American (White, Caucasian)

_____ Hispanic-American (Latino, Latin-American)

_____ Native-American (American Indian)

_____ Other (If so, what ethnicity? ______________________________)

3. Age: _________

4. Height: ___________________ Weight: ________

5. Do you have any visual disabilities? Yes_____, No_____

6. Do you have any hearing disabilities? Yes_____, No_____

EXPERIENCE

7. Experience of working in a delivery or service industry? Yes____, No____

If yes, how long have you been working? ________ years _______ months

8. Experience of lifting / lowering training or lectures relate to this topic? Yes____, No____

If yes, what kind of training or lectures? ____

9. Are you comfortable with using computer? (e.g., using mouse to navigate) Yes____, No____

10. Past injury experience related to lifting and lowering? Yes____, No____

If yes, are you still uncomfortable with it? Yes____, No____

APPENDIX B: Prequalification self-check score sheet

PREQUALIFICATION SELF-CHECK SCORE SHEET

Instructions: As you go through the web-based training in the Prequalification phase, there will be nine "Check for Understanding". After each question, feedback will be provided informing whether you have selected the correct / incorrect answer. Please check on the below table.

Slide #	Correct	Incorrect
22		
25		
27		
29		
31		
37		
38		
39		
40		

APPENDIX C: Screenshots of workshop 2 assessment questions

Check for Understanding #2

When lifting / lowering, you should always try to keep the load away from the body.

A. True
B. False

When you have selected your answer, write it on the answer sheet.

Check for Understanding #3

A way to reduce the weight of the load is to distribute the load using objects in the environment. Provide an example of distributing the load using surrounding objects or environment. (Short answer)

When you have selected your answer, write it on the answer sheet.

Check for Understanding #6

Lift with your _______ to avoid low back injuries.

A. Hands
B. Waist
C. Legs
D. Knees
E. Arms

When you have selected your answer, write it on the answer sheet.

Check for Understanding #7

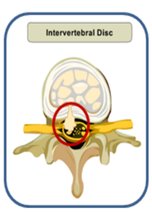
When the package is small enough, it is safe to move quickly to lift the package up from the ground.

A. True
B. False

When you have selected your answer, write it on the answer sheet.

Check for Understanding #8

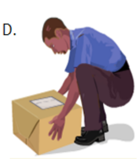
What activities can be the cause of herniated displaced disk (slipped disk)? (Select as many)



A.


B.


C.


D.


E.


F.


When you have selected your answer, write it on the answer sheet.

Check for Understanding #14

In the following image, a person is twisting his waist while handling a package. What should he do to avoid twisting?
(from the eight keys of lifting and lowering) (Short answer)

When you have selected your answer, write it on the answer sheet.

APPENDIX D: Workshop 2 self-check answer sheet

WORKSHOP 2 SELF-CHECK ANSWER SHEET

Instructions: There will be fourteen questions that ask you about what you have learned during Workshop 2. Please write down the answers in the below table.

No.	ANSWER
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	

APPENDIX E: KLM Peer Review Checklists

No.	Does the candidate do the following?	Answer	
1	Check for weight beforehand	Yes	No
2	Get close to the object (in Power Zone + avoid end-range)  	Yes	No
3	Get a firm grip and grasp opposite corners	Yes	No
4	Feet position- shoulder width apart, one foot slightly ahead of the other 	Yes	No
5	Lift with a smooth, steady motion	Yes	No
6	Keep the natural curve of the back – bend at the knees  	Yes	No

7	Avoid twisting – move the feet 	Yes	No
8	When carrying, maintain balance	Yes	No
9	Scan the walk path	Yes	No
10	Use equipment or get help when necessary  	Yes	No

APPENDIX F: Training performance task instructions

Participant A

1. Begin by walking towards the package

2. Lift the package, which is labeled as “Package #1”.

3. Carry the package to the designated location, which is labeled as “**A**”

4. Lower the package at the designated location

5. Lift the same package (package #1)

6. Carry the package to its original location

7. Place the package at the original location

Participant B

1. Begin by walking towards the package

2. Lift the package, which is labeled as “1”

3. Carry the package to the designated location “**B**”

4. Lower the package

5. Lift the same package (#1)

6. Carry the package to its original location

7. Place the package at the original location

APPENDIX G: Reaction questionnaires

Instruction: Please type in the most relevant option (numbers) in the parentheses.

1. How satisfied are you with quality of this training course overall? (   )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

2. How satisfied are you with the quality of course materials? (   )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

3. How satisfied are you with exam coverage and importance of material tested? (   )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

4. How satisfied are you with feedback you received as result of course testing? (   )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

5. How satisfied are you with match of course objectives with your idea of what would be taught? (   )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

6. How satisfied are you with the relevance of the course content to your job? (   )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

7. How satisfied are you with course's emphasis on most important information? ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

8. How satisfied are you with the extent to which the course prepared you to perform current job tasks more effectively? ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

9. How satisfied are you with the length of training course? ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

10. How satisfied are you with the pace of the course material presented? ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

11. I have an overall good feeling about how the training program was carried out ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

12. The training program allowed me to develop specific skills that I can use on the job ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

13. The training program was, overall, very effective ( )
  1. Strongly disagree
  2. Disagree
  3. Neither
  4. Agree
  5. Strongly agree

14. The training program is very useful ( )

1. Strongly disagree
2. Disagree
3. Neither
4. Agree
5. Strongly agree

15. My time was well spent ( )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

16. The session objectives were met ( )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

17. I would recommend this program to others who have the opportunity ( )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

18. I learned something I can apply immediately to my work ( )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

19. The course provided me with new ways of thinking about my job ( )
    1. Strongly disagree
    2. Disagree
    3. Neither
    4. Agree
    5. Strongly agree

APPENDIX H: Informed Consent Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
GRADO DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

Informed consent form for participants of Research Involving Human Subject

Title of the Project: Finding the Predictive Validities of Different Types of Learning Outcomes in a Lifting and Lowering Training Program

Principal Investigator: Yoon Suk Lee, Graduate Student, ISE

Faculty Advisor: Dr. Tonya L. Smith-Jackson, Associate Professor, ISE

I. PURPOSE OF THIS RESEARCH

The purpose of this study is to understand the transfer of training in a lifting and lowering training program by finding the predictive validities of different types of learning outcomes to training performance.

II. PROCEDURES

If you choose to participate in this study, you will be asked to sign one informed consent form (this document). You will keep a copy for yourself. We will then ask you to complete a demographic form, which helps us to categorize our data to make comparisons as needed. Then you will be asked to begin the training program. After completing the training, you will be required to perform a lifting and lowering task for training performance assessment. You will be video and audio taped occasionally during the experiment for further data analysis. The experiment will last about 90 minutes. However, you are welcome to take breaks as needed.

III. RISKS

There will be some level of risks for participating in this study since the participants have to perform physical activities of lifting and lowering packages. However, every effort will be made to prevent potential injuries by pre-screening of injury experiences, and breaks as needed.

IV. BENEFITS OF THIS RESEARCH

You will probably gain knowledge and skills to safely handle packages. Your participation will also contribute to improve designing training programs that are related to lifting and lowering.

This research has been approved by the VT Institutional Review Board, beginning 16 July 2007 until 15 July 2008.

V. EXTENT OF ANONYMITY AND CONFIDENTIALITY

This research will assure your confidentiality. However, anonymity can not be guaranteed, because we will need to have your signatures on the Informed Consent document. However, this document will be kept in a locked cabinet for 5 years and your name will not be released. At the end of the 5-year period, we will destroy the documents. Your name will not be associated with the content of this observation, but you will be assigned a three- digit number to protect your privacy. Your number is _______.

All data will be collected by the researchers only. No one other than the researchers will have access to the data. We will use digital camcorders so we can analyze your lifting and lowering postures later. This is important so we don't miss critical information. As you will be given a 3-digit number above, we ask that you refer to yourself by this 3-digit number and please do not refer to any other participants by name. All responses will be coded so as not to include the name of the participant. The information you provide will have your name removed and only a three- digit participant number will identify you during analyses and any written reports of the research.

This study is being conducted solely for research and development purposes, and the resulting data and interpretations will also be the part of the researcher's academic work. Consistent with these academic purposes, any results would be freely publishable. However, to protect your identity, personal names will not be used in any published works. We are willing to share drafts of reports with you before submitting them for publication.

VI. COMPENSATION

There is no monetary compensation for participating in this study. However, if you were recruited by the VT Psychology SONA system, you will receive 2 credits for participation.

VII. FREEDOM TO WITHDRAW

Participation in the study is voluntary and the decision about whether you wish to participate is strictly your own. You may discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. Withdrawal from the study will not result in any adverse effects.

VIII. APPROVAL OF RESEARCH

This research project has been approved by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering (College of Engineering).

IX. PARTICIPANT'S RESPONSIBILITIES

Upon signing this form below, I voluntarily agree to participate in this study. I have no restrictions to my participation in this study. I also agree not to discuss this procedure with anyone until I have been formally debriefed.

X. PARTICIPANT'S PERMISSION

I have read and understand the Informed Consent and conditions of this study. I understand that the experiment will be videotaped to support data recording. All of my questions have been answered. I agree to participate in this project.

______________________________ ______________________

Participant's Signature Date

Should I have any questions about the research or its conduct, I may contact:

Principal Investigator: Yoon Suk Lee Phone: (540) 230-0236
Graduate Student, Grado Department of Industrial and Systems Engineering
Email: yoonlee@vt.edu

Faculty Advisor: Tonya L. Smith-Jackson Phone: (540) 231-4119
Associate Professor, Grado Department of Industrial and Systems Engineering
Email: smithjack@vt.edu

Dr. David M. Moore, Chair, IRB Email : moored@vt.edu Phone: (540) 231-4991

APPENDIX I: Internal Review Board (IRB) approval letter

VirginiaTech

Office of Research Compliance
Institutional Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, Virginia 24061
540/231-4991 Fax 540/231-0959
e-mail moored@vt.edu
www.irb.vt.edu
FWA00000572 (expires 1/20/2010)
IRB # is IRB00000667

DATE: October 17, 2007

MEMORANDUM

TO: Tonya L. Smith-Jackson
Thurmon E. Lockhart
Brian M. Kleiner

FROM: David M. Moore

Approval date: 7/16/2007
Continuing Review Due Date: 7/1/2008
Expiration Date: 7/15/2008

SUBJECT: **IRB Amendment 3 Approval:** "Driver Service Provider (DSP) Training System: Ride-Alongs", OSP #415319, IRB # 06-369

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on July 16, 2007. You subsequently requested permission to amend your IRB application. Since the requested amendment is nonsubstantive in nature, I, as Chair of the Virginia Tech Institutional Review Board, have granted approval for requested protocol amendment, effective as of October 16, 2007. The anniversary date will remain the same as the original approval date.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.
2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.
3. Report promptly to the IRB of the study's closing (i.e., data collecting and data analysis complete at Virginia Tech). If the study is to continue past the expiration date (listed above), investigators must submit a request for continuing review prior to the continuing review due date (listed above). It is the researcher's responsibility to obtained re-approval from the IRB before the study's expiration date.
4. If re-approval is not obtained (unless the study has been reported to the IRB as closed) prior to the expiration date, all activities involving human subjects and data analysis must cease immediately, except where necessary to eliminate apparent immediate hazards to the subjects.

As indicated on the IRB application, this study is receiving federal funds. The approved IRB application has been compared to the OSP proposal listed above and found to be consistent. Funds involving procedures relating to human subjects may be released. Visit our website at www.irb.vt.edu for further information

cc: File
T. Coalson 0118

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution