SPECIAL REPORT 2010-001

ADAPTIVE TRAINING CONSIDERATIONS FOR USE IN SIMULATION-BASED SYSTEMS

SEPTEMBER 2010

Carla R. Landsberg
KAEGAN Corporation

Wendi L. Van Buskirk
Randolph S. Astwood Jr.
Naval Air Warfare Center Training Systems Division

Alyssa D. Mercado
Anthony J. Aakre
KAEGAN Corporation

NAVAL AIR WARFARE CENTER TRAINING SYSTEMS DIVISION
12350 Research Parkway
Orlando, FL, 32826-3275

J. OWEN, Head
Human Systems & Engineering Division

M. ZETTLER, Director
Research & Engineering Department

M. BERGONDY-WILHELM
Director, Research & Technology

REPORT DOCUMENTATION PAGE			*Form Approved* OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**			
1. REPORT DATE *(DD-MM-YYYY)* 10/14/10	**2. REPORT TYPE** Special Report		**3. DATES COVERED** *(From - To)* 05/2009 – 10/2010
4. TITLE AND SUBTITLE Adaptive Training Considerations for use in Simulation-based Systems			**5a. CONTRACT NUMBER** N0001410WX20227 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Carla R. Landsberg, Wendi L. Van Buskirk, Randolph S. Astwood Jr., Alyssa D. Mercado, Anthony J. Aakre			**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Air Warfare Center Training Systems Division 12350 Research Parkway Orlando, FL 32826-3275			**8. PERFORMING ORGANIZATION REPORT NUMBER** 2010-001
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research 875 North Randolph Street, Suite 1425 Arlington, VA 22203-1995			**10. SPONSOR/MONITOR'S ACRONYM(S)** ONR **11. SPONSORING/MONITORING AGENCY REPORT NUMBER**
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			
13. SUPPLEMENTARY NOTES Questions concerning this report should be referred to: NAVAIRWARCENTRASYSDIV (4651), 12350 Research Parkway, Orlando, FL 32826-3275.			
14. ABSTRACT In this report, we examine theoretical and empirical papers that describe adaptive training (AT). When applied effectively, AT has the potential to improve training interventions, making them more successful and efficient. An extensive search was conducted to locate and review articles pertaining to AT. Our review of the research suggests that there are a number of different models to select from when developing an AT system. However, there are also a number of considerations that one must take into account prior to, during, and after the development of an AT training system. Future research will need to determine the relative effectiveness of different approaches and which variables are most useful for adapting training.			
15. SUBJECT TERMS Adaptive training, adaptive guidance, individual differences			

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	SAR	81	Carla Landsberg **19b. TELEPONE NUMBER (*Include area code*)** (407) 380-4331

THIS PAGE INTENTIONALLY LEFT BLANK

Special Report 2010-001

ACKNOWLEDGEMENTS

This work was sponsored by Dr. Ami Bolton and the Office of Naval Research (Contract No. N0001410WX20227). We would like to thank Ms. Natalie Steinhauser, Mr. Andy Mikutel, Ms. Miliani Jimenez, Ms. Gabriella Severe, and Dr. Karen Holness for their contributions to the annotated bibliography.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Problem

Training research has sought to assess and evaluate the effectiveness of various instructional techniques and methods. In the current report, we examine theoretical and empirical papers that describe one such training method, adaptive training (AT). In AT, aspects of the instruction are manipulated in order to create an optimal learning experience for the trainee. When applied effectively, AT has the potential to improve training interventions, making them more successful and efficient. This is particularly important because of the significant costs, in terms of both time and money, which are associated with training systems.

Objective

The objective of this report is to reveal the "state of the science" in the area of AT. Specifically, this report sought to determine the models and components of AT systems, the learner characteristics that could be used to adapt instructional content, and the methods and procedures that have been successfully implemented within AT systems. In this report, we review various parameters of AT and identify guiding principles for what is and is not effective.

Approach

An extensive search was conducted to locate empirical, theoretical, and review articles pertaining to AT. Our search resulted in over 6,000 articles. However, our inclusionary criteria left us with only 34 articles; these were separated into three main sections: adaptive training review and theory, adaptive training assessment, and individual differences.

Findings

The AT review suggests that there are a number of different models to select from when developing an AT system. However, there are also a number of considerations that one must take into account prior to, during, and after the development of an AT system. The review also found that, in general, AT seems to be an effective training method. Lastly, it seems that trainee characteristics can be used to effectively adapt training.

Conclusions

Despite almost four decades of research and evidence illustrating the utility of AT, there are still a number of questions left unanswered. Future research will need to determine what variables to adapt training on, as well as if variables should be adapted during task performance, after the task is complete, or both. Furthermore, empirical evaluations that compare the effectiveness of various AT systems to each other are needed, as well as research to determine which individual difference variables are most useful and how these variables interact. Other considerations for future research are discussed.

TABLE OF CONTENTS

ADAPTIVE TRAINING: DEFINITION AND STUDY GOALS

The Department of the Navy has recently placed an increased importance on advanced training methods and technologies in order to avoid the potentially disastrous consequences of ineffective training. For example, the Naval Aviation Enterprise Naval Warrior Performance Science and Technology Objective for training and education seeks to develop training technologies to maximize transfer of training from the classroom and trainer to the operational environment. In this report, we examine theoretical and empirical papers that describe one such advanced training method, Adaptive Training (AT). In AT, some aspect of the instruction, such as order of presentation, type of presentation, type of information, and/or task difficulty are varied to create an optimal learning experience for a particular learner. In other words, AT methods are "educational interventions aimed at effectively accommodating individual differences in students while helping each student develop the knowledge and skills required to learn a new task" (Park & Lee, 1996, p.651). Kelley (1969) describes AT as "training in which the problem, the stimulus, or the task is varied as a function of how well the trainee performs" (p.547). This definition implies that training is adapted based on the performance of the trainee, however, this is not always the case. A more broad definition, espoused by Shute and Zapata-Rivera (2008) describes an adaptive system as one that "adjusts itself to suit particular learning characteristics and the needs of the learner" (p. 269). This definition of AT implies that training can be adapted on other variables besides performance. Specifically, their definition suggests that adaptation of instructional content can occur as a result of user traits or characteristics, as well as learning needs that arise during instruction.

The goal of this special report is to reveal the "state of the science" in the area of AT. Specifically, what are the models and components of AT systems, what learner characteristics could be identified and used as decision points to adapt instructional content, and what methods and procedures have been used successfully? The literature described in the following sections illustrates various methods that can be used to adapt training, each with its own benefits and shortcomings. The literature also highlights that there is more to consider than whether or not to adapt training. For instance, training system designers and instructors need to consider how to adapt training and what variables to adapt for each learner. Additionally, it is important to consider the assessment techniques used to evaluate learner characteristics and performance before and/or during training because these will have a profound impact on the effectiveness of AT. The study of AT is important because when AT is applied effectively, it has the potential to improve training interventions, making them more successful and efficient. This is particularly important because the development and use of training systems can be costly in terms of both time and money. In each section of this report, we review different parameters of AT and provide guidelines and suggestions for what is and is not effective.

However, before continuing onto the review of AT models and research, it is important to briefly address what is *not* considered AT and to clarify constructs that are often confused with AT in order to scope the review that follows. For example, adaptive automation is "a form of automation that is flexible or dynamic in nature" (Scerbo, 1996, p. 37). In these types of systems, automation can take over a task when the human operator"s workload demand is high, or when performance of the task falls below an acceptable level. Similarly, adaptive aiding is automation that is applied to help a user perform a task when task demand is high (Rouse, 1988). In this case, the adaptation is used to aid the user with the task rather than taking over performance of the task outright. In comparison to AT, adaptive automation and aiding can be used in the operational environment to combat high levels of workload, whereas AT is a training technique that would be used to prepare trainees for the operational environment. Finally, the construct of adaptivity and training for adaptivity is probably most often confused with AT. Training for adaptivity aims to teach a learner to adapt to task or environmental changes and/or to transfer their skills to new tasks (Allworth & Hesketh, 1999). This is different from AT where the goal is to tailor the training to the learner, not train the learner to tailor their behaviors to the task. Therefore, for the purpose of this report, *AT is defined as training interventions whose content can be tailored to an individual learner's aptitudes, learning preferences, or styles prior to training and that can be adjusted, either in real time or at the end of a training session, to reflect the learner's on-task performance*. Adaptive automation, adaptive aiding, and adaptivity training do not fit this definition, and were, therefore, excluded from our analysis.

An extensive search was performed to locate empirical, theoretical, and review articles pertaining to AT. We conducted a search of the PsychINFO, Defense Technical Information Center, Google Scholar, and Human Factors and Ergonomics Society (HFES) Training and Individual Difference Technical Groups databases using the key terms: adaptive training, computer-based training, adaptive instruction, adaptive learning, adaptive E-learning, adaptive feedback, and expert/novice differences. This search resulted in 6,044 articles. We excluded studies that covered adaptive guidance, adaptive automation, and adaptivity. Additionally, we limited our review to articles that were published in English and based on healthy participant samples. In regards to AT assessment, we also excluded articles that did not include empirical data on system effectiveness. These criteria left us with 34 articles which were separated into three main sections: adaptive training review and theory, adaptive training assessment, and individual differences.

ADAPTIVE TRAINING REVIEW AND THEORY

In the past, AT took the form of humans adapting instruction based on verbal or nonverbal cues they received from students - for example, if the student answered a question incorrectly, had a puzzled look on their face, looked bored, or was not paying attention, the instructor could observe those signs and adjust training appropriately. In fact, Bloom (1984) found that students who received one-on-one tutoring performed two standard deviations higher than students in a typical classroom setting (which he referred to as the "2 Sigma Problem"). Bloom suggested that because one-on-one tutoring is too costly, researchers must "find methods of instruction that are as effective as one-to-one tutoring" (p. 15). With the advent of advanced computer systems, it is now possible for computers to adapt instruction to the needs of the learner. Therefore, AT systems may be one type of instructional technology that could potentially solve the 2 Sigma Problem.

According to researchers, there are several different ways in which training can be adapted to the needs of the learner (Mödritscher, Garcia-Barros, & Gütl, 2004; Park & Lee, 1996; Shute & Towle, 2003; Shute & Zapata-Rivera, 2008). The most widely used method is the *macro-adaptive approach* (Park & Lee, 1996). In general, when using the macro-adaptive approach, students are presented different instructional modules and are allowed to progress at their own pace. Typically, students have to master a module before proceeding. Therefore, evaluation of student performance takes place to determine whether the student can move on to the next module or if he/she has to partake in remediation activities. Likewise, the module that the student starts with is determined based on prior performance such as previous course grades. In some macro adaptive systems, students are also allowed to determine the sequence in which they progress through different topics. The macro-adaptive approach represents the most basic type of AT. There is not a lot of variation in the instructional material presented to students and no need for sophisticated algorithms to diagnose performance or to determine the presentation of instructional content.

The second approach to instructional adaptation, *Aptitude Treatment Interaction (ATI)*, allows for adaptation based on the learner‟s current aptitudes or individual differences. The idea is if instruction is more suited toward the characteristics of the student, optimal learning will take place. Some individual difference variables that have been related to performance and have been considered for use in AT systems include intellectual ability, cognitive styles, learning styles, and prior knowledge (Park & Lee, 1996; Shute & Towle, 2003). However, before implementing this type of approach, AT developers should consider that ability levels may change over time and become less optimal for determining adaptations in real time. Additionally, individual difference characteristics may be task dependent (especially for individual differences such as self-efficacy, motivation, etc). Finally, in order to be able to effectively use individual differences as an adaptation variable, researchers, instructors, and AT system developers also need valid and reliable measurement of these constructs.

The third type is the *micro level approach* in which continual, on task performance measurement (e.g., errors, response times, etc.) forms the basis for adaptations. Park and Lee (1996) suggest that this approach is the most receptive to student needs and is the approach that most closely represents one-on-one tutoring. In order to apply this approach, sophisticated algorithms/mathematical models for real time performance measurement and diagnosis are required. Additionally, after a diagnosis of performance has been made, algorithms are needed to deliver the appropriate instructional content to the student. Therefore, a weakness of this approach is that development of real time, diagnostic algorithms and instructional models are resource intensive and costly.

While ATI, micro-adaptive, and macro-adaptive have been the traditional approaches used for AT, Park and Lee (1996) suggest using a hybrid approach which they call the *two-step approach*. The two-step approach integrates the ATI and micro approaches. More specifically, this approach is based on the idea that pre-task aptitudes are less predictive of future performance over time than on-task performance. Therefore, prior to training, the ATI approach could be used to place the learner in the appropriate level and/or to determine the instructional content to use at the beginning of instruction. Then, during task performance, a micro-adaptive approach could be used to continuously diagnose student performance and prescribe interventions during instruction. While this approach does have promise, it has not been used in practice and its effectiveness is yet to be determined.

When developing an AT system using one of the more sophisticated approaches such as the micro-adaptive or two-step approach, what components should be included? In his seminal article, Kelly (1969) suggests that an effective AT system should include three components: real time continuous performance measurement, a variable that will be adjusted, and the adaptation logic (how or how much the adaptive variable will be adjusted based on continuous performance measurement). The adaptive variable in Kelly"s system was task difficulty which was adjusted to match the skill level of the trainee. In other words, the task got harder as the trainee"s performance improved. More recently, Shute and Zapata-Rivera (2008) expanded the AT system components to include an instructional model. These authors suggest an effective AT system should include a continuous "Capture-Analyze-Select-Present" cycle. The capture process includes gathering information (cognitive ability, motivation, prior performance levels, etc.) about the learner to develop a student or learner model. The analyze process involves updating the student model to include their current state in relationship to the domain. This could include capturing their current performance on the task, or determining their current level of attentional resources, fatigue, or motivation. The select process is used to deliver the instructional content that is most appropriate based upon the parameters in the student model. Shute and Zapata-Rivera (2008) suggest that this process is "often required to determine how and when to intervene" (p. 281). Finally, after completing the select process, the present process is used to provide the instructional content to the learner through the use of different media and/or software algorithms.

In addition to which approach to take and which components to use, AT system developers should also consider a host of other issues when developing AT systems, for instance timing of feedback. Training could be adapted during task performance, as well as after a scenario or training module has been completed. Task difficulty could be changed during a task, or the type of feedback presented in a debrief could be adapted after a task is over so that specific performance deficiencies could be identified. Additionally, different measures of performance (e.g., number of errors, number of correct decisions, reaction time, etc.) could be used to adjust the adaptive variable (e.g., task difficulty). However, AT developers need to decide which performance measures will be most predictive of individual skill level for the specific task they are trying to train (i.e., tradeoffs between speed and accuracy). Finally, AT developers and researchers should also consider performing empirical evaluations to assess the effectiveness and cost-benefit ratio of AT systems. These evaluations are critical to determine if AT systems are cost effective solutions to one-on-one tutoring. While these are not all the issues that AT developers and researchers must consider, it does highlight that many other questions besides "should I adapt or not?" need to be addressed. Table 1 presents a summary of the models and system components described above, as well as other AT considerations.

Table 1: Summary of Adaptive Training Theory and Review

Adaptive Training Techniques/Models

- Research suggests several different models of AT:
    - Macro-adaptive
    - Micro-adaptive
    - Aptitude Treatment Interaction (ATI)
    - Two Step Approach: combines the micro-adaptive and ATI
- Further research is needed to determine which approach is most optimal and under which circumstances including task characteristics, as well as budget and time constraints.

Adaptive Training System Development

- There are 3 components of an AT system: (1) real time, continuous performance measurement, (2) the variable that will be adjusted, and (3) the adaptation logic (how the adaptive variable will be adjusted based on continuous performance measurement).
- An effective AT system should include a Capture-Analyze-Select-Present cycle
- There are four components of an effective adaptive e-learning system:
    - Content model
    - Learner model
    - Instructional model
    - An adaptive engine
- Strategy choice changes over time and strategy implementation can change based on the difficulty of the problem. This can make it difficult to develop diagnostic algorithms and determine instructional content for AT over long periods of performance.

General Adaptive Training Considerations

- Adaptation can be used to select or edit training scenarios, as well as create a trainee‟s de-brief or after action review.
- Real time performance can be used to adapt task difficulty and there is initial support that shows improved trainee performance.
- Several different measures of performance can be used to adapt training (e.g., error rates, reaction times, accuracy, etc.). Therefore, care must be taken to determine which performance measure will be used to change the adaptive variable.
- Researchers must be able to accurately measure individual difference constructs in order to adapt training to the characteristics of the learner.
- When adapting training, researchers should be aware of how the information is presented and the capabilities of the trainee.
- It is important to know students/trainees prior knowledge in order to predict future learning.
- Interviews with SMEs indicated that their instruction was typically adapted on task and state variables (e.g., anxiety or fatigue) rather than trait variables.
- It is critically important that formal, empirical evaluations are performed to assess the effectiveness and cost-benefit ratio of these systems.

ADAPTIVE TRAINING ASSESSMENT

As mentioned previously, AT provides the trainee with a personalized experience with a training system that can adjust to an individual‟s knowledge and skill level. However, the empirical question of whether AT is a more effective or efficient way to train still remains. Of the empirical research that has been performed, in general, it has been shown that AT is more efficient than fixed difficulty training (Gopher, Williges, Williges, & Damos, 1975; Johnson & Haygood, 1984; Johnson, Haygood, & Olson, 1982; Mané, 1984; Norman & Matheny, 1972; Romero, Ventura, Gibaja, Hervas, & Romero, 2006). Additionally, when trainees utilized programs that adapted in real time, it was found that they performed better than with partial and non-AT (Tennyson & Rothen, 1977). Trainees also showed an increase in motor skills with AT (Cote, Williges, & Williges, 1981; Johnson & Haygood, 1984). Further, research has found that adapting difficulty based on performance can lead to better performance on transfer tasks (Cote et al., 1981; Johnson & Haygood, 1984). It has also been suggested that tasks should steadily increase in difficulty, going through stages of easy, moderate, and hard (Kelley, 1969). Additionally, Kelly (1969) suggests that learning takes place once the trainee receives the correct level of difficulty.

Despite the increase in performance (i.e., effectiveness), AT may not be the most efficient type of training. For instance, one study found that participants who received AT required more trials to achieve the desired criterion (Norman, & Matheny, 1972). Romero et al. (2006) found that participants trained with AT completed tasks in less time than those who trained without, but did not find a significant difference in performance. However, in another study, despite requiring more time to master the material, those trained adaptively were found to have higher posttest perceptual motor performance (Anthony, Goettl, Derek, & Ramirez, 1999).

When developing an AT system, it is important to consider the various factors that may be utilized prior to and during a training scenario (e.g., adaptive guidance, task difficulty, feedback, etc.). Research has shown that participants, who received guidance matched with their performance level, made better decisions than those who received no guidance (Bell, Kanar, Liu, Forman, & Singh, 2006; Bell & Kozlowski, 2002). Additionally, guidance increased self-efficacy and, when coupled with instructional framing, helped participants learn and perform better (Bell et al., 2006).

Training system designers and instructors can also consider adapting the content of feedback. Maddox, Love, Glass, and Filoteo (2008) found that detailed feedback was helpful when applied to rule based tasks and that less detailed feedback was beneficial in information integration tasks. Additionally, detailed feedback marked an improvement early on for rule based tasks and later during information integration (Shute & Towle, 2003). Research has also found interactions between feedback content and aptitudes. For example, Maddox et al. (2008) also found that students who exhibited more exploratory behavior performed better when the feedback they received did not reveal explicit relationships between variables. However, students who showed less exploratory behavior performed better when their feedback was more explicit. The following table provides a summary of the guidance we discovered in regards to assessment techniques in the AT literature.

Table 2: Summary of Assessment Literature

General Adaptive Training Assessment
• In general, support was found for AT over fixed training, however, conflicting results were found in one study (Mané, 1984). • Adapting training based on real time performance scores led to better performance than partially AT (adapted on aptitude and prior achievement only) and non-AT. • Adapting training based on prior knowledge or performance may be a more efficient way to train. • Adapting task stimuli (e.g., task difficulty) was better for transfer performance than operator response variables (e.g., joystick sensitivity). • It is important to not increase the level of difficulty such that the task becomes so difficult that one can‟t perform well regardless of how many attentional resources are devoted to the task.

• Participants in an AT condition required more trials to reach criterion than non-adaptive conditions. • Allowing trainees to progress through training based on performance criteria led to higher posttest perceptual-motor performance than groups that were forced through training components. • Adaptation based on individual-level performance criteria is more effective than adaptation based on group-level performance criteria. • Workload can be assessed in real time with physiological measures to adapt training.
Adaptive Guidance
• Participants who received guidance on what to study, which was tailored to their level of performance, made better decisions than participants who did not receive guidance (but were allowed to review any material they wanted). • A combination of adaptive guidance and framing of instructions ("you must" review X) led to improved performance for high ability participants and participants with low motivation to learn compared to a traditional learner control condition and an adaptive guidance condition with less imposing instructions ("you could" review X). • Adaptive guidance increased self-efficacy in early training trials.
Adaptive Task Difficulty
• AT has been more effective than fixed difficulty when training motor skills. • Adapting task difficulty based on performance led to improved transfer performance over fixed task difficulty. • When adjusting task difficulty, you should not move too quickly from easy to difficult. You should adjust slowly, moving from easy to moderately difficult to hard. • Adjusting primary task difficulty based on secondary task performance increased performance on the secondary task but decreased primary task performance. • Adjusting primary task difficulty based on primary task performance increased performance on the primary task but decreased secondary task performance.
Feedback
• Feedback may need to be adapted based on characteristics of the task. • More detailed feedback was beneficial for rule-based tasks while less detailed feedback was beneficial for information-integration tasks. Further, detailed feedback was beneficial early in training for rule-based tasks and beneficial for information integration tasks later in training. • An ATI has been found between exploratory behavior and training content (i.e., amount of information provided in feedback). Learners who exhibited more exploratory behavior performed better when the feedback they were given did not explicitly explain relationships between variables, i.e. they had to be induced. Those who showed less exploratory behavior did better when they were given explicit rules for the relationships.

INDIVIDUAL DIFFERENCES

Throughout the psychological and educational literature there is the premise that the influence of individual differences on learning processes and learning outcomes is dependent upon the learning environment or situational variables. This is known as aptitude-treatment interaction (ATI), where aptitude refers to a person‟s knowledge, skills, and personality traits and treatment refers to the environment that supports the learning (Cronbach & Snow, 1977). This premise focuses on the match between individual differences and the learning environment and how changes in the learning environment are best suited for specific patterns of aptitudes (Shute & Towle, 2003). As noted by Shute and Towle (2003), "the goal of ATI research is to provide information about learner characteristics that can be used to select the best learning environment for a particular student to optimize learning outcome[s]" (p. 106). Research suggests that intellectual abilities, cognitive styles, learning styles, prior knowledge, anxiety, achievement, motivation, and self-efficacy are important predictors of learning (Mödritscher, Garcia-Barrios, & Gütl, 2004).

In our review of the AT research, we classified individual differences into 1 of 3 broad categories: cognitive ability, working memory capacity, and prior knowledge. Cognitive ability refers to an individual‟s capacity to perform higher mental processes of reasoning, remembering, understanding, and problem solving (Bernstein, Penner, Clarke-Stewart, & Roy, 2006). The review of the literature suggests that cognitive ability can be broken down into distinct abilities that are used to accomplish various types of tasks. These abilities include, but are not limited to, verbal ability, spatial ability, attention (focused, sustained, divided, and alternating), motor skills, and perceptual skills. In most cases, research suggests that individuals with higher levels of cognitive ability yield superior performance in work and academic settings, as well as in everyday life (Beier & Oswald, 2009).

The second individual difference category found in the AT literature was working memory capacity (WMC). WMC can be considered a subset of cognitive ability and can be thought of as a short-term system involved in the control, regulation, and active maintenance of a limited amount of information with immediate relevance to the task at hand (Miyake & Shah, 1999). WMC signifies that an individual‟s working memory can store and process only a limited number of items at one time (Batka & Peterson, 2005). As an individual difference variable, this suggests that some people have more of this construct and some have less (DeCaro, Thomas, & Beilock, 2008). In general, research has found that individuals with more WMC perform better on a variety of tasks (e.g., reasoning, problem solving, comprehension, etc.).

The third and final individual difference category found in the AT literature was prior knowledge. Prior knowledge can be considered the relevant knowledge that participants have acquired before the training intervention or task (Shute & Zapata-Rivera, 2008; Park & Lee, 1996) and includes other individual difference variables such as expertise and general aptitude. Like all individual difference variables, the amount of prior knowledge varies between individuals (e.g., experts vs. novices). In addition, research suggests that prior knowledge is the most salient factor in regard to predicting future learning (Shute & Zapata-Rivera, 2008).

It is important to note that through our AT research review we found information on other individual difference constructs that did not fit into our three overall categories (e.g., motivation and self-efficacy). These constructs were not included in our review for a variety of reasons. For example, some individual difference variables had little to no empirical evidence regarding their effects within an AT setting. Additionally, some constructs are extremely variable over time and can change during a single training session (Park & Lee, 1996). For those variables which had empirical evidence, the articles had a number of methodological flaws and measurement issues that prevented us from including the results in our guidelines. Without sound empirical evidence we were not able to provide practical conclusions or guidelines from these studies. The following table provides a summary of the guidance we discovered in regards to individual differences in the AT literature.

Table 3: Summary of Individual Differences Literature

Components of Cognitive Ability
• High cognitive ability participants perform poorly in the lab, but well in the field. Further research must be performed to investigate interactions between cognitive ability and task characteristics (including whether devoting attentional resources will increase performance). • High cognitive ability trainees perform poorly after changes in complex tasks occur as compared to low cognitive ability participants. Researchers suggest that high cognitive trainees appear to suffer greater performance decrements because they have "more to lose" due to their already high level of performance. • Multimedia instructional content could be adapted based on spatial ability. When receiving instruction with text and pictures, high spatial ability participants outperformed high spatial ability participants who received text only, as well as low spatial ability participants. • Using a spatial aid led to higher performance for high spatial orientation participants than low spatial ability participants. This suggests that it may be beneficial to adapt training content based on spatial orientation. • It may be important to account for differing levels of verbal comprehension because trainees with high verbal comprehension typically outperform low verbal comprehension trainees on declarative knowledge tasks. While low verbal comprehension trainees also report greater cognitive effort than high verbal comprehension trainees. • The visualizer-verbalizer construct has several different dimensions: cognitive ability, cognitive style, and learning preferences. These different dimensions may moderate performance in multimedia environments and it could be beneficial to adapt instructional content based on how individuals score on the visualizer-verbalizer construct. • Verbalizing task strategies was beneficial for male trainees overall compared to no verbalization groups. However, verbalization was only beneficial for females later in training.
Working Memory Capacity (WMC)
• Trainees with a higher WMC perform better on multimedia tasks than participants with lower WMC. • Trainees with a lower WMC perform better with simultaneous presentation of animation and narration than with sequential presentation. • Trainees with higher WMC learn rule-based tasks faster than low WMC trainees. • Trainees with low WMC learn information integration tasks faster than high WMC trainees. Researchers suggest that trainees with a lower WMC rely on shortcuts while high WMC use more complex strategies which may be why it takes them longer to learn information integration tasks.
Prior Knowledge, Aptitude, Expertise
• Prior knowledge interacts with feedback interventions such that participants with little prior knowledge on the task benefited from the presentation of feedback (versus no feedback). • Experts and novices differ in their strategy choices, with experts using the most effective strategies more often. This suggests that training on strategy choice development and implementation should be adapted based on expert/novice differences.

Self-efficacy
• Collaborative training with a more experienced partner is more beneficial for low self-efficacy participants than individual training. • Collaborative training only weakens rather than removes the negative effect of low self-efficacy.

CONCLUSION

There are several AT approaches, each with associated costs and benefits to learning. For example, one can use an ATI approach, a micro-adaptive approach, or a combination of approaches. Regardless of which AT approach is taken, instructional system developers and instructors must always focus on the ultimate goal of all training systems. That is, providing instruction that enhances the acquisition of attitudes, concepts, knowledge, rules, or skill that results in improved performance (Goldstein, 1991). However, while most training approaches focus on altering the individual to fit the training, AT seeks to alter itself to meet the trainees‟ needs. This distinction is important because altering the training system itself seems more efficient and effective than trying to change trainees‟ attitudes, aptitudes, or skills. This may be because training that adapts to the trainee is similar in style to one-on-one tutoring, which has been shown to be more effective than traditional classroom learning (Bloom, 1984).

Despite almost 40 years of research on AT, our review has illustrated a need for more empirical research in this domain. There were numerous articles we excluded that reported on AT system development. However, these studies did not report on evaluation criteria. In order for AT systems to move out of the lab, researchers and developers must report on their effectiveness. Nevertheless, AT has demonstrated its potential to provide more effective and efficient training. A few studies have shown that AT is superior to fixed difficulty training (Gopher, Williges, Williges, & Damos, 1975; Johnson & Haygood, 1984; Johnson, Haygood, & Olson, 1982; Mané, 1984; Norman & Matheny, 1972; Romero, Ventura, Gibaja, Hervas, & Romero, 2006), but little agreement can be found on the best way to adapt training, or on which variable to adapt for each learner. Therefore, future research is needed to determine if training variables should be adapted during task performance, after the task is complete, or both and what variables to use to adapt the training. Further, there is a lack of studies directly comparing different types of AT. Therefore, empirical evaluations are needed that assess the cost-benefit ratio and effectiveness of different AT systems. There is also evidence to suggest that individual differences are useful for adapting training. However, more research is needed to determine which individual difference variables are most useful, and how these variables interact. Our conceptualization of AT suggests that instructors and AT system designers can adapt task variables (e.g., task difficulty) and training content/material. However, future research should investigate the differences between adapting task variables versus training content/material. Lastly, research must be performed to determine which performance measures are most predictive for the particular skills and tasks that are being trained.

THIS PAGE INTENTIONALLY LEFT BLANK

REFERENCES

References marked with an asterisk indicate studies included in the annotated bibliography. Article summaries are provided in the Appendices (page numbers are provided in parentheses at the end of the reference).

Allworth, E. & Hesketh, B. (1999). Construct-oriented biodata: Capturing change-related and contextually relevant future performance. *International Journal of Selection and Assessment,* 7, 97-111.

*Anthony, M. K., Goettl, B. P., Derek, K., & Ramirez, K. G. (1999). Adaptive remediation in complex skill training. *Proceedings of the Human Factors and Ergonomics Society, 43*, 1146-1150. **(B-2)**

*Batka, J. A. & Peterson, S. A. (2005). The effects of individual differences in working memory on multimedia learning. *Proceedings of the Human Factors and Ergonomic Society, 49,* 1256-1260. **(C-2)**

*Beier, M. E. & Oswald, F. L. (2009). Adaptability and complex performance: Can you really be too smart for your own good? Manuscript submitted for publication. **(C-4)**

*Bell, B. S., Kanar, A., Liu, X., Forman, J. & Singh, M. (2006). Adaptive guidance: Effects on self-regulated learning in technology-based training. Unpublished manuscript, Center for Advanced Human Resource Studies, Cornell University, Ithaca, NY. Retrieved from http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1455&context=cahrswp. **(B-4)**

*Bell, B. S. & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*(2), 267-306. **(B-6)**

Bernstein, D. A., Penner, L. A., Clarke-Stewart, A., & Roy, E. J. (2006). *Psychology* (7th ed.). New York: Houghton Mifflin.

Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4-16.

*Chapman, R., Ryder, J., Szczepkowski, M., & Benton, D. (2006). Full-cycle adaptive instructorless training. *Proceedings of the Human Factors and Ergonomics Society, 5*, 2664-2668. **(A-2)**

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd edition). Mahwah: Lawrence Erlbaum Associates, Inc.

*Cote, D. O., Williges, B. H., & Williges, R. C. (1981). Augmented feedback in adaptive motor skill training. *Human Factors, 23*, 505-508. **(B-8)**

Cronbach, L. J., & Snow, R. E. (1977). Attitudes and instructional methods. New York: Irvington.

*Crooks, W. H., & Roscoe, S. N. (1973). Varied and fixed error limits in automated adaptive skill training. *Proceedings of the Human Factors and Ergonomics Society, 17*, 272-280. **(B-9)**

*Day, E. A., Boatman, P. R., Kowollik, V., Espejo, J., McEntire, L. E., & Sherwin, R. E. (2007). Collaborative training with a more experienced partner: Remediating low pretraining self-efficacy in complex skill acquisition. *Human Factors, 49*(6), 1132 – 1148. **(C-6)**

*DeCaro, M. S., Carlson, K. D., Thomas, R. D., & Beilock, S. L. (2009). When and how less is more: Reply to Tharp and Pickering. *Cognition, 111,* 397-403. **(C-8)**

*DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*, 284-294. **(C-10)**

Erkstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Goldstein, I. L. (1991). Training in work organizations. In M. D. Dunnett & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 2, pp. 507-619). Palo Alto, CA: Consulting Psychologists.

*Gopher, D., Williges, B. H., Williges, R. C., & Damos, D. L. (1975). Varying the type and number of adaptive variables in continuous tracking. *Journal of Motor Behavior, 7(3)*, 159-170. **(B-11)**

*Johnson, D. F., & Haygood, R. C. (1984). The use of secondary tasks in adaptive training. *Human Factors, 26*, 105-108. **(B-12)**

*Johnson, D. F., Haygood, R. C., & Olson, W. M. (1982). Yoked design and secondary task in adaptive training. *Proceedings of the Human Factors and Ergonomics Society, 26*, 21-24. **(B-13)**

*Kelley, C. R. (1969). What is adaptive training? *Human Factors, 11*(6), 547-556. **(A-3)**

*Kline, K. A. & Catrambone, R. (2009). The influence of spatial ability on multimedia learning. *Proceedings from the 53rd Annual Meeting of the Human Factors and Ergonomics Society,* 1249-1253. **(C-11)**

*Krause, U., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning & Instruction*, *19*(2), 158-170. **(C-12)**

Lang, J. W. B. & Bliese, P. D. (2009). General mental ability and two types of adaptation to unforeseen change: Applying discontinuous growth models to the task-change paradigm. *Journal of Applied Psychology, 94*, 411-428.

*Lemaire, P. & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children‟s learning of multiplication. *Journal of Experimental Psychology: General, 124*(1), 83-97. **(B-15)**

*Ludwig, J., Ramachandran, S. & Howse, W. (2002). Developing an adaptive intelligent flight trainer. *In Proceedings of the Industry/Interservice, Training, Simulation and Education Conference (I/ITSEC 2002).* **(B-16)**

*Maddox, W. T., Love, B. C., Glass, B. D., & Filoteo, J. V. (2008). When more is less: Feedback effects in perceptual category learning. *Cognition*, 108, 578-589. **(B-17)**

*Mané, A. M. (1984). Acquisition of perceptual-motor skill: Adaptive and part-whole training. *Proceedings of the Human Factors and Ergonomics Society, 28*, 522-526. **(B-18)**

*Mayer, R. E. & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95(4),* 833-846. **(C-14)**

Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 312-320.

*McGregor, M. U., Schunn, C. D., & Saner, L. D. (2007). Expertise as effective strategy use: Testing the adaptive strategies model in the ill-structured domain of leadership. *(Technical report 1204). Arlington, VA: US Army Research Institute for the Behavioral and Social Sciences.* **(C-15)**

Miyake, A. & Shah, P. (Eds.) (1999). Models of working memory. New York: Cambridge University Press.

*Mödritscher, F., Garcia-Barrios, V. M., & Gutl, C. (2004). The past, the present and the future of adaptive e-learning: An approach within the scope of the research project AdeLE. *Proceedings of the Interactive Computer Aided Learning Conference,* Villach, Austria. **(A-4)**

Mulgand, S. S., Asdigna, M., Zacharias, G. L., Krishnakumar, K., & Dohme, J. A. (1995). An intelligent tutoring system for simulation-based helicopter flight training. *Proceedings of the 1995 Flight Simulation Technologies Conference.*

*Norman, D. A. & Matheny, W. G. (1972). Need for within-trail feedback as a function of task similarity in adaptive training of manual control. *Proceedings of the Human Factors and Ergonomics Society*, *16*(1), 136-139. **(B-20)**

*Pak, R., Rogers, W. A., & Fisk, A. D. (2006). Spatial ability subfactors and their influence on a computer-based information search task. *Human Factors*, 48(1), 154-165. **(C-16)**

*Park, O. C. & Lee, J. (1996). Adaptive instructional systems. In Jonassen, D. (Ed.). *Handbook of research for educational communications and technology* (pp.651-684). New York: MacMillan Publishers. **(A-6; C-18)**

Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology, 46*, 128-148

*Romero, C., Ventura, S., Gibaja, E. L., Hervas, C., & Romero, F. (2006). Web-based adaptive training simulator system for cardiac life support. *Artificial Intelligence in Medicine, 38*(1), 67-78. **(B-21)**

Rouse, W. B. (1988). Adaptive aiding for human/computer control. Human Factors, 30(4), 431-443.

Scerbo, M. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M Mouloua (Eds.). *Automation and Human Performance: Theory and Applications* (37-63). Mahwah, NJ: Lawrence Erlbaum Associates.

Schaafstal, A., Hoeft, R. M., & van Schaik, M. (2002). Training a team with simulated team members. *Proceedings of the Human Factors and Ergonomics Society, Training*, 2026-2029.

*Scielzo, S., Cuevas, H. M., & Fiore, S. M. (2005). Investigating individual differences and instructional efficiency in computer-based training environments. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting,* 1251-1255. **(C-19)**

*Shute, V. & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist, 38*(2), 105-114. **(A-8)**

*Shute, V. J. & Zapata-Rivera, D. (2008). Adaptive technologies. In J. M. Spector, D. Merrill, J. van Merriënboer, & M. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (3rd Edition) (pp. 277-294). New York, NY: Lawrence Erlbaum Associates. **(A-10)**

Snow, E. R. (1980). Aptitudes, learner control and adaptive instruction. *Educational Psychologist, 15,* 151-158.

*Tennyson, R. D., & Rothen, W. (1977). Pretask and on-task adaptive design strategies for selecting number of instances in concept acquisition. *Journal of Educational Psychology, 69*(5)*,* 586-592. **(C-21)**

Tharp, I. J. & Pickering, A. D. (2009). A note on DeCaro, Thomas, and Beilock (2008): Further data demonstrate complexities in the assessment of information integration category learning. *Cognition,* 111, 411-415.

Tyroller, M. (2005). Effekte metakognitiver prompts beim computerbasierten statiskiernen [Effects of metacognitive prompting on computer-based learning of statistics]. Doctorial Dissertation, University of Munich, Germany.

Williges B. H. & Williges, R. C. (1977). Learner-centered versus automatic adaptive motor skill training. Journal of Motor Behavior, 9, 325-331.

Williges, B. H., Williges, R. C., & Savage, R. E. (1977). Models for automated motor skills training. *Proceedings of the Human Factors and Ergonomics Society 21st Annual Meeting,* 18-22

*Williges, R. C., & Williges, B. H. (1978). Critical variables in adaptive motor skills training. *Human Factors, 20*, 201-214. **(A-12)**

*Wilson, G. F. & Russell, C. A. (2007). Performance enhancement in and uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors, 49*(6), 1005-1018. **(C-23)**

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A
ADAPTIVE TRAINING REVIEW AND THEORY

Chapman, R., Ryder, J., Szczepkowski, M., & Benton, D. (2006). Full-cycle adaptive instructorless training. *Proceedings of the Human Factors and Ergonomics Society, 5*, 2664-2668.

Training Intervention(s): AAR, Practice
Type of Study: Review

This article discusses the Synthetic Teammates for Realtime Anywhere Training and Assessment (STRATA) adaptive training tool. The tool is currently being used in an F/A-18 simulation to teach Close Air Support (CAS). It trains teamwork and cross-platform skills in the simulation to individuals without the need of an instructor and provides complete training from briefing to after-action review.

Training scenarios are selected using intelligent software functions and are based on each user's profile. The profile is used to determine the training goals of a session and includes the user's level of proficiency, amount of practice, and currency of exposure to the topics. Next, scenario templates are instantiated, which generate a training brief and produce performance measures tailored to the specified goals. The appropriate scenarios are then completed and performance is recorded. A guided after-action review is completed by a cognitive agent and is specific to each user‟s performance.

When choosing a scenario template, STRATA initially selects scenarios in which the user has performed weakly. Next, it picks scenarios for which the user‟s performance has yet to be assessed. Lastly, STRATA will implement template scenarios that have not already been presented to the user, which reduces redundancy in training scenarios.

Events in the scenarios are linked to training goals. Since the training goals are defined by observable performance measures, evaluation of performance can occur automatically after the scenario is complete. During the after-action review, segments of the scenario are played back and the cognitive agent indicates instances of appropriate and inappropriate actions. The user can choose to let the cognitive agent determine the order and content of the review (adaptive guidance) or freely explore the contents themselves through the use of the graphical user interface (GUI) in the briefing outline. The GUI includes the list of training goals and indicates their success or failure in achieving them, as well as a scenario replay tool, a scenario communications log, and a scenario map.

The advantage of this system is that it allows full-cycle instructorless training that is adapted to the trainee. Additionally, the authors state that the approach and infrastructure used in STRATA can be applied in other training domains that use simulation. Adaptive training has the potential to be beneficial to the learner, but the effectiveness of the STRATA adaptive training system was not assessed in this article. Also, the description of the system and its underlying framework was vaguely described. Therefore, future work should focus on evaluating the training effectiveness of this system, determining if adaptation to prior performance is beneficial and determining performance differences when adaptive guidance is used as compared to free exploration.

Kelley, C.R. (1969). What is adaptive training? *Human Factors, 11*(6), 547-556.

Adaptive Training Intervention(s): Theory building & developing an adaptive training system
Type of Study: Theoretical

The following theoretical article discusses the important characteristics of an adaptive training system. In this article, adaptive training is defined as "training in which the problem, the stimulus, or the task is varied as a function of how well the trainee performs" (p. 547). The author argues that adaptive training is effective because learning takes place when the appropriate level of difficulty is provided to the trainee. Since an adaptive training system is manipulated to accommodate the learner based on his/her skill/knowledge level, it is portrayed as the best way for a trainee to learn. For example, it is suggested that when a task is too easy, a trainee learns very little from performing the task. Likewise, if a task is too difficult the trainee may become overwhelmed and benefit little from the training.

The author proposes three components of an adaptive training system: (1) continuous performance measurement, (2) the adaptive variable (i.e., task variable that will be adjusted), and (3) the adaptive logic which automatically changes the adaptive variable as a function of how the trainee performs on the task. The author discusses key points that need to be considered when developing an adaptive training system. The following five points were considered critical aspects of an adaptive training system: (1) selecting an adaptive variable, (2) measuring performance, (3) adaptive logic, (4) error standard and difficulty level, and (5) knowledge-of-results display (i.e., feedback).

The first aspect of an adaptive training system that a designer should consider is the selection of an adaptive variable. The author provides a list of tasks that can be manipulated. For instance, adaptivity can be adjusted by changing the simulated environment, the level of stress applied to participants, altering displays, and changing the work load required. The author stressed that regardless of the adaptive variable selected, the level of difficulty should follow the easy to moderately difficult to hard pattern.

The second important aspect of an adaptive training system is how performance should be measured. It is suggested that the most important aspect of performance measurement is to do so reliably. The less reliable the measurement system, the more likely errors will occur in the selection of the adaptive variable difficulty. The third aspect of an adaptive training system that should be considered is the adaptive logic. The adaptive logic is the mechanism which ties the adaptive variable to performance measurement. The adaptive logic can take different forms. It can be a mathematical equation, depicting the relationship between performance and the adaptive variable of interest. The adaptive logic can also be expressed in terms of adjustment rules. For example, the level of difficulty could be altered based on a predetermined level of performance.

The designer should also take into account the error standard and difficulty level of an adaptive training system. The author notes that when a tight standard error or difficulty is set, the adaptive system is much easier than when a looser error standard is set. At strict/tight error tolerance levels, trainees must perform at an easier level of difficulty until they have achieved a higher level of performance before the system progresses to a more difficult level. At larger error tolerance levels, trainees will progress more quickly to a more difficult training system level, but they will not be operating at a high level of performance. The final aspect of an adaptive system is the knowledge-of-results display (i.e., feedback). It is suggested that it is important to provide the trainee with knowledge about his/her progress.

The author concludes by warning the reader that adaptive training is "by no means a panacea for the problems of training" (p. 555). If a training system is poorly developed, making it adaptive will not make it a better training system. The author also warns that a training system can be made less effective by being made adaptive in an incorrect manner. Lastly, it is suggested that for each adaptive training system, decisions must be made in regards to the adaptive variable(s), performance measurement, and the adaptive logic connecting both. Each adaptive training system must be developed and evaluated based on its own individual merit.

Mödritscher, F., Garcia-Barrios, V. M., & Gütl, C. (2004). The past, the present and the future of adaptive E-learning: An approach within the scope of the research project AdeLE. *Proceedings of the Interactive Computer Aided Learning Conference*, Villach, Austria.

Adaptive Training Intervention(s): None
Type of Study: Review

The goal of the AdeLE project is to develop a new framework for adaptive e-learning. Within this context the article provides a review of adaptive e-learning in terms of its history and implementation. Historically, the four different strategies that have been used to adapt learning are the macro-adaptive, aptitude-treatment interaction, micro-adaptive, and the constructivistic-collaborative approaches. The macro-adaptive approach allows different alternatives to be selected for a few main components such as learning objectives, levels of detail, and type of delivery system. This approach provides different instructional alternatives based on the student‟s learning goals, general abilities, and achievement levels.

In the aptitude-treatment interaction (ATI) approach, students receive different kinds of instruction based on specific characteristics. Research has suggested that the most important learner characteristics are intellectual abilities, cognitive styles, learning styles, prior knowledge, anxiety, achievement, motivation, and self-efficacy. In this approach, the level of control that the student has over the learning experience can also be manipulated. Snow (1980) described three different levels of control. A student can have full or partial control over a task scenario or control over pace for fixed tasks within the learning material. However, only a few instances regarding the benefits of e-learning have been researched and studies have shown that the success of different levels of learner control is strongly dependent on the students‟ aptitude.

The micro-adaptive approach is used to identify and remedy specific learning problems encountered during instruction. In order to fine tune instruction for students, researchers have used techniques which monitor student behavior and performance by looking at variables such as response errors, latencies, and emotional states. Many micro-adaptive instructional models have been developed which use particular theories of learning in their implementation. This style of adaptive e-learning, which is similar to having a one-on-one tutor present, contains two main processes: the diagnostic process and the prescriptive process. The diagnostic process assesses learner characteristics and the prescriptive process optimizes the learning experience by changing the order and content of instruction in accordance with students‟ ability and recent performance.

The constructivistic-collaborative approach allows the learner to be active in their learning experience. The student constructs his/her knowledge through experiences with the learning material. In this approach, context, learning activities, cognitive structures, and time on task are important factors to consider. Also, the student‟s motivation should be considered when deciding how to deliver material on a topic. Five characteristics that have been found to make collaborative learning more effective are participation, social behavior, performance analysis, group processing and conversation skills, and primitive interaction.

Each of the four approaches mentioned above has corresponding systems which were designed to implement them. The first systems developed were those designed to apply the macro-adaptive approach. From the early 1900s to the 1960s macro-adaptive approaches were commonly used in classrooms. They allowed students to control the pace of their learning and allowed instruction to be personalized for each student based on current and past performance. More effective macro-adaptive systems called Computer-managed Instructional (CMI) systems were developed in the 1980s. These systems could recommend appropriate content based on an individual student‟s learning needs and allowed instructors to control the students learning experience. Intelligent Tutoring Systems (ITS) are micro-adaptive systems that are built to simulate the experience between an instructor and a student. These systems are able to assess performance, current state, and reasoning strategies and then use those assessments to provide the student with personalized learning material. A two level model was proposed in which the micro-adaptive approach was combined with the use of aptitude variables. In this model the system would use the student‟s attributes to determine the conditions of instruction while also adapting the instruction to the student by changing such conditions as display time, sequence, format of examples, etc. (Tennyson et al, 1998). Adaptive Hypermedia Systems (AHS), developed in the 1990s, combine adaptive interfaces with hypermedia systems. With these systems there are two kinds of adaptation: content-level and link-level. In content-level adaptation, the content is displayed in a different format or with different sequencing. In link-level adaptation, students can follow links to content, and certain links appear only at certain times, for instance when the student has mastered the content from a prerequisite link. These systems are easy to create and widely used, however, they are not supported by empirical or theoretical research. Some researchers have also suggested that, if there are errors in programming, students will be led to content that is confusing or irrelevant. For these types of systems to be effective, it is necessary to assess their current learning state. The approaches and current systems described above contain implications for requirements of adaptive instructional systems that are currently being developed. Theoretically, an ideal adaptive instructional system should contain a CMI component and a Learning Management System (LMS). The CMI component combines tools for creating content and managing the curriculum. The LMS component would be created by combining an AHS system with ITS technology. The authors suggest that the instructional system should include web-based interface. Using AI techniques, the system could be adapted in accordance with each student‟s current learning performance while taking into account different learning theories for content presentation. The authors conclude that any new adaptive system should integrate aspects of all four adaptive instructional approaches in order to be fully effective.

Park, O. C. & Lee, J. (1996). Adaptive instructional systems. In D. H. Jonassen (Ed.). *Handbook of research for educational communications and technology* (pp. 651–684). New York: MacMillan Publishers.

Adaptive Training Intervention(s): None
Type of Study: Theoretical

This book chapter describes the history of adaptive training, components of adaptive training systems, and describes three traditional approaches to providing adaptive instruction. Additionally, the authors present an alternative approach for providing adaptive instruction. Park and Lee (1996) explain that adaptive training has a long history, but that this history mainly focused on one-on-one tutoring. More recently, with the advent of computers, other systems of adaptation have been implemented in both educational formats and in industry via computer models. The authors state that because of this influx of high powered, yet affordable, computer systems and increased accessibility, new and more efficacious types of adaptive training are available. However, there is not a large amount of research on the effectiveness of different types of adaptive training systems and approaches.

According to the authors, there are three main types of adaptive training approaches: macro level, Aptitude Treatment Interaction (ATI), and micro level. The authors describe these models in detail, as well as provide strengths and weaknesses of each. In general, macro-adaptive systems provide minimal individual adaptation. In these types of systems, instruction is typically adapted based on general achievement levels (determined prior to instruction) and curriculum structure is student paced (i.e., student determines how to proceed through the units/modules). Macro-adaptive models are the most widespread type of adaptive training systems and are used in many classrooms. However, the authors mention that adaptation methods are primitive and unsystematic as adaptations can differ widely based on the teachers implementing the systems. The second approach to instructional adaptation, Aptitude Treatment Interaction (ATI), allows for adaptation based on learners‟ current aptitudes or individual differences. The idea is that if instruction is more suited toward the characteristics of the student, optimal learning will take place. The authors provide a review of the literature on different types of individual difference variables that have been related to performance and, thus, could be used for adaptation (e.g., intellectual ability, cognitive styles, learning styles, prior knowledge, etc.). They mention some weaknesses such as the ability level may change over time and become less optimal for determining adaptations. Additionally, the characteristics may be task dependent. The third type is the micro level approach in which there is continual, on task performance measurement (e.g., errors, response times, etc.) that forms the basis of adaptations. The authors describe several different types of micro-adaptive systems, as well as different computational models that can be used for performance diagnosis. A strength of the micro model as compared to the macro model is that adaptations are based on analyses of performance data and are more systematic. A weakness with this approach is that development of real time, diagnostic algorithms are difficult to develop and costly.

To eliminate some of the weaknesses with the individual models, the authors propose a Two-Step approach which integrates the ATI and micro approaches. This approach is based on the theory that pre-task aptitudes have less predictive value of future performance than on task aptitude measurement. First, the ATI approach would be used to place the learner in the appropriate level and to determine the instructional approaches to take upon beginning instruction. Second, the authors recommend utilizing the micro level approach to continuously diagnose student performance and prescribe interventions during instruction. That is, after the initial testing provided by the ATI, the following intervening adaptations will be based off of on task analysis as opposed to pre-task analysis or post-task analysis only. A strength of the Two-Step approach is that it eliminates the unstable nature of adapting instruction based solely on individual difference characteristics and capitalizes on the strengths of the micro level approach. Like the micro approach, these types of systems can be time consuming and costly to build. Park and Lee conclude by providing a model which describes the components that a Two-Step adaptive training system would need.

The authors suggest several areas of future research in adaptive training. For instance, they suggest extensive experimental research to determine the effectiveness of different approaches of adaptive training because many systems are based off of theoretical assumptions which have not been experimentally verified. They also suggest future research should determine which individual difference characteristics would be optimal for adaptation and how these characteristics will change over time. Previous adaptive training systems have focused on declarative and procedural knowledge type tasks which performance is easier to measure and misconceptions are easier to diagnose. Therefore, Park and Lee suggest future research looks into performance measurement and diagnosis techniques for more complex tasks such as decision making and problem solving. There were limitations with Park and Lee‟s review. First, there is a very limited explanation of Park and Lee‟s model provided. They provide a diagram of their suggested model, however, the explanation of the model is limited and ambiguous in its definitions of stages. Additionally, there are no specific procedures or technical guidelines mentioned for actual implementation of their approach. In general, they omitted a discussion of a variable that may have a relationship with the effectiveness of some adaptive training systems. Specifically, there should be some mention of the frequency in which adaptation should take place during instruction (e.g., the effectiveness of frequent versus minimal adaptation).

Shute, V. & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist, 38*(2), 105-114.

Adaptive Training Intervention(s): None
Type of Study: Theoretical

In the context of e-learning research, the term aptitude-treatment interaction (ATI) is used to describe the relationship between the individual differences of a learner (such as incoming knowledge and skills, cognitive ability, and personality traits) and different learning environments or interventions. The goal of ATI research is to optimize the e-learning experience for each individual by identifying characteristics or attributes of the learner that can be used to individualize training. In the past it has been hard to identify ATIs because of data that was "noisy", replete with extraneous and confounding variables. An empirical study conducted by Shute (1993) sought to measure one trait that was proposed to interact with type of learning environment: exploratory behavior. Participants were placed in one of two conditions, either rule application or rule induction. In the rule application condition the participants were given feedback which identified all relevant variables, as well as the relations between them. In the rule induction condition, the feedback that was given to the participants identified the relevant variables, but the relationship between them was not provided, and therefore had to be deciphered by the participants. Exploratory behavior was measured for both conditions and all participants were given four post-tests that measured knowledge and skills that were learned from the training program. Analysis of the results revealed no differences in exploratory behavior between the two groups, however, an aptitude-treatment interaction was found between exploratory behavior and learning environment. Participants who exhibited more exploratory behavior performed better if they were assigned to the rule induction condition, while participants who exhibited less exploratory behavior performed better if they were assigned to the rule application condition. The results of this study have important implications for designers of e-learning programs. If individual traits that are related to learning can be identified and mapped to associated learning environments, training could be adapted for any learner.

Shute and Towle (2003) identified four components that are necessary for designing an effective e-learning program. The first component, called the *content model,* contains all knowledge that is domain-related. The authors call this a "knowledge map" which stipulates what should be taught and assessed. In addition, it provides the basis for assessment and diagnosis of learning problem areas, as well as how to improve. The content that is used must be arranged so that the delivery system can adapt it to different users. To make this possible the authors suggest the use of learning objects (LOs). The authors describe learning objects as small learning units that can be put together and reused; this includes materials such as videos, procedures, tutorials, stories, assessments, and simulations. Arranging these pieces together creates a cohesive learning experience. Current practice is to arrange the LOs before they are presented, but an adaptive system would need to arrange them during the training in order to customize content to individuals. This can be accomplished by dividing the LOs into sub-collections that each represent a different skill or category of knowledge. Knowledge structures can be established to assess the status of an LO, find problem areas and fix them and determine when it is appropriate to move on to the next LO. Different nodes in the knowledge structure represent different kinds of knowledge, and recommendations are made on how different kinds of knowledge are introduced and assessed. The authors argue that instruction on basic knowledge should be straightforward and its assessment should involve making sure the student can recognize or reproduce the information (a definition, formula, rule, etc.). For procedural knowledge (which defines step-by-step information or relations among steps) instruction should involve more hands-on practice, so that the learner can practice the skill. Learners who are assessed in procedural knowledge should be able to complete the procedure, or apply a rule. Another knowledge type is conceptual knowledge which connects basic and procedural knowledge into a "big picture". Learners should be taught this kind of information with analogies or diagrams that draw all of the information together in a clear manner. Conceptual knowledge assessment should require the user to use basic and procedural knowledge in novel situations, predict outcomes, or make strategies.

The second component necessary for e-learning is the *learner model*. The learner model contains information about the individual learner‟s knowledge and progress in relation to the content model (knowledge map) and also includes important aspects about the learner so that instruction can be individualized. Information that comes into the system from assessments completed by the learner is used to determine what should be taught next. This makes the learning experience adaptive to each individual using the system. Two types of information can be assessed in order to adapt the system to a learner: domain-dependent or domain-independent. Assessment of domain-dependent information is based on data from pretests, as well as performance outcomes. This allows the system to determine which LOs have been sufficiently learned and which ones require further instruction. Assessment that is domain-independent is based on information about the learner, such as relevant cognitive or personality traits. This allows the system to decide the best format and order of presentation of LOs for each learner. Specific learner traits may be linked to certain kinds of content delivery modes, formats, or sequences. An example given by the authors is that learners who are high on inductive reasoning skills learn better when practice and examples are presented prior to the concept while learners low on inductive reasoning learn better when presented with the concept first.

The third necessary component for e-learning is the *instructional model* which decides the best possible learning path and decides how to present the learning material to a given learner. The authors base their suggestions for the design of this model on Gagne‟s “events of instruction”. In Gagne‟s model, the instruction must first capture the learner‟s attention, let them know what the course objectives are, and stimulate retrieval of prior knowledge. Then the learning material is presented and guidance is provided. Next, performance is assessed and feedback is given to the learner. Finally, the performance is evaluated and generalization of the material is promoted. Other general guidelines for instructional design suggested by Shute and Towle include the use of multiple representations to illustrate a concept, designing to elicit manipulation and creation of relevant representations, the use of a final activity that helps the learner integrate the material into a “big picture” representation. Finally, the system should be intuitively designed in order to reduce the amount of time the student spends learning how to use it. Two approaches to student modeling can be used in order to decide specifically which LO or LOs should be selected for a given learner: SMART and BIN. The SMART approach (Student Modeling Approach to Responsive Tutoring) separates low-level knowledge and skills into basic knowledge, procedural knowledge and conceptual knowledge. If an assessment of an LO results in scores below an acceptable level, additional instruction is presented. The other approach involves the use of Bayesian Inference Network**s** (BINs) which are used to estimate the learner‟s aptitude in relation to the material. Both of the approaches seek to assess how close the learner is to mastering the material and the source of any problems that are encountered by the student. After the learner model has assessed the mastery of the learner, it recommends to the adaptive engine what should be presented to the learner next.

The last component necessary for an e-learning program is the *adaptive engine* which integrates information from the other models in order to present the e-learning content to the learner. First, it chooses a topic to present based on the learner‟s progress and then it selects which LOs for that topic should be presented, as well as what order the LOs should be presented for each individual (based on ATIs or performance on assessments). The adaptive engine continues to present LOs until that particular topic has been mastered. The system can choose any topic from the pool which has not yet been presented, provided that any prerequisites for that topic have been mastered. The order in which each LO is presented can be determined by comparing the student model and the learners progress so far with the content model. Each LO receives a priority number which will determine where in the presentation it appears.

Current e-learning systems are little more than traditional learning materials translated to an online format. Shute and Towle suggest that adapting learning content based on cognitive abilities will lead to more efficient, effective and enjoyable learning experiences. They suggest that rather than adapting content to meet constraints of the delivery system or adapting the interface to meet the needs of different users (such as those with disabilities), researchers need to concentrate on adapting learning based on the mappings between user characteristics and instructional content.

Shute, V. J. & Zapata-Rivera, D. (2008). Adaptive technologies. In J. M. Spector, D. Merrill, J. van Merriënboer, & M. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (3rd Edition) (pp. 277-294). New York, NY: Lawrence Erlbaum Associates.

Adaptive Training Intervention(s): Review of multiple interventions and techniques
Type of Study: Review article

This book chapter provides a review of adaptive training technologies and adaptive systems. The authors suggest that the goal of an adaptive system is "to create an instructionally sound and flexible environment that supports learning for students with a range of abilities, disabilities, interests, backgrounds, and other characteristics (p. 278)." Further, they suggest that the success of this goal is reliant upon (1) how well you can identify and measure these different characteristics and (2) how you can use that information to improve learning and performance. The chapter provides definitions of different terms used in adaptive systems, discusses why adaptation should occur, provides a framework for adaptive technologies, and summarizes areas of future research.

Adaptive systems are defined as systems that are adjusted to particular needs and characteristics of the learner. This could include real time assessment and delivery of instructional content or modifying multimedia content based on learner characteristics. On the other hand, adaptive technologies are computational devices that are used to control adaptation. Adaptive technologies can consist of both "hard technologies" (e.g., CPUs, eye tracking equipment, etc.) and "soft technologies" (e.g., software, adaptation algorithms, student models, etc.). The authors provide a summary of different soft technologies that can be used to capture and diagnose performance, as well as to select adaptation strategies. These soft technologies include quantitative modeling, qualitative modeling, cognitive modeling, machine learning algorithms, overlay methods, Bayesian networks, pedagogical agents, and student models. Likewise, the authors present a summary of different hard technologies that can be used as input and output devices in the adaptive system. These hard technologies include biologically based devices (e.g., eye tracking, scan patterns), speech capture devices, and hand gesture capture devices.

The authors cite several reasons why adapting instructional content can be beneficial. First, research suggests that prior knowledge is the most important factor to predict future learning. Second, demographic and sociocultural differences can affect learning and performance. Lastly, students tend to differ on affective states, such as motivation and fatigue, and these variables have been shown to influence performance. The authors reiterate that the success of any adaptive system to improve learning and performance "requires accurate diagnosis of learner characteristics (p. 280)" and that this diagnosis can form the basis for the adaptation of instructional content. Based on this thesis, the authors present a framework for the development of adaptive systems called the Four-Process Adaptive Cycle. The four components of this model are capture, analyze, select, and present. The capture process entails collecting information about the student such as prior knowledge, cognitive ability, gender, etc. The analyze process entails diagnosing and making inferences about a learners current state. The select process is used to determine how and when to intervene in student learning. Finally, the present process entails providing the optimal instructional content to the learner.

The authors informally surveyed five experts in the field and asked them two questions: (1) What to adapt (e.g., what learner characteristics should adaptations be based upon; what instructional interventions should be adapted?)? and (2) How to adapt (e.g., what techniques should be used to develop and analyze student and instructional models)? Regarding learner characteristics, the experts identified cognitive abilities, metacognitive skills, affective states (e.g., motivation, attention), personality variables, and learning styles as candidate adaptation variables. Additionally, the experts identified feedback type, feedback timing, sequencing, and scaffolding as candidate adaptive instructional variables. Regarding question 2, the experts suggest the best adaptive approaches are probabilistic learner models, concept mapping, unsupervised machine learning, and matching instructional support to cognitive ability.

Several areas of future research were suggested including measurement of valid student data (especially when self-report information is used), determining which adaptive variables provide the highest benefit, and determining which data to use to develop the most useful learner models. Additionally, one major challenge is to maximize the benefits while minimizing the costs of employing adaptive systems. The cost of developing these types of systems is quite high and the benefits are unknown due to the lack of formal evaluations. Therefore, the authors suggest that it is critically important that formal, empirical evaluations are performed to assess the effectiveness and cost-benefit ratio of these systems. They further suggest that "building adaptive systems and not evaluating them is like building a boat and not taking it in the water" (p. 291).

Williges, R. C. & Williges, B. H. (1978). Critical variables in adaptive motor skills training. *Human Factors, 20*, 201-214.

Adaptive Training Intervention(s): None
Type of Study: Review

The authors note that adaptive training procedures are typified by the automatic change in an adaptive variable according to some computer logic. The changes in the adaptive variable(s) are based on the trainee‟s performance which is continuously monitored. The authors suggest that there has been conflicting evidence regarding the effectiveness of adaptive training over traditional fixed difficulty training in the area of motor skills. Therefore, the purpose of this paper was to review and evaluate results from a series of studies regarding critical variables in adaptive motor skills training. These critical variables include performance measurement procedures, the selection of the adaptive variable, and the selection of the adaptive logic. In the current review, adaptive training refers to motor learning tasks where the difficulty of the training varies in response to how well the trainee is performing. If the trainee is within the specified error range, the task difficulty increases until the criterion is reached. However, if the trainee is outside the error range, the task difficulty is decreased.

The specific articles reviewed were Gopher, Williges, Williges, and Damos (1975), Williges and Williges (1977), and Williges, Williges, and Savage (1977). Each study used the same pursuit tracking task for training and transfer of training. Participants completed a series of 3-minute trials where they used a control stick to follow the movements of a symbol, X, displayed on a computer screen. In addition, there were two performance feedback bars that displayed tracking accuracy and task difficulty. When the tracking accuracy line was within the error tolerance level, the task difficulty line increased. Conversely, when tracking accuracy was out of tolerance the task difficulty line decreased. This procedure continued until the participant was able to perform the task while sustaining both feedback bars at or above the exit criterion for a particular period of time. After a short rest period, each participant completed a 7-minute transfer session similar to the training except that the feedback bars were not present. During the transfer task the level of difficulty changed automatically after each minute.

The authors note that before any adaptive variable can be chosen, a continuous measure of performance is required. Further, they suggest that researchers must choose between a number of performance measures (e.g., absolute error or root mean square) and as the task dimensions increase, one must choose between separate scoring for the adaptive procedure or some type of combined scoring (e.g., averaging of vector scores). After reviewing the three articles above, the authors identified a number of performance measurement issues that must be addressed. They suggest future research should evaluate the approach of adapting dimensions simultaneously or separately during multidimensional tracking. They also suggest that future research must determine the effects of various methods for combining task dimensions before operational use of adaptive procedures in complex simulator environments. For example, in an aircraft simulator, a number of system and task variables are combined to determine effective performance, therefore, performance measurement considerations are critical.

Williges and Williges (1978) remark that another critical variable to consider during adaptive training is the selection of the adaptive variable used to manipulate task difficulty. A review of the empirical evidence suggests that the rate of „adaptation‟ is enhanced through the use of stimulus related variables versus response related variables. However, the authors note that this finding needs to be supported with other stimulus variables and transfer effects need to be studied. Lastly, they suggest that the effects of different types and numbers of adaptive variables need to be assessed through factorial manipulations so that the separate, as well as interacting effects can be determined.

Williges and Williges (1978) reviewed the literature regarding the selection of the adaptive logic scheme. They note that this process is the core of any adaptive training program. They suggest that regardless of the variables utilized in the adaptive logic, the logic scheme can be viewed as one of several general models of motor skills training. By evaluating the general characteristics of these models, implications can be extracted for the design of efficient adaptive logic systems. The authors note that the traditional approach to motor skills training is a fixed-difficulty model where trainees are given the criterion task and their error decreases as training advances. But, this model does not recognize individualized instruction thereby making the training either too hard or too easy for some trainees. Williges and Williges suggest that although this logic system will provide a variety of task difficulty profiles it may not be flexible enough. Instead, adaptive training should provide individualized motor skills training utilizing a closed-loop model where aspects of the trainee‟s performance are monitored and a logic system is used to adjust the task difficulty by individually adjusting task difficulty for a wide range of skill levels.

The authors reviewed two studies and evaluated the effectiveness of a fixed difficulty model, an adaptive model, a learner-centered model where the trainee determines changes in task difficulty, and an open-loop training model where task difficulty is initially set low and then shifts to the criterion level after one trial. In these studies, the learner-centered and the adaptive training models are both closed-loop models. Results indicated that the closed-loop models resulted in less error during training and resulted in better transfer than an open-loop model. The authors suggest that future research should not focus on comparing open versus closed-loop models but rather optimizing a closed-loop model for training motor skills.

Lastly, the authors make a number of recommendations for future research regarding the adaptive logic scheme. First, they note that a learner-centered approach should be considered by training system designers until more sophisticated automatic logic schemes are developed. Additionally, future research should investigate the use of a changing adaptive logic scheme versus the traditional fixed logic scheme for adaptive training. They also suggest that research is needed to determine the individual difference variables that are associated with various adaptive logic strategies. For example, research by Pask (1976) indicated that matching trainee characteristics with the proper instructional strategy is critical in order to maximize training. It is also recommended that new adaptive logic schemes need to be developed based on current human learning models. These models can be used to describe the student‟s performance so that the adaptive logic can select instructional strategies based on hypothesized learning states of the trainee. Research must also evaluate the role of feedback during motor skills training. Finally, Williges and Williges recommend that future research investigate the potential interactive effects between feedback and various adaptive schemes. The authors note that "all of these investigations are required to provide the necessary data base before sophisticated and cost-effective adaptive logic schemes optimized for individual differences can be developed" (p. 211).

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B
ASSESSMENT OF ADAPTIVE TRAINING SYSTEMS AND TECHNOLOGY

Anthony, M. K., Goettl, B. P., Derek, K., & Ramirez, K. G. (1999). Adaptive remediation in complex skill training. *Proceedings of the Human Factors and Ergonomics Society, 43*, 1146-1150.

Adaptive Training Intervention(s): Multiple Emphasis on Components (MEC), MEC-Adaptive Remediation, and Verbal elaboration
Type of Study: Empirical

The purpose of this study was to explore the feasibility of implementing a Multiple Emphasis on Components (MEC) and adaptive remediation version of MEC during training to improve complex skill acquisition. Using the MEC protocol, participants are forced to focus on selected task components rather than global task performance. The MEC-Adaptive Remediation (MEC-AR) protocol was developed for this study in order to accommodate learning differences between individuals. Here, using the same selective component emphasis as MEC, participants were able to advance to successive components after reaching a proficiency criterion on the current component. Verbal elaboration also has a track record of being a successful training intervention. Therefore, verbal elaboration was investigated in this study as well. The Space Fortress (SF) video game was used for this study using a 3 (total score, MEC, MEC-AR) X 2 (verbal elaboration or no elaboration) between subjects design. The authors hypothesized that MEC-AR would improve skill acquisition more than the MEC and total score conditions.

The objective of the SF video game is to destroy a shooting space fortress, identify friend and foe mines, and aim for the highest total score (composed of ship control, ship velocity, appropriate response to mines, and battle points) possible. Participants control the ship‟s flight and missile firing capability with a joystick. For this study, 118 participants (61 males and 57 females) with an age range of 18-40 (M = 23.67) were recruited through temporary employment agencies. Participants were screened to ensure that they had no experience with SF and to determine that there were no differences in general cognitive abilities prior to training.

The experiment lasted for 3.5 days. On the first day, participants completed demographic forms and cognitive ability measures, were randomly assigned to one of six conditions, reviewed written and verbal SF instructions, and completed a baseline pretest module (four 3-minute trials). Participants completed six training modules on the second day and seven training modules on the third day, each training module consisting of two 4-trial practice sessions and one 2-trial test session. Participants completed retention and transfer modules and were debriefed on the final day. The total score group emphasized total game score for all training modules, whereas the MEC groups emphasized a single component during practice trials in each module. This single component was switched every three training modules in the following order: (1) ship velocity, (2) ship control, (3) speed of mine recognition, (4) battle points, and (5) total score. The study was composed of 13 training modules and the article does not discuss this setup in detail. It should be noted that MEC-AR participants were permitted to advance to succeeding components if they reached an acceptable score on the current component. Participants were required to attain criteria on three out of four practice trials in a training session. A brief remediation session, administered by experimenters, was given in between modules to MEC-AR participants not attaining criteria by the end of each training module. Emphasis was placed on total score on the task during test trials for all participants.

The ship control, $F(2, 101) = 3.78$, $p = .026$, and ship velocity, $F(2, 101) = 3.44$, $p = .036$, tasks support the authors‟ hypotheses that MEC-AR condition would yield the most learning and highest performance. Emphasis condition scores, ranked from highest to lowest, for the ship control and ship velocity tasks were (1) MEC-AR, (2) MEC, (3) total score, and (1) MEC-AR, (2) total score, (3) MEC, respectively. There was no significant main effect of emphasis on the battle points task. However, there was a significant interaction between practice, emphasis, and gender, $F(26, 178) = 1.58$, $p = .046$. Results revealed that women in the total score group improved more than women in the MEC groups, while women in the MEC-AR group started off worse than the MEC group but caught up at module 6. On the other hand, men in the MEC-AR group showed a minor learning advantage over men in the MEC and total score groups late in training. It is unclear from the article whether these results were from training, retention, and/or transfer.

The elaboration condition showed mixed results. Participants in the no elaboration condition performed better than those in the elaboration condition, $F(1, 101) = 4.19$, $p = .043$. However, a significant practice x gender x elaboration interaction was found, $F(13, 89) = 1.98$, $p = .031$, showing that elaboration had minimal effect on the improvement of males, and females showed a larger improvement under no elaboration. The battle points task showed a significant emphasis x elaboration interaction, $F(2, 101) = 3.93$, $p =.023$, in which the elaboration condition led to better performance for the MEC-AR participants and no elaboration led to better performance for the total score participants. Lastly, a significant practice x gender x elaboration interaction was found for the ship control task, $F(13, 89) = 1.89$, $p = .042$. This showed that elaboration had minimal effect for men and that women in the elaboration condition started off worse than no elaboration but caught up by the fifth module. It is unclear from the article whether these results were from training, retention, and/or transfer.

Overall, as expected, the results indicated that the MEC-AR condition had a significant positive impact on skill acquisition for some task components, specifically for the perceptual-motor based components (ship control and velocity). It should be noted that although participants were randomly placed into one of six training conditions, they were not completely balanced for gender across all six training conditions. This makes the conclusions based on gender differences tentative. The authors also point out that the effects of social interaction between participants and experimenters in the MEC-AR condition were not completely controlled. In other words, extra attention given to MEC-AR participants by human experimenters may have influenced results in a way that an intelligent tutoring system would not have. Results were mixed for the verbal elaboration condition, but there was evidence that show elaboration is beneficial later in training for females. However, it is possible that a ceiling effect influences this conclusion. It also appears that the total score component was emphasized for one training module, whereas the other components were emphasized for three training modules each (5 single components, switch every 3 modules, requires 15 training modules for equal amount of single component emphasis). Future studies should ensure that emphasis be placed on a single component for an equal amount of training modules across the board. The potential benefits of verbal elaboration should be investigated more thoroughly in future research.

Bell, B. S. & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*(2), 267-306.

Adaptive Training Intervention(s): Adaptive guidance, Learner control
Type of Study: Empirical

The authors examine adaptive guidance, an instructional strategy that is designed to assist trainees in making effective learning decisions. The article begins with an in-depth background regarding adaptive guidance which is defined as a training strategy that provides trainees with diagnostic and interpretive information that aids them in making effective decisions. In order to examine the benefits of adaptive guidance, 277 undergraduate students (56% female) were placed into either a learner control condition (control group) or an adaptive guidance condition. The experimental platform was TANDEM, a PC-based radar-tracking task. Trainees were familiarized with the task and then completed nine 9-minute training trials. Each trial consisted of a cycle of 2 minutes of online task manual study, 5 minutes of hands on practice, and 2 minutes of feedback review. Following the ninth trial trainees completed a 10-minute generalization (i.e., transfer) task which was more complex and difficult than the previous trials. Participants in both conditions received descriptive feedback on the same elements of performance and were given control over the content, pace, and sequence of the learning process. However, trainees in the learner control condition did not receive any guidance.

The trainees in the adaptive guidance condition were presented with guidance based on their level of performance. Cut-off scores were determined by pilot data and were set below the 50th percentile, between the 50th-85th percentile, and above the 85th percentile. Individuals below the 50th percentile were told they needed to focus on necessary skills and strategies. Suggestions were provided regarding what the trainee should be practicing and studying to improve. Individuals who fell between the 50th and 85th percentile were informed that they had reached a level of minimum performance, but still needed to improve and were given suggestions for what to practice. Trainees who scored in the 85th percentile or above were told they had mastered the skill or strategy and that they should concentrate on improving deficient areas. They were then provided with evaluative information from prior performances which was used to give individualized suggestions regarding what they should study and practice in order to further improve.

Following the third and ninth trials, self-efficacy, on-task cognition, practice sequence, study sequence, basic and strategic knowledge, and performance were assessed. Basic performance consisted of the number of correct and incorrect decisions. Strategic performance was comprised of the number of high priority targets that were processed, the number of times trainees zoomed out, and the number of times an attempt to identify an invisible perimeter was attempted. The same approach was used to assess basic and strategic performance during the generalization task. Cognitive ability was also assessed and used as a covariate during data analysis.

Using a repeated measures MANCOVA, with cognitive ability as a covariate, it was found that both cognitive ability, $F\ (8,267) = 12.95$, $p < .01$, $ή^2 = .28$, and adaptive guidance, $F\ (8,267) = 24.94$, $p < .01$, $ή^2 = .43$, had significant effects on performance. Follow up hierarchical regression analyses found that adaptive guidance significantly increased individuals" self-efficacy early in training, $\beta = .12$, $p < .05$ (H1), but contrary to the hypothesis, decreased self-efficacy later in training, $\beta = -.12$, $p < .05$. Conversely, trainees who did not receive guidance reported lower self-efficacy early in training and higher self-efficacy later in training. Contrary to beliefs, adaptive guidance did not have a significant impact on trainees" on task cognition (H2). In addition, trainees in the guidance condition also followed a better sequence when studying (H3) and practicing (H4) the material and this effect was observed both early and later in training. Trainees in the adaptive guidance condition spent over 25% more time studying the relevant training topics and practiced almost twice as many of the relevant training topics than those in the control condition. It was also found that adaptive guidance had a significant positive effect on trainees" basic knowledge early in training ($\beta = .15$, $p < .01$) and strategic knowledge later in training, $\beta = .15$, $p < .01$ (H5 and H6). Adaptive guidance significantly improved trainees" performance for basic tasks early in training, $\beta = .21$, $p < .01$ (H7) and strategic performance later in training, $\beta = .34$, $p < .01$ (H8). The final hypothesis (H9) was partially supported as adaptive guidance had a significant positive effect on strategic performance ($\beta = .30$, $p < .01$), but not basic performance during the generalization trial.

The results of this study lend support to the theory that adaptive guidance helps individuals make better learning decisions, above and beyond cognitive ability, and as a result increases the individual‟s knowledge and performance. Mechanisms for how adaptive guidance, and through what processes, influences learning and performance are discussed. One limitation of the design of the study was that adaptive guidance was given at only three levels of performance. Another limitation was the pace that the individuals were allowed to move through the material; because of the need for equivalence across situations the pace was the same for all individuals. A finer granularity of adaptation based on performance and alternate forms of trainee pacing is needed for future studies. A final weakness of this study is the generalizability of the results; the individuals participating in the study consisted of young college-age adults which may not be representative of the general population.

Bell, B. S., Kanar, A., Liu, X., Forman, J. & Singh, M. (2006). Adaptive guidance: Effects on self-regulated learning in technology-based training. Unpublished manuscript, Center for Advanced Human Resource Studies, Cornell University, Ithaca, NY. Retrieved from http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1455&context=cahrswp

Adaptive Training Intervention(s): Adaptive guidance
Type of Study: Empirical

Learner controlled training programs give users more control over their learning experience, but research has shown that learners do not always make good use of that control (Bell & Kozlowski, 2002) and, oftentimes, overlook important information. Adaptive guidance systems can provide information about what to study and practice, when, and how much, while still allowing the student to maintain a sense of control over their learning. The authors wished to examine how much control the learner should be given when using adaptive guidance systems. To this end, adaptive guidance systems which were autonomy-supportive or controlling were compared. The autonomy-supportive system used instructions with wording such as, "you could" or "if you choose" while the controlling guidance system used instructions such as "you must" or "you had better". The authors hypothesized that the learners who received adaptive guidance would acquire more basic and strategic knowledge and perform better than those in the learner control condition and that those who received the controlling guidance would gain more knowledge and perform better than trainees in the autonomy-supportive condition. Additionally, the authors hypothesized that the adaptive guidance would have a positive effect on the learning and performance of individuals with high ability, but no effect or a negative effect for individuals of low ability. They also predicted that in the adaptive guidance conditions, motivation to learn would be positively related to knowledge and performance, but that it would be unrelated to the learner control condition.

One hundred and thirty undergraduate students (59% males) participated in a study which used a radar-tracking simulation called TANDEM. The trainees were randomly assigned to one of the three conditions: autonomy-supportive guidance, controlling guidance, or learner control (the control group). Cognitive ability and motivation were measured at the beginning of the experimental session. Cognitive ability was measured using participants" SAT or ACT scores which were standardized using national means and standard deviations. The researchers used a seven-item test based on a five-point Likert scale to measure motivation. The training took place in a three hour session with groups of one to four participants at a time. There were nine 10 ½ -minute cycles followed by a transfer task. After the third cycle, the participants took a test to assess their basic knowledge and after the ninth cycle they took another to assess their strategic declarative knowledge. After the second test, they were given the transfer task which was more difficult and longer than the scenarios they were exposed to during training. In the learner control condition, participants were given a randomized list of topics and were told that they could decide which topics to study and in what order. For the experimental conditions the same list was given in an ordered sequence that allowed trainees to study the basic material first and then build on that knowledge. After each cycle all participants received feedback, but the participants in the adaptive guidance conditions also received personalized information on what they could/must (based on condition) study and practice to improve their performance. The guidance that was given was based on three levels of performance (low, medium and high) determined by cutoff scores which were established in pilot testing. The researchers evaluated performance for basic and strategic aspects of the task during the third and ninth cycles respectively.

The first hypothesis was partially supported. Analysis of the data revealed that participants who were in the controlling guidance condition had significantly higher levels of basic declarative knowledge ($\beta = .17$, $p < .05$), performed better during training for both basic and strategic aspects of the task ($\beta = .19$, $p < .05$ and $\beta = .50$, $p < .01$ respectively) and had better strategic performance during the transfer task ($\beta = .40$, $p < .01$) compared to the participants in the learner control condition. Autonomy supportive guidance did not have any significant effect on basic outcomes, but trainees in this condition performed significantly better than those in the learner controlled condition in strategic performance during training ($\beta = .19$, $p < .05$) and transfer ($\beta = .17$, $p < .05$). In regards to the second hypothesis, the analysis revealed that the trainees in the two experimental conditions did not differ significantly on any basic outcomes or on strategic knowledge, but the trainees in the controlling guidance condition did exhibit significantly higher levels of strategic performance for both the training and the transfer tasks ($t = 3.55$, $p < .01$ and $t = 2.33$, $p < .05$ respectively; no degrees of freedom reported) than those in the autonomy-supportive guidance condition. The third hypothesis predicted an interaction between adaptive guidance and cognitive ability that would be more prevalent in basic knowledge and performance developed earlier in the training. Results showed that guidance did not interact with cognitive ability to affect any strategic outcomes. However, controlling guidance and cognitive ability interacted significantly to affect basic knowledge ($\beta = .27$, $p < .05$), as well as basic performance during the transfer task ($\beta = .26$, $p < .05$). As predicted, controlling guidance had a positive effect on high ability trainees (in regards to basic knowledge and performance), but did not have any effect on low ability trainees. Lastly, an interaction was found between autonomy-supportive guidance and motivation to learn which affected trainees‟ strategic performance during training ($\beta = .25$, $p < .05$) and basic performance during the transfer scenario ($\beta = .26$, $p < .05$). In the guidance conditions there was a positive relationship between motivation to learn and performance, while in the learner control condition there was a negative relationship observed. In addition, controlling guidance and motivation interacted significantly to effect basic performance during training. The interaction revealed that controlling guidance benefited participants with low motivation to learn, but it did not help participants who had high levels of motivation.

Overall, the results support the idea that trainees benefit from adaptive guidance in a learner controlled environment. They also suggest that the type of adaptive guidance used affects performance and learning. One possible problem with the study is the use of cash money used as rewards for correct answers during training. Up to $100 dollars in prizes were available to win which may have interfered with the intrinsic motivation being measured in the study. Another possible limitation of this study is that the participants were all undergraduate students, mostly between 18 and 21, who participated in the study for course credit. Using a sample with non-students may give a different picture of the effects.

Cote, D. O., Williges, B. H., & Williges, R. C. (1981). Augmented feedback in adaptive motor skill training. *Human Factors, 23*, 505-508.

Adaptive Training Intervention(s): Adaptive task difficulty, Outcome feedback
Type of Study: Empirical

This paper reports the results of two studies in which the effects of both adaptive task difficulty and outcome feedback, presented in different modalities, on task performance were evaluated. The authors note that during adaptive training, trainee‟s responses are compared to some criterion level of performance and some aspect of the system (e.g., task difficulty) is modified in order to maintain a constant level of performance throughout training. However, one of the weaknesses of this approach is that trainees are not usually provided with knowledge of their results, especially within motor skills training. It is suggested that, in order to overcome this deficiency, feedback should be given during adaptive training.

Two experiments were conducted using the same two-dimensional pursuit tracking task in which participants had to follow the movement of a pursuit symbol. In Experiment 1, 24 right-handed males participated in a 2 (training type: fixed vs. adaptive training) x 2 (feedback: visual feedback bars vs. no feedback) between subjects design. In Experiment 2, 96 right-handed males participated in a 2 (training type: fixed vs. adaptive training) x 4 (feedback: auditory feedback, visual feedback bars, auditory and visual feedback, vs. no feedback) between subjects design. The feedback bars presented task difficulty and tracking accuracy information on the participant‟s display screen as they performed the task. As task accuracy or difficulty increased, the feedback bars would move upward and disappear once criterion performance was reached. Auditory feedback was indicated via a 2000-HZ tone and 400-Hz beeps representing task accuracy and difficulty respectively. Whenever the criterion performance was achieved the tones stopped. The criterion performance level for both the adaptive and fixed training conditions was the maintenance of the tracking symbol within 1.8 cm of the pursuit symbol while the pursuit symbol moved at the maximum level of speed. Both experiments were divided into a 3 minute training session and a transfer session. In the adaptive training condition, task difficulty was defined in terms of the speed of the pursuit symbol. After the training session, participants were presented with a 6 minute transfer task identical to the training task except that no feedback was provided and three levels of tracking difficulty were presented for 2 minutes each. The difficulty levels (i.e., speed of the pursuit symbol) included the training exit criteria difficulty level, 0.5 times the exit criteria difficulty level, and 1.5 times the exit criteria difficulty level. Presentation of the difficulty levels within the transfer task was counterbalanced.

Results for Experiment 1 revealed no significant effects during the training session (i.e., no differences in training time due to training type, presence of visual feedback, or the interaction between the two). However, during the transfer task results revealed that those who received the adaptive training had fewer errors, $F(1, 20) = 9.24$, $p < .01$) than those who received fixed difficulty during training. No effect for feedback type was found. Results for Experiment 2 were similar to Experiment 1. There were no effects during the training session, however, during the transfer task those who received the adaptive training performed significantly better ($p < .05$) than those who received fixed difficulty during training (the authors did not provide a statistical value with this p-value). Again, there was no significant effect for feedback type or any of the interactions. These results suggest that the results of Experiment 1 were not due to visual information overload. In total, the results from both experiments suggest that changing the level of task difficulty during training is superior to fixed training. One limitation of this study is that the type of feedback may have not been useful to trainees because it provided little to no information that was not already available from the task itself. This suggests that the use of outcome feedback during adaptive training may not be necessary when performance feedback is clearly evident from the task. Other limitations include the relatively short training period (3 minutes) and the generalizability of the results based on the simple experimental task and the restricted sample (i.e., right handed males).

Crooks, W. H. & Roscoe, S. N. (1973). Varied and fixed error limits in automated adaptive skill training. *Proceedings of the Human Factors and Ergonomics Society, 17*, 272-280.

Adaptive Training Intervention(s): Fixed and variable error limits; changes in acceleration control
Type of Study: Empirical

The authors investigated the use of allowable error limits as a function of an adaptive training variable, and the impact of these limits on task performance. An aircraft training task was used, which required participants to keep a horizon line on a simulated aircraft display "horizontal" using a joystick. The computer program randomly changed the pitch and roll axes to alter the angle of the horizon line. The adaptive variable for this task was percentage of acceleration control, α, or "the ratio of acceleration to velocity or rate control exerted by the output of the handstick controller upon the pitch and bank attitude presented on the simulated aircraft attitude indicator" (p. 273). The experiment was designed such that α progressed from 0% to 80% through a sequence of 20 step changes.

The error limit, or the percentage of possible error, could increase or decrease incrementally using predetermined formulas after α step changes occurred. The error limit was also used to modify α for the adaptive training conditions based on the participant‟s performance. Each participant‟s mean absolute performance error, |ē|, for both the pitch and roll axes, were calculated every five seconds. If the participant‟s |ē| exceeded the existing error limit for that five second interval, then α decreased one step. If the reverse was true, then α increased.

The experiment used five training condition groups, using combinations of acceleration control and error limit: (1) Adaptive acceleration control/Error limits decreasing (becoming more restrictive), (2) Adaptive acceleration control/Error limits constant, (3) Adaptive acceleration control/Error limits increasing (becoming less restrictive), (4) Adaptive acceleration control (increasing only)/Error limits constant (i.e. if |ē| > error limit, α does not decrease), and (5) Constant acceleration control / No error limits (α set at 80%). Groups 4 & 5 were considered the training control groups. Participants completed the training task until the exit criterion was reached or 60 minutes had elapsed. The training task was presented in five minute intervals, with two minute breaks in between. The authors hypothesized that participants in Group 1 would complete the experiment in less time than Group 3. They also hypothesized that Group 2 completion times would fall between those for Groups 1 and 3.

After the training task, participants performed a transfer task using a different simulator that only modified the roll axis. Participants worked on the transfer task until an exit criterion was reached or 30 minutes had elapsed. One week after the training task, participants returned to the laboratory to complete a skill retention task. The retention task was completed until the exit criterion from the training task was reached. The dependent variable for all three tasks was time to achieve the exit criteria. A between subjects design was used. Eighty-two participants were assigned, using an unspecified method, to the five training groups and a transfer task control group (Group 6).

A Kruskal-Wallis one way ANOVA by ranks revealed a significant difference between the median times to attain the exit criteria for the five training groups, $H(4) = 10.86$, $p < .05$. To explore the details of this finding, all time scores that exceeded 60 minutes were removed, producing a more normal data set. An ANOVA (unspecified type) revealed significant differences between the training groups, $F(4, 54) = 4.83$, p value not reported. Scheffé‟s test on all of the mean time scores revealed that the mean for Group 1 was significantly higher than either of the training control group means (Groups 4 & 5), $p < .05$; Scheffé values were not reported. The authors report no significant difference between the average for Groups 1, 2, & 3 and Groups 4 & 5. The authors also note that the difference between Group 1 and the average of Groups 2 & 3 approached significance. No additional clarification of these results was provided.

A Kruskal-Wallis one way ANOVA by ranks on the transfer task data revealed no significant differences, between Groups 1-5, who performed both the training and transfer tasks, and Group 6, the transfer control group. A similar test on the retention task data revealed no significant differences between the five training groups. However, the median times on the training task (1492.5 seconds) were greater than those for the retention task (85.0 seconds). A Wilcoxon matched-pairs signed ranks test revealed a significant difference between the task times, $z = 6.57$, $p < .001$. The authors note that median retention task times were also lower (165 seconds) for those participants who went over 60 minutes on the training task. The authors speculate that "some cognitive or verbal reorganization occurred" (p.277) between the two tasks.

The authors admit that their hypotheses were not supported and attribute the reason primarily to an interference paradigm associated with the use of acceleration control as an adaptive variable. The authors explain that "optimum control with an acceleration system usually entails small, quick control movements, often in the form of small „jerks‟" (p.276). Therefore, Group 1 participants had higher mean times because they could make large movements early in the experiment, since the error limit was very wide and the acceleration was small, but were at a disadvantage when the acceleration increased and smaller movements were required.

Theoretically, random assignment ensures groups are equal at pretest, however, the assignment to conditions was not explained. Therefore, any systematic differences in groups could be attributed to selection bias. Another key weakness of this paper was a second literature review after the results section on the limitations of using human error rates as a training performance measure. It would have been better to use this information to formulate hypotheses for the experiment, since the adaptive conditions intentionally limited the amount of error participants could make. Finally, the authors did not provide details on the experimental procedure for Group 6 and hypotheses for Groups 4, 5, or 6, which made the interpretation of the results for these groups more difficult.

Gopher, D., Williges, B. H., Williges, R. C., & Damos, D. L. (1975). Varying the type and number of adaptive variables in continuous tracking. *Journal of Motor Behavior, 7(3)*, 159-170.

Adaptive Training Intervention(s): Fixed or variations of the frequency of forcing function, control stick sensitivity, and ratio of acceleration to rate control
Type of Study: Empirical

It has been suggested that, theoretically, any aspect of a task that affects behavior could be selected as an adaptive variable. However, some researchers have argued that the use of certain "adaptive" variables may actually interfere with learning. For example, Osgood‟s (1963) transfer surface theory suggests that pairing different responses to similar stimuli hinders learning and results in negative training transfer. Alternatively, the transfer surface theory suggests that changing the stimulus and requiring a similar response does not result in negative transfer. This study was designed to use adaptive variables that are either stimulus or response variables to determine if the original learning and transfer performance using adaptive techniques are compatible with the transfer surface theory. The purpose of this study was to investigate the implications of adapting stimulus and response variables, the effect of varying the number of adaptive variables, and the evaluation of adaptive techniques in training, transfer, and retention.

Forty-eight students (36 males and 12 females) participated in a two-dimensional pursuit tracking task for five 3-minute training sessions with 3 minute breaks between periods. Each period was started at the final level of the previous period. Three types of adaptive variables were manipulated during the task. These variables were force function frequency (stimulus aspect of tracking task), control stick sensitivity (response aspect of tracking task), and the ratio of acceleration to rate control (response aspect of tracking task). Each was manipulated using a small-step adaptive logic. Including a control condition with no adaptive variables, participants were randomly assigned to one of eight conditions: three conditions in which only one variable was adapted, three conditions in which two variables were adapted, and one condition in which all three adapted during training. Performance was evaluated during learning, transfer, and retention. In both transfer and retention, the participants performed a tracking task which periodically changed in terms of task demands. A 30 minute break separated the training from the transfer scenario. The transfer scenario was comprised of four 2 minute sections. The retention task was performed approximately one week after the transfer scenario and was equivalent to the transfer task.

Each adaptive variable was analyzed separately. Results indicated that the highest rate of adaptation in frequency occurred when frequency was the only adaptive variable. The rate of adaptation in acceleration was greatest early in training when frequency was also adapted, $F(3, 120) = 3.07$, $p < .04$. Adaptation in control stick sensitivity (gain) increased when another variable also adapted compared to gain adapting alone, $F(3, 20) = 3.43$, $p < .04$. . These results support the transfer surface theory and imply that stimulus rather than response variables should be manipulated during adaptive training when the goal is to maximize "adaption". Participants trained with frequency adapting showed more stable performance during the transfer task, $F(1, 40) = 6.54$, $p < .02$ and $F(1, 40) = 11.18$, $p < .01$ (on the horizontal and vertical axis, respectively) than participants with fixed variable training. This finding suggests that adaptive training facilitates trainees‟ adjustments to unstable or changing conditions. There were no statistically significant differences among the various conditions during retention, however, an analysis of root mean square (RMS) error showed a significant interaction between the gain and task section on the horizontal axis, $F(3,120) = 7.80$, $p < .001$. One limitation of this study is the small sample size per experimental condition (six participants per condition) limiting the statistical power. In addition, another limitation of this study was that it used a simple tracking task that may not be generalizable to more complex military tasks (e.g., radar tracking tasks). Lastly, a direct comparison of the eight experimental conditions was not possible because each differed with respect to the specific dependent measure utilized.

Johnson, D. F., & Haygood, R. C. (1984). The use of secondary tasks in adaptive training. *Human Factors, 26*, 105-108.

Adaptive Training Intervention(s): Adaptive task difficulty
Type of Study: Empirical

The study examined the use of adapting task difficulty during training and its effect on performance. The authors define adaptive training as "training in which task difficulty is varied systematically as a function of how well an individual trainee performs" (p. 105). In addition, this study examined the usefulness of conducting adaptive training by adapting primary task difficulty on the basis of secondary task performance. It is noted that in many complex tasks, the primary task is one of many tasks necessary for adequate performance. Therefore, trainee improvements within complex tasks are reflected in both primary task performance and other task performance (e.g., secondary and unexpected task demands). The authors suggest that adapting primary task difficulty based on secondary task performance should better facilitate the acquisition of secondary task performance because increases in primary task difficulty will be reduced until improvements in the secondary task are seen.

The experimental design was a 2 (training type: adaptive versus fixed difficulty) X 2 (adaptive criterion: primary versus secondary task) X 2 (gender) X 10 (trials) mixed model. Sixty-four participants engaged in a task (primary task) in which they were required to keep a simulated automobile within the boundaries of a moving roadway. The secondary tasks included a left-right shadowing task in which participants had to indicate which side of a vertical line an asterisk was located by pushing one of two buttons. Primary task difficulty was adapted by increasing or decreasing the width of the roadway based on either primary or secondary task performance. Task difficulty in the control condition was set at the final value achieved and maintained during Trial 10 by the *previous* adaptive participant (yoking method). This was done to ensure that the final task difficulty for both training groups was equal. Instructions to participants stressed attention to the primary task. Therefore, there was no way for participants to determine *when* adaptation was based on secondary task performance.

Results from an ANOVA indicated that across all 10 trials and the final trial (Trial 10) participants whose task difficulty was adapted performed significantly better than those whose task difficulty was fixed, $F(1,56) = 10.19$, $p < 0.01$ and $F(1,56) = 17.79$, $p < 0.001$, respectively. These results support the notion of the usefulness of adapting training difficulty versus fixed difficulty during training. Furthermore, across all 10 trials and the final trial, participants whose task difficulty was adapted based on primary task performance performed better than those whose task difficulty was adapted based on secondary task performance, $F(1,56) = 8.81$, $p < 0.01$and $F(1,56) = 4.91$ $p < .03$, respectively. Secondary task performance was greater for those participants whose primary task difficulty was adapted based on secondary rather than primary task performance, $F(1, 28) = 21.42$, $p < 0.001$. On one hand, these results suggest that primary task performance can be hindered by the use of a secondary task adaptation criterion. On the other hand, these results suggest that secondary task performance can be facilitated by adapting the difficulty level of the primary task on the basis of the secondary task. Lastly, no gender differences were found. In this study both the primary and secondary tasks were fairly simple in nature limiting the generalizability of these effects to more complex tasks. Therefore, it is unclear as to whether this methodology can be easily modified to fit more complex tasks.

Johnson, D. F., Haygood, R. C., & Olson, W. M. (1982). Yoked design and secondary task in adaptive training. *Proceedings of the Human Factors and Ergonomics Society, 26*, 21-24.

Adaptive Training Intervention(s): Adaptive vs. Fixed Level of Difficulty
Type of Study: Empirical

This study examined the feasibility of using a yoked design and adaptation on secondary task performance in adaptive training over a fixed level of difficulty. Most research comparing adaptive with fixed training schedules have been based on an arbitrarily chosen fixed level of difficulty. The absence of a clear rationale in choosing a difficulty level may have led to use of an inappropriate difficulty level for an experiment. The authors point out that a plausible solution to this problem is to yoke participants, by "setting fixed training difficulty level to the final difficulty level achieved by the previous adaptive subject" (p. 21). This ensures that the difficulty level is equal for adaptive and fixed training participants, a situation that is ideal for comparing the two training schedules for a set number of trials. The first purpose of this study was to test the feasibility of applying a yoked design to adaptive training in order to compare an adaptive and fixed training schedule. The second purpose of this study was to test the feasibility of using secondary task performance as a means for providing adaptive criterion used to gain a better understanding of the primary task difficulty level.

The study consisted of two experiments. For the first experiment, participants were 88 randomly selected introductory psychology students. The design was a 2 (training condition: adaptive vs. fixed) x 2 (adaptive criterion: primary vs. secondary task performance) x 2 (adaptive variable: speed of target roadway vs. width of target roadway) x 2 (gender: male vs. female). Each of the eight experimental groups consisted of the same percentage of males and females. The participant‟s primary task was to keep a simulated vehicle within the boundaries of a moving roadway. As a secondary task, participants were asked to indicate, by pressing one of two buttons, whether a number (from 1 to 8) shown on the screen was odd or even. The secondary task was continuously available (i.e., a new number shows on the screen after the participant‟s response). Instructions placed emphasis on the primary task. The experiment consisted of 10 trials, 3 minute duration for each trial, with an average 1 minute break between trials.

Two variables (rate and width adaptation) were used for altering the primary task difficulty level. Here, the difficulty level was changed by increasing or decreasing horizontal movement and width of the roadway. Participant performance was based on a percent time on target (TOT) measuring scheme for the preceding 20 second interval. For the primary task, criterion cutoff for advancing to a more difficult level was 90% TOT and 60% TOT for downgrading to an easier level. Criterion cutoffs for the secondary task were 15 or more correct responses and 5 or more correct responses in the last 20 seconds. Because adaptive participants started with an easy task and advanced to a more difficult task, each participant‟s raw TOT score was adjusted by dividing the total TOT by the difficulty level for each trial.

Analysis of all 10 trials and trial 10 (essentially same difficulty for adaptive and fixed participants) alone show that participants that were adaptively trained performed better than fixed-level participants, $F(1, 72) = 11.02$, $p < .001$. This supports the idea of using adaptive training as a more effective alternative to fixed-level training. Participants adapted on secondary task performance (and yoked controls) performed better than participants adapted on the primary task, $F(1, 72) = 8.85$, $p < .005$. The authors did not clearly state whether or not all adaptive participants (for both primary and secondary tasks) were yoked. The rate adaptation variable accounted for most of this superiority by generating a significant interaction of adaptive criterion (primary vs. secondary task performance) and adaptive variable (rate vs. width), $F(1, 72) = 10.63$, $p < .002$. Participants with rate adaptation (and yoked controls) performed better than the width adaptation participants. However, this result is not interpretable because the difficulties of the various rate and width levels of the roadway were unknown. The difference between the adaptive and fixed group was much larger for width adaptation compared to rate adaptation, $F(1, 72) = 5.23$, $p < .025$. Finally, the performance of males was superior to females, $F(1, 72) = 19.44$, $p < .001$.

It was discovered that the difficulty level of the rate-adapted task increased geometrically with an increase in a participant‟s difficulty level. Also, the odd-even number shadowing task may not have provided the most receptive measure of residual capacity. Due to these facts, a second experiment was performed, similar to experiment one with the following exceptions. First, the experiment was performed with width adaptation as the only variable. The rate of the roadway was held constant at a slightly higher rate than the value used for the width adaptation participants of experiment one. Next, the secondary task was altered to a spatial (left-right) shadowing task where an asterisk appeared to the left or right of a vertical line and participants were required to press one of two buttons indicating the asterisk‟s location. Similar to the secondary task of experiment one, the appearance of the asterisk was continuous and reset after each participant‟s response. Participants consisted of 64 randomly selected introductory psychology students with an equal percentage of males and females in each group. This experiment consisted of four groups. The setup was a 2 (training conditions: fixed vs. adaptive) x 2 (adaptive criterion: adaptation on primary vs. secondary task) x 2 (gender: male vs. female) factorial. Finally, the trial duration was changed from 3 minutes to 2.5 minutes.

According to the results of trial 10, performance of adaptive participants was superior to fixed-level participants, $F(1, 56) = 17.79$, $p < .001$. There was not a significant difference in male and female performance for the 10th trial, but on average for all 10 trials, males performed better than females, $F(1, 56) = 17.46$, $p < .001$. On the contrary from experiment one, participants adapted on primary task performance (and yoked controls) performed better than participants adapted on secondary task performance. The authors did not clearly state whether or not all adaptive participants (for both primary and secondary tasks) were yoked. This difference is most likely due to interference from the primary task as a result of shared processing resources between primary and secondary tasks. This is supported by a lower secondary task performance rating and lower level of difficulty reached by secondary-adapted participants throughout experiment two. The lower level of difficulty means that participants could never receive higher performance ratings, regardless of how well they performed on the primary task, until the automation process freed up processing capability. Based on these results, this could not be achieved from 10 trials.

The results of the two experiments show that both yoked control and secondary task adaptation can be used as alternatives to past methods that use a fixed-level of difficulty and adapt only on primary task performance. It is unclear as to whether or not these results would be replicated using different primary and/or secondary tasks. Therefore, future research should investigate the external validity of these results using different types of tasks. It may be worthwhile to see how other variables affect the results. For example, varying the speed of the vehicle instead of the actual roadway or adding obstacles (cars and pedestrians) may be possible alternatives. Furthermore, assuming that a participant performed 10 successive trials, the validity of the results may be threatened due to a participant gaining familiarity of the experiment with each trial. Experiments in the future should take these factors into consideration.

Lemaire, P. & Siegler, R.S. (1995). Four aspects of strategic change: Contributions to children‟s learning of multiplication. *Journal of Experimental Psychology: General, 124*(1), 83-97.

Adaptive Training Intervention(s): None
Type of Study: Empirical-Non-experimental (longitudinal investigation with no manipulation)

The authors conducted a longitudinal investigation on one classroom of French second grade students ($N = 20$, $M_{age} = 97$ months) in order to examine the changes that occurred in strategy use, choice, and execution during the year that they learned single-digit multiplication. Measurements were taken three times during the year, once in the beginning of the year (about one week of instruction), once in the middle of the year (about four months of instruction), and once at the end of the year when they had a great deal of experience with multiplication (about six months of instruction). The investigators measured the speed and accuracy with which the students completed the problems, as well as what strategy was used to obtain the answer. Lemaire and Siegler (1995) postulated that changes in speed and accuracy reflected changes in four dimensions of strategic change: which strategies are used, when each strategy is used, how the chosen strategy is executed and how the strategy is chosen. A computer simulation called the adaptive strategies choice model (ASCM) was used to calculate a number of predictions regarding the four dimensions of strategy use, as well as the relation between earlier and later performance of individual students. For instance, the model predicts that students will progressively increase the amount that they use the retrieval strategy and the more effective backup strategies as compared to the ineffective ones. It also predicts that more difficult problems are more likely to be solved with backup strategies than easier ones and also that the correlation between the difficulty of the problem and the percentage of backup strategy usage will increase with experience

Analysis of the data yielded five strategies used by the students. The three most common were retrieval, repeated-addition, and saying "I don‟t know", and two strategies which were more rare were counting sets of object (such as tally marks), and writing the problem down. All 20 children used at least two of the strategies in each of the three sessions and, in each session, each of the five strategies was used by at least one child. Over the three sessions, changes in which strategies were used increased both speed and accuracy. In the first session, the three main strategies were used in between 30 and 40% of the trials, and by the third trial the fastest approach, retrieval, was used on 90% of the trials. Accuracy of retrieval increased from the first to third session, with errors decreasing from 27% in the first session to 9% in the third. Over time, the students‟‟ choice in strategy became more adapted to the difficulty for each problem. On all three sessions, retrieval was used for the easiest problems and the backup strategies were used for the more difficult ones. As the sessions progressed, the students were able to identify each type of problem with more precision allowing them to choose the best strategy more quickly.

The results of the study support the idea that changes in speed and accuracy that occur as a part of learning are a result of three out of the four underlying changes in strategy use: use of more efficient strategies, improved execution of strategies, and more adaptive choices among strategies. However, the small sample size used, as well as the narrow sample chosen for the study warrant further research be done to test the generalizability of the results to a larger and more diverse population. While the authors did not explicitly investigate issue related to adaptive training, their analyses showed that strategy choice changes over time and strategy implementation can change based on the difficulty of the problem. These findings have implications for developing diagnostic algorithms and determining instructional content for adaptive training.

Ludwig, J., Ramachandran, S. & Howse, W. (2002). Developing an adaptive intelligent flight trainer. *In Proceedings of the Industry/Interservice, Training, Simulation and Education Conference (I/ITSEC 2002).*

Adaptive Training Intervention(s): Intelligent Tutoring System
Type of Study: Case Study, Interview

In current simulator training, students are paired with an instructor pilot (IP) who guides student learning while also making sure students learn to perform skills correctly. To eliminate the need to have an IP present for all flight training a new flight trainer was developed. The Intelligent Flight Trainer (IFT) is an adaptive flight trainer (for helicopters) which combines the role of a flight simulator with the role of the IP. To this end, an intelligent tutoring system (ITS) was included in the IFT in order to provide the same kinds of instruction that would be provided by the IP. The ITS in the trainer has two components called the helper and the advisor. The helper component adjusts the flight model of the trainer in accordance with the skill level of the student. For example, students who are less experienced tend to make large and impulsive movements. Because of this the flight model is adjusted to be less responsive until student"s skill level increases. The advisor component of the trainer "talks" to the student and can take on four roles. The first role, called the tutorial role, teaches the student basic procedures, the performance monitoring role supervises the student"s performance and offers suggestions, the third role monitors flight control manipulation and comments on how the student uses the controls and finally the advisory role tells the students how they can control and correct their flight.

A one participant case study was used to evaluate the performance of the IFT. The student had basic knowledge of helicopter flight, but minimal flight experience. The student performed generally poor on the task and the helper role did not help her acquire skills, however, the level of control assistance did help the student hover and the advisor successfully guided the student through traffic (Mulgand et al., 1995). Although this evaluation had many limitations such as small sample size, lack of a control group, massed rather than distributed training, and lack of measurement of transfer to real helicopters, it did provide some support for a variable flight model and the ITS of the trainer.

One limitation of the IFT is the lack of attention to attributes that contribute to students" current performance. The IFT trainer bases the realism of the flight model on the student"s current performance without taking into account variables such as past performance and state and trait attributes. The authors introduced an adaptive instructional system (AIS) as an extension to the ITS that could consider these variables and use them to adapt instruction to individual students. The construction of the AIS would be focused on trait (e.g. preferred learning style) and state (e.g. anxiety) variables rather than past performance. In order to begin the development of the AIS, one civilian and five Army IPs were interviewed. The authors wanted to learn how IPs adapted their instruction to students and what role they played during training by gathering information about specific instances where the IP altered their training to help a student. When these instances, or episodes, were analyzed the authors found three areas of adaptive instruction that would be relevant to the development of the AIS: anxiety/confidence, fatigue and visual cues. Support was not found for trait-based instruction during the interviews. The interviews helped to shed some light on what should be adapted for the AIS, and the next step of how to adapt these factors is a point for future research.

Maddox, W. T., Love, B. C., Glass, B. D., Filoteo, J. V. (2008). When more is less: Feedback effects in perceptual category learning. *Cognition*, *108,* 578-589.

Adaptive Training Intervention(s): Feedback, Collaborative learning
Type of Study: Empirical

This article differentiates between rule-based and information-integration category learning and their underlying neural circuits. These circuits are called the rule system (specializing in rule-based learning) and the procedural system (specializing in information-integration type learning). The authors investigated how performance on each type of learning is affected by minimal or full feedback. Using the rule system, participants learn and reason in an explicit manner utilizing working memory; this process is likened to hypothesis testing. Conversely, the procedural learning system, which is best for information-integration type tasks, does not use working memory and operates by associating regions of perceptual space and actions that lead to rewards. Therefore participants perform better on procedural learning tasks when immediate feedback is provided. Further, the rule-based neural system is used for tasks that involve simple verbalizable rules, whereas the procedural neural system seems to be activated when the rules become more complex. Additionally, there is belief that the rule system is used first and, only when the results are incorrect, does the rule system "allow" the procedural system to take over (i.e. the rule system decides when the procedural system should be used). The authors hypothesized that the participants utilizing rule-based structures would do best with full-feedback and that those using procedural structures would perform better with minimal feedback.

One hundred-sixteen participants were randomly assigned to one of four conditions: Information-Integration-Full-Feedback ($N = 32$), Information-Integration-Minimal-Feedback ($N = 30$), Rule-based-Full-Feedback ($N = 27$) and, Rule-Based-Minimal-Feedback ($N = 27$). Nine participants were excluded for not meeting minimal learning scores. All participants received six 100-trial blocks where they were asked to categorize a stimulus on a computer screen in either category A, B, C, or D after judging their length and orientation. For the Rule-Based groups, stimulus lines could be categorized based on rules such as "respond A to short, shallow angle lines, B to short, steep angle lines, C to long, shallow angle lines and D to long steep angle lines" (p. 580). For the Information-Integration groups, the rules which dictated category belonging were not verbalizable. However, none of the groups were given explicit rules regarding how to categorize the stimuli. Either full or minimal feedback was given after each trial. Both types of feedback informed the participant whether their answer was correct or incorrect, however, the full feedback also provided the participant with the correct answer.

A 2 (I.I. vs. R. B.) X 2 (full vs. minimum feedback) X 6 (blocks) ANOVA found that the main effect of block, $F(5,515) = 153.03$, $p < .001$, as well as the feedback X block interaction was significant, $F(5,515) = 6.30$, $p < .001$. The category structure X feedback condition interaction was also significant, $F(1,103) = 9.90$, $p < .01$, which supported the authors" prediction that the rule-based groups who received full feedback would perform better on posttest of all six blocks ($M = .73$) than those who received minimal feedback ($M = .67$), $t(48) = 2.09$, $p < .05$. For information-integration tasks, full feedback resulted in poorer performance ($M = .65$) in comparison to minimal feedback ($M = .71$), $t(55) = 2.37$, $p < .05$. Also, the authors expected that the effects of feedback would weaken with experience for rule-based groups (by the third block the effect was not significant) but would strengthen for information-integration groups [after the third block the effects were significant, block 4 ($p < .05$), 5 ($p < .05$), and 6 ($p < .01$)]. Overall, participants utilizing the rule-based system still outperformed those using procedural systems in both category structures (I.I. and R.B) when aided with full feedback, but only in early learning. However, after the acclimation period this may have hindered the transfer of decision making to the procedural system in the Information-Integration-Minimal-Feedback condition. In addition, it appears the authors measured acquisition performance rather than actual learning effects. Lastly, the authors mention that there are other systems that may contribute to category learning that are not discussed in the paper (e.g., hippocampal learning system).

Mané, A. M. (1984). Acquisition of perceptual-motor skill: Adaptive and part-whole training. *Proceedings of the Human Factors and Ergonomics Society, 28,* 522-526.

Adaptive Training Intervention(s): Changes in speed of external task elements
Type of Study: Empirical

The purpose of this study was to compare and contrast the effectiveness of three different training strategies: part task, adaptive, and whole. The Space Fortress video game was used, allowing participants to (a) move their ship around the display using a joystick, while firing shots at and avoiding shots from the enemy fort, and (b) identify mines as friend or foe and destroy foes. Game points (the dependent variable) were added by hitting either the fort or the enemy mines; points were deducted when own ship was destroyed. Four experiment groups were created based on the three training strategies: Group 1 – participants completed a series of part training (PT) tasks, then transitioned to a fixed (whole) training task (FT) for the remainder of the experiment trials, Group 2 – participants completed an adaptive mine speed training task (AT10) until criterion was reached, then transitioned to FT, Group 3 – participants completed a different adaptive training task (AT5), then transitioned to FT, and Group 4 (control group) – participants completed only FT for the duration of the experiment. The author hypothesized that AT would be better than FT, with a transfer effectiveness ratio greater than one.

For the FT group, participants practiced the game at its normal speed (20 units). In the AT5 condition, the mine‟s speed decreased from 20 units to 5 units whenever own ship was hit, and the speed of the fort and own ship consequently decreased at specified rates. In the AT10 condition, mine speed decreased to 10. In both conditions, when a participant destroyed a foe mine or a fort, the speed of the mine, fort, and own ship increased positively at the same rates. The PT condition consisted of three subtasks: (1) the button press motor subtask, which mimicked the foe identification task, (2) the letter appraisal subtask, where participants differentiated foe mine letter designations, and (3) the ship control perceptual-motor subtask, where participants maneuvered own ship at slow speeds.

One hundred right-handed, male university students with normal to normal-corrected vision were recruited. Sixty participants who completed three blocks of the aiming task and achieved an unspecified performance criterion were assigned to either the PT, one of the two AT, or the FT group and were matched on performance. It is unclear as to why the other 40 participants were not included in the experiment. Participants completed five blocks for the first session (except the PT condition which had 4 blocks at the start), seven for the second, eight for the third, then any additional blocks. Training ended when own ship was destroyed as much as the fort was destroyed or three training sessions (minimum of 20 blocks) occurred. To allow for better comparison with FT data, the mine speed was set at 20 units for three probe blocks – the 5th, 12th, and 20th blocks across the first three training sessions.

An analysis comparing the ability of participants to keep own ship out of the fort‟s line of fire was conducted. Of the four test conditions, the average number of ship turns was significantly worse for the AT5 participants, $F(3, 40) = 3.07$, $p < .05$. Essentially, when the fort moved slower, participants would better maneuver away from it. However, when it moved faster, there was not enough time to shoot and react. ANOVA‟s revealed significant differences between the three probe blocks and the training conditions, $F(3, 40) = 3.50$, $p < .05$, and between the average learning curves among the four training conditions, $F(3, 40) = 3.80$, $p < .05$. When compared to FT participants over the first 20 training blocks, PT participants had, on average, more points than FT participants, $t(20) = 2.76$, $p < .05$, and completed fewer blocks, $t(20) = 1.85$, $p < .05$. As depicted in a graph, PT participants had the best learning curve overall. Therefore, the author concluded that the PT condition was more effective compared to FT, with a transfer effectiveness ratio greater than one. However, the author does not state if the hypothesis for the transfer effectiveness ratio between the AT and FT conditions was supported. While not statistically significant, a comparison of the average scores for the three probe blocks for the AT5, AT10 and FT conditions revealed (a) AT5 was not better than FT, (b) AT10 had higher scores on these blocks than the FT, and (c) AT10 took fewer blocks than FT to reach criterion. A second data set was created, using similar AT10 and FT scores from identically numbered blocks; no definition of a similar score was provided. The authors concluded that AT10 participants performed better than FT participants, $t(20) = 2.03$, $p < .05$.

Finally, a stepwise regression was performed, using the scores from the 20th block as the dependent variable. The screening score accounted for 21.33% of the variance, PT accounted for 6.86% and AT10 accounted for 2.51%, with multiple R^2 accounting for 30.71% of the variance. The AT 5 condition did not increase the multiple R, and was not included in the equation. A second stepwise regression was performed using the total number of blocks used to completion; some participants finished more than 20 blocks during the experiment. In this analysis, the screening point and AT10 accounted for 5.34% of the variance and AT10 alone accounted for 4.04% of the variance.

A strength of this study was the discovery of a negative transfer effect for the AT5 condition. However, there were several weaknesses in this study. First, the author fails to hypothesize about the performance of the PT participants in relation to the AT and FT participants. This made interpretation of the results for the PT group more difficult. Second, the experimental design and description of dependent variables was unclear at times which limits our confidence in the validity of the results and conclusions based on the results. Additionally, the authors used stepwise regression to analyze the scores from the 20th block. Stepwise regression has a number of statistical issues (e.g., results can differ depending on whether variables were added or subtracted either forward or backward) and should be avoided (Cohen, Cohen, West, & Aiken, 2003). Therefore, regression results should be reanalyzed utilizing a theoretical reasoning for the adding and/or subtracting predictor variables.

Norman, D. A. & Matheny, W. G. (1972). Need for within-trail feedback as a function of task similarity in adaptive training of manual control. *Proceedings of the Human Factors and Ergonomics Society, 16*(1), 136-139.

Adaptive Training Intervention(s): Between vs. Within Trial feedback
Type of Study: Empirical

This article focuses on adaptive training and the effects of within-trial feedback of difficulty level during the course of training trials and transfer task performance. Past studies by the authors have shown adaptive training to be superior to conventional training, even though difficulty levels were not displayed within those experiments. Kelley (1969), on the other hand, advocates for continuous display of task difficulty level (i.e., task-inherent feedback) because it will provide progress information to the trainee. The authors‟ objective was to investigate if feedback about difficulty level will affect the number of trials to reach criterion in an adaptive training environment.

The experiment compared four randomly assigned groups of participants on the number of trials it took to reach a criterion level of performance. The experiment utilized 30 volunteers (college student and Navy recruits), all males between the ages of 18-25, and capable of using Jaeger No. 1 binoculars. These groups included: 1) a control group which practiced the transfer task to criterion and received between trials summary feedback, 2) a control group which practiced a fixed difficulty training task to criterion and transfer task practice with both receiving between trials summary feedback, 3) an experimental group (#1) practiced the adaptive training task to criterion with between trials summary feedback followed by practice on the transfer task also with between trials summary feedback, and 4) an experimental group (#2) that received within trial feedback of difficulty level and summary feedback at the end for both training and transfer tasks. The number of training and transfer trials was measured in addition to the number of failures and cost.

The training task utilized in this experiment was a simulated head-up display with a side arm controller in order to emulate the flight of an aircraft while the transfer task was representative of an actual aircraft. Six participants had to pass each of the conditions within a 20 trial allotment otherwise, they were rejected. Each trial was two minutes long and feedback scores for past trials was displayed for ten seconds immediately afterwards. After four trial blocks, two minute breaks were given, which were signaled by tone pulse (i.e., breaks-one tone pulse, end of break-two tone pulses). For within trial feedback, a vertically lit meter displayed the difficulty level on the screen. The study also developed an adaptive logic formula to allow a one degree pitch angle error for the adaptive training conditions.

The authors used the Mann-Whitney U test to analyze the data and results showed there was no significant reduction in training trials to criterion, $U = 12.5$ for the within trial feedback as expected. Instead, significance appeared during trials to criterion on the transfer task, $U = 5.5$, $p < .03$ and for cost, $U = 2$, $p < .004$. Results also indicated zero transfer under the no-feedback and non-adaptive condition due to the dissimilarities in tasks. Participants in experimental group #1 seemed to need more trials to pass. This study provides premise for further studies to investigate if within trial feedback of difficulty level is more effective when training and transfer tasks are dissimilar. The within trial feedback condition was less expensive, had zero failures, and the average transfer trials were less than those in the other conditions.

Unfortunately, the training task used in this study was too dissimilar to the transfer task to understand if within trial feedback would provide better results than past studies have shown without it. Another consideration is the small sample size for each condition. This study also brings to light the importance of having a training simulator accurately portray the transfer task because it is unclear if the results of the other conditions were due to the feedback or the lack of fidelity.

Romero, C., Ventura, S., Gibaja, E. L., Hervas, C., & Romero, F. (2006). Web-based adaptive training simulator system for cardiac life support. *Artificial Intelligence in Medicine, 38*(1), 67-78.

Adaptive Training Intervention(s): Adaptive/Non-Adaptive Training
Type of Study: Empirical

The use of simulation-based training has not been as widespread in the medical community as it has been in other communities such as aviation and the military. Therefore, the authors discuss the development and evaluation of an adaptive, web-based training system for cardiac life support. While the major focus of this article is on the development of the system architecture, the authors also present a preliminary evaluation of the adaptive training system. Specifically, the authors were interested in determining if students using an adaptive system could "obtain better or equal scores in less time (p.76)" as compared to students using a non-adaptive system.

Fifty-eight students participated in a 3-day cardiac life support course. After completion of the course, students used either the non-adaptive or adaptive training system to receive additional training in cardiac life support. In the non-adaptive version of the system, students could navigate freely through content, take classic (i.e., not computer adaptive) tests, and complete case studies. In the adaptive system, students were guided through course content, test questions, and case studies based on their current level of performance. For example, difficulty of case studies (e.g., novice, medium, expert) would be tailored to student‟s performance on tests and case studies encountered previously in the training. Comparisons were made between the groups on performance variables such as time spent interacting with the system, performance scores, and the number of items completed in the knowledge test and the case studies.

The results indicate there was a significant reduction in the total number questions and case studies used in the adaptive version (F = 37.8, F= 54.4, p<.05, respectively; NOTE: Degrees of freedom were not reported). As such, there was also a significant reduction in time needed to complete the tests and case studies (F= 3.4 & F = 36.9, respectively; NOTE: neither degrees of freedom nor p values were reported). Despite the decrease in time spent, the final scores obtained for each group on the tests and case studies were statistically equivalent.

While there was no difference between the adaptive and non-adaptive groups in performance, it did require less time to complete. This result indicates that adapting training based on prior knowledge or performance may be a more efficient way to train. However, this interpretation should be taken with caution due to the lack of detail regarding the experimental design, methods, and data analytic techniques used. For instance, it is unclear if participants were randomly assigned to experimental groups or if there were an equal number of participants in each group. Additionally, there is a lack of detail regarding the training the participants underwent before being assigned to one of the two conditions. Finally, the authors claim that their adaptive system can improve performance when their results showed no differences between the groups on knowledge tests and case studies.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C
INDIVIDUAL DIFFERENCES

Batka, J. A. & Peterson, S.A. (2005). The effects of individual differences in working memory on multimedia learning. *Proceedings of the Human Factors and Ergonomic Society, 49,* 1256-1260.

Adaptive Training Intervention(s): Multimedia presentation
Individual Difference Variable: Working memory
Type of Study: Empirical

The current study investigated individual difference in working memory and its effect on learning outcomes. The authors wanted to replicate the findings of Mayer and Moreno (2003) who found support for three basic multimedia design principles: contiguity, redundancy, and modality. The contiguity principle suggests that related visual and auditory material should be presented simultaneously rather than in succession. The redundancy principle discourages the presentation of similar information when the current material provides an effective explanation. Theoretically, violation of these principles would increase cognitive load and learning demands. Finally, the modality principle suggests that learning can be enhanced when information is provided in both the visual and auditory formats simultaneously. The authors of the current study hypothesized that individuals with larger working memory capacities would perform better in a multimedia learning activity than those with smaller working memory capacities. Furthermore, the authors suggest that Mayer and Moreno‟s (2003) multimedia design principles would be less applicable for individuals with larger working memory capacities because they should be less affected by an increase in cognitive load.

One hundred and eighty college students completed a computer-based trainer on weather formations (i.e., how hail forms). The total duration of the task was 2 minutes. Participants were randomly assigned to one of four levels of multimedia presentation; animation (visual) and narration (auditory) simultaneously (AN), animation and narration sequentially (A/N), animation and on-screen text simultaneously (AT), and finally, simultaneous presentation of animation, narration, and text (ANT). The contiguity principle suggests that learning should be better for those individuals in the AN condition than in the A/N. The redundancy principle suggests that learning should be better in the AN condition than the ANT condition. Lastly, the modality principle suggests that learning should be better in the AN condition than the AT condition. The dependent variable consisted of four problem solving questions (transfer questions) on hail formation (a total of 8 possible points). Turner and Engle‟s (1989) operation span (OSPAN) task and Shah and Miyake‟s (1996) letter rotation task were used to assess working memory.

Correlational analyses indicated that participants with higher working memory structures (measured via OSPAN) performed significantly better on the transfer questions ($r = .158$; $p < .05$). Initial results using an ANOVA did not find support for Mayer‟s principles. However, after re-analyzing the data using working memory capacity (i.e., OSPAN) as a second factor, results supported the contiguity principle for low working memory capacity participants. A significant interaction was found between condition (AN vs. A/N) and OSPAN on the transfer questions, $F(1, 42) = 6.03$, $p < .05$, $ή^2 = .14$, indicating that low working capacity individuals performed better in the AN condition than in A/N. Significant interactions were not found for the other principles. Although for the modality principle, the interaction approached significance, $F(1, 45) = 3.95$, $p = .054$, $ή^2 = .09$. Analyses using the letter rotation scores revealed no significant differences. In general these results suggest that individuals with larger working memory capacities tended to perform better on multimedia tasks than those who are lower. In addition, it seems that high working memory capacity individuals do not benefit from Mayer‟s principles. These results lend support to the idea that both individual differences (i.e., cognitive capabilities) and the cognitive demands of the task must be considered when developing a multimedia training task. Limitations of the study include the short hail task, the study population, and the use of ANOVA during the data analyses. The hail formation task may not generalize to more complex tasks (e.g., submarine navigation). In addition, the generalizability of this study to other populations could be an issue because college students may have higher levels of working memory capacity than the general population. Finally, it may have been more beneficial for the authors to use multiple regression to analyze the interactive effects of multimedia presentation and working memory capacity on learning. For example, using regression would have allowed the use of the OSPAN measure without having to split it into high and low quartiles. Creating the split causes a loss of systematic variance and could be a reason for the non-significant interactions.

Beier, M. E. & Oswald, F. L. (2009). Adaptability and complex performance: Can you really be too smart for your own good? Manuscript submitted for publication.

Adaptive Training Intervention(s): None
Individual difference variable(s): Cognitive ability
Type of Study: Review/Theoretical

Research has demonstrated that, in general, higher levels of cognitive ability yield superior performance in work and academic settings, as well as in everyday life. However, some research has also shown that higher amounts of cognitive ability can be detrimental for adaptation to changes that occur in complex tasks. The authors of this paper believe that these findings could be explained when analyzed in the context of cognitive theory. They note that there are three research paradigms in which contradictions regarding general cognitive ability arise. The first of these is the "pressure to perform" paradigm. Studies performed in this area have found that, on difficult tasks, participants with higher cognitive ability are more likely to suffer performance decrements when pressure is high compared to participants with low cognitive ability. Because of this, researchers have claimed that those with lower cognitive ability are more adaptable to stressful situations. However, Beier and Oswald (2009) disagree with this assertion and offer a different explanation of the results. They suggest that the differences observed in these studies are the result of two different types of tasks: resource-limited and resource-insensitive tasks. Resource-limited tasks are tasks in which performance could be improved if more attentional resources were dedicated to their execution. However, because attentional resources are limited, performance decreases when the attentional demands of the task increases or when attention is diverted. On the other hand, resource-insensitive tasks are either very easy or very difficult, so that increase in attentional resources would not improve performance. When easy tasks are used, both high and low cognitive ability participants perform well because easy tasks are resource-insensitive for both groups. More difficult tasks, however, are resource-limited for high cognitive ability participants, but resource-insensitive for low cognitive ability participants. Because of this, participants with lower cognitive ability do not suffer performance decrements; they perform poorly regardless of the amount of resources devoted to the task. For the participants with higher cognitive ability, more difficult tasks are resource-limited. A performance decrement occurs because they can perform the task well when they are able to devote sufficient resources to the task (low pressure), but poorly when they cannot (high pressure). It would appear that high cognitive ability participants suffer a larger performance decrement, however, this is due to their superior performance when pressure is low, whereas low ability participants perform poorly regardless of pressure condition.

The second paradigm is the "proceduralized learning" paradigm. The authors describe procedural tasks as those which are processed automatically without the need for attentional resources and whose rules cannot be stated explicitly. In contrast, rule-based tasks have simple rules which are easily verbalized. Researchers who use this paradigm hypothesize that high ability people are more likely to develop and follow rules and strategies which interfere with performing these tasks. On the other hand, they postulate that those with lower cognitive ability will apply automatic or procedural processes which allow them to be more flexible. One study noted by the authors found that while participants with high cognitive ability performed better on a rule-based task, they were outperformed by those with lower cognitive ability on a procedural task. Beier and Oswald suggested that the nature of the task and the mastery criterion chosen led to these results. The mastery criterion for each task was eight correct categorizations in a row. Beier and Oswald propose that participants with lower cognitive ability were able to use automatic processing to reach the mastery criterion faster than participants with high cognitive ability could figure out the underlying rules of the task. They suggest that if the mastery criterion was higher, the results for the procedural task would be reversed. Indeed, another study performed using a higher mastery criterion found this reversed effect where those with higher cognitive ability outperformed those with lower cognitive ability. They also suggested that the procedural task used in these studies was resource-limited for those with higher cognitive ability, but resource-insensitive for participants with lower cognitive ability.

The third paradigm in which those lower in cognitive ability seem to perform better is "skill acquisition and adaptive performance." Lang and Bleise (2009) performed a study using a complex task which also included an adaptive component. They were interested in two different kinds of adaptation: transition adaptation and reacquisition adaptation. Transition adaptation occurs directly after a change is made and defines how quickly the participant can adapt to those changes to minimize decrements in performance. This kind of adaptation is measured relative to the learning level at the end of the original task. Reacquisition adaptation, which is the recovery rate after a change, is measured using the learning rate after the changes occur while controlling for the rate of learning before the change. Lang and Bleise found that after a change in the task, participants with higher cognitive ability suffered a greater performance decrement that those lower in cognitive ability. In other words, transition adaptation was lower for higher cognitive ability participants. They concluded that higher ability participants did not adapt to change, as well as participants with lower abilities. Beier and Oswald believe that there is a better explanation of these results. According to theory on skill acquisition, performance of novel tasks depends on cognitive ability, while more practiced tasks become automatic and depend less on cognitive abilities. The authors propose that higher cognitive ability participants had "more to lose" than lower cognitive ability participants because they had learned more of the original task in the first place. They also suggest that the drop in score would be greater for higher ability participants because their initial scores were higher before the task was changed and then once the change occurred their scores were comparable to those of the lower ability group. Another argument they make is that because the task changed, the scoring system changed, and it is therefore inappropriate to compare scores across tasks. Lastly, the authors suggest that to test the hypothesis that higher ability participants are less adaptable, an experiment should be performed where participants of both high and lower ability must adapt to inconsistent rules across trials.

The authors suggest that future research carried out in these areas should focus not only on individual differences in cognitive ability, but the nature of the tasks being used. They also suggest that a wider range of task difficulties should be utilized; most of the studies in these areas have used either very easy or very difficult tasks. In addition, the authors recommend that future studies examine a wider range of cognitive ability, as much of the research done has dichotomized participants into low and high groups which may have distorted the results that were obtained. Despite their critiques, the authors note that the paradigms highlighted do offer avenues for future research. They suggest, for example, that future research should explore the tradeoffs between automatic processing and adaptability. While the authors disagree with Lang and Bleise‟s (2009) interpretation of their results, they believe the methodological framework developed is useful for future research on adaptability.

Day, E. A., Boatman, P. R., Kowollik, V., Espejo, J., McEntire, L. E., & Sherwin, R.E. (2007). Collaborative training with a more experienced partner: Remediating low pretraining self-efficacy in complex skill acquisition. *Human Factors, 49*(6), 1132 – 1148.

Adaptive Training Intervention(s): Collaborative training
Individual Difference Variable(s): Self-efficacy
Type of Study: Experimental

Collaborative learning is characterized as an "instructional technique in which individuals are placed in small groups or pairs while learning a specific task and are encouraged to communicate with their partners by sharing ideas and working together toward a common goal" (p. 1133). The researchers wanted to examine the effects of collaborative learning on skill acquisition and self-efficacy. Specifically, they wanted to see whether collaborative learning would have an effect on skill acquisition and mid and post-training self-efficacy. The authors hypothesized that collaborative training would be more beneficial for individuals with low pre-training self-efficacy than for individuals with high pre-training self-efficacy. They postulated that the differential influence of collaborative training on skill acquisition would be mediated by mid-training self-efficacy and practice performance.

Two hundred twenty-three right-handed males who were enrolled in an introductory psychology course participated in the experiment. Some of the participants were randomly assigned to one of the two experimental conditions (i.e. individual or collaborative) while the others were assigned to be more experienced partners for participants in the collaborative condition (i.e., quasi-experimental design). Twenty-four participants dropped out of the study which made it necessary to reassign some participants to the individual condition leaving 87 participants in the individual condition, 68 in the collaborative condition, and 68 as more experienced partners. Participants played six 10-game sessions of a computer game called Space Fortress. The last two games of each session were test games. Participants in the individual training condition performed all practice and tests by themselves. In the collaborative training condition, participants worked with a more experienced partner during the practice games of sessions 1 and 2 (for the first practice day) but performed the test games without aid from their partners and trained individually for sessions 3 through 6. Self-efficacy was measured before session 1 and after sessions 2 and 6.

The results indicated that participants in the collaborative condition had greater skill acquisition than participants in the individually trained condition $F(6, 546) = 2.68$, $p < .05$, $\eta^2 = .04$. Additionally, participants with high pre-training self-efficacy had greater gains in skill acquisition from baseline to the end of training than participants with low pre-training self-efficacy $F(6, 546) = 5.80$, $p < .001$, $\eta^2 = .06$. Collaborative training also yielded different practice effects for participants in the low versus high pre-training self-efficacy groups. Participants with low pre-training self-efficacy that trained with an expert experienced higher performance levels during practice than did participants with low pre-training self-efficacy that trained individually $F(2, 182) = 6.18$, $p < .01$, $\eta^2 = .06$.

Results also showed that collaborative training had benefits for participants that reported low self-efficacy prior to training. The low pre-training self-efficacy group had substantially larger changes in self-efficacy scores than the high pre-training self-efficacy group, $F(2, 182) = 11.50$, $p < .001$, $\eta^2 = .11$. Lastly, "practice scores mediated the differential influence on individual skill acquisition that collaborative training had on individuals with low pre-training self-efficacy relative to individuals with high pre-training self-efficacy" (p. 1143) ($\Delta R^2 = .22$, $p < .001$) which the authors argue is the reason for the differential influence of the collaborative training on skill acquisition.

The level of expertise of the partners in the collaborative condition may limit the generalizability of these results. In the current study, participants in the collaborative training condition worked with a more experienced partner. Past research suggests that a team members" level of experience can play a role in the number of training opportunities and the instances that team members can practice teamwork (Schaafstal, Hoeft, &Van Schaik, 2002). Future research should address the question of whether collaborative research is beneficial when two novice trainees are paired together. In addition, the generalizability of these finding is limited due to the sample used (i.e., right-handed males). The results of this study suggest that collaborative training may benefit individuals with low self-efficacy, however, it did not increase their skill acquisition to the level of those with high pre-training self-efficacy. In other words, collaborative training only weakened, rather than removed, the negative effects of low pre-training self-efficacy on skill acquisition.

DeCaro, M. S., Carlson, K. D., Thomas, R. D., & Beilock, S. L. (2009). When and how less is more: Reply to Tharp and Pickering. *Cognition, 111,* 397-403.

Adaptive Training Intervention(s): None
Individual Difference Variable(s): Working memory capacity
Type of Study: Experimental

In their 2008 study, DeCaro, Carlson, Thomas, and Beilock showed that those with less working memory (WM) outperformed individuals with higher WM on an information-integration task. "It is not always a positive correlation between individual differences in WM and performance is not always positive" (p. 397). Individuals higher in WM capacity learned rule-based categories faster, but learning was slower for information-integration categorization tasks. They postulated that individuals lower in WM had less capacity to test and switch among rules and would learn information-integration categorization tasks quicker.

Recently, Tharp and Pickering (2009) questioned participants‟ use of procedural learning during information-integration categorization tasks. They showed that classification of the information-integration stimuli could be done correctly 75% of the time using "simple explicit rules" (p. 398). Their study found that individuals used these rules and scored high on the information-integration tasks. However, individuals who achieved eight consecutive correct responses (CCR), a criterion, were unable to maintain accuracy during the next eight trials (16 CCR). Individuals benefited from the simple rules when lower performance levels were desired but suffered when demands were higher. The individuals using these rules had lower WM capacity, prompting DeCaro et al. to reassess their previous study. In the present study, DeCaro et al. hypothesize "if low WM individuals are using simple strategies that circumvents complex hypothesis testing in WM but does not produce optimal category learning, then they should not excel on the information-integration task in question at a stricter 16 CCR learning criterion" (p.398).

Thirty undergraduate students who scored in the upper and lower quartiles of two measures of WM were selected. Participants completed four different category learning sets comprised of two rule-based and two information-integration category structures. The authors presented the same tasks used in DeCaro et al. (2008), but built 13 learning blocks and randomly sampled all 16 stimuli within each block without replacement. This gave them a maximum total of 208 trials. Rule-based and information-integration category learning were assessed and four learning scores were taken for each individual: the number of trials needed to reach 8 CCR and 16 CCR (averaged across the two rule-based categories) and the number of trials needed to reach 8 and 16 CCR averaged across the two information-integration categories.

A 2x2 ANOVA was used to analyze the eight CCR criterion. A main effect was found for category structure, $F(1, 22) = 65.77$, $p < .001$. The results also indicated a significant WM X category structure interaction, $F(1, 22) = 4.11$, $p = .05$. At eight CCR, DeCaro et al.‟s 2008 results were duplicated; having higher WM produced faster learning of rule-based categories, relying on explicit hypothesis testing. The "less-is-more" effect was replicated as those with less WM learned the information-integration categorization tasks faster. When testing at 16 CCR criterion, a main effect was found for category structure, $F(1, 25) = 565.30$, $p < .001$ and WM, $F(1, 25) = 8.16$, $p < .01$. However, there was no significant interaction between WM and category structure. In rule-based tasks, high WM participants outperformed those with low WM and also learned the information-integration categories faster. These findings align with the critics‟ thoughts that training at eight CCR does not reflect a reliance on learning.

Strategies were assessed over the 16-block trials and results showed that those with low WM favored one dimension rules (considered a shortcut) and though participants with high WM also used them, it was not as frequently and they switched between strategies. Participants with high WM also increased their use of optimal strategies, $F(1, 14) = 13.03$, $p < .01$, where those with low WM did not, $F(1, 11) = 3.32$, $p > .05$. Low WM participants were more likely than high WM to use one dimensional rules $M = 0.36$, 95% CI [0.30, 0.43]. High WM participants were more likely to use three dimensional rules $M = 0.23$, 95% CI [0.20, 0.26] and the optimal strategy $M = 0.22$, 95% CI [0.15, 0.29]. To summarize, low WM individuals depended on simple one-dimension rules and were quicker to attain eight CCR than high WM individuals. High WM individuals initially used explicit rules but they did not coincide with early learning as those with high WM were less inclined to continue using one-dimension rules that led to early success.

As both DeCaro 2008 and 2009 showed, high WM participants need more time to achieve eight CCR than low WM participants. Their "less is more" approach was found to be true in low WM participants achieving early success but less was not more when a more stringent criterion was required. DeCaro et al. (2009) now interpret the less is more hypothesis by crediting simple one-dimension rules instead of procedural-based responding. Tharp and Pickering (2009) were accurate in their assertion that less is not more when dealing with low WM individuals and the rules they rely on.

The results of this study show that adaptive training may benefit from knowledge of individual differences in not just cognition, but working memory. Adaptive training could be successful in teaching tasks since it accounts for individual differences and also allows for variations in working memory.

DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*, 284-294.

Adaptive Training Intervention(s): None
Individual difference variable(s): Working memory capacity
Type of Study: Empirical

Research has shown that, in general, people with more working memory (WM) capacity perform better on complex tasks than those with lower WM capacity. The authors wanted to demonstrate situations in which this pattern is reversed, when those with lower WM outperform individuals with higher WM. They examined how differences in WM capacity affect the ease in which participants learn different category structures. The category structures chosen were rule-based and information-integration categorization. Rule-based categorization tasks have rules which can be explicitly stated and are thought to be reliant on available WM resources. Information-integration categorization tasks, on the other hand, do not have rules that are easily verbalized and require participants to "integrate stimulus values across multiple dimensions prior to making a categorization decision" (p.286). According to the authors, these types of tasks are more reliant on the participant making associations between category members and correct responses (stimulus-response mappings). Participants who are presented with a categorization task try to figure out the underlying explicit rules while simultaneously developing stimulus-response mappings. When the participant finds the correct method for the task, that method "wins". In other words, explicit learning dominates if a rule is found and procedural learning takes over if one is not. The authors hypothesized that participants with higher WM would have lower performance scores on an integration-information task. According to the authors, participants with higher WM would spend longer trying to find an explicit rule, which would impede the ability of procedural processes to take over.

Seventy-one undergraduate students participated in an experiment where they performed both types of categorization on a computer. The stimulus chosen was a square with embedded shapes. It varied in background color (yellow or blue), embedded figure shape (circle or square), embedded symbol color (red or green) and number of embedded symbols (one or two). In the rule-based task, the participants were required to place the squares into categories based on one dimension. In the information-integration task, three dimensions were used. Each dimension was coded with a value of either +1 or -1 and labeled X, Y, or Z. For this task, the rule was "If value(X) + value(Y) + value(Z) > 0, respond category A, otherwise respond category B" (p.289). Participants completed four sets of 200 trials each. There were two sets of the rule-based task and two of the information-integration task. After the categorization task, participants filled out a questionnaire which assessed the amount of pressure they felt to perform well on the task. Additionally, they completed two measures that assessed WM.

The authors measured how many trials it took for participants to reach the mastery criterion (eight correct trials in a row). Based on a regression analysis they found a significant interaction between WM capacity and category structure ($\beta = .61$, $t = 3.11$, $p < .01$), such that participants with higher capacity took significantly less trials to learn the rule-based task ($\beta = -.24$, $t = 12.02$, $p < .05$), but significantly more trials to learn the integration-information task ($\beta = .27$, $t = 2.36$, $p < .03$). The results indicated that pressure to perform was not correlated to WM scores.

The results suggest that higher WM capacity is detrimental when performing tasks that do not include an easily discovered set of rules. The authors argue that participants with higher capacities use complex computational processes that hinder the procedural learning processes that must occur during information-integration tasks. However, it is possible that the mastery criterion chosen for this study was too low. Beier and Oswald (2009) suggest that participants with lower WM capacity found shortcuts to reach the criterion, but that a higher mastery criterion would allow the participants with higher WM capacity to discover the underlying rule, and thus change the pattern of results. DeCaro (2009) supported this hypothesis with the finding that when the mastery criterion was raised to 16 correct answers in a row, the pattern of results reversed; participants with higher WM capacity outperformed those with lower WM capacity.

Kline, K. A. & Catrambone, R. (2009). The influence of spatial ability on multimedia learning. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting,* 1249-1253.

Training Intervention(s): Multimedia presentation
Individual Difference Variable(s): Spatial ability
Type of Study: Empirical

This paper reports the results of an investigation of the moderating effects of spatial ability on multimedia learning. It has been suggested that multimedia instruction, which presents instructional content through both text and pictures (or animation), results in better learning than traditional text-based instruction. The reason for this multimedia effect has been explained in regards to Paivio"s dual coding theory and Baddeley"s separate working memory processing channels. For instance, if one channel is overloaded, you can reduce the load by presenting instructional content in the alternate channel which increases the likelihood that the student can process the information. Further, previous research has shown that spatial ability has moderated the effect of simultaneous (narration combined with animation) and successive (narration followed by animation) presentation of multimodal (auditory and visual) content such that students with high spatial ability benefited from simultaneous rather than successive presentation of instruction. The current study investigates if the moderating effect of spatial ability would exist with unimodal, simultaneous presentation of instruction (e.g., if the visual channel is overloaded with the concurrent presentation of text + pictures). The authors hypothesize the following: When learning material with a spatial component, (a) individuals with high spatial ability would benefit from the simultaneous presentation of content, and (b) low spatial ability individuals will not be able to integrate visual and verbal information in working memory and will not benefit from simultaneous presentation.

Twenty-one participants (10 males, 11 females) volunteered for the study. The researchers implemented a 2 (lesson format: text only vs. text + pictures) x 2 (time of testing: immediate vs. delayed one week) within-subjects design. Participants completed lessons, with a spatial component (such as how air pumps work), that were presented either in text or text plus pictures. Formats and lesson orders were counterbalanced. Immediately following the lessons, participants completed a knowledge test with questions assessing recall of the lessons. One week later, participants were required to complete an additional set of recall questions. Spatial ability was assessed using both the cube rotation task and surface development task developed by the Educational Testing Service. A median split was used to group participants into low and high spatial ability groups. Results showed a main effect of lesson format indicating improved performance from text + pictures than from text only on the immediate test, $F(1, 17) = 8.87, p < .01$. There was a significant interaction with only one of the measures of spatial ability. Specifically, the interaction between cube rotation performance and instructional format on performance of the immediate test was significant indicating that highly spatial participants who received text + pictures performed better on the immediate test than both high spatial participants who received text only and low spatial participants, $F(1, 17) = 7.83, p < .05$ (Note: no post hoc analyses were reported). Neither a significant 3-way interaction (lesson format x time of testing x spatial ability) nor a significant 2-way interaction with the alternate measure of spatial ability was found.

The authors conclude that their study supports prior research which suggests multimedia instruction is more effective than text-based instruction. Further, they conclude that their hypothesis regarding the moderating effects of spatial ability was partially supported. However, there are a few limitations and areas of concern especially with regard to statistical conclusion validity. First, the hypothesis regarding low spatial individuals was unclear. Based on their description, I could not determine if they were hypothesizing the null (e.g., no performance differences for low spatial individuals based on lesson format) or if they had a directional hypothesis (e.g., low spatial individuals who receive text will outperform those who receive text + pictures). It is unclear why the authors degraded the spatial ability data from a continuous variable to a dichotomous variable. Lastly, they did not report on the main effect of instructional format on the delayed test nor post hoc analyses for the 2-way interaction.

Krause, U., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning & Instruction*, *19*(2), 158-170.

Adaptive Training Intervention(s): Cooperative Learning, Example-based Feedback
Individual Difference Variable(s): Prior knowledge
Type of Study: Empirical

The purpose of this study was to determine how cooperative learning and example based feedback affects objective performance and subjective belief of performance. The study utilized the e-learning statistics program Koralle, which is designed to increase competence in statistics, specifically correlation analyses. The two interventions employed were social context of learning (cooperative vs. individual) and feedback intervention (available vs. unavailable). Feedback consisted of ratio correct, correct answers with elaboration, and corrective information for errors in the learners" incorrect answer(s). Furthermore, the effects were evaluated objectively and subjectively. The subjective effects measured the learners" perception of both performance and competence.

Koralle is geared towards learners with prior knowledge in statistics and is designed to help students on often missed key points within statistics (e.g., linearity and effects of outliers). Within the Koralle program, participants answer a problem and then worked examples are provided, allowing the learner to compare and contrast their answers with the correct answer. Therefore, these worked examples are a form of feedback, but mainly for students with prior knowledge in statistics. A previous study using Koralle (Tyroller, 2005) found that student scores were higher on tests of correlation analysis when Koralle was used as compared with students who only attended a lecture of the same content. However, deficits in deep knowledge structure, especially in subjects that participants had little prior knowledge on, were evident. The authors think this may be partially due to e-learning being a solitary process such that a lack of social interaction may have adverse effects on learning. They posit that those with little prior knowledge do not compare their answers with the worked examples, thus receiving no or inaccurate feedback. Previous studies have shown that, in general, cooperative learning aids in higher-order thinking and learning (Cohen, 1994). Research suggests that cooperation should be centered on elaboration of the material, especially when computers and groups are involved. Research also indicates that elaborate feedback increases performance more so than ratio of correct answers or correct answer feedback, individually. However, research also suggests that students do not always take full advantage of the information provided by elaborated feedback (See Hancock, Thurman, & Hubbard 1995; Krause, 2002). Therefore, the authors suggest using cooperative learning because groups will use elaborate feedback more effectively than individuals. The authors hypothesized that (1a) cooperative learning and feedback will promote objective learning outcomes when compared to individual learning and no feedback, (1b) feedback should be especially beneficial to those students who work in groups, (1c) feedback should be more beneficial for students with little prior knowledge, (2) groups will be more successful in problem solving than individuals and, (3)cooperative learning and feedback will enhance students" perceived performance and competence.

One hundred and thirty-seven participants (*M* = 23.82, *SD* = 5.08, 105 females) were asked to voluntarily e-mail their grades and background knowledge in statistics. A 2 (individual vs. cooperative) X 2 (feedback available vs. not available) fully between subjects experiment with pre-and post-test was used. Students were randomly assigned to each of the four conditions. However for the collaborative learning groups, homogeneous groups were created based on the students" background knowledge and grades, making full randomization not possible. The use of homogenous ability groups in this research was established because heterogeneous ability groups may have high ability students dominating. Thirty-five students worked individually and 102 students worked in pairs (51 dyads). The dyads solved problems interactively and afterwards all participants dealt with six problem solving tasks on correlation analysis. Following this, students received the worked example. Those in the additional feedback conditions also received six multiple-choice tests with adaptive, elaborated (i.e., they were given the correct answer with a description, as well as if it was right or wrong) feedback. The dyads took the multiple-choice tests together and outcome feedback was provided based on their group performance. Solutions in the learning phase (during task) displayed the group"s performance, but posttest (post task) results demonstrated individual outcomes. The participants" perceived performance and competence were assessed with a one sentence question.

A 2 (social context) X 2 (feedback) ANOVA showed a significant effect of feedback using post test scores $F(1, 133) = 32.91$, $p < 0.001$. However, the main effect of social context (collaboration) was not significant. Therefore, partial support was found for hypothesis 1a. Students who received feedback and worked alone scored higher on the posttest in comparison to dyads that received feedback, $F(1,133) = 5.03$, $p < 0.05$, failing to find support for hypotheses 1b. Using a 2 (prior knowledge) X 2 (social context) X 2 (feedback) ANOVA, results showed that students with little prior knowledge benefited from the feedback intervention, $F(1,129) = 5.94$, $p < 0.05$, finding support for hypothesis 1c. Hypothesis 2 was supported; groups were more successful than individuals in problem solving, $F(1, 82) = 6.61$, $p < 0.05$. Finally, the authors found partial support for hypothesis 3. Social context had a significant effect on perceived performance, $F(1, 133) = 13.85$ $p < 0.001$, and perceived competence, $F(1, 131) = 4.80$, $p < 0.05$. However, feedback did not have a significant effect on students" perceived performance or competence.

Results from the current study show obvious discrepancies between subjective and objective learning outcomes. Possible reasons for using cooperative learning may be to aid in the individual"s internal desire to continue learning, as well as a group understanding of the material, protecting from individual interpretations. However, participants did not produce better individual results when in groups and the authors suggest that micro-analyses may help to further understand the quality of the cooperation process in regards to individual learning. Additionally, all participants (including the no feedback groups) in the current study received the worked example. Therefore, the authors may have found more support for the effects of feedback if they used a true no feedback control group for comparison. Finally, the authors state that future research should investigate motivational and emotional reactions to feedback.

Mayer, R. E. & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*(4), 833-846.

Adaptive Training Intervention(s): None
Individual Difference Variable(s): Visualizer or Verbalizer
Type of study: Empirical- Non-experimental (questionnaire/survey study with no manipulation)

The authors presented a battery of 14 measures to 95 college students in order to explore the hypothesis that some people learn better using text/written words(verbal learners) while others learn better through pictures (spatial learners). The authors wanted to see if the verbalizer-visualizer dimension could be described using three facets: cognitive ability, cognitive style and learning preference. Cognitive ability refers to whether the individual has high or low ability for spatial tasks such as mental rotation while cognitive style refers to how the individual thinks, for example do they think using words or images more often. How an individual prefers to learn, whether through the use of illustrations or through the use of text explanations falls under the category of Learning Preference.

Both subjective and objective tests were used to assess different aspects of the verbalizer-visualizer dimension. The first three measures were the students" scores on the SAT math and verbal sections and their score on a vocabulary test. For the vocabulary test the students were given a target word and asked to find five synonyms in a list within a time limit. These tests measured general cognitive ability. The next three tests were measures of spatial ability. These included two timed tasks, a card rotations test and paper folding task, and a Verbal-Spatial Ability rating questionnaire which was a self-rated measure of spatial and verbal ability on a scale from 1 to 5. There were four tests in the study that sought to measure the cognitive style of the participants. On the Verbalizer-Visualizer Questionnaire (VVQ) the students would rate their level of agreement (from 1: Strongly Agree to 7: Strongly Disagree) with statements related to verbal and visual styles of thinking. The second measure used was an original measure called the Santa Barbara Learning Style Questionnaire which was similar to the VVQ but asked questions more specifically related to learning preference. A computer-based test called the Cognitive Styles Analysis was used to measure cognitive style by calculating response times to statements about visual or verbal information. Lastly, the Verbal-Visual Learning Style Rating rated students" preference for visual or verbal learning on a 7-point scale. The last group of tests measured multimedia learning preference and consisted of two subjective questionnaires and two computer based tasks. The Learning Scenario Questionnaire asked about preference to specific learning scenarios and the Multimedia Learning Preference Questionnaire asked the participants to rate their preference for either text or pictorial learning material in a paper version of a multimedia lesson. The Multimedia Learning Preference Test was a computer-based version of the previous questionnaire. Students could choose either text or pictorial help and also rate which one they found more useful.

Correlational analysis of the scores showed that the first 3 measures correlated highly with one another (r values ranged between .30 - .63, $p < .01$), as did the 3 tests of spatial ability (r values ranged between .23 - .52, $p < .05$), 3 of the 4 measures of cognitive style(r values ranged between .19 - .74, $p <$ .05; the Cognitive Styles Analysis measure did not correlate well), and the 4 measures of learning preference (r values ranged between .25 - .59, $p < .01$). An Exploratory Factor Analysis was done and showed that each of the measures loaded most highly on 1 of 4 categories: Cognitive Style (factor loadings ranged .483-.831), General Achievement (factor loadings ranged .479-.976), Learning Preference (factor loadings ranged .425-.979) and Spatial Ability (factor loadings ranged .399-.744). This supported the hypothesis that the dimension is multifaceted. Sex differences analyses showed no differences between males and females on the measures of cognitive style or learning preference. However, regarding spatial ability, males scored significantly higher than women on the Card Rotations Test, $t(93) = 2.62$, $p = .01$, and the Paper Folding Test, $t(93) = 2.13$, $p = .04$, and marginally higher on the Verbal-Spatial Ability Rating, $t(93) = 1.75$, $p = .08$. For general ability tests, men reported scoring significantly higher than women on SAT Verbal, $t(78) = 2.30$, $p = .02$ and SAT Math, $t(78) = 3.28$, $p=$ .002, and scored marginally higher than women on the Vocabulary Test, $t(93) = 1.89$, $p = .06$.

McGregor, M. U., Schunn, C. D., & Saner, L. D. (2007). *Expertise as effective strategy use: Testing the adaptive strategies model in the ill-structured domain of leadership*. (Technical report 1204). Arlington, VA: US Army Research Institute for the Behavioral and Social Sciences.

Adaptive Training Intervention(s): None
Individual Difference Variable(s): Prior knowledge
Type of Study: Non-experimental (Free-response study with no manipulation)

The adaptive strategies model (ASM) developed by Lemaire and Siegler (1995) predicts that experts and novices will vary in four different dimensions of strategy usage: Existence, Choice, Base rate and Execution. This means that experts will know better strategies than novices (Existence), know which strategy is best for a given situation (Choice), use the strategies that are effective most often (Base-rate) and use commonly known strategies more effectively (Execution). Although the model was tested and validated in well measured and defined constructs (i.e., math and spelling), it had not been validated for constructs that were ill-defined (such as Leadership). For the purpose of this study the authors used the definition espoused by the US Army which describes leadership as "the process of interpersonal influence in which direct and indirect means are employed to get others to accomplish the organization"s goals, where influence is achieved by providing purpose, direction and motivation." (p.3)

Platoon leaders (Lieutenants and Captains, N=66) from the US Army were chosen to participate as experts in this study. The intermediates were undergraduate students (N=23) enrolled in the Reserve Officer Training Corps (ROTC) and the novices were undergraduates (N=39) with no military experience. All participants were given a packet of leadership scenarios encountered regularly by leaders of Platoons and asked to write a short response. The responses were categorized according to what leadership style/strategy was demonstrated and compared to the ideal responses that were identified by experts. No difference was found between groups on the dimension of Existence, however, this was postulated to be the result of the normal acquisition of leadership knowledge that people possess by adulthood. Significant differences were found among groups for strategy choice. Intermediate participants chose better strategies than novices, $t(198) = 2.14$, $p < .05$, and experts chose better strategies than intermediates, $t(272) = 2.57$, $p < .01$, and novices, $t(330) = 5.99$, $p < .0001$). For base rate, experts were more sensitive than novices, $t(103) = 2.18$, $p < .03$. There were no significant differences between intermediates and novices and intermediates and expert participants. A separate study was conducted to test the fourth dimension of strategy usage, execution. In this experiment, Army Lieutenants and Captains (N=26) rated the overall quality of responses obtained by all three groups in experiment 1. Statistically significant differences were found between the three groups. Intermediate and expert participants were rated significantly higher than novices, $t(18) = 3.12$, $p < .01$ and $t(18) = 2.46$, $p < .05$ respectively. No significant difference was found between intermediates and experts, although there was a trend in the direction hypothesized by the ASM. Overall, the ASM model was able to predict differences between experts and novices in 3 out of 4 dimensions: Choice, Base-rate and Execution.

A few alternative explanations for the results obtained in the study exist. It is possible that the differences in strategy use between groups were the result of different internal representations. However, this difference is difficult to measure, especially in the ill-defined domain chosen for this study. As an alternative, it may be possible to change and/or strengthen them by studying the strategies they produce. Lastly, there were a few aspects of expertise in leadership that were not addressed by the study. These were recognition, representation and motivation. A leader must recognize a situation to be a problem in order to start the problem solving process, yet in the study the situations were represented as problems already and no interpretation was necessary. The way the problem situation is framed will also have an impact on how the situation is handled. However, problem representation was not directly measured in this study.

Pak, R., Rogers, W. A., & Fisk, A. D. (2006). Spatial ability subfactors and their influence on a computer-based information search task. *Human Factors, 48*(1), 154-165.

Adaptive Training Intervention(s): None
Individual Difference Variable(s): Spatial ability
Type of Study: Experimental

The authors presented a series of 32 internet search tasks to undergraduate students to determine the influence of spatial ability on computer-based tasks. General spatial ability refers to spatial visualization, the "ability to manipulate or transform the image of spatial patterns into other arrangements" (Ekstrom, French, Harman, & Dermen, 1976) and spatial orientation, the "ability to perceive spatial patterns or to maintain orientation with respect to objects in space" (Ekstrom et al., 1976). Prior research suggests that these subfactors are psychometrically different, but previous literature has not compared the two. The authors suggest that navigating webpage space may be related to spatial abilities. They hypothesized that a more challenging navigational aid (i.e., a webpage map) would be beneficial for those with high spatial abilities and that a less challenging navigational aid (i.e., step-by-step directions) would be less connected with the level of spatial ability.

After a telephone prescreening, 101 participants (48 woman and 58 men, $M_{age} = 19.9$) who were all intermediate level computer users completed the 2 day study. Participants were given standardized tests measuring cognitive ability, attention, psychomotor speed, hearing, and vision. The participants also completed mouse training. In the map condition, information was organized to indicate relative positioning in the webpage and therefore more spatial awareness was required. The step-by-step condition was likened to written driving directions and gave explicit instructions requiring less spatial ability. To keep an equal level of ability distribution, participants were assigned to a condition based on their scores on a spatial visualization test. For each condition, participants searched within eight topic areas on modified webpages. Each area included two levels of difficulty which were manipulated by changing the amount of steps indicated by the navigational aid. Each topic domain included four tasks with two tasks at each difficulty level. Tasks were statements instructing participants to find the webpage that contained a specific bit of information (i.e., "When is the last day to drop a course in spring semester 2002?"). Depending on their condition, participants had either the map or the step-by-step navigational aid to guide them through the domains. Performance was measured by task completion time and number of errors made. An error was made when a participant navigated to an incorrect webpage and each time the "back" and "start over" buttons were used.

Results indicated that participants using the step-by-step navigational aid were faster in search task completion than those using the map aid, $F(1, 99) = 15.6$, $p < .05$. In the map-based condition, task completion times were significantly related to psychomotor speed ($r = .36$, $p < .05$), verbal ability ($r = -.46$, $p < .05$), crystallized intelligence ($r = -.38$, $p < .05$), attention ($r = -.38$, $p < .05$), and spatial orientation ability ($r = -.32$, $p < .05$). The step-by-step condition yielded a significant correlation only with perceptual speed ($r = -.31$, $p < .05$) and verbal ability ($r = -.37$, $p < .05$). The authors performed a hierarchical regression analysis in order to determine which variables uniquely predicted task completion time. The authors found that for the map condition, psychomotor speed and perceptual speed were significantly predictors in step 1, $R^2 = .142$, $p = .035$. There was a significant change in R^2 when working memory, crystallized intelligence, and attention were added, $R^2 = .306$, $p = .032$. Spatial orientation was statically significant when added in step 3, $R^2 = .383$, $p = .031$, however, adding spatial visualization in step 4 did not account for any additional variance. In the step-by-step condition, speed variables were the only factors predictive of performance, $R^2 = .135$, $p = .031$. Error rate was low and was not used in the analyses to determine performance differences.

Spatial orientation was significantly predictive of performance in a text-based search task, but only when navigational demands were high, such as in the map condition. The study shows that the correlation between spatial abilities and computer-based task performance relies on the spatial ability measure being used and the distinctiveness of the task. Generally, navigational aids that reduce demands for spatial abilities will yield more efficient searches and in the case of a demanding aid, those with higher spatial ability orientation will outperform those with lower spatial orientation ability.

In the write-up of the study, the authors seem to misinterpret the regression scores for working memory and claim that this variable is not significant. However, the regression table, included by the authors, indicates that working memory accounted for a significant amount of variance in performance (i.e., task completion time). Lastly, the step-by-step condition seemed easier to navigate than the map condition due to the unfamiliar and visually complicated layout of the map. This could account for why participants completed the tasks faster in the step-by-step condition compared to the map condition.

Park, O. C. & Lee, J. (1996). Adaptive instructional systems. In D. H. Jonassen (Ed.). *Handbook of research for educational communications and technology* (pp. 651–684). New York: MacMillan Publishers

Adaptive Training Intervention(s): None
Individual Difference Variable(s): Numerous variables
Type of Study: Review

In their book chapter, Park and Lee (1996) provide a review of the literature of different types of individual difference variables that have been related to performance and, thus, could be used for adaptation (e.g., intellectual ability, cognitive styles, learning styles, prior knowledge, etc.). In their brief review of cognitive ability, the authors list the various types of cognitive abilities (e.g., verbal ability, spatial ability, and deductive and logical reasoning) and suggest that these variables have moderating relationships with instruction. For example, research suggests that low cognitive ability students should receive instruction that is structured and less complex, while high cognitive ability students will benefit from less structured and more complex instruction.

Regarding cognitive styles, the authors suggest that field dependent versus field-independent and impulsive versus reflective cognitive styles will be most useful for adaptive instruction. Field independent/field dependent refers to the tendency to approach the environment analytically as opposed to globally. Field independent students tend to be influenced by intrinsic motivators and are able to distinguish figures as discrete from their backgrounds. Field dependent students are characterized as being more influenced by extrinsic motivators and have difficulty distinguishing figures that are embedded in backgrounds. Research suggests that field independent students perform best with deductive instruction and field dependent students perform best with example-based instruction. Reflective individuals tend to take a longer time making decisions while impulsive individuals take less time examining problems and come to decisions more quickly than reflective individuals. Like field dependent/independent styles, impulsive students tend to be influenced by extrinsic motivators while reflective students tend to be influenced by intrinsic motivators. Further, research has shown that intrinsically motivated students outperform extrinsically motivated students.

Learning styles refers to student preferences for the presentation of instructional material. The authors list several types of learning style such as holist/serialist and verbal/spatial/haptic learners. These style preferences can have implications for designing instruction. However, there is limited research to support the value of adapting instruction based on learning styles.

Prior knowledge is one of the greatest predictors of task achievement. Further, the authors state that “because prior achievement measures relate directly to the instructional task, they should provide a more valid and reliable basis for determining adaptations than other aptitude variables” (p. 656). In fact, previous research has shown that the greater prior knowledge in a domain, the less instructional support that is needed.

Park and Lee mention several other individual difference variables such as anxiety, self-efficacy, and locus of control. They suggest that these variables could be used in an Aptitude Treatment Interaction (ATI) approach. However, there has been little research suggesting how they could be used as adaptive variables. They also suggest future research should determine which individual difference characteristics would be optimal for adaptation, how these characteristics will change over time, and how to reliably measure these variables.

Scielzo, S., Cuevas, H. M., & Fiore, S. M. (2005). Investigating individual differences and instructional efficiency in computer-based training environments. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting,* 1251-1255.

Adaptive Training Intervention(s): Learning Check Strategy (LCS)
Individual Difference Variable(s): Verbal ability
Type of study: Empirical

The purpose of this study was to determine if a learning intervention called the learner check strategy (LCS) could increase the instructional efficiency (IE) and effectiveness of a computer-based training intervention. Instructional efficiency is high when the cognitive effort required for a training task is low relative to the performance gains achieved. In the study, students in the LCS condition were required to use concepts and terms learned during training to fill in blank question stems and then answer the completed question. The authors hypothesized that students in the LCS condition would perform better on a knowledge assessment test than students in the control condition (H1) and that using the LCS would lead to higher instructional efficiency (H2). They also postulated that the LCS would have a greater effect on students with low verbal comprehension ability (VCA) (H3).

Thirty-six undergraduate students (17 males, 19 females; $M_{\text{age}} = 20$) were randomly assigned to one of two experimental groups (i.e., LCS and control). All of the participants completed a computer-based tutorial on a simulated command and control military task (i.e., DDD) which consisted of two modules (declarative and strategic). At the end of the modules, participants in the LCS condition received question stems such as "How would you use ____ to ____?" to complete and then answer. The control group did not generate or complete these questions. Knowledge was assessed using multiple-choice questions which were created from both modules of the tutorial. To measure IE, participants filled out questionnaires regarding their perceived cognitive effort during the tutorial and assessments. Students" VCA scores from the Guilford-Zimmerman Aptitude Survey were used in a median-split to create a high and a low VCA group.

Hypotheses 1 was not supported (i.e., LCS training did not lead to greater performance on knowledge assessment). Regarding Hypothesis 2, the authors found a significant main effect of training condition on IE, $F(3, 24) = 3.23$, $p = .04$, indicating that LCS had high IE while the control condition resulted in lower IE. In other words, greater performance was achieved with less cognitive effort for those trainees in the LCS condition than in the control condition. The authors incorrectly claim to have found support for Hypothesis 3. However, after closer inspection of the data it appears the overall F-test for the interaction between LCS and VCA was not significant. The authors then performed post-hoc analysis on the data and made recommendations based on these analyses. This is a violation of statistical conclusion validity. Before one can make recommendations based on the post hoc analysis in ANOVA the overall F-test must *first* be significant. The authors did find that high VCA trainees outperformed low VCA trainees on declarative knowledge, $F(1, 28) = 9.97$, $p < .01$, and overall knowledge assessment, $F(1, 28) = 10.09$, $p < .01$. The authors also found that low VCA trainees reported greater cognitive effort than high VCA trainees when responding to the declarative knowledge questions, $F(1,26) = 8.67$, $p < .01$, and overall knowledge assessment, $F(1,26) = 8.67$, $p < .01$. Lastly, the authors found that instructional efficiency was greater for high VCA trainees than low VCA trainees when responding to the declarative knowledge questions and overall knowledge assessment.

This study may have been limited by the use of ANOVAs to analyze the data. It may have been more beneficial for the authors to use regression instead. For example, using regression would have allowed the use of the VCA measure without having to do a mean split. Creating the mean split causes a loss of systematic variance and could be a reason for the non-significant interactions. It is also possible that students in the LCS condition spent more time training than those in the control condition. In order to increase the generalizability of these findings, future research should replicate this study using task performance on the DDD computer simulation instead of performance on a multiple-choice knowledge test. While this study does not directly address adaptive training, results from this study support the notions that when compared to low VCA individuals, high VCA individuals: (1) demonstrate greater knowledge, (2) report less cognitive effort, and (3) demonstrate greater instructional efficiency when responding to knowledge assessment questions.

Tennyson, R. D. & Rothen, W. (1977). Pretask and on-task adaptive design strategies for selecting number of instances in concept acquisition. *Journal of Educational Psychology, 69*(5), 586-592.

Adaptive Training Intervention(s): Full Bayesian Adaptive Model, Partial Bayesian Adaptive Model
Individual Difference Variable(s): Aptitude and performance scores
Type of Study: Experimental

Previous research has supported the idea that presentation of multiple examples of a concept is beneficial to students during learning. The authors propose that individual differences can be used to determine the number of examples of a concept that should be presented to maximize effectiveness for each student. The aim of this research was to examine a design strategy used to determine how many examples should be used based on students" attitudes and aptitudes before and during a task. The authors hypothesized that a fully adaptive strategy which used aptitude, prior achievement, and learning scores to select the number of examples would lead to better performance than a partially adaptive strategy which used aptitude and prior achievement scores alone. They also predicted that both the full and partial adaptive models would lead to better performance than the non-adaptive strategy.

Sixty-seven undergraduate students were randomly assigned to be in either the full adaptive, partial adaptive or non-adaptive experimental condition. The participants were trained on two legal concepts associated with court decision making and were then required to take on the role of a judge in a civil case. Using computers, the participants had to make rulings on objections made by the defendants or the plaintiff"s lawyer based on the concepts they learned. Four court cases were used for the training period where students would receive feedback after their rulings were given and three cases were used for both the pretest and posttest. In the fully adaptive condition, the number of instances (examples of each concept) presented was based on pretest scores which were then altered based on performance during the task. In the partially adaptive condition only pretest information was used to determine the number of instances of each concept presented. In the non-adaptive condition each student received the same number of instances. The aptitude measure was used as a covariate.

Results indicated that the fully adaptive strategy was more effective than the partially adaptive strategy, $U(1, 1, 64) = .66$, $p < .001$, and that both were more effective than the non-adaptive strategy, $U(1, 1, 65) = .63$, $p < .001$. Analysis on posttest scores revealed that students in the fully adaptive treatment group had scores that were on average four points higher than those in the non-adaptive group. They also indicated that the students in the partially adaptive group had scores that were on average two points higher than the non-adaptive group. Student-Newman-Keuls and least significant difference contrast tests revealed that the fully adaptive, partially adaptive, and non-adaptive strategy means were significantly different at the .05 and .01 level respectively. Further analyses conducted on time in finishing the task revealed a similar pattern of significant results. Participants in the fully adaptive group completed the learning task four minutes faster than those in the partially adaptive group and nine minutes faster the non-adaptive group while participants in the partially adaptive group completed the task five minutes faster than those in the non-adaptive group. Student-Newman-Keuls and least significant difference contrast tests revealed that the fully adaptive, partially adaptive, and non-adaptive strategy means were significantly different at the .01 level. Effect sizes and degrees of freedom for contrast tests were not provided.

These results indicate that, after controlling for trainee aptitude, fully adaptive and partially adaptive training strategies are more effective and efficient (i.e., faster completion times) than non-adaptive training strategies. One issue to note is that students in the fully adaptive condition were provided with their mastery test once mastery of the task was predicted, whereas non-adaptive condition participants had to wait until every instance was presented. This may account for the differences in learning time between the groups. One of the limitations of this study is that a student sample was used, which can lead to diminished generalizability of the results to adult or working populations. Along the same lines, this study showed the effects of adaptive training on immediate performance. Therefore, the generalizability of these results on retention or transfer tasks still remains.

Wilson, G. F. & Russell, C. A. (2007). Performance enhancement in and uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors, 49*(6), 1005-1018.

Adaptive Training Intervention(s): None (adaptive aiding)
Individual Difference Variable(s): Workload
Type of study: Empirical

The goal of this study was to determine if physiological measures could be used to provide adaptive aiding in real time during scenario-based training. More specifically, the authors wanted to determine if aiding thresholds determined by individual operator performance would lead to higher performance improvement than aiding thresholds determined by group level performance. Ten participants performed a simulated uninhabited air vehicle (UAV) task in which they were required to perform bombing missions with four UAVs. Participants performed both easy and difficult scenarios over a 3-4 day period. A within-subjects design was used in which all participants completed 7 experimental conditions: (1) No aiding-Individual, (2) No aiding-Group, (3) Aiding-Individual, (4), Aiding-Group, (5) Random aiding-Individual, (6) Random aiding-Group, and (7) Leave on aiding.

All participants practiced the task for approximately 10 hours. During this training period, participant‟s completed several difficult tasks. During these tasks, a difficulty level for each participant was determined by increasing the speed of UAVs until the operator was able to complete 25-30% of weapon releases. This vehicle speed was then designated as participant‟s individual performance level for difficult scenarios. Additionally, the mean of these speeds was used for the group performance level. The testing portion of the experiment required participants to complete 4 easy and 4 difficult scenarios in each of the 7 aiding conditions. Aiding thresholds were determined using artificial neural networks (ANNs) that detected participant‟s periods of high and low mental workload based on several psychophysiological measures (e.g., 5 channels of EEG, electrocardiograph, and electro-oculograph). When the ANNs detected a state of high workload, aiding was provided. Aiding consisted of decreasing the speed of UAVs and providing vehicle health information on the participant‟s display. In the No aiding conditions, no aiding was provided and either individually or group determined speeds were used. In the aiding groups, aiding was determined by the ANNs and either individual or groups speeds were used. In the random aiding groups, aiding was presented randomly and either individual or groups speeds were used. In the Leave on aiding group, aiding was determined by the ANNs and the aiding remained on until weapons were released (instead of when the ANNs detected low workload).

For the difficult scenarios, results showed that physiologically determined aiding based on individualized criteria significantly improved performance, $F(7,63) = 4.90$, $p < .0001$. Specifically, the Aiding-individual condition led to higher accuracy than both No aiding conditions and the Random aiding-Individual condition. However, there were no statistical differences between the Aiding-individual and Aiding-Group, Random aiding-Group, and the Leave on aiding conditions. While this article does not address adaptive training per se, it does lend support to following: (1) a trainee‟s workload can be assessed with physiological measures in real time to deliver adaptive instruction and (2) adaptation based on individualized criteria vice group-derived criteria may lead to enhanced performance.

DISTRIBUTION

NAVAIRWARCENTRASYSDIV (Competencies 4.6, 4.6A, 4.6T, 4651, CO, XO, TD)

Dr. Ami Bolton
Office of Naval Research
Warfighter Performance Department, Code 342
875 N. Randolph Street
Arlington, VA 22203-1995

Dr. William "Kip" Krebs
Office of Naval Research
Warfighter Performance Department, Code 342
875 N. Randolph Street
Arlington, VA 22203-1995