A CONTROLLED RESOURCE APPROACH TO UNDERSTANDING THE EFFECTS OF FEEDBACK ON LEARNING

A Dissertation
Presented to
The Academic Faculty

By

Anne Collins McLaughlin

In Partial Fulfillment
of the Requirements for the Degree
Ph.D. in the
School of Psychology

Georgia Institute of Technology

August, 2007

A CONTROLLED RESOURCE APPROACH TO UNDERSTANDING THE EFFECTS OF FEEDBACK ON LEARNING

Approved by:

Dr. Wendy A. Rogers, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Ute M. Fischer
School of Literature, Communication, and Culture
Georgia Institute of Technology

Dr. Richard Catrambone
School of Psychology
Georgia Institute of Technology

Dr. Arthur D. Fisk
School of Psychology
Georgia Institute of Technology

Dr. Randall W. Engle
School of Psychology
Georgia Institute of Technology

Date Approved: 11 June, 2007

ACKNOWLEDGEMENTS

I dedicate this work to all who have advised, supported, and put up with me. This includes Tom, for always making everything better, Maribeth for a safe place to rest, and my lab mates for a sanity check. I thank my parents for taking the time to understand what I'm doing, even if it took a number of years to teach them I'm not "that kind of psychologist." Last, I thank Wendy and Dan for showing me what I can handle.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

It is a testament to the complexity of learning that one hundred years of research on feedback has not produced universal prescriptions for training. Results are split in two directions; those recommending more feedback during training and those recommending less. There are numerous theories that explain and predict certain feedback effects, but none explain the mixed findings in the literature. This has resulted in: a) no singular theory and b) little understanding of other factors that might affect the mechanism of feedback.

Previous research has shown that cognitive load, or the degree to which a task demands attentional resources, is an important component in training. According to cognitive load theory, decreasing cognitive load during the acquisition phase will lead to increased learning after a retention interval (Sweller, 1988). In other words, enhanced performance in the acquisition phase facilitates retention.

An alternative theory is that learners have a "challenge point," where at least some cognitive load facilitates learning. In this case, very low task loads during the acquisition phase result in *reduced* learning (Bjork, 1994; Elshout, 2006; Guadagnoli, 1999). Examples of challenge points come from studies that resulted in the paradox of retention and transfer, where high performance levels in acquisition dropped sharply on retention and transfer tests. Lower acquisition performance, however, remained steady through retention and transfer (see Schmidt & Bjork, 1992, for a review). It is thought that the increased task load during the acquisition phase was responsible both for the lower performance during training and the higher performance at retention. This

relationship between support during acquisition and performance at retention/transfer is not well understood, but there are some general patterns in the literature.

One methodological pattern is that studies showing learning benefits for reducing task load (via supportive feedback) used populations with low working memory capacity (i.e., remedial students, older adults) and/or very complex tasks (i.e., requiring numerous stages of processing). Similarly, studies showing benefits for increasing task load (by offering little support via feedback) used either high capacity populations (e.g., college students) and/or simple tasks (e.g., a single degree of freedom motor task). Previous studies included these various populations and tasks, but none systematically manipulated how these variables might affect a learner’s use of feedback (Table 2).

Because of this apparent link between feedback and learner resources, the following series of studies systematically manipulated the cognitive load of the experimental task and measured the working memory capacities of the learners. The overall question was whether forcing the learner to self-evaluate would result in more or less learning of a rule-based cognitive task and how this effect might be moderated by the working memory capacity of the learner and the load of the task to be learned.

The results generally showed that reducing task load resulted in more learning. It was expected that high working memory capacity learners might learn more when difficulties were introduced for a simple task (via less supportive feedback). Instead, all groups not only learned more when receiving more support, the high working memory capacity learners appeared more able to utilize the additional feedback. Instead of providing their own support when feedback was minimal or lacking, high working memory capacity participants seemed best able to make use of the information provided in supportive feedback. Low working memory capacity participants seemed unable to either provide their own support when feedback was minimal or lacking, but also were not as able to make use of the information provided in more supportive feedback.

The results of these studies do not suggest that there are no cases where less feedback support is better, but for a cognitive, rule-based logic task, providing more support for performance in acquisition resulted in more learning across working memory capacity groups and cognitive task loads.

The contribution of the current series of studies is an explanation of *why* and *how* appropriate level of feedback support can change based on the working memory capacity of the learner and demands of the task. Feedback can either impose a load upon the learner to self-evaluate or provide support for acquisition performance. Though learners may benefit from feedback neither too high *nor* too low, the current results indicate that additional feedback is most useful to those with the attentional resources available to utilize it.

Feedback effects are a complex phenomenon; there are not only questions of type of feedback necessary to improve performance for a certain group, but there are questions of *why* that feedback improved performance. The mechanism of the beneficial effects of feedback seemed to differ by the usefulness of feedback to a particular group.

CHAPTER 1: THEORETICAL FEEDBACK MECHANISMS

Feedback research has one of the longest histories in psychological literature. Feedback was originally viewed in a stimulus-response paradigm (Thorndike, 1911), but is now typically defined and studied as augmented, external (to the organism) *information* meant to promote learning (Kluger & Denisi, 1996; Salmoni, Schmidt, & Walter, 1984). Some of the various roles for feedback include confirmation of correct actions and correction of incorrect actions (Kulhavy, 1977) and energizing and motivating other processes that promote learning, such as practice and attention (Salmoni et al., 1984). However, there is still the unanswered question of how feedback may best aid learning.

At present, there are two schools of thought in the feedback literature. They may be categorized as supporting “more” or “less” feedback support of performance in acquisition. More and less are in quotation marks due to the many ways they have been defined. For example, more might mean “more content in each feedback presentation,” or “feedback given more frequently,” or even “the type of content in the feedback was more prescriptive of what to do.” The ideas of “more” or “less” feedback have only previously been considered for individual feedback variables, such as content or frequency, and the definition was often literal. For example, more content contained in feedback simply indicated the units of information. Even studies observing human tutors found some tutors offer “more” or “less” feedback to their students (*cf.*, (Fox, 1991; Merrill, Reiser, Ranney, & Trafton, 1992). However, considering content only in terms of units of information did not capture the larger picture of “more” and “less” feedback support.

Consider the varieties of feedback content. Prescriptive feedback content specifies the correct action to take in a task whereas conceptual feedback content may remind the learner of the greater context of an error. Prescriptive and conceptual feedback may contain the same *amount* of information content, but because prescriptive feedback is more directive, it would be classified as “more” supportive of performance than conceptual information. (See Table 1 for a list of how studies defined conditions of more or less feedback.)

Table 1. Conceptualization of Amount of feedback Support in Studies of Content, Frequency, and Timing.

"More" feedback Support	"Less" feedback Support
Directive/concrete information	Conceptual/abstract information
More pieces of information [1]	Fewer pieces of information[1]
Higher absolute frequency[2]	Lower absolute frequency[2]
Higher relative frequency[2]	Lower relative frequency[2]
Feedback every trial[2,3]	Summary feedback after a number of trials[2,3]
Performance feedback[1]	Knowledge of results (KR) [1]
Knowledge of correct response [1]	Answer until correct[1]
Local information[1]	Goal information[1]
Immediate feedback[3]	Delayed feedback[3]
Part-task feedback[1,3]	Whole-task feedback[1,3]

Note: Labels for more or less feedback were derived from experiments comparing differing levels of feedback. For a more complete review of these comparisons, see (McLaughlin, Rogers, & Fisk, 2005) [1]content variable, [2]frequency variable, and [3]timing variable.

After a review of the feedback literature, it became apparent there were similarities between findings for the feedback variables of content, frequency, and timing. There appears to be an underlying principle to these feedback variables: how to assign feedback depends on the information processing requirements of the feedback and how the feedback changes the information processing requirements of the task. It is not enough to state that for a motor task feedback should occur once every 10 trials or once every 15 trials, or whatever specific prescriptions a study reveals. It is the concept of *feedback processing requirements during the task* that will ultimately determine the amount of learning in the task.

To be used, feedback must be processed and understood. Conversely, categorizing feedback into its processing demands also helps to explain the phenomena that learning can decline as feedback increases. In other words, *amount* and *frequency* of feedback are merely subcomponent descriptors: the most predictive descriptor is the amount of resources consumed by the feedback and how feedback changes task demands. This is linked to the amount and frequency of the feedback.

Feedback Support

For many years it was accepted that the more feedback provided during acquisition of a skill, the more learning would occur. This time-honored belief was challenged after numerous results inconsistent with this view. Feedback researchers examined feedback parameters individually (such as content, timing, and frequency) and manipulated the amounts of each. Researchers occasionally found that more feedback resulted in poorer performance of a task, compared to those who received less feedback on their performance during acquisition (e.g, Ho & Shea, 1978; Taylor & Noble, 1962). These findings were prevalent enough that they could not be ignored as anomalous results. Thus, in 1984, Salmoni, Schmidt, and Walter published a review article describing feedback's impact on learning: the guidance hypothesis.

According to the guidance hypothesis, feedback guides and motivates learning. Feedback can act as an energizer to prolong or increase practice and also serve as a guide to correct performance. Thus, the variables that actually increase learning are practice and attention; feedback only energizes and enables these processes rather than being an integral part of learning itself.

Thinking of feedback as guiding information helped to organize prior mixed results (Blackwell & Newell, 1996; Salmoni et al., 1984). Many studies of feedback measured performance rather than learning. Learning is commonly defined as a permanent change in long-term memory and must be demonstrated through a retention or transfer test (Brosvic, Dihoff, Epstein, & Cook, 2006; Carlson, Sullivan, & Schneider, 1989; Salmoni et al., 1984; Schmidt & Bjork, 1992). If one measures performance at the end of acquisition with feedback still present, those are performance data, not learning data. It is unknown whether any permanent change to long-term memory has occurred. Once the historical literature was divided into performance data and learning data, the following pattern emerged: more feedback resulted in better *performance* whereas less feedback resulted in more retention and transfer (*learning*) (Salmoni, et al.)

Once differences in feedback effects were ascribed to learning versus performance, researchers included measures of retention or transfer in their tests (e.g., Schmidt & Bjork, 1992; Schmidt & Wulf, 1997; Winstein & Schmidt, 1990). Indeed, these studies tended to find that less feedback during acquisition resulted in higher retention after a retention interval. Table 2 provides a sample of studies claiming that reducing feedback resulted in greater learning.

In motor tasks, “less” feedback was generally represented via less *frequent* feedback. In these tasks, a schedule of infrequent feedback (i.e., feedback every 15 trials vs. Feedback after every trial) resulted in better learning of a motor skill. Feedback frequency manipulations compared low frequency of feedback to feedback on every (or almost every) trial (e.g., Schmidt & Bjork, 1992; Schooler & Anderson, 1990; Schroth, 1992). Feedback on every trial generally improved performance in acquisition, but less frequent feedback benefited retention performance. In studies of feedback frequency, trial feedback corresponded to more frequent information while end-of-block summary feedback corresponded to less frequent feedback (Schmidt & Bjork, 1992; Schmidt & Wulf, 1997; Schooler & Anderson, 1990; Schroth, 1997). Previous research has shown that summary feedback depresses performance in acquisition, but improves demonstration of learning via a retention test, (Schmidt, Young, Swinnen, & Shapiro, 1989).

Table 2. Sample of Research Investigating Effects of Feedback on Learning

Studies finding more feedback better for learning	**Studies finding less feedback better for learning**
Adams, 1971	Birenbaum & Tatsuoka, 1987
Blackwell & Newell, 1996	Cope & Simmons, 1994
Bohlmann & Fenson, 2005	Goodman, Wood, & Hendrickx, 2004
Dihoff, Brosvic, Epstein, & Cook, 2005	Goodman & Wood, 2004
Droit-Volet & Izaute, 2005	Ho & Shea, 1978
Clariana, 1990	Magill & Hall, 1990
Farquhar & Regian, 1994	Nicholson & Schmidt, 1991
Maddox, Ashby, & Bohil, 2003	Schmidt, Lang, & Young, 1990*
McKendree, 1990	Schmidt & Wulf, 1997
Roper, 1977	Schmidt et al., 1989
Schmidt, Lang, & Young, 1990*	Schooler & Anderson, 1990
Thorndike, 1931	Schroth, 1997a, 1997b
Whyte, Karolick, Nielsen, & Elder, 1995	Sherwood, 1988
Wishart, Lee, Cunningham & Murdoch, 2002*	Winstein & Schmidt, 1990
Wulf, Shea, & Matschiner, 1998	Wishart, Lee, Cunningham & Murdoch, 2002*
Vollmeyer & Rheinberg, 2005	Wulf, Schmidt, & Deubel, 1993

Note: * Indicates presence in both columns due to multiple findings.

Feedback timing manipulations revealed a depressing effect on performance in acquisition with delayed feedback, but an improvement in retention performance compared to immediate feedback (e.g., Schooler & Anderson, 1990; e.g., Schroth, 1992). This effect was strengthened by instructing learners to evaluate the correctness of their responses during the interval before feedback appeared (Kulhavy & Stock, 1989; Swindell & Walls, 1993). This effect disappeared when learners performed unrelated tasks in the interval before feedback was presented (Brackbill, Bravos, & Starr, 1962; Kulik & Kulik, 1988; Pashler, Zarow, & Triplett, 2003).

Feedback content was usually studied by comparing different amounts of information, levels of specificity, or comparing different types of information (e.g., Cope & Simmons, 1994; Davis, Carson, Ammeter, & Treadway, 2005; Goodman & Wood, 2004; Smith & Ragan, 1993). At times, more feedback content depressed learning. In cognitive tasks, such as the addition of signed numbers, learning a computer language, or problem solving, "less" generally corresponded to less information contained in the feedback content or the feedback concerned a specific procedure in the task rather than an over-arching concept (Table 2). Even in observational studies of human tutors, successful tutors attempted to provide as little information as possible to the learner (Merrill, Reiser, Merrill, & Landes, 1995).

Currently, the idea that less feedback promotes learning is well accepted (Schmidt & Bjork, 1992; Schneider, Healy, & Bourne, 2002), but mixed findings continue to occur (Table 2). Most studies find some amount of feedback to be beneficial for learning, whether it be a small amount (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995) or a large amount (e.g., Schroth, 1997). Though many agree feedback is generally necessary to guide performance, it is unknown how to identify too much or too little feedback during acquisition. After all, recent studies found *more* feedback in acquisition was not only beneficial, but actually necessary for learning (e.g., Wishart, Lee, Cunningham, & Murdoch, 2002).

There are a number of theories that explain the phenomena of "less is more" when it comes to feedback. The first of these is similar to transfer appropriate processing, as illustrated by the paradox of retention and transfer.

The Paradox of Retention and Transfer

More learning has been demonstrated after a retention interval for instruction methods that showed depressed performance in acquisition compared to instruction methods with high initial performance. This was termed a paradox because there was no ready explanation as to why performing a task more poorly while learning it should result in better retention and transfer than performing it well. This is also commonly called the effect of contextual interference (Shea & Morgan, 1979). One theory that explains this finding is that training conditions which most closely approximate what will be cognitively expected of the learner in retention or transfer result in the most learning (Schmidt & Bjork, 1992), similar to transfer-appropriate processing (Bransford & Schwartz, 1999). One example would be that if recall will be expected in retention, acquisition that forces recall will result in more learning than a form of training that does not. In this view, feedback which promotes the most similar cognitive processes in acquisition to what will be expected in retention or transfer will result in the most learning.

Lower Limits on Learning

Creating Challenges

The zone of tolerable problemacity is defined as an internal threshold dependent on individual abilities (including transitory effects and motivation at the time.) (Elshout, 2006; Snow, 1989) This threshold is where the person learns most optimally (Figure 1). Similarly, a recent paper suggested each learner had a "challenge point," or an optimal point when feedback should be provided for a certain level of task difficulty (Guadagnoli & Lee, 2004). The zone of tolerable problemacity is the flip-side of cognitive load theory (Sweller, 1988) in that people do not learn tasks as well that are far below their threshold.

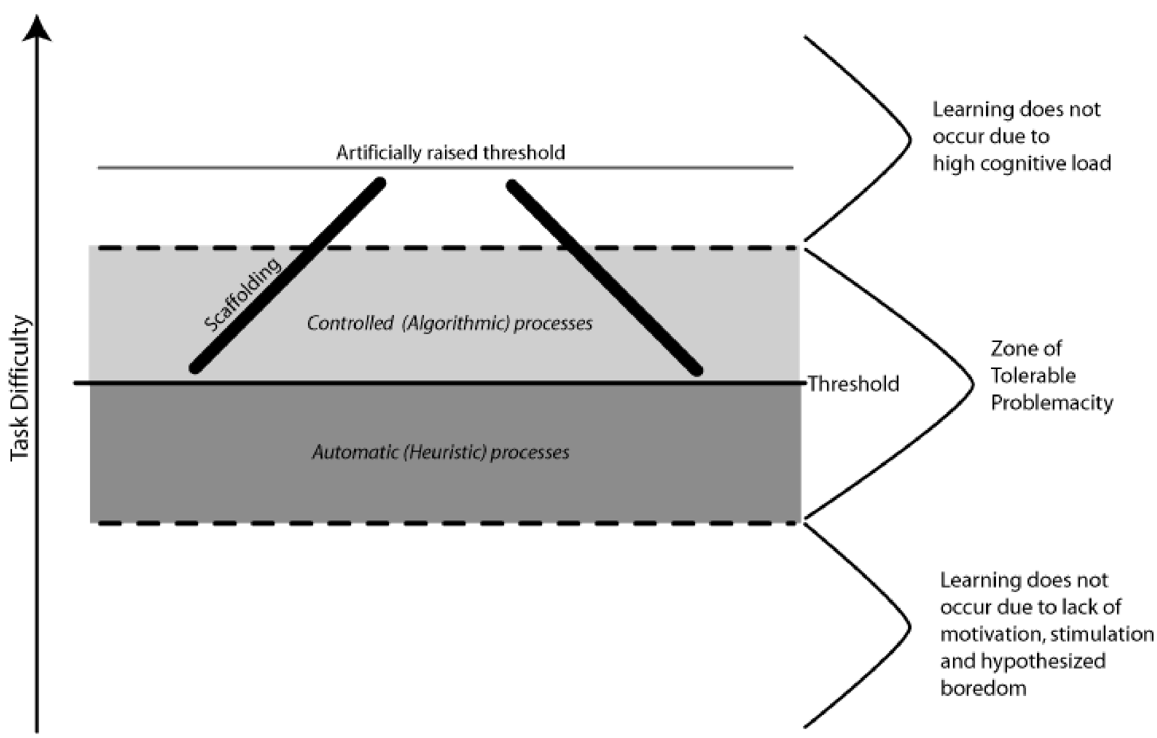


Figure 1. Conceptualization of the zone of tolerable problemacity, cognitive load theory, and the action of scaffolding on learning.

Though it may seem that having all tasks operate far below threshold would unilaterally improve performance, it may also harm it both in acquisition and retention. Ostensibly, poor performance under minimal demand occurs because tasks far below threshold are “boring” and motivation is reduced enough to decrement performance (Snow, 1989) but this assumption is fairly untested and there is room for other explanations. For example, the zone of tolerable problemacity is linked to control theory (Cabanac & Russek, 2000), which has become more popularly used in recent theories of feedback (Kluger & Denisi, 1996) and learning (Szalma, Hancock, Dember, & Warm, 2006).

Control Theory

Control theory, from physics and computer science, states that the organism or system of interest has a point of activity at which it prefers to exist (Szalma et al., 2006). This point may be different for different times and tasks, but the system will try to return to this point of optimal stimulation. The threshold of problemacity is such a point. Far below the threshold, the system seeks stimulation to return to threshold and this stimulation may be in the form of distracters or other activities (mental or physical.) Far above the point, the system tries to reduce activity to the optimal point, either by eliminating goals or being unable to perform the task (Kluger & Denisi, 1996).

However, knowing the human cognitive system tends to return to a steady-state of stimulation would be a descriptive fact (if it is a fact). A description is only the beginning of understanding why or how such a point would exist. It also does not specify under what conditions that point could move, or how those conditions might be maneuvered to produce optimal learning.

However, these theories did not completely explain why some studies still found benefits for increased feedback. After all there were several theories which predicted just the opposite.

ACT-R

Anderson's ACT-R cognitive architecture espoused the "model tracing approach," where the most learning should occur when a learner is held as closely as possible to an expert model of performance, as this would strengthen the correct neural pathways for performance of a skill (Anderson, 1993). The ACT-R cognitive architecture design principles specify "minimize working memory load" (Anderson et al., 1995, p. 180) to uphold the highest performance possible during training. In other words, feedback which promotes the most accurate performance in acquisition should also result in the most accurate performance on retention and transfer tests. Increased performance in acquisition is associated with more feedback.

Cognitive Load Theory

Cognitive load theory (Sweller, 1988) suggests that learning is always enabled through a reduction in the need to integrate elements of the task during training. Cognitive load experiments provided numerous examples, including the benefit of worked-examples over performing a task oneself and eliminating redundant problems where the reduction in extraneous (due to the task) cognitive load, results in a reduction of acquisition time as well as an increase in performance on tests of retention and transfer (Sweller & Chandler, 1991). Feedback that reduces the most extraneous cognitive load should result in increased performance during acquisition as well as higher performance on retention and transfer tests, compared to feedback that does not reduce (or increases) cognitive load.

Scaffolding

Scaffolding means to make a task easier during acquisition. This may be done by removing parts of the task during learning, dividing the task into sub-components, or segmenting the task into smaller temporal chunks. In relation to the threshold of problemacity, scaffolding may be thought of as temporarily raising an individual’s threshold. When a person’s threshold is artificially raised, a task that would have been too demanding can be learned. The scaffold is removed as learning and performance raise the individual’s actual threshold closer to the requirements of the task (and therefore within the “zone”).

Characterizing the Problem of Feedback Support

The first step in understanding when it is beneficial to provide more feedback is to identify what makes the studies that found more feedback better different from those that found less feedback better. This was a difficult problem in the feedback literature as few researchers specifically considered variables other than feedback (Kluger & Denisi, 1996). However, it was possible to identify the type of task as described in the method section of each study and classify it according to the type of task (motor, psychomotor, cognitive, etc.) processing required (controlled, automatic) and the amount of resources needed for this processes. This was done at a general level, but revealed some important differences.

When looking at the type of task, there were examples of the benefits of increased feedback from a variety of learning domains. In the extensively studied domain of motor learning, Wulf, Shea, and Matschiner (1998), found that more trials with feedback present resulted in better learning of a complex slalom-skiing maneuver than feedback present on fewer trials. Similarly, Blackwell and Newell (1996) found that more frequent feedback aided learning of a single-limb movement timing task and Wishart, Lee, Cunningham, and Murdoch (2002) found that older adults were unable to learn a bimanual coordination task with feedback every five trials, but *were* able to perform and retain the task when given feedback on every trial. In a simple striking task, retention improved when feedback was provided after 15 trials. However, in a complex striking task, more frequent feedback (after every trial) improved retention performance (Guadagnoli & Dornier, 1996).

In the domain of cognition, McKendree (1990) tracked geometry proof learning under three conditions: no feedback, feedback on whether a step was correct (known as knowledge of response feedback [KR]), or feedback reminding the learner of the eventual goal. In essence, the last condition contained the "most" feedback, as the learner knew *receiving the feedback* meant an error just occurred, and also was given extra information about how to progress through the rest of the proof by being reminded of the eventual goal. Thus, goal feedback also contained the other feedback condition of KR. The participants in the goal feedback condition showed the most retention of geometry proof learning (McKendree, 1990).

In the verbal learning domain, increasing the precision of feedback (therefore increasing the information) improved performance (Roper, 1977). Learning of a complex procedural task (LOADER problems) improved with more informative feedback (Farquhar & Regian, 1994). Participants better learned a rule-based and information integration task with more feedback during acquisition (Maddox, Ashby, & Bohil, 2003) and remedial students who received more information in their feedback learned more than those told to keep answering until they got the correct answer (Clariana, 1990). Though these recent studies finding benefits for increased feedback during acquisition were not the majority, they were certainly present and must be accounted for.

Patterns of Variables

Task Load

Thus, when of the methods of those studies supporting “less” feedback to those supporting “more” feedback were compared, there were differences in their experimental tasks. Specifically, the tasks placed different loads on the learner. Most of the experimental tasks from the “less is better” camp were either single degree-of-freedom motor learning tasks or cognitive tasks involving few steps and little mental effort. The studies that found *more* feedback necessary for learning typically used high-load tasks, such as movement coordination and balancing, or required information integration. Thus, the first step in differentiating studies advocating more or less feedback for learning was not according to learning domain (such as verbal tasks, motor tasks, etc.), but by the cognitive demands of the experimental task.

Learner Characteristics

However, not all of the studies listed in Table 2 followed this pattern. In some cases, more feedback was beneficial for what appeared to be a cognitively undemanding task (e.g., Droit-Volet & Izaute, 2005; Vollmeyer & Rheinberg, 2005). In these cases where the task seemed fairly simple (low cognitive or motor load), but the data showed that more feedback improved learning; it was the participant population that differed. The participant populations for studies with low-load tasks tended to be cognitively reduced in some aspect necessary for the task, such as having lower WMC than the average learner. Thus, these populations had fewer cognitive resources to apply toward learning. Example populations included remedial students (Clariana, 1990), older adults (Wishart et al., 2002), special needs students (Brosvic et al., 2006), and children (Wulf & Shea, 2002 for a review).

Thus, more feedback was beneficial to learning either when the task had high cognitive demand or the learner was cognitively limited in some way. The benefit of more or less feedback seemed to be cognitive resource-based. Those who *lacked* resources needed for the task benefited from more feedback as did those confronted with a taxing task: both groups, whether cognitively limited or not, needed more feedback when tasks were above a certain demand level. These same patterns were seen across the feedback parameters of frequency, timing, and content. Thus, it appears that available cognitive resources played an important role in determining appropriate feedback for training.

Interim Summary

In brief, numerous parameters for feedback have been studied. The results of these variables fall under the concept of their *amount of support*. Feedback parameters that give more support for acquisition performance seem more appropriate for a highly demanding task or for learners with fewer cognitive resources. Parameter settings that give little support for acquisition performance seem more appropriate for a task with low demand or for learners with large available cognitive resources.

Available Resources

Working Memory Capacity

> "The most important characteristic of complex learning is that students must learn to deal with materials incorporating an enormous number of interacting elements. In conceptual domains, there are many interacting knowledge structures that must be processed simultaneously in working memory in order to be understood. In skill domains, there are many interacting constituent skills that must be coordinated in working memory for a coherent performance."

-van Merriёnboer and Sweller, 2005, p. 156

The idea that working memory is important for learning was well stated by van Merriёnboer & Sweller (2005). Although there are other individual characteristics that may relate to learning (e.g., motivation, experience level learning (See Snow & Swanson, 1992 for a review), working memory capacity is a highly predictive characteristic. Working memory capacity (WMC) can be defined as the amount of attentional processes available to an individual as well as the ability to focus and allocate these processes (Feldman-Barrett, Tugade, & Engle, 2004). This capacity and has been linked to success in many learning domains and situations (Engle & Kane, 2004; Woltz, 1985), and to learning success in the stages of skill acquisition that require controlled processes (Ackerman, 1988).

WMC also tends to remain steady within an individual and during learning of a task (Baddeley, 2000). Though there has been research demonstrating changes in WMC (Olesen, Westerberg, & Klingberg, 2003), these increases took place over the course of five weeks of specifically training WMC. For the most part, measuring WMC can indicate how quickly an individual will learn many types of tasks (Turner & Engle, 1989) and is a good indicator of the amount of controlled processes available to a learner.

Other reasons why WMC could indicate the appropriate feedback for a task links back to the previous chapter discussing differences in feedback for tasks with low or high cognitive demands. The possible interaction of feedback and task loads can only take place *within* a learner. When we speak of the load on a learner, we must acknowledge that individual capacity dictates acceptable load. Task loads that are of high cognitive load for one learner might be low for another, due to extensive practice, transferable knowledge, or working memory capacity.

Being able to choose feedback according to one cognitive predictor would be of great benefit to our understanding of the mechanisms involved in learning. WMC is a powerful predictor of performance on many types of tasks. Knowing an individual's WMC may help predict how that person will learn a new task, but will not necessarily indicate how to help the individual learn more quickly. For example, it is probably safe to assume for any task requiring cognitive resources, the individual with higher WMC will learn more quickly than the low WMC individual. However, how can we improve the learning of both individuals? We need to understand not only how well a person might learn a task on his or her own, but how to improve that learning through appropriate feedback.

Interim Summary

Working memory capacity is a valuable predictor of learning ability and relatively easy to measure. Because manipulating feedback changes the cognitive load on the learner, it seems likely WMC would have an effect on the appropriateness of a certain feedback to a learner (illustrated in Figure 2).

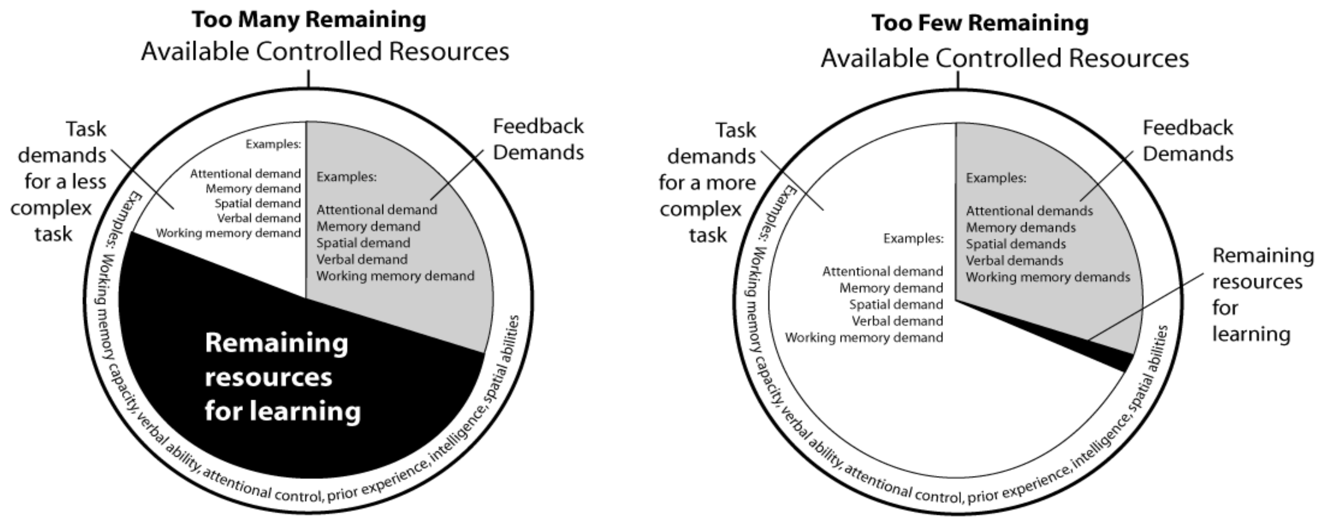


Figure 2. Model interaction of feedback variables, learner characteristics, and task loads within the learner characteristic of controlled processes (cognitive resources).

Summary: Interaction of Learner Resources, Task Load, and Feedback

This introduction specified how previous studies investigated feedback efficacy. Often, these efforts added to our knowledge of feedback outcomes, but did not completely explain feedback mechanisms. Schmidt and Bjork (1992) brought several areas of research together to understand how training that depressed acquisition performance could actually *improve* retention performance. They concluded that it was important in training to have learners perform the same cognitive processes in acquisition they needed for a retention test. For example, if learners need to retrieve information or steps from long-term memory, a feedback condition that promotes retrieval enhances learning. This includes reduced feedback support, thus encouraging (or forcing) the learner to produce internal feedback.

However, other than saying “some amount of feedback is necessary,” training concepts have not accounted for what happens when a learner could not perform the required processes (such as retrieval). Learners in this situation are unable to take advantage of the training and can experience large amounts of practice with little improvement. This links back to research concerning cognitive load (Sweller, 1988).

A possible way to reduce task load during learning is the amount of feedback support. More supportive feedback could be seen as similar to simplification and scaffolding, and likely produce the same effects. Last, it may be possible to have too little load during initial learning of a task. Under-loaded learners also do not learn well (e.g., Hancock & Warm, 1989), and too much support during acquisition results in less retention (e.g., Schooler & Anderson, 1990). If there can be too much or too little feedback support in a task, choosing and designing feedback becomes balancing act between learner and task load.

In summary, we have the problem of giving feedback that promotes the same cognitive activities that will be needed when demonstrating retention, except when the feedback combined with the task overloads the learner. We know that learners may differ in their working memory capacities. Thus, if the task combined with feedback overloads the learner the task load can be reduced by increased feedback support.

If the task and feedback under-load the learner, many researchers blame lack of motivation for poor learning in this case (e.g., Kanfer & Ackerman, 1989). Others ascribe motivation to the larger picture of a "steady state" biological system that requires a certain amount of stimulation (e.g., Cabanac & Russek, 2000; e.g., Hancock & Ganey, 2003). Conceivably, feedback could be designed to *increase* the load on the learner to maintain a steady-state and/or increase motivation. The desired end result would be a balance of steady-state maintenance for different capacity learners.

CHAPTER 2: OVERVIEW OF THE EXPERIMENTS

A series of four experiments examined how learner working memory capacity affected the amount of feedback support needed to learn tasks with differing demands. Particular contributions of these experiments were manipulation of task load and controlling the learner characteristic of working memory capacity. Traditionally learner ability has not been measured, particularly in studies with manipulated task loads. These interactions of learner, task, and feedback variables should result in the ability to prescribe feedback based on knowledge of WMC and task loads. The results of this study produce principles to indicate initial prescriptions for feedback; actual feedback may be refined by testing for a specific training system.

The purpose of these experiments was to understand and predict the effects of feedback support level for learners with differing WMC for learning of tasks with differing levels of controlled-process demand. Measures were taken at acquisition, after multiple retention intervals, and on transfer tests. Retention intervals were selected to be long enough to dissipate temporary effects of acquisition, such as motivation, fatigue, or boredom (see Salmoni et al., 1984; Schmidt & Bjork, 1992). The primary measures were in retention and transfer, since performance in the acquisition phase is more likely to be affected by transitory experimental manipulations (Schmidt & Bjork). The extent of controlled process demand in the experimental task was manipulated through the use of a logical reasoning task.

Task

Logical Reasoning Tasks

In a logical reasoning task, participants follow pre-determined rules to make a decision about an outcome. An example of a logical reasoning task is the Logic Gate Task (LGT), used in numerous studies of cognitive skill acquisition (e.g., Carlson & Yaure, 1990; Schneider, 1985).

There are many reasons to study the effects of feedback in a logic task. First, such tasks relate to troubleshooting proficiency, a highly demanded skill both in the workplace and at home (Kyllonen & Woltz, 1989). Troubleshooting extends from setting a digital clock to wiring a house for electricity. Second, logic tasks may be heavily weighted to have a high cognitive and low motor component. Though there is reason to believe motor and cognitive skill acquisition operate under similar principles (Burton, Moore, & Magliaro, 2004; Goldstone, 1998; Newell, 1991; Rosenbaum, Carlson, & Gilmore, 2000), there have been fewer studies of feedback in cognitive skill acquisition. Using a task with a high cognitive component would help to generalize the findings from the motor learning literature. Similar results to studies of motor tasks support the theory that common mechanisms underlie both types of learning (Goldstone;.Newell; Rosenbaum et al.). Else, the results would help determine parameters of difference between motor skill and cognitive skill learning. Further, prior studies of feedback in motor skill acquisition have not manipulated task load nor studied learner WMC. Thus, this research not only extends the motor learning literature to the cognitive arena, but may identify new interactions that apply to motor and cognitive skill acquisition.

Third, participants adopt a “common, serial strategy” for performing the LGT (Carlson et al., 1989) that allowed us to assume all participants likely perform similar cognitive operations. This can be difficult to assume in other high-load tasks, as people may use heuristics (Rothrock & Kirlik, 2003) or adopt variable strategies. It is easier to make a case for feedback effects on a certain cognitive process when one may be fairly certain all participants used similar cognitive processes.

Last, rule-based logic tasks can impose low or high working memory load depending on the number of interacting elements provided to a participant. A high-load rule-based logic task may be compared against itself in low-load form. This informs the literature of the effects of load using the same task and should help to provide a general principle for feedback design. These were the reasons for choosing a rule-based logic task for this series of experiments.

Logic Gate Task

The LGT is a logical decision making task in which gates operate on binary inputs (see Carlson et al., 1989, for more detail). One example is the "AND" gate, where if both of two binary inputs are 1s, the output of the gate will be 1 (see Figure 3).

A.

B.

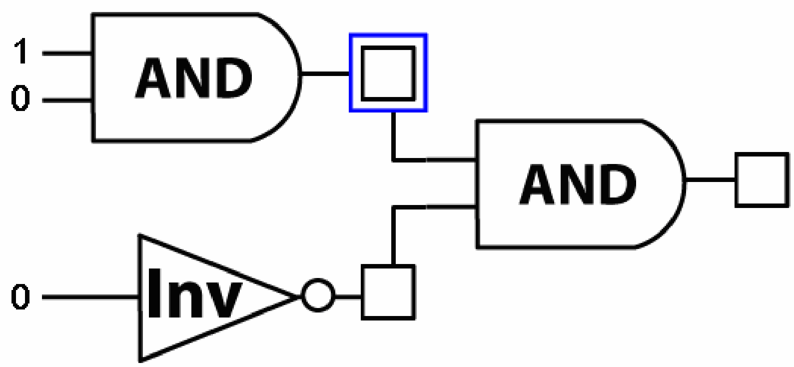


Figure 3. Example A demonstrates simple operations performed through logic gates, requiring only memory for gate operation to produce answer. Example B contains high-load example of gate task, where outputs from leftmost gates must be used as inputs in rightmost gate.

Logic gates change task load according to the number of gates in a task and the operations they perform. For example, Figure 3a demonstrates a simple version of the LGT. Binary inputs are compared through the gate a single time. Figure 3b demonstrates a complex logic-gate task. One can see the high element interactivity of the differing sets of binary inputs. The results of the leftmost gates must be held in working memory and compared through further gates until reaching the single output on the right.

Previous studies using the LGT measured accuracy only on the final gate output. Participants held the intermediate answers in working memory to solve the final gate, but provided no record of these answers. The downside of such a task was that there was no way to know why a final answer was incorrect. Potential reasons were as follows: 1) Any gate in the problem could have been answered incorrectly by the participant. This may have lead to cascading errors with subsequent gates. 2) A participant may have remembered a gate output incorrectly and used the wrong value as the input for subsequent gates. Thus, failure may have been due to incorrect knowledge of gate operation *or* failures of working memory. The current study attempted to solve this methodological problem programmatically as follows.

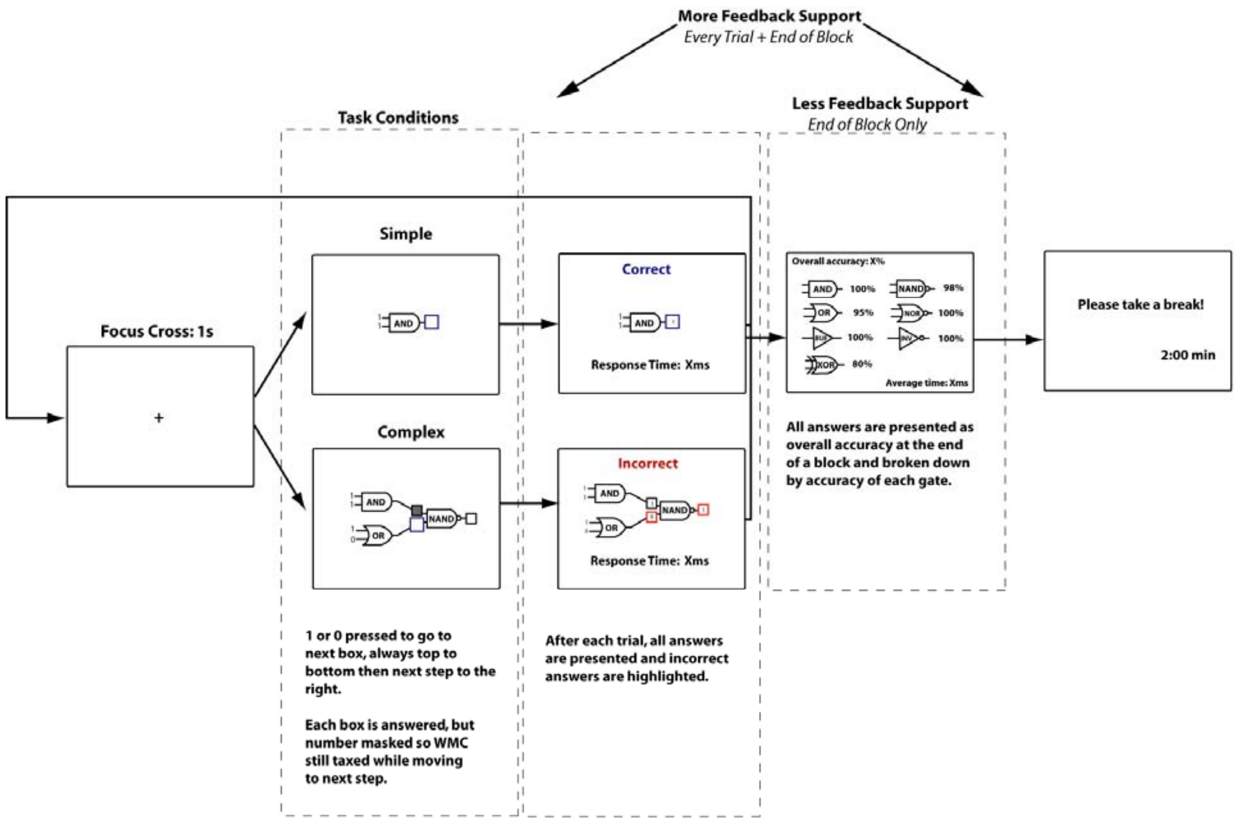


Figure 4. Feedback manipulation, providing more or less support during acquisition.

In the complex LGT, each gate must be answered to provide inputs for the final output gate. Constraining the program in this way captured where the participants made errors. Their answer for each gate was recorded, but not presented on the screen. Thus, participants still held the gate outputs in working memory to answer the next gate step. The only difference between this method and previous methods was that each gate required a physical response. Figure 4 demonstrates the order of operations in the task. As can be seen in the complex task condition, an answer was entered for the first gate (then masked) and the prompt has moved to the next gate. This continued until all gates were completed, then one of the two feedback conditions appeared.

Two pilot studies were carried out in addition to the four main experiments to determine trial time limits and the feasibility of feedback condition as a within-participant variable (Appendix D). Task loads were manipulated between participants due to the possibility of the simple version of the LGT "training" the participant for the complex version or vice versa. Feedback support was manipulated within participants where some gates received high feedback support and others received minimal support.

Feedback conditions. In the more supportive feedback condition in Experiments 1, 2, and 3, individual gate errors in the logic gate problem were revealed to the participant on every trial (Figure 4). In the less supportive condition, feedback came only at the end of a block, where the average accuracy for each gate was displayed to give the same information as the more supportive condition but in summary form. In the fourth experiment, increased information and immediacy was added to the more supportive feedback condition.

Thus, no matter what the support condition, the type of information given was the same: the eventual goal state and which responses did not correspond to that goal state. The differences were in the amount of learner resources (controlled processes) needed to understand the feedback, use the feedback, and perform the task. Other differences included the amount of information and the frequency with which that information was presented.

Task load conditions. Task load was operationally defined as the number of mental computations required to solve a single trial of the Logic Gate Task. The low-load condition required solving one logic-gate problem whereas the complex condition required solving three logic-gates per trial wherein the answer to the third gate depended on the output of the first two gates. These levels of task load are referred to as “simple” and “complex” throughout this document.

***Working memory capacity groupings*.** Working memory capacity was determined via the automated operation span (Ao-span) test (Turner & Engle, 1989; Unsworth, Heitz, Schrock, & Engle, 2005). Low and high WMC groups were defined as the first and fourth quartiles of performance from the Aospan scores from Unsworth, Heitz, Schrock and Engle (2005).

Measures

***Criterion test performance*.** The criterion test of logic gate definitions required participants to match each gate with the appropriate rule (Kyllonen & Woltz, 1989). This declarative test was given prior to any procedural experience of performing the LGT. All participants reached criterion in 4 trials or fewer. The high WMC group took significantly fewer trials to reach criterion on the test of matching logic gates and their definitions across experiments; the means and significance tests may be found in the tables for each experiment.

Independent variables were the cognitive load of the logical decision making task, the amount of feedback support provided in acquisition, and the quasi-independent variable of WMC. Participants aged 18-35 were screened to be in the first and fourth quartiles for WMC via the operation-span task compared to prior samples of the Atlanta community. Amount of feedback support (low or high) was operationally defined as being low or high on the three feedback parameters: content, frequency, and timing (Figure 5).

Dependent measures were performance at acquisition and multiple tests of retention and transfer. Performance on these tests was measured as mean accuracy and reaction time for each block. Accuracy was measured as number of logic gates answered correctly divided by total gates presented, and was the primary measure of learning. Response time was the time required for each trial and was measured, but not considered primary. Response time was limited for each of the task load conditions to make accuracy the primary measure. Time spent looking at feedback on each trial and time spent studying summary feedback were also measured.

Such division of results into acquisition and retention for analysis is common and necessary, due to problems equating performance in acquisition to retention or transfer (Schmidt & Bjork, 1992).

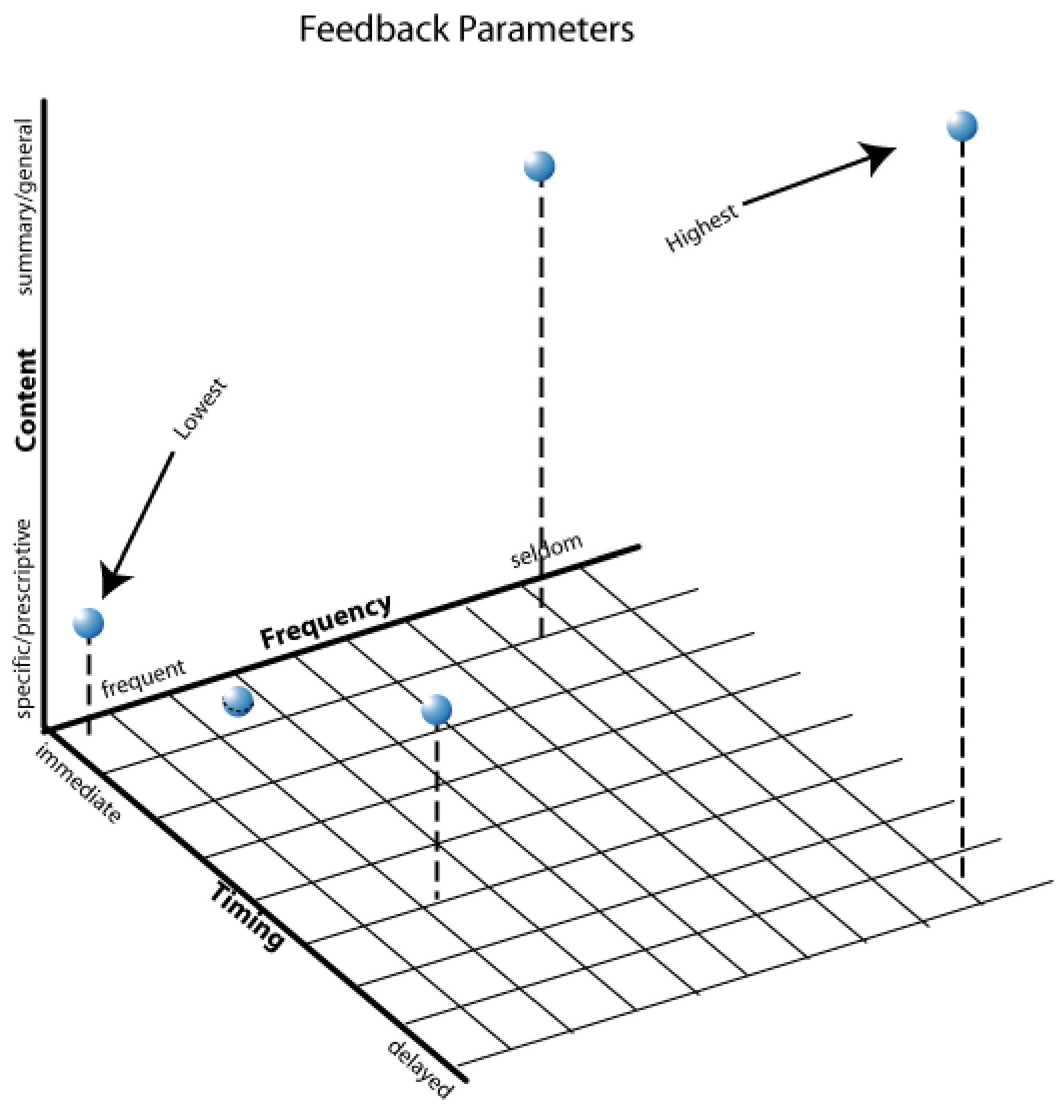


Figure 5. Three parameters of feedback illustrating the concept of feedback support

Retention Intervals

The purpose of a retention interval between acquisition and test was to dissipate the temporary effects of acquisition (Schmidt & Bjork, 1992). The retention intervals were chosen after considering intervals used by previous studies of feedback and research specifically on retention intervals. One criticism of retention intervals used in research is that they are not long enough to reveal differences (Fisk, Hertzog, Lee, Rogers, & Anderson-Garlach, 1994). Indeed, some studies of feedback used intervals as short as 10 minutes (e.g., Lee & Carnahan, 1990; Wrisberg & Wulf, 1997; Wulf & Schmidt, 1994; e.g., Wulf, Shea, & Matschiner, 1998), 48 hours (Rogers, 1996; Schmidt et al., 1989), and some as long as 30 days (Meyer, 1998). Intervals of 48 hours or more tended to be for studies of well-learned material. However, the current experiments investigated the first stage of learning where decay may occur more quickly and is not yet considered well-learned (Fitts & Posner, 1967). Retention intervals of 10 minutes and one week were chosen to dissipate the temporary effects of acquisition.

Ensuring controlled processing. The learning stages in these experiments required controlled processing resources and have not moved to automaticity or dependence on perceptual-motor speed (Ackerman, 1988). From previous research, it appears that learners are still in the first stage of learning with 400 or fewer trials of the LGT (Carlson et al., 1989). Thus, exposure to 200 trials of the LGT should ensure learners use controlled processing to solve the gates. It was important that participants' performance in acquisition and retention be resource-limited, not data-limited. This was ensured by putting time constraints on each of the task loads.

Tests

10-minute retention. The first retention test occurred ten minutes after initial acquisition trials ended. This 10-minute interval was filled with demographics questionnaires and the Shipley vocabulary test.

1-week retention. The second retention test occurred seven days after the initial acquisition session.

***Content transfer*.** The Content Transfer test was to solve logic gates using “L” (for low) and “H” for high in place of the 1’s and 0’s learned previously (Kyllonen & Woltz, 1989). Logic gate inputs were also converted to Ls and Hs.

***Load transfer*.** The Load Transfer test consisted of solving gates in the opposing load condition from acquisition: high load learners transferred to the low load task and vice versa.

Overview of the Analyses

The following analyses were carried out in each of the four experiments. Data were divided into acquisition and retention/transfer sessions for analysis, each answering the question of how feedback influenced accuracy. Simple task accuracies were computed as correct or incorrect for each trial. Complex task accuracies were computed as correct if the last gate in the problem were answered correctly. Session 1 consisted of pre-training (10 trials), acquisition with feedback present (10 blocks), immediate test (1 block), and 10-minute retention (1 block). Session 2 consisted of 1-week retention (4 blocks), the Content Transfer task (4 blocks), and a Load Transfer task (2 blocks). Effects of WMC group and Task Load are also reported, even when there was no interaction with feedback, to demonstrate there were indeed differences between the groups.

A one-way analysis of variance (ANOVA) was conducted to evaluate the relationship between WMC, Task Load, and Feedback Condition for each phase of learning and on transfer tests. Independent variables, their levels, and the dependent measures are described in the section for each experiment.

Although response time may also be an indicator of learning, accuracy was assumed to be the most valid indicator of learning for a cognitive task requiring controlled processing resources. Response time was limited to make accuracy the primary dependent measure. Accuracy proportions were subjected to an Arcsine transform prior to analysis to approximate a normal distribution (Stuart & Ord, 1994); however graphs are presented untransformed to reflect actual accuracy.

An alpha level of .05 was used for all ANOVAs with marginal effects discussed up to alpha = .07. *A priori* contrasts were performed between feedback types for retention and transfer tests with an alpha level of .05. Type I error was controlled in any post hoc contrasts by reducing alpha to .01.

CHAPTER 3: EXPERIMENT 1 – FEEDBACK SUPPORT, LEARNER RESOURCES, AND TASK LOAD

Overview

Experiment 1 explored the role of feedback support in acquisition as it related to learner characteristics and task load. Participants were grouped according to working memory capacity and attempted to learn either a simple or complex logical decision making task. On some gates, they received feedback designed to highly support their performance and for other gates they received feedback that provided little support for performance during acquisition.

Participants received one acquisition session then performed retention tests after various delays. These tests were followed by transfer tests. Accuracy and reaction time on these tests were the dependent variables in the study used to represent learning of the LGT.

Research Questions

Experiment 1 was designed to test the effects of feedback support on learning including the possible moderating effects of learner working memory capacity and task load.

Hypotheses

It was expected that learning would vary according to the support of the feedback provided, but this effect would be different for each of the WMC groups depending on task loads. If cognitive load theories of instruction were supported, all groups should learn more in the highest feedback support condition: KCR. On the other hand, if challenge increases learning, in a simple task, high WMC individuals should learn more with less supportive feedback (summary feedback) whereas the opposite should be true for low WMC individuals (KCR). However, in a complex task, high WMC individuals should also benefit most from KCR.

Method

Participants

Eighteen high- and eighteen low-working-memory capacity ("high WMC" and "low WMC") young adults (18-35 years of age) were recruited from the community and from a database collected by the Attention and Working Memory Laboratory at the Georgia Institute of Technology. Participant demographics and characteristics are presented in Table 4. No participants had experience with logic gates (assessed by a screening survey and prior experience questionnaire (Appendix A). Participants were compensated for their time either at the rate of $10/hour or 1 credit per hour.

Table 4

Participant Characteristics for Experiment 1

	Low WMC n = 18					High WMC n = 18					
	Simple Task		Complex Task			Simple Task		Complex Task			
	M	SD	M	SD	*t*	M	SD	M	SD	*t*	*t*
						General demographics					
Gender	3 males 6 females		5 males 4 females			3 males 3 females		3 males 6 females			
Highest level of education[1]	5.38	1.60	5.13	0.99	0.38	5.14	1.21	4.88	1.36	0.40	0.55
Age	26.00	4.72	23.43	2.76	1.26	22.14	4.60	21.75	5.04	0.16	1.80
Handedness[2]	1.78	0.44	1.89	0.33	-0.60	1.89	0.33	2.00	0.00	-1.00	-1.05
Near Vision[3]	23.33	7.07	20.00	0.00	-1.00	20.00	0.00	20.00	0.00	-1.00	1.37
Far Vision[3]	19.44	9.63	14.67	4.18	1.36	15.33	3.50	16.00	4.72	-0.34	0.68
Shipley Vocabulary Score[4]	28.33	6.18	30.44	4.39	-0.84	31.56	2.19	31.78	3.07	-0.18	-1.63
Simple Reaction Time[5]	307.33	77.17	287.33	40.60	0.69	260.00	21.92	264.56	21.49	-0.45	2.31*
Choice Reaction Time[5]	346.89	67.34	334.67	34.64	0.48	292.33	17.76	306.78	35.29	-1.10	2.95*
Digit Symbol Substitution[6]	69.22	14.42	60.22	18.36	1.16	73.67	13.95	75.67	7.35	-0.38	-2.12*
Reverse Digit Span[6]	6.00	1.12	7.56	2.40	-1.76	10.00	2.00	10.56	2.46	-0.53	-5.02*
Ao-span absolute[7]	16.00	8.51	19.67	8.67	-0.90	60.00	7.62	63.22	8.09	-0.87	-16.05*
Ao-span Total[7]	40.22	11.46	39.89	13.22	0.06	68.67	3.32	71.11	2.80	-1.69	-10.18*

Table 4 (continued).

	Low WMC n = 18					High WMC n = 18					
	Simple Task		Complex Task			Simple Task		Complex Task			
	M	SD	M	SD	*t*	M	SD	M	SD	*t*	*t*
LGT-specific demographics											
Number of times to pass matching test[8]	2.22	0.83	2.00	0.71	0.61	1.44	0.73	1.22	0.44	0.78	3.43*
Are you familiar with any of the logic gates?	0.22	0.44	0.22	0.44	0.00	0.11	0.33	0.11	0.33	0.00	0.88
Have you ever taken a computer science course that required programming in a a computer language?	0.22	0.44	0.11	0.33	0.60	0.44	0.53	0.44	0.53	0.00	-1.84
If yes, have you ever used operators such as AND, OR, XOR, or NAND in programming?	n = 2		n = 1			n = 4		n = 4			
	1.00	0.00	0.00	--	--	0.75	0.50	0.75	0.50	0.00	-.25
Have you ever used operators such as AND, OR, XOR, or NAND in a library or web search?	0.33	0.50	0.33	0.50	0.00	0.67	0.50	0.67	0.50	0.00	-2.06*
I now feel confident about my ability to solve logic gates. (Start)	4.56	1.42	4.56	1.01	0.00	3.89	2.09	4.56	1.42	-0.79	0.66
I now feel confident about my ability to solve logic gates. (End)	3.33	1.73	4.11	0.78	-1.23	3.67	1.58	4.67	1.22	-1.50	-0.94
I am motivated to do my best (Start)	6.67	0.50	6.67	0.71	0.00	5.67	1.00	6.33	0.71	-1.63	2.61*
I was motivated to do my best. (End)	6.67	0.50	5.33	1.50	2.53	5.33	1.50	5.44	1.59	-0.15	1.31

Table 4 (continued).

	Low WMC n = 18					High WMC n = 18					
	Simple Task		Complex Task			Simple Task		Complex Task			
	M	SD	M	SD	*t*	M	SD	M	SD	*t*	*t*
	CREATE Battery demographics[9]										
Number Comparison	52.60	10.90	50.33	12.90	0.267	62.50	14.46	68.00	10.63	-0.73	-2.67*
California Verbal Learning Test	10.00	6.60	9.67	2.89	0.081	14.25	2.36	13.43	1.81	0.651	-2.26*
Meaningful Memory	15.25	5.19	12.60	3.36	0.931	17.50	2.07	17.44	3.05	0.039	-2.66*
Alphabet Span Absolute	4.25	0.65	4.40	0.55	-0.378	5.33	0.75	5.22	0.75	0.28	-3.29*
Alphabet Span Total	27.75	8.22	31.20	6.57	-0.702	43.17	12.21	44.56	9.28	-0.25	-3.72*
Letter Sets	16.40	6.88	14.67	3.51	0.398	24.50	3.00	24.71	2.43	-0.13	-4.70*
Information (WAIS-III)	21.75	3.59	20.40	3.36	0.581	21.33	4.18	19.11	2.37	1.322	0.72

Note. * indicates significant differences between groups ($p < .05$).
[1]Education rated on an ordinal scale; available in Appendix A. [2]Handedness scored as 1 for left, 2 for right. [3]Vision scored as Snellen acuity. [4]Shipley, 1986 [5]reaction time in milliseconds [6]Wechsler, 1997a [7]Unsworth, et al., 2005 [8]Criterion test of logic gate rule memorization 9Full details on tests in Appendix C.

Excluded Participants

Nine high WMC participants and one low WMC participants were excluded from the analyses. Three of the nine high WMC participants did not return for the second session. The other participants were removed randomly to ensure equal numbers of participants per condition and counterbalance. Their exclusion did not change the pattern of results; analyses performed with unequal numbers of participants are available upon request. These participants were not included in the thirty-six listed above.

Materials

Ability Tests

Participants completed tests of perceptual speed, long-term memory, verbal ability, spatial/visualization, working memory capacity, reasoning/induction, and crystallized intelligence. A complete list of the tests used is in Appendix C. This battery was developed by the CREATE group to understand age-related changes in abilities (Czaja, Sharit, Charness, Rogers, & Fisk, 2002).

Instructions

Instructions consisted of a written walkthrough of logic gate operation followed by a single sheet of logic gates and their rules. Participants were allowed five minutes to memorize these rules before taking a matching test of gates and their rules (Appendix A). Any errors on this test were explained to the participant, and the participant took the criterion test again until achieving perfect performance.

Questionnaires

All questionnaires may be found in Appendix A. Questionnaires included a screening survey to determine previous exposure to logic gates, participant demographics, prior-experience with logic gate-like tasks, the criterion test of logic gate definitions, pre-acquisition opinions of the task, the test of logic gate declarative learning, and an exit interview.

The screening survey assessed whether participants could solve logic gates prior to participation in the study; all only those who marked “no idea” or answered all logic gate questions incorrectly were included. The prior-experience questionnaire was designed to not only discover if a participant specifically knew the LGT, but whether the participant was familiar with the *type* of task. Similar tasks included searching the internet using Boolean logic, electrical wiring, philosophy classes and the open ended question of whether logic gates seemed similar to anything they learned before.

The test of logic gate declarative learning was similar to the test of definitions, but given *after* the retention test and was an open-ended test to determine if the participant could write the correct rule for each gate.

The exit interview contained general questions about the participant’s experience in the study, including distraction and comfort level. The interview also collected information as to how capable the participant felt completing the LGT after retention and transfer tests.

Equipment

The experimental task was performed on IBM-compatible computers (3.2 GHz Pentium 4, 1 GB RAM). Screen size was 19” with a resolution of 1024 x 768 pixels and a refresh rate of 85 Hz.

Experimental task. A computerized version of the LGT presented acquisition trials, feedback, retention tests, as well as the Content and Load Transfer tests. This program was used for all experiments. The high feedback support condition consisted of half of the gates, where participants received feedback after every trial and in summary form at the end of a block of 20 trials. The other five gates (low support) received feedback only in summary form at the end of a block (according to counterbalance). The feasibility of manipulating feedback support level within participants was examined in Pilot Study 2.

Equal numbers of each gate were presented in the same order for each participant. Thus, participants in the simple version of the task received the same logic gates, presented in the same order and participants in the complex group received the same gates in the same order. Equal numbers of each gate appeared in every block of trials to allow comparison between blocks. Order was determined by ensuring no trials of the same gate followed each other within a block.

Feedback Conditions

Summary Feedback

This comprised the low-support feedback condition. In this condition, gates received no feedback during a block. At the end of the block, a summary screen appeared with a picture of each gate and the percent accuracy for that gate in the previous block. Participants could tell from this information for which gates they had the lowest accuracy. This feedback was infrequent, abstract, and delayed.

Knowledge of Correct Response (KCR)

This comprised the high-support feedback condition. In this condition, feedback of correct or incorrect was presented each time a gate was answered. These gates also received summary feedback at the end of a block. This feedback was frequent, prescriptive, and immediately presented.

Learning Phases

Acquisition

Acquisition occurred over 10 blocks of 20 trials each. Participants received feedback on the gates in the two feedback conditions.

Immediate test. This block of trials occurred immediately after block 10 in acquisition, but no feedback was presented.

Retention Tests

Retention tests occurred ten minutes after the immediate test and one week after acquisition.

Transfer Tests

Transfer tests included changing the content (the inputs and outputs) of the gates and changing the task load (from simple to complex or vice versa). These tests always occurred after the 1-week retention test. Because the Content transfer test was of primary interest, it always occurred before the Load transfer test.

Counterbalance effects. Gates were divided into counterbalance groups according to data acquired in Pilot Study 1. There were no effects or interactions of counterbalance group with the other variables of interest. Thus, counterbalance groups were combined for the main analyses.

Procedure

All participants completed the CREATE battery of ability tests in addition to the Ao-span. Participants completed this battery in two sessions prior to the acquisition and retention sessions for the LGT.

Session One

Before Logic Gate acquisition, participants signed an informed consent then took the screening survey if they had not done so previously. Participants completed the digit-symbol substitution and received instruction on logic gates, followed by the Prior Experience Questionnaire.

Because of the decision to include specifically low WMC individuals, the task was simplified so that memorization of gate shape was not required. The label for each gate always appeared on the gate, negating the need to also recall the name of the gate when solving a logic gate problem.

Participants were given 5 minutes to study the logic gates and their definitions before answering a matching test. Once the participants demonstrated memorization of the gate definitions on the matching test (criterion), the experimenter demonstrated solving logic gates in the LGT. This consisted of pointing out the keys to use for answering (1 and 0 on the numeric keypad) and pointing out screen elements that contained information (how errors were displayed, how to interpret feedback both after a trial and after the end of a block). Practice, or pre-training, consisted of participants solving ten trials in their task load condition, using the rule-sheet from which they learned the definitions. These ten trials specifically exposed the participant to each of the ten gates.

For the acquisition phase, participants performed 10 blocks of 20 trials each. Each trial consisted of: 1) A screen presented the incoming signal on the left side of the logic gates. 2) The participant pressed a numeric key to indicate the outgoing signal as transformed by the logic gates. Figures 6 details task and feedback presentation. Time constraints were placed on answering gate problems to ensure participants were not performing at data-limited levels. These time constraints (2.89s for simple and 10.35s for complex) were determined via pilot testing and consisted of the mean of Georgia Tech student performance plus one standard deviation. At the end of acquisition, participants completed the immediate test block of trials, then the Shipley vocabulary test and a 10-minute retention block.

Session Two

All participants returned after one week (all within a range of 4 hours of their first session time) to perform 80 trials of the LGT with no feedback, breaks every 20 trials. Then participants completed the content and Load Transfer tests. Last, participants filled out the exit interview and received verbal/written debriefing and payment for their time. Participants who had not yet attended the CREATE individual testing session did so at the end of the Logic Gate study following the protocol for CREATE.

Design

The study was a 2 (WMC group: low WMC, high WMC) x 2 (task load: simple, complex) x 2 (feedback condition: Summary Feedback [low support], KCR [high support]) factorial. Feedback Condition was a within-participant factor, Task Load a between-participant factor, and WMC Group a between participants grouping variable. The dependent variable of interest was accuracy of performance, however response time was also measured.

Analyses

A repeated measures analysis of variance (ANOVA) was performed on the main dependent measure of interest: accuracy. Between participant variables included Working

Memory Capacity and Task Load while Feedback was manipulated within participants. The variable of Block was within-participants when present.

Additional Analyses

Performance on the ten gates was highly variable due to differences in the difficulties of the gates. For example, the BUF gate often showed performance accuracy of higher than 90%, while the XNOR gate hovered above chance (50%). A post hoc analysis was carried out where the gates were divided into levels of easy, medium, and difficult according to prior research on logic gates as well negation procession of the gate (for example, a negation gate such as NAND will be more difficult than AND). Gates may be grouped into three levels of difficulty: Easy (BUF, INV, HIGH, LOW), Medium (AND, OR, XOR), and Difficulty (NAND, NOR, XNOR). A look at the actual accuracies of the gates across Experiment 1 confirmed that the gates did conform (for the most part) into the predicted general categories of difficulty.

Thus, additional analyses using Gate Difficulty as a between participant variable were performed for each learning stage and significant results involving Gate Difficulty and feedback are presented in the following results section.. Because of the mixture of gates in the complex task condition, this post hoc analysis was only carried out on the simple task.

There was some evidence that participants responded well within the time allotted per trial. Thus, a doubly-multivariate analysis was performed on all data due to the non-commensurate dependent variables of accuracy and response time, measured synchronously. In almost all cases where accuracy showed an effect that effect was also significant at the multivariate level. Those which were not are marked with a (†). If the test from the doubly multivariate analysis was significant, a univariate test of that effect on each dependent variable was performed. These results and all ANOVA tables are available in Appendix E, however most differences in response time were due to task load and did not interact with feedback.

A Post Hoc Caveat

The simple and complex LGTs were designed to vary by load; however, there was evidence that the time constraints put on the tasks made the simple task more difficult than the complex task. Thus, though the tasks differed in complexity and load levels by design, the simple task cannot be thought of as universally easier than the complex task. Across experiments participants in the simple task condition were more likely to time-out on a trial than were those in the complex task condition.

In conclusion, the tasks reflect simple and complex versions of the LGT, but in interpretation it may be better to think of them as different tasks rather than an extreme groups manipulation of complexity and difficulty for the same task.

Results

Acquisition

Beginning and End of Acquisition

Acquisition was defined as the blocks of trials with feedback present. There were 10 acquisition blocks. The first and last blocks of acquisition were analyzed to measure performance changes with feedback still present. It was expected that performance would improve from block 1 to block 10 and that it would improve differentially according to feedback condition. Performance with KCR was expected to exceed that of Summary Feedback by the last block of acquisition (Block 10). Table 5 provides a list of group means and contrast analyses.

Feedback effects and interactions. As expected, there was an interaction of Block x Feedback, $F(32,1) = 6.97$, $p = .013$, $np^2 = .18$, where accuracy was the same for both feedback types at block 1 (.80, .81), but by block 10 KCR was more accurate than Summary Feedback ($t = -2.91$, $p = .006$). The main effect of Feedback was marginally significant, where KCR was more accurate than Summary Feedback (.90 vs .97) $F(32,1) = 3.84$, $p = .059$, $\eta p^2 = .11$.

Other effects. There was a main effect of Block, $F(32,1) = 65.76$, $p < .001$, $\eta p^2 = .67$, where participants were more accurate by block 10 than on the first block. High WMC participants were more accurate than low WMC participants, $F(32,1) = 10.47$, $p = .003$, $np^2 = .25$. There was an interaction of Block x WMC where high WMC participants improved their performance more from beginning to end of acquisition than did the low WMC participants, $F(32,1) = 5.87$, $p = .021$, $np^2 = .16$.

Table 5

Accuracy Levels for Experiment 1 divided by Feedback Condition

		Beginning of Acquisition					End of Acquisition					10-minute retention				
		Mean accuracy		Standard		*t*	Mean accuracy		Standard		*t*	Mean accuracy		Standard		*t*
WMC	Task Load	End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial	
Low	Simple	0.61	0.44	0.20	0.16	2.72*	0.62	0.73	0.12	0.15	-1.75	0.57	0.74	0.16	0.17	-2.49*
	Complex	0.38	0.49	0.14	0.15	-1.63	0.57	0.68	0.22	0.27	-1.27	0.52	0.60	0.16	0.19	-1.12
High	Simple	0.59	0.61	0.13	0.14	-0.39	0.76	0.84	0.19	0.12	-1.25	0.76	0.84	0.22	0.12	-0.97
	Complex	0.50	0.51	0.21	0.18	-0.13	0.79	0.88	0.19	0.13	-1.44	0.76	0.86	0.19	0.14	-1.17

		1-week retention					Content transfer					Load transfer				
		Mean accuracy		Standard		*t*	Mean accuracy		Standard		*t*	Mean accuracy		Standard		*t*
		End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial	
Low	Simple	0.57	0.72	0.17	0.10	-2.56*	0.53	0.65	0.16	0.15	-1.55	0.61	0.59	0.16	0.16	0.14
	Complex	0.57	0.66	0.13	0.14	-2.17*	0.58	0.56	0.14	0.14	0.34	0.62	0.73	0.17	0.14	-2.01
High	Simple	0.69	0.80	0.17	0.10	-2.17*	0.69	0.75	0.16	0.08	-1.33	0.71	0.81	0.21	0.13	-1.84
	Complex	0.71	0.78	0.16	0.14	-0.90	0.74	0.72	0.15	0.12	0.46	0.83	0.82	0.11	0.09	0.38

Note. * indicates significant differences between groups ($p < .05$).

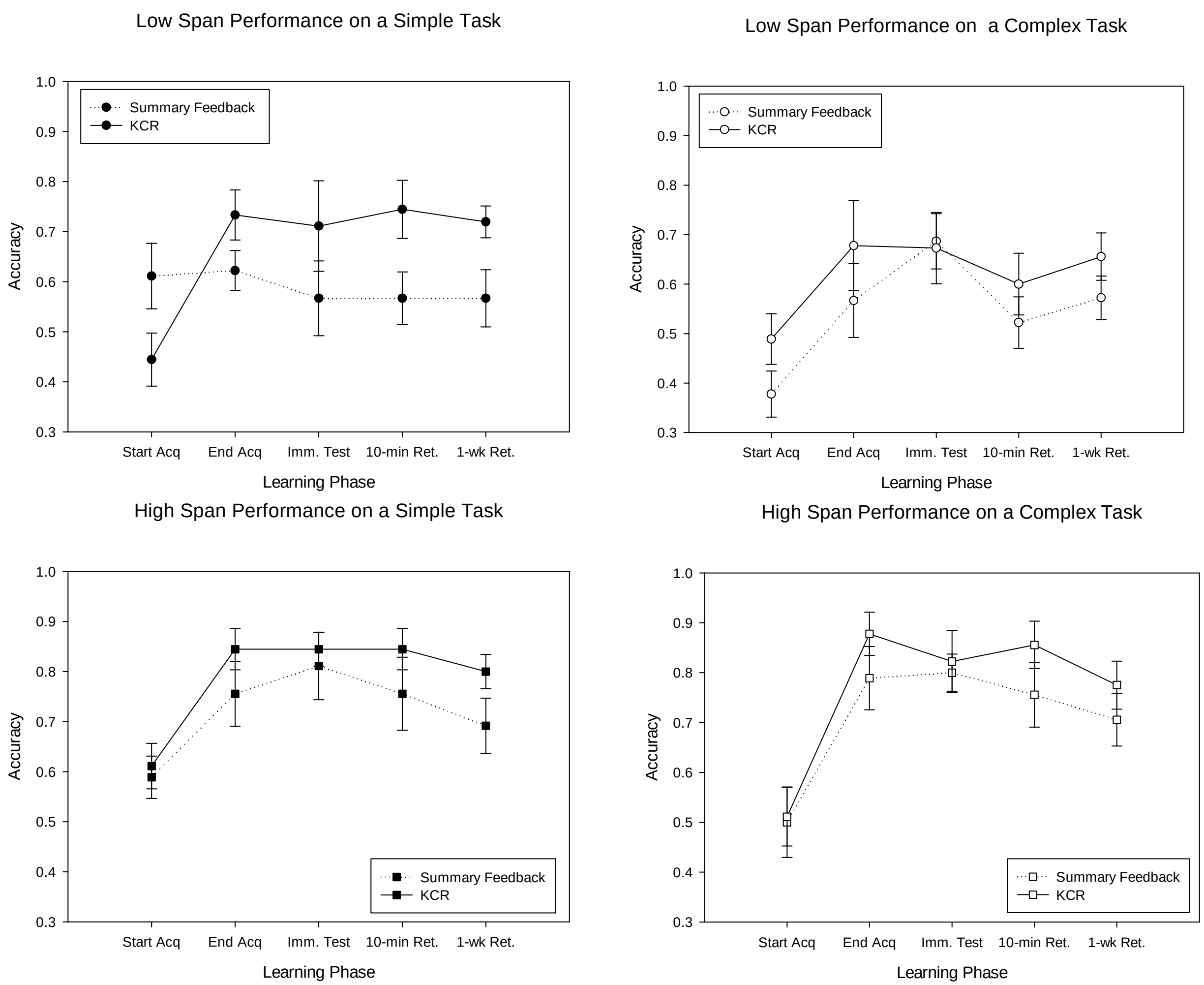


Figure 6. Experiment 1 performance under feedback support conditions divided by WMC and Task Load.

Acquisition summary. The predictions were supported (Figure 6). On gates where participants received high feedback support, they performed more accurately and improved more across blocks. High WMC individuals performed more accurately than low WMC individuals and improved more across acquisition.

Immediate Test

Participants completed a single block of trials with no feedback present immediately after the 10 blocks of acquisition. Immediate test was not a test of learning or retention as there was no retention interval and any temporary effects of feedback could not be assumed to have dissipated. It was expected that the effects of acquisition would still be present (fatigue, motivation from the feedback, etc) and results would be similar to block 10 performance. Results are presented as an indicator of performance without feedback in place to better compare end-of-acquisition performance to retention and transfer.

Feedback effects and interactions. There were no effects or interactions concerning Feedback at immediate test.

Other effects. Working memory capacity contributed main effects to performance where high WMC participants performed more accurately than low WMC participants, $F(32,1) = 12.88$, $p = .001$, $np^2 = .29$.

Retention

There were two separate retention tests. The first retention test occurred 10 minutes after the immediate test. The second retention test occurred one week after the acquisition session. Retention data were analyzed both as post scores (analyzing the retention score only and ignoring the baseline) and as change scores (where differences between sessions were analyzed). Both analyses are presented because although change scores provide some of the most important information concerning retention, they do not provide any ordinal information concerning the feedback conditions. For example, accuracy may decline for a certain condition by 25% but still have accuracy far above the opposing condition.

Cognitive load theory, (Sweller, 1988; Sweller & Chandler, 1991; van Merrienboer, Kirschner, & Kester, 2003) predicted all groups would demonstrate better performance in retention on gates that received the highest feedback support in acquisition, due to a lower cognitive load while learning the task. These predictions fit within the ACT-R cognitive framework (Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004). Explicitly, all participants should perform more accurately retention tests when provided with KCR as opposed to Summary Feedback.

If learners need appropriate challenge when learning, (Elshout, 2006; Guadagnoli & Lee, 2004; Szalma et al., 2006), high WMC participants should learn the low-support gates better than the high-support gates in the simple task condition and would demonstrate higher accuracies for those gates in retention. Low WMC participant performance would benefit universally from additional feedback support. Because it was assumed load was relative to the complexity of the task and the WMC available to the learner, high WMC participants in the complex task should also benefit from increased feedback support, just as low WMC participants should for both task loads.

10-minute Retention

Feedback effects and interactions. There was a main effect of Feedback, $F(32,1) = 7.67$, $p = .009$, $np^2 = .19$, where KCR produced higher accuracies than did

Summary Feedback.

Other effects. High WMC participants were more accurate than low WMC participants, $F(32,1) = 16.17$, $p < .001$, $np^2 = .34$. There were no effects of Task Load on accuracy in this phase (all *p's* > .05).

Summary. These results support the idea that more supportive feedback in acquisition results in more learning demonstrated at retention. The next test occurred after one week and presumably after greater decay of temporary information.

1-week Retention Post Scores

Feedback effects and interactions. There was a main effect of Feedback where both WMC groups had higher accuracies on the gates where they received KCR a week earlier, $F(32,1) = 13.68$, $p = .001$, $np^2 = .30$, Figure 7.

Other effects. High WMC participants were more accurate at 1-week Retention than low WMC participants, $F(32,1) = 7.38$, $p = .011$, $np^2 = .19$.

Summary. These results support the idea that more supportive feedback in acquisition results in more learning demonstrated after a retention interval where decay could have occurred. The next test was designed to measure the amount of decay across the week interval.

Retention Change Scores

A percent change score was computed to further examine 1-week Retention using the following formula:

(1-week accuracy – 10-minute accuracy)/10-minute accuracy

A positive change indicated an improvement in performance from 10-minute retention to 1-week Retention and a negative score indicated a drop in performance. A change score of zero meant the information from the first session was perfectly retained across the week (Figure 8).

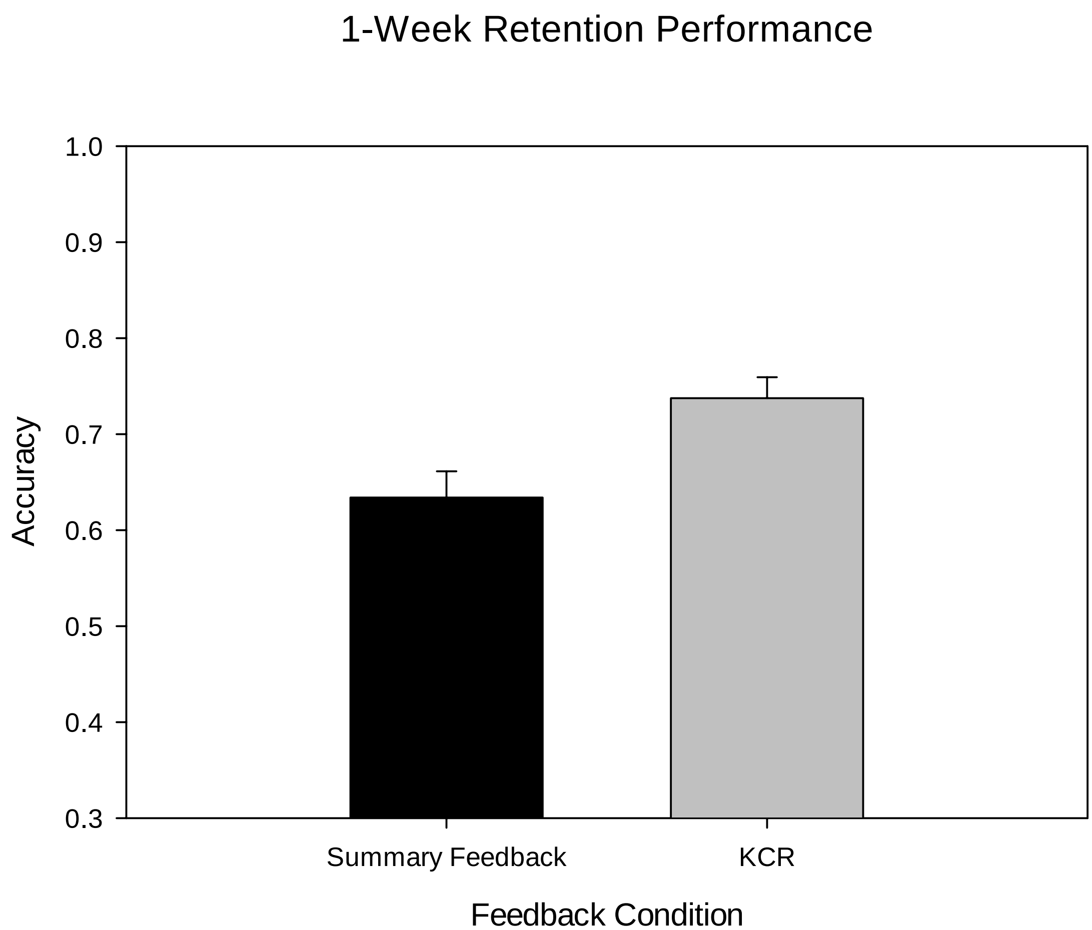


Figure 7. Experiment 1-week Retention performance.

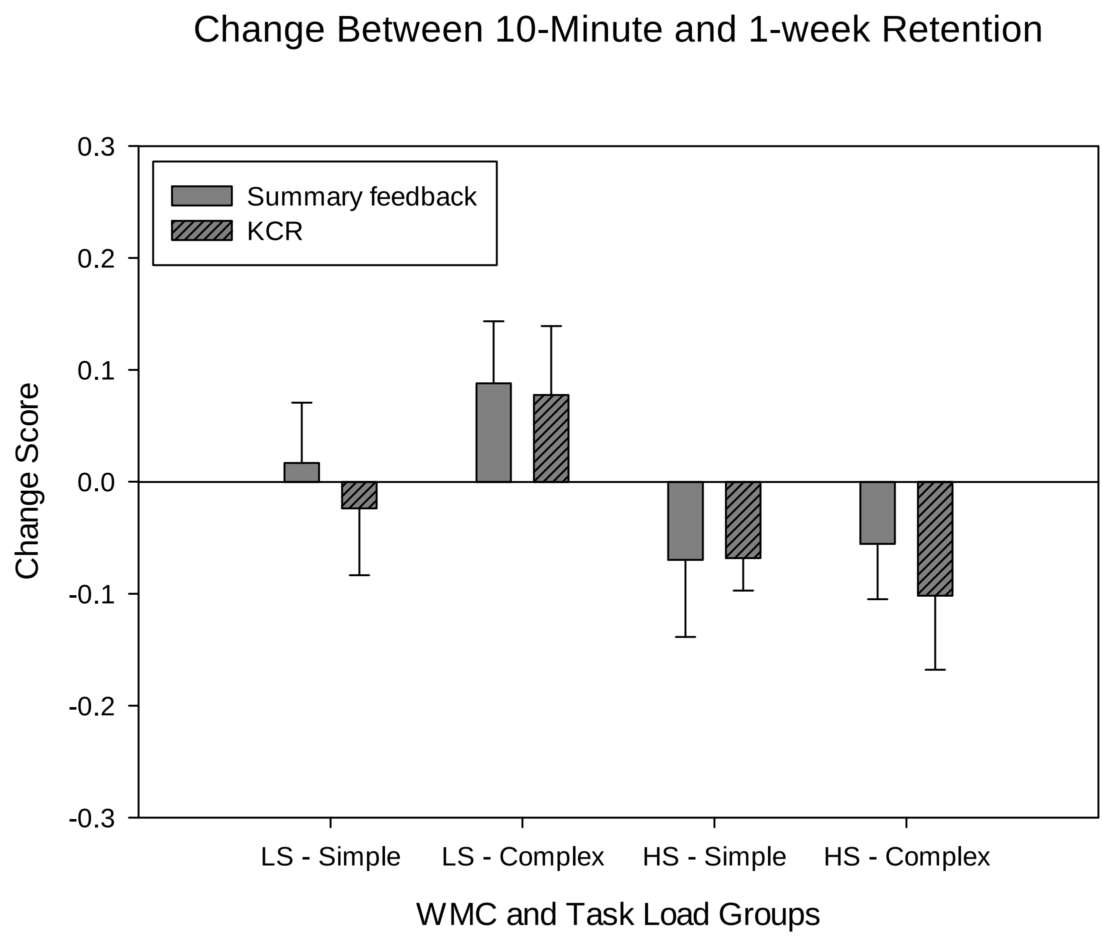


Figure 8. Change scores between 10-minute and 1-week Retention for Experiment 1.

Feedback effects and interactions. There were no effects of Feedback on the change scores from 10-minute to 1-week Retention. When change scores were compared to zero, only the high WMC participants in a simple task showed a significant decline, which was under the KCR feedback condition, t = -2.35, p = .047.

Other effects. There was an effect of WMC group on percent change scores in that high WMC participants showed decline or no change from a week earlier while low WMC participant performance showed no change, $F(32,1) = 11.43$, $p = .002$, $np^2 = .26$. A possible explanation for these data is that high WMC participants were initially at a higher accuracy level a week previous compared to low WMC participants and thus had more to "lose". Low WMC participant performance was not significantly different from chance (50% accuracy) at either 10-minute or 1-week retention when provided with Summary Feedback (t = 1.20, p = .246; 1.98, p = .063), however they were significantly higher than chance in the KCR condition (t =3.42, p = .003; 5.63, p < .001).

Retention Summary

The retention test results from Experiment 1 support the cognitive load theory of learning. Feedback that reduced load while learning a novel cognitive task resulted in better retention. This retention did not differ by the resources available to the learners (WMC) nor by the load placed on them via the complexity of the task. Learners performed more accurately when given KCR as opposed to Summary Feedback and maintained that edge from the end of acquisition through 1-week Retention tests. Thus, there were no differences in the change scores from 10-minute retention to 1-week Retention.

Transfer

All transfer tests occurred after the 1-week Retention interval. Tests included Content Transfer, where different inputs and response keys were substituted in the LGT, and Load Transfer, where those initially in the simple task condition transferred to the complex task and vice versa.

Predictions for the performance of the different WMC and feedback support groups were the same as for the tests of retention. Content Transfer was intended to be the primary transfer measure and thus always followed immediately after the test of 1-week Retention. Load transfer followed Content Transfer. Post and Change scores were analyzed for both transfer tests.

Content Transfer Post Scores

Feedback effects and interactions. Content Transfer scores indicated a marginal interaction of Feedback and Task Load, $F(32,1) = 3.325$, $p = .078$(m), $np^2 = .09$, where there were no differences between Feedback conditions for the complex task and participants benefited from KCR in the simple task. High WMC participant scores were significantly higher than chance (all p’s < .05). Only low WMC participants in the simple task receiving KCR were higher than chance on the content transfer test ($t = 2.76$, $p = .025$), (Figure 9).

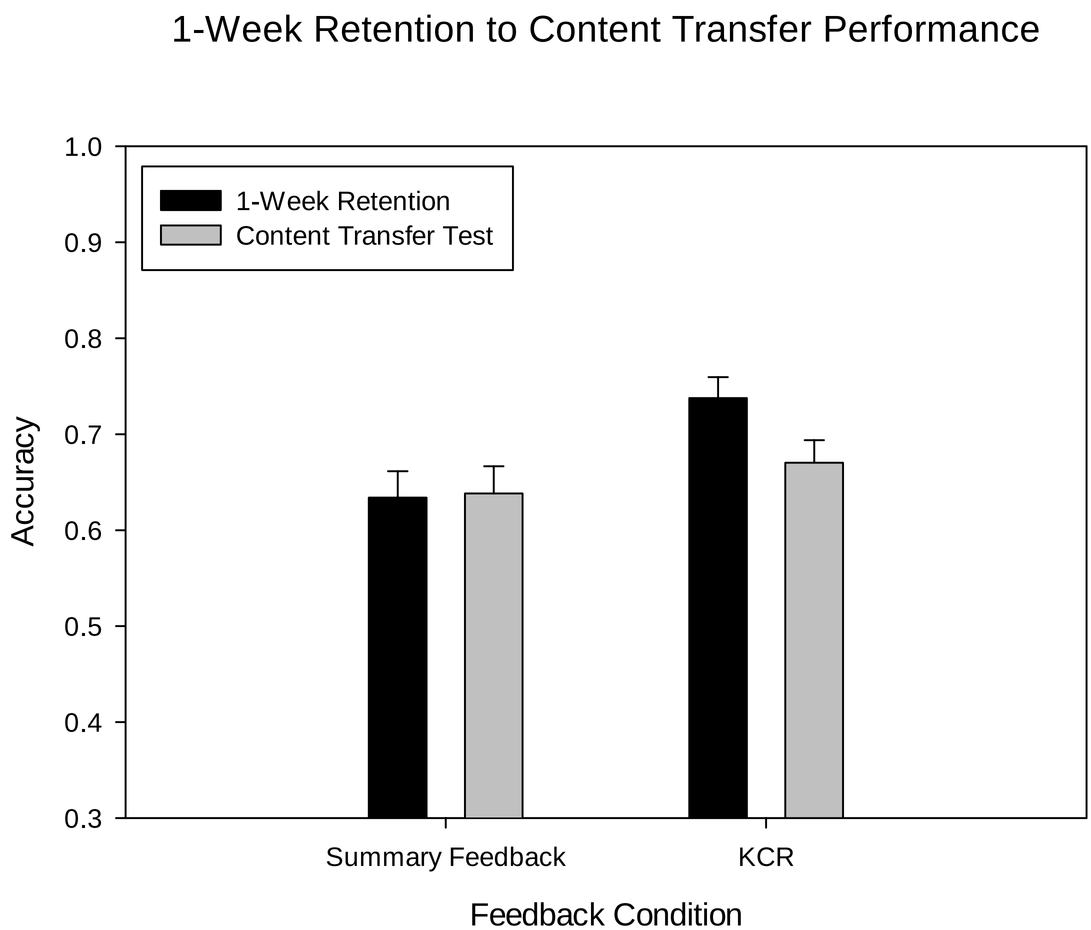


Figure 9. Experiment 1 Content Transfer Test performance data.

Other effects. High WMC participants were significantly more accurate than low WMC participants for the Content Transfer task, $F(32,1) = 14.64$, $p = .001$, $np^2 = .31$.

Content Transfer Change Scores

A change score was computed for the Content Transfer task using the equation: (Content Transfer accuracy – 1-week Retention accuracy)/1-week Retention accuracy Percent change was related to 1-week Retention accuracy rather than Immediate Test or 10-minute Retention because it was assumed non-feedback related changes would have dissipated during the weeklong retention interval, making the 1-week Retention the closest match to the Content Transfer task. Gate Difficulty interacted with Feedback for these scores; thus the difficulty analysis is reported.

Feedback effects and interactions. There was a main effect of Feedback on the Content Transfer change score, $F(32,1) = 5.27$, $p = .028$, $np^2 = .14$, where KCR performance declined more than Summary Feedback performance. Contrasts revealed that no groups differed significantly from 0% change except the low WMC participants in a complex task who received KCR feedback (Figure 10).

Gate difficulty. The content transfer test showed the first interaction of Feedback and Gate Difficulty. When the simple task data were analyzed using difficulty as a within-participant variable, the following effects emerged. There was a marginally significant interaction of Gate Difficulty and Feedback where performance improved via Summary Feedback for the easy gates and either declined or did not change for the medium or difficult gates, $F(32,2) = 2.51$, $p = .097$(m), $np^2 = .14$. KCR either declined or did not change for all levels of Gate Difficulty.

Summary for Content Transfer. When there was a benefit for one feedback condition, it favored KCR, despite KCR showing more decline from 1-week Retention. Though it may appear that Summary Feedback trials transferred better, since performance did not change from retention to transfer, the overall lowness of their accuracies must be considered. In both 1-week Retention and Content Transfer, low WMC participant accuracies under Summary Feedback were not significantly different from chance. Thus, the KCR feedback allowed for some transfer of skill while Summary Feedback did not.

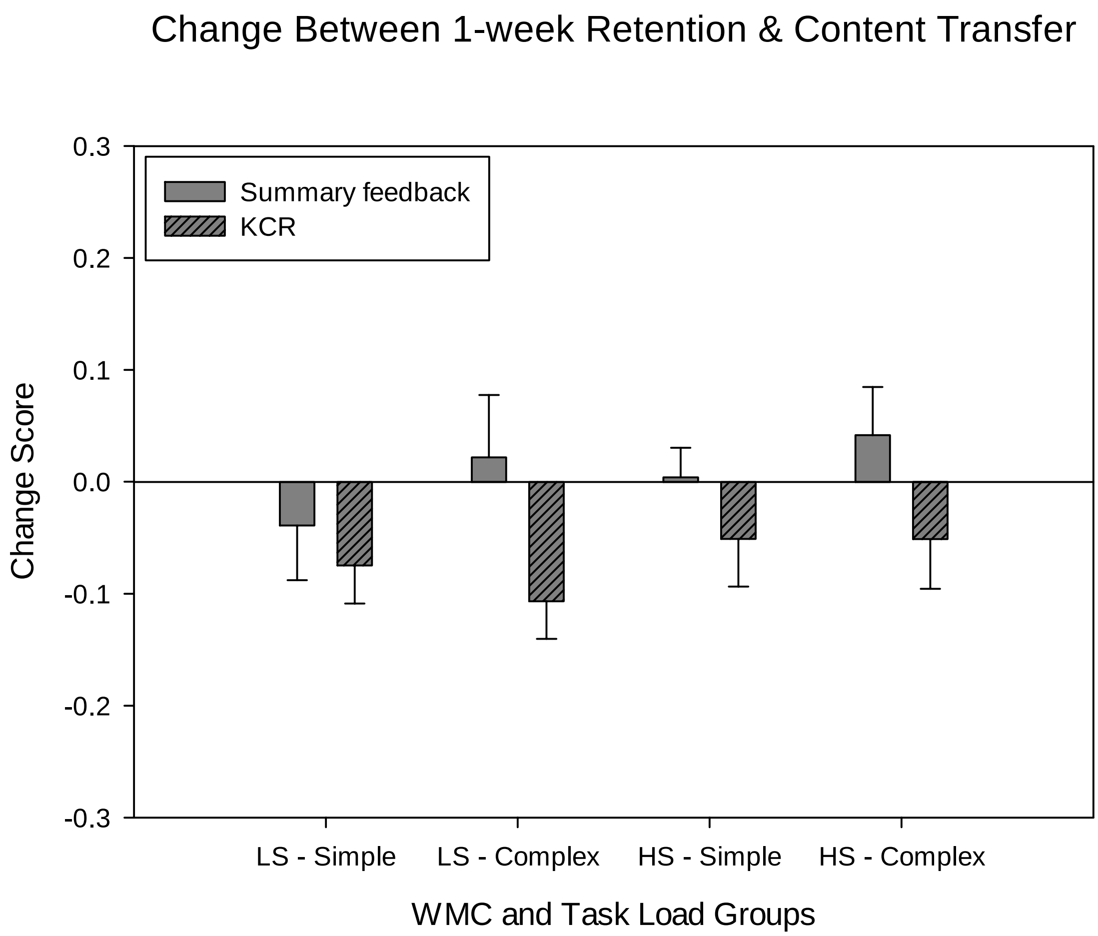


Figure 10. Change scores between 1-week Retention performance and Content Transfer Test.

Load Transfer

The Load Transfer task consisted of switching the between participant variable of task load. Those who acquired the task in the simple condition answered gates in the complex condition and vice versa. Because the main interest is feedback, the goal of this test was not to see whether the simple task transferred better to the complex task or vice versa. Therefore the analyses are overall post scores and change scores with each feedback condition.

Feedback effects and interactions. The Post score data indicated an interaction of

Feedback x Task Load, where those switched from the simple task to the complex task were more accurate with KCR feedback while there was no difference between feedback types for those switching from complex to simple task, $F(32,1) = 4.32$, $p = .049$, $np^2 = .12$, (Figure 11). Contrasts indicated that low WMC participant performance was significantly different from chance only when switching from a complex to a simple task and having received KCR, $t = 4.44$, $p = .002$. High WMC participant performance was higher than chance in all Task Load and Feedback conditions (all *p*'s > .05).

Gate difficulty**.** When gates were divided into easy, medium, and difficult, there was an interaction of difficulty and feedback, where the easy gates had similar accuracies for both feedback types while the medium gates were more accurate with KCR than with Summary Feedback, $F(31,2) = 5.20$, $p = .012$, $np^2 = .26$. The difficult gates had the lowest accuracies and under the Summary Feedback condition were not significantly different from chance.

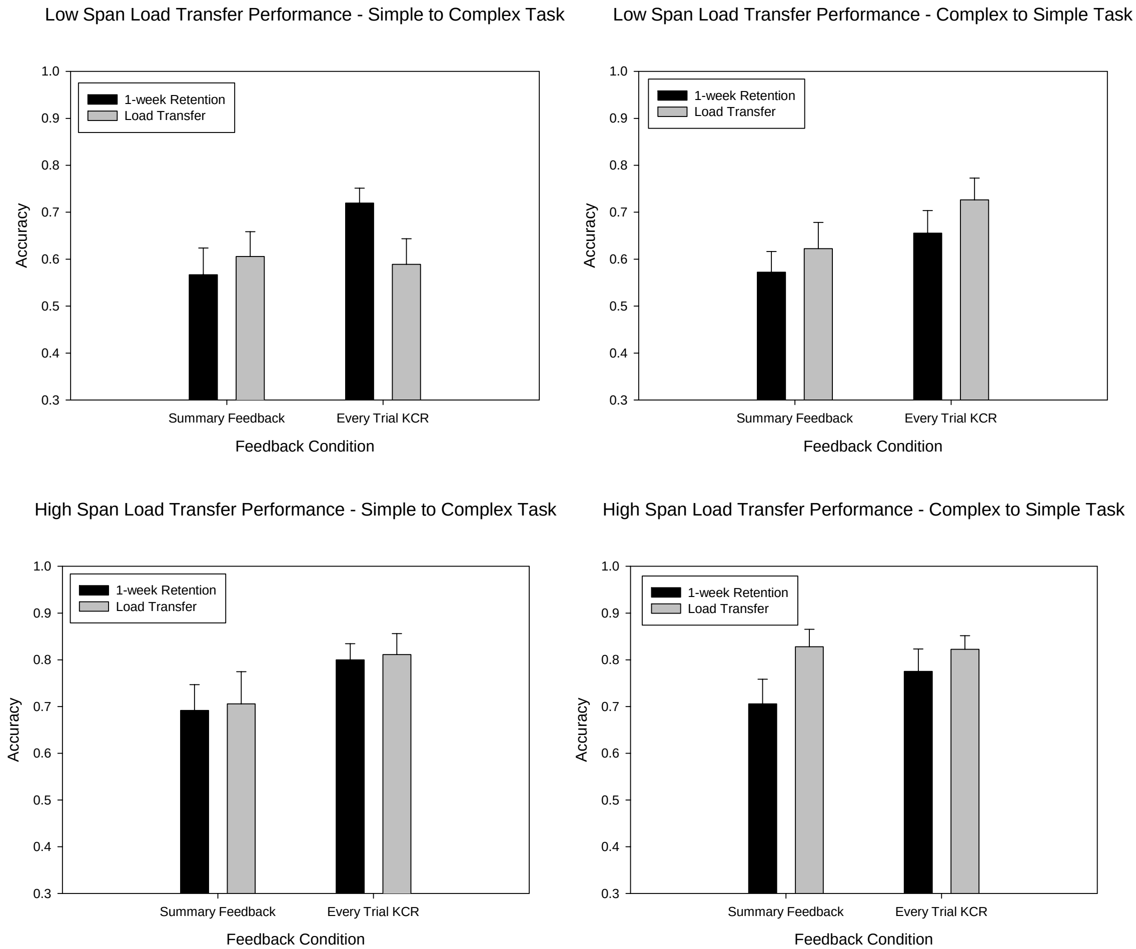


Figure 11. Experiment 1 Load Transfer Test performance

Load Transfer Change Scores

Feedback effects and interactions. The percent change analysis for the Load Transfer test compared to 1-week Retention revealed a three way interaction of feedback, WMC, and Task Load, $F(32,1) = 4.17$, $p = .049$, $np^2 = .12$, (Figure 12). Though no change was present for most groups, low WMC participants switching from simple to complex task load showed a negative change from their 1-week Retention performance on gates that received KCR, $t = -3.74$, $p = .006$. High WMC participants switching from the complex to simple task load showed a *positive* change for Summary Feedback, $t = 3.61$, $p = .007$. There was a marginal main effect of Feedback where accuracy changed more for Summary Feedback gates than for KCR gates, $F(32,1) = 3.35$, $p = .077$(m), $np^2 = .10$.

Other effects. As might be expected, there was a main effect of Task Load where participants switching from complex to simple improved their performance more than those switching from the simple task to the complex, $F(32,1) = 6.948$, $p = .013$, $np^2 = .18$.

Gate difficulty. When gates were divided into difficulty levels, feedback condition interacted with Gate Difficulty, $F(31,2) = 3.80$, $p = .034$, $np^2 = .20$. Though in the previous analysis Summary Feedback showed a gain for high WMC participants compared to KCR, it appears that this was only true for the easy and difficult gates. It is possible that part of this effect is due to their lower performance for Summary Feedback at 1-week Retention on difficult gates: any correct gates at Load Transfer would be an improvement in the Change Score. Another post hoc explanation is that high WMC participant improvement on the easy gates was due to crosstalk between the feedback conditions. When the task load was lessened, (complex switching to simple), they might have been able to apply their knowledge of the KCR gates to the previously unanswerable Summary Feedback gates.

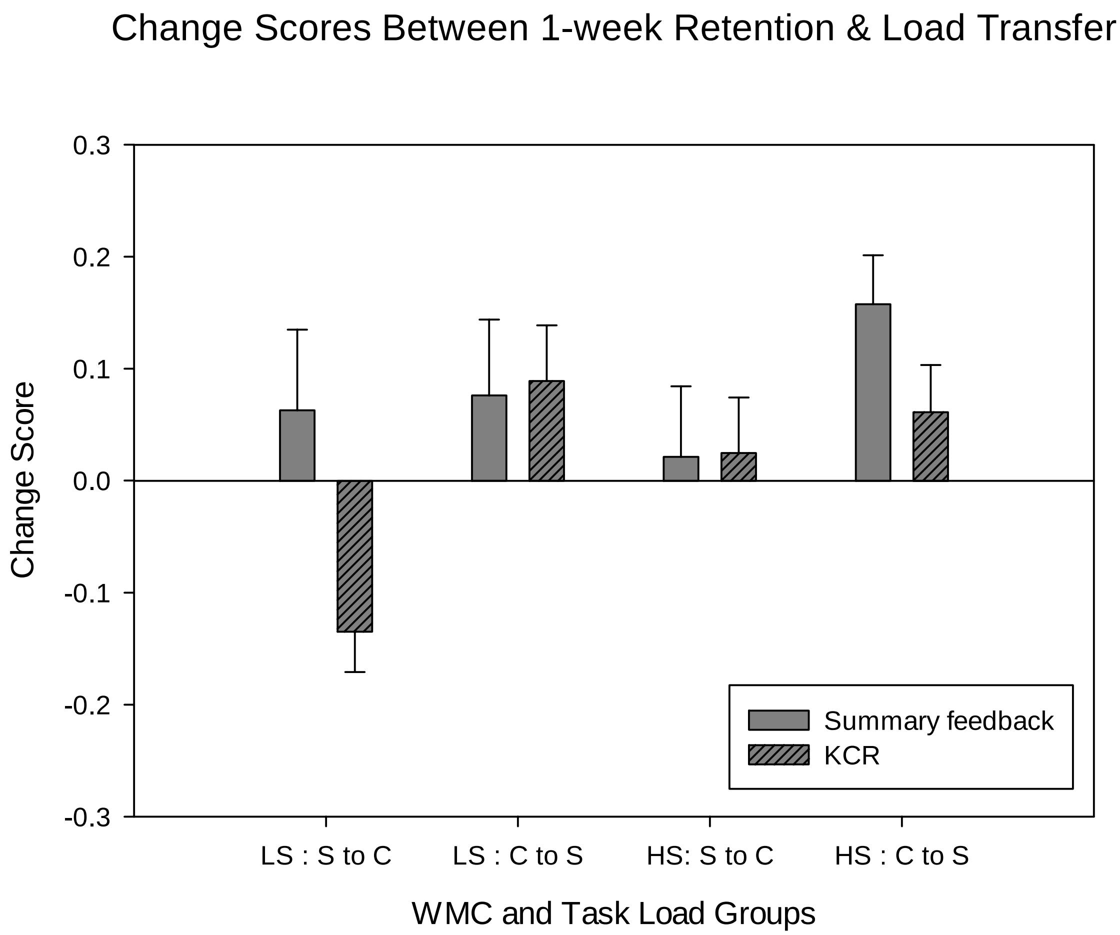


Figure 12. Change Score from 1-week Retention to Load Transfer Test performance for Experiment 1.

Load transfer summary. The highest transfer performance occurred for those switching from the simple task to the complex task, especially on gates where KCR was provided. The change score analysis, however, indicated that even the higher accuracies for KCR gates declined compared to retention test performance. Summary Feedback gates increased in accuracy at Load Transfer, possibly due to the increase in allowed response time and cross-talk between the feedback conditions.

Transfer Summary

For both tests of transfer, KCR gates showed higher accuracies than Summary Feedback gates. Indeed, in many cases performance was not significantly different from chance under the Summary Feedback condition, particularly for the low WMC group. In combination with the results of the retention tests, there was a clear benefit for KCR Feedback over Summary feedback for the different task loads and WMC groups.

Subjective Data

Participants were interviewed at the end of the study to determine what they noticed about the feedback and what strategies they used. Almost universally participants reported they thought feedback occurred randomly and did not correspond to specific gates. They also reported that they tended to ignore the Summary Feedback at the end of each block, (however the average time spent looking at the Summary Feedback during acquisition was ~13 seconds and did not differ between WMC groups.) This suggests that neither high nor low WMC participants were aware of the feedback manipulation and may have only paid attention to gates where feedback was present. This might account for the difference between the feedback conditions if participants ignored the gates that received Summary Feedback in favor of the more highly supported KCR gates.

DISCUSSION

The main research question in Experiment 1 was to discover the effects of feedback support on learning including the possible moderating effects of learner working memory capacity and task load. It was expected that learning would vary according to the support of the feedback provided, but this effect would be different for each of the WMC groups depending on task loads. Two competing theories of instruction predicted differing results. If cognitive load theories of instruction were supported, all groups should learn more in the highest feedback support condition: KCR. On the other hand, if challenge increases learning (after decreasing performance in acquisition), in a simple task, high WMC individuals should learn more with less supportive feedback (summary feedback) whereas the opposite should be true for low WMC individuals (KCR). However, in a complex task, high WMC individuals should also benefit most from KCR.

Key findings were that KCR was superior to summary feedback in most measures. High WMC participants demonstrated learning under both feedback conditions, though KCR was usually statistically superior. Low WMC participants not only demonstrated higher accuracies under KCR, at most times they performed no differently than chance under the summary feedback condition. These general findings were true for both levels of task load.

The results of this study support the idea that a reduced load results in more learning (as measured through retention and transfer). This is in accordance with cognitive load theory. When there were differences between WMC groups, it tended to be that high WMC participants were better able to use the KCR support to improve their performance in acquisition, retention, and transfer. Feedback not only supported performance, it was a tool to be used by those able to take advantage of it. This can be captured by the idea that the “rich get richer,” or those who are already capable of learning a task improve the most when given additional aid. Low WMC participants also benefited from the KCR feedback, though not as much as the high WMC participants. When provided with Summary Feedback, low WMC participant performance often did not differ from chance.

Despite the overall finding that KCR helped participants learn the LGT better than Summary Feedback, there were several questions raised by this first study. 1) Did high WMC participants benefit especially from the KCR feedback because they had deficient strategies in place for learning via the Summary Feedback? If they knew explicitly that some gates would receive different types of feedback and that the Summary Feedback would be the only source of information on half the gates, would their accuracies for Summary Feedback gates meet or exceed the KCR gates? 2) Did low WMC participants not benefit from the Summary Feedback because they were unable to learn from the minimal amount of information provided or because they were not allowed enough practice with the feedback in place? 3) If Summary Feedback were so minimal that both WMC groups were not able to judge their own performance, perhaps the feedback conditions of Summary Feedback and KCR did not truly represent “low” and “high” feedback support. It was possible Summary Feedback represented “zero” support and KCR “low” support. These three questions were addressed in the following experiments.

CHAPTER 4: EXPERIMENT 2 - STRATEGY USE & FEEDBACK SUPPORT

Overview

Experiment 1 examined how learner working memory capacity affected the amount of feedback support needed to learn tasks with differing demands. A general result of this study was that participants benefited from higher feedback support no matter what the task load or learner WMC. One possible explanation was that high WMC participants may not have generated strategies for learning the low-support gates. Although high-WMC learners are more likely to spontaneously generate strategies (Hertzog & Robinson, 2005), there was no guarantee that they did so in Experiment 1. If they did not *use* the low-support feedback for those gates it would not be surprising that they showed no benefit for low-support feedback at retention. This would relegate the low support condition to “nonexistent” rather than low.

Remaining Questions

Thus, it was important to determine if the lack of benefit for low-support feedback was due to lack of learning strategy. In Experiment 2, participants were told explicitly which gates would receive low feedback support and which would receive high. They were then told where to look on the Summary Feedback to track their performance in the two feedback conditions. Next, they were instructed to use the Summary Feedback to plan their responses to gates on the next block of trials. Last, they were left with the instruction sheet listing which gates received what feedback. All break-screens emphasized that they should be using the break time to think about the gates for the next block of trials.

The goal of these manipulations was first to make the amount of feedback support transparent to the learner. They would know which gates got which support level and that some gates would have to be learned via only summary feedback. Further, having high working memory span participants learning the simple version of the logic gate task represented the best-case scenario for learning via summary feedback. This study provided insight into whether the results of Experiment 1 represented strategy differences or a feedback design issue.

Hypotheses

It was predicted that, if low Summary Feedback performance in Experiment 1 was due to lack of strategy generation and use by high WMC learners, the participants in this experiment would learn the gates more thoroughly via Summary Feedback than with KCR. If the lacking performance with Summary Feedback in Experiment 1 were due to a deficiency in the feedback itself, then increased strategy use should have little effect in bolstering learning via that feedback compared to KCR.

Method

Participants

Four high WMC young adults (18-35 years of age) were recruited from a database collected by the Attention and Working Memory Laboratory at the Georgia Institute of Technology. Working memory capacity was determined in the same manner as Experiment 1. Participant demographics and characteristics are presented in Table 6. No participants had experience with the LGT (assessed by a screening survey and prior experience questionnaire), however some participants reported experience with programming or Boolean searches. Participants were compensated at the rate of

$10/hour.

Table 6

Participant Characteristics for Experiment 2

	n = 4				
	M	SD		M	SD
General demographics			LGT-specific demographics		
Gender	1 male 3 females		Number of times to pass matching test[8]	1	0
Highest level of education[1]	5.25	0.50	Are you familiar with any of the logic gates?	0	0
Age	21.5	4.51	Have you ever taken a computer science course that required programming in a a computer language?	0.5	0.58
Handedness[2]	1.75	0.50		n = 2	
Near Vision[3]	20	0.00	If yes, have you ever used operators such as AND, OR, XOR, or NAND in programming?	0.5	0.71
Far Vision[3]	12.25	1.50			
Shipley Vocabulary Score[4]	32.25	2.22	Have you ever used operators such as AND, OR, XOR, or NAND in a library or web search?	0.5	0.58
Simple Reaction Time[5]	267.333	31.21	I now feel confident about my ability to solve logic gates. (Start)	5.25	0.96
Choice Reaction Time[5]	296.75	29.77	I now feel confident about my ability to solve logic gates. (End)	3.5	1.91
Digit Symbol Substitution[6]	77.75	11.53	I am motivated to do my best (Start)	6.25	0.96
Reverse Digit Span[6]	9.5	1.29	I was motivated to do my best. (End)	5.25	2.36
Ao-span absolute[7]	60.75	5.62	MIA Strategy Subscale Score	2.66	0.32
Ao-span Total[7]	70.25	2.99			

Note . [1]Education rated on an ordinal scale; available in Appendix A. [2]Handedness scored as 1 for left, 2 for right. [3]Vision scored as Snellen acuity. [4]Shipley, 1986 [5]reaction time in milliseconds [6]Wechsler, 1997a [7]Unsworth, et al., 2005 [8]Criterion test of logic gate rule memorization [9]Dixon, et al., 1988; available in Appendix A.

Materials

Ability Tests

Participants completed the same ability tests as Experiment 1 with the addition of the Metamemory in Adulthood questionnaire (MIA) (Dixon, Hultsch, & Hertzog, 1988) and the Preferred Learning Styles (PLS) questionnaire (Hertzog & Dixon, 1994) Appendix A). These measures were added due to the interest in strategy adoption and use for this experiment.

Experimental Task

The experimental task was identical to Experiment 1 with two changes. First, the gates on the summary screen were re-ordered into columns reflecting whether they were in the KCR or Summary Feedback conditions. Second, "Please remember to pay careful attention to the gates you have had difficulty with. It is very important to use the feedback provided at the end of each block." was added to the 1-minute break screens between blocks.

Feedback Conditions

Feedback support conditions were identical to Experiment 1. Low support was given via summary feedback at the end of each block and high support was given by KCR on every trial. Gates were counterbalanced into these conditions as in Experiment 1.

Instructions

Instruction was identical to Experiment 1 with "explicit" instruction added after practice on the gates. This explicit instruction identified to every participant the gates that would receive KCR Feedback and which would receive Summary Feedback and that the only way to understand and improve performance on the Summary Feedback gates would be to examine the Summary Feedback carefully. Further, they were told a good strategy for improving performance would be to identify problem gates via the Summary Feedback and plan their responses for the next block. They kept the explicit instruction sheet with them during the task (Appendix A).

Questionnaires

All questionnaires may be found in Appendix A.

Procedure

Session One

The procedure was identical to Experiment 1, excepting the additional explicit instructions regarding the feedback conditions. Also, participants completed the additional questionnaires about their strategy use and PLS for the gates prior to the LGT and then a second PLS after the end of the session.

Session Two

The procedure was identical to Experiment 1. Participants completed the retention test, content and Load Transfer tests. Last, participants filled out the exit interview and received verbal/written debriefing and payment for their time.

Design

The experiment examined differences for the levels of the within-participant variable of Feedback Condition (2: Summary Feedback, KCR). Participants all performed the simple LGT on the assumption that high WMC participants would best be able to make use of the Summary Feedback for a low load task. Dependent measures were identical to those in Experiment 1.

Results

The main question for this study was whether the high WMC participants used different strategies for the different feedback types and whether explicit strategy use would result in higher retention of Summary Feedback gates. If high WMC participants did not use the Summary Feedback in Experiment 1, it would explain why their performance suffered at retention and transfer. Because they were given strategies and instruction for the current study, their Summary Feedback performance should improve and perhaps exceed KCR performance at retention and transfer in Experiment 2. However, if participants were unable to use the Summary Feedback in Experiment 1 because it did not contain enough information, then performance in this study should be similar to Experiment 1. Because of the small number of participants ($n = 4$), error bars are not included on graphs and only trends reported in the performance data. Means are provided for each learning phase in Table 7.

Table 7

Participant subjective experience and motivation for Experiment 2

	Questions							
	If you had to pick one and only one strategy to learn the logic gate task, what would it be?		I am motivated to do my best (1-7, where 1 was "not motivated at all" and 7 was "extremely motivated")		What strategy did you use for learning the no-feedback gates?	How did you use the end of block feedback, with the summary of each gate percent correct, to learn	Did you feel able to use the end of block feedback?
Participant	Pre-acquisition	Post-acquisition	Session 1	Session 2			
1	Attentive reading	Attentive reading	6	7	I would examine the feedback at the end of the set.	I examined this feedback more closely than the feedback I got during the trials.	I felt it was more useful than the feedback during the trials. It gave me time to examine all the gates I had trouble with. The feedback I got during the trials tended to frustrate me. I think the end of trial feedback improved my accuracy much more than the every trial feedback.
2	Rote repetition	Focal attention	7	2	I didn't.	I didn't.	No! Too overwhelmed by lack of time to think at all.
3	Rote repetition	Rote repetition	7	7	I tried to remember everything using ones. So I made a rule up for each one dealing with the number 1 and what it would equal.	I tried to remember the rules and figure out different inputs and answers for each gate I did poorly on.	Yes, because I knew which ones I needed to try and remember more.
4	Focal attention	Focal attention	5	5	During the break, I would constantly re-remind myself of the rules of those I missed.	I only looked at the ones I missed briefly and tried to memorize the rules again.	Not too much help. Only helped with the no-feedback rule.

Strategy Use

Strategy use was measured using the strategy subscale of the Metamemory in Adulthood (MIA) questionnaire (Dixon et al., 1988). The results were compared to means from a sample of 476 participants of the same age group (Ponds & Jolles, 1996). The high WMC participants in this experiment reported a mean score of 2.66 (SD=.32) on a 5-point Likert scale where a higher score indicated more use of strategies. This is lower than the average reported by Dixon, et al. One possible explanations is the high WMC of these participants: they may not need memory strategies as much as a normal population and therefore report less use.

Participants indicated their preferred learning strategies (Hertzog & Dixon, 1994) for learning the LGT before acquisition and again in the 10-minute interval before the first retention test. Rote repetition was one of the least effective memory strategies listed (Trow, 1928) and two of the four participants reported a preference for rote repetition in learning the LGT, and one of those maintained that preference at the end of acquisition (Table 8).

Table 8

Accuracy Levels for Experiment 2 divided by Feedback Condition

Beginning of Acquisition				End of Acquisition				10-minute retention			
Mean accuracy		Standard deviation		Mean accuracy		Standard deviation		Mean accuracy		Standard deviation	
End of block	Every trial	End of block	Every trial	End of block	Every trial	End of block	Every trial	End of block	Every trial	End of block	Every trial
0.45	0.50	0.17	0.16	0.78	0.73	0.32	0.10	0.85	0.90	0.24	0.08

1-week retention				Content transfer				Load transfer			
Mean accuracy		Standard deviation		Mean accuracy		Standard deviation		Mean accuracy		Standard deviation	
End of block	Every trial	End of block	Every trial	End of block	Every trial	End of block	Every trial	End of block	Every trial	End of block	Every trial
0.76	0.79	0.15	0.13	0.72	0.74	0.13	0.15	0.67	0.69	0.13	0.11

Performance and Learning

Graphs of participant performance may be found in Figures 13, 14, & 15. These data were most informative when participants were examined individually (Figure 16). The next section provides a profile of each participant, performance data, and their qualitative comments concerning strategy use and preferred feedback condition.

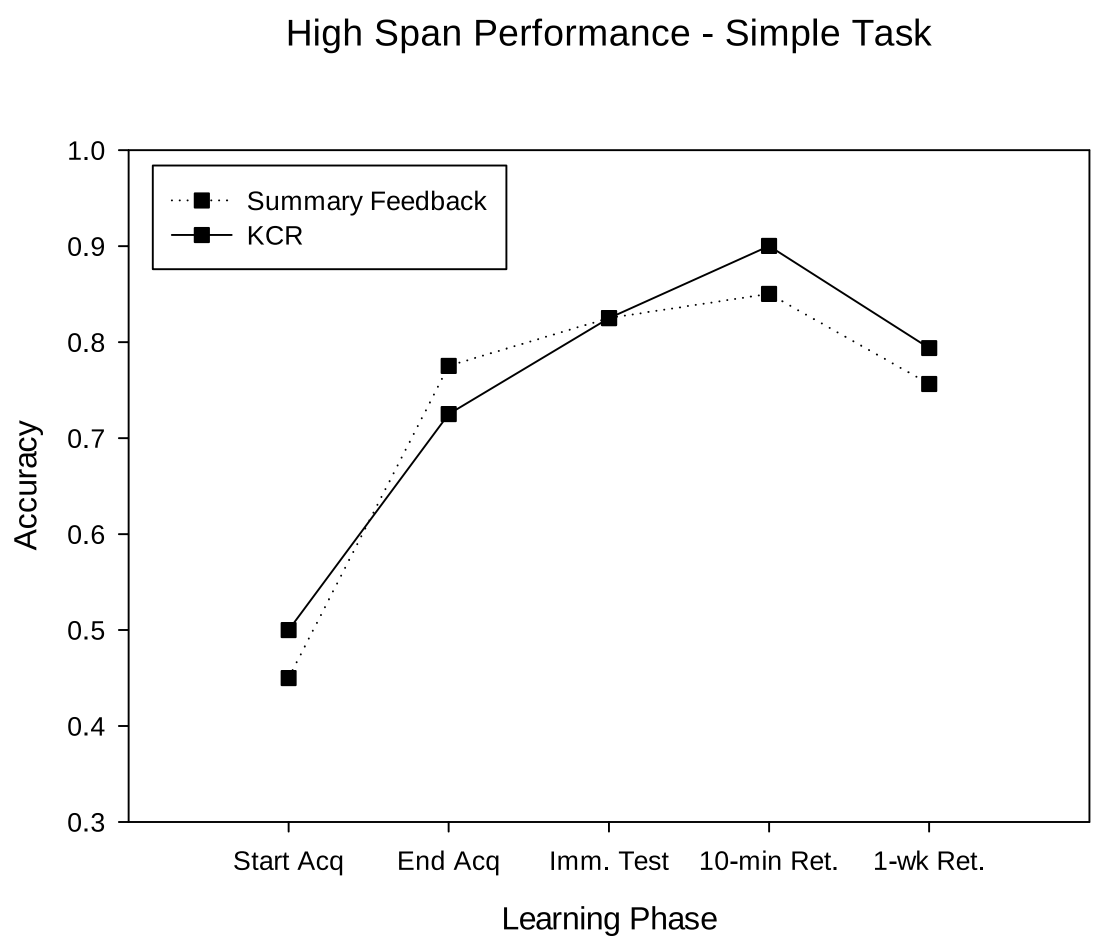


Figure 13. Experiment 2 learning phases.

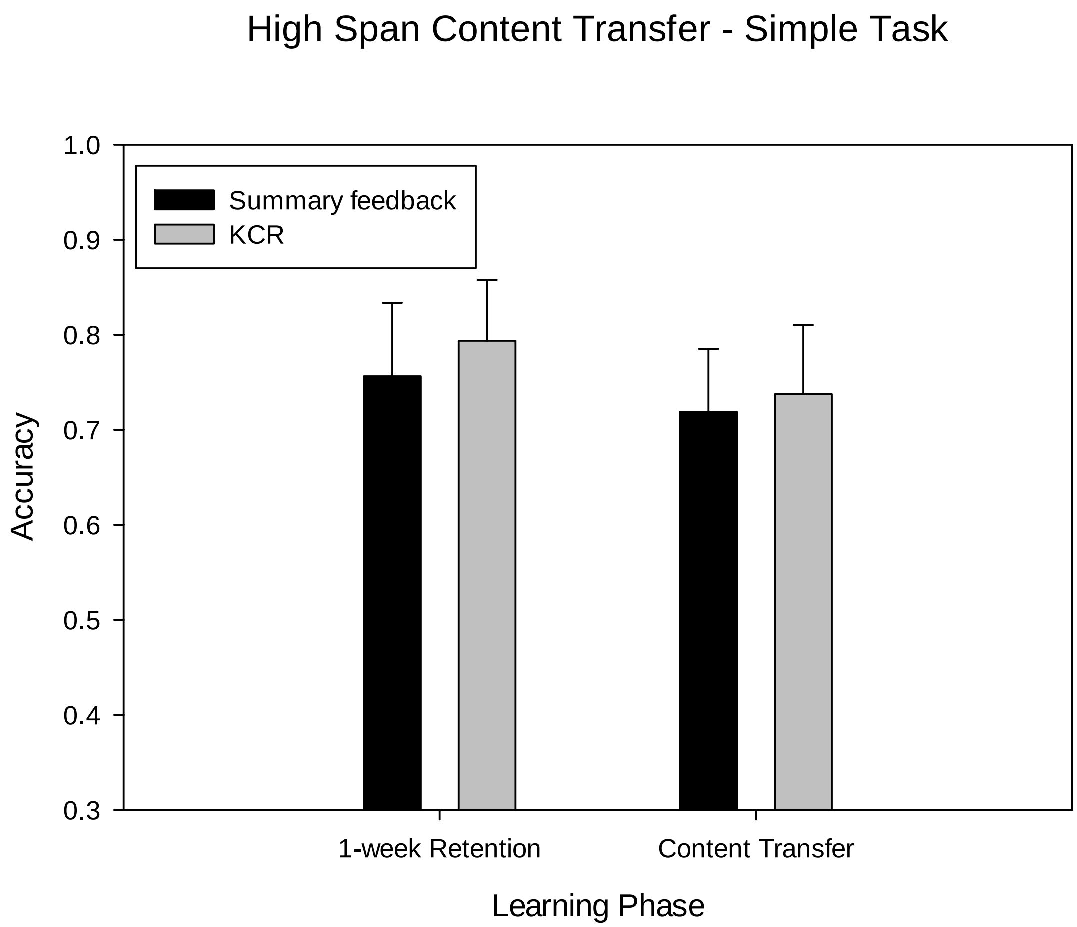


Figure 14. Experiment 2 1-week Retention and Content Transfer Test accuracies

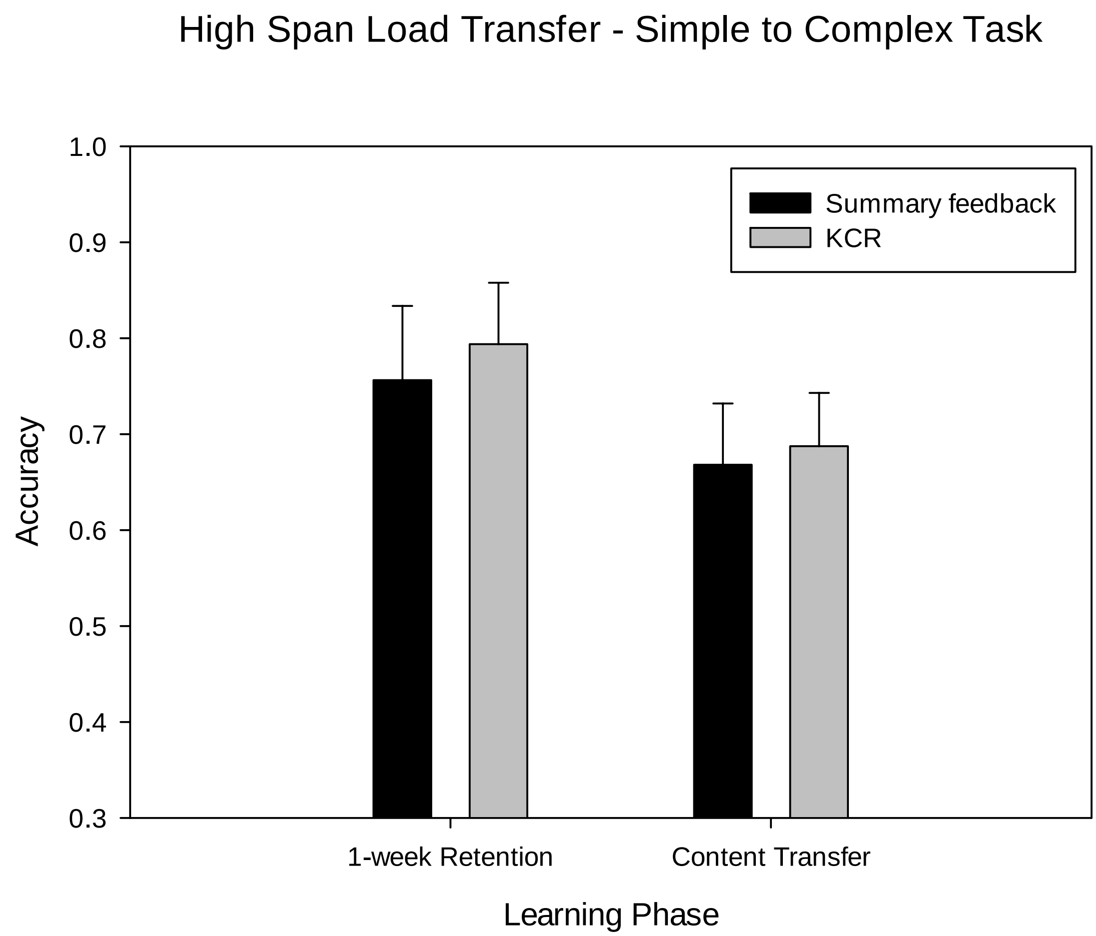
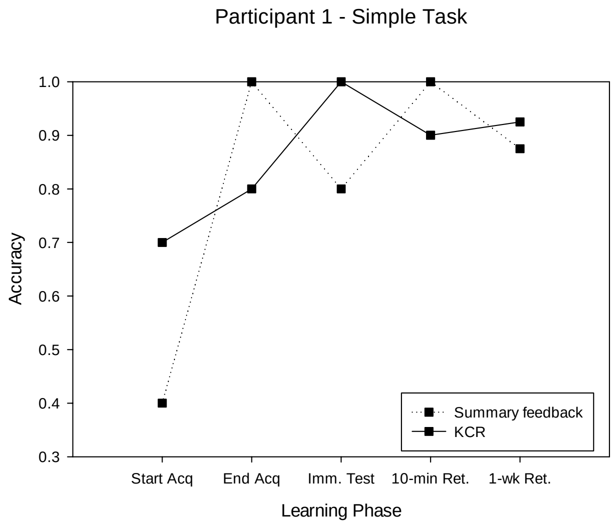
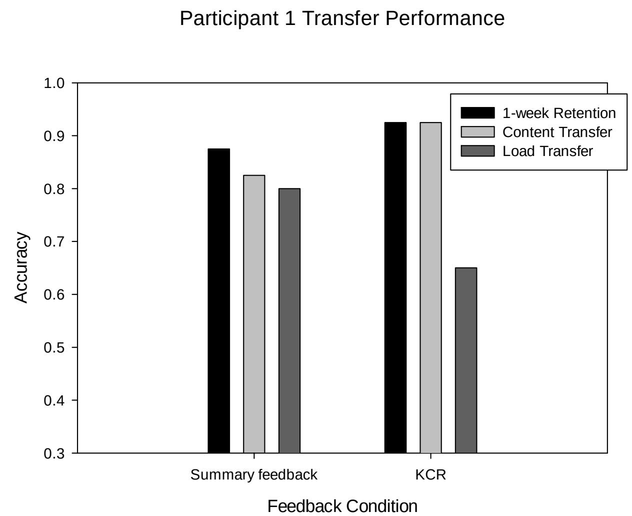
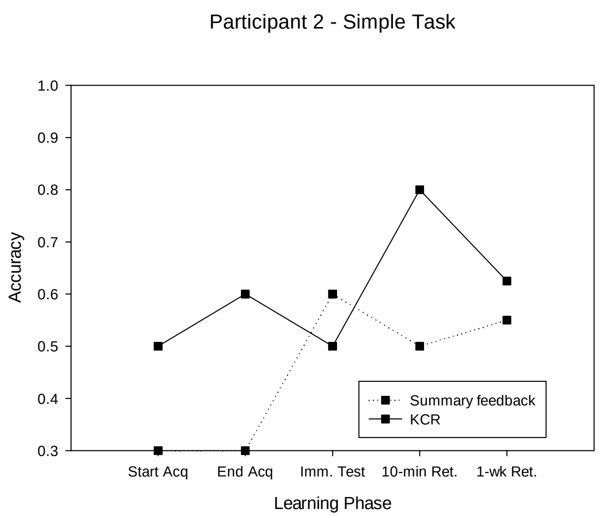
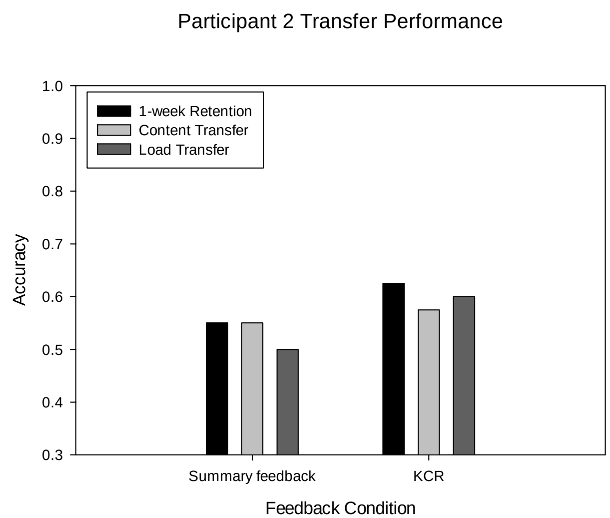


Figure 15. Experiment 2 1-week Retention and Load Transfer Test performance.


Participant 1 - Simple Task
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Start Acq
End Acq
Imm. Test
10-min Ret.
1-wk Ret.
Learning Phase
Summary feedback
KCR


Participant 1 Transfer Performance
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Summary feedback
KCR
Feedback Condition
1-week Retention
Content Transfer
Load Transfer


Participant 2 - Simple Task
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Start Acq
End Acq
Imm. Test
10-min Ret.
1-wk Ret.
Learning Phase
Summary feedback
KCR


Participant 2 Transfer Performance
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Summary feedback
KCR
Feedback Condition
1-week Retention
Content Transfer
Load Transfer

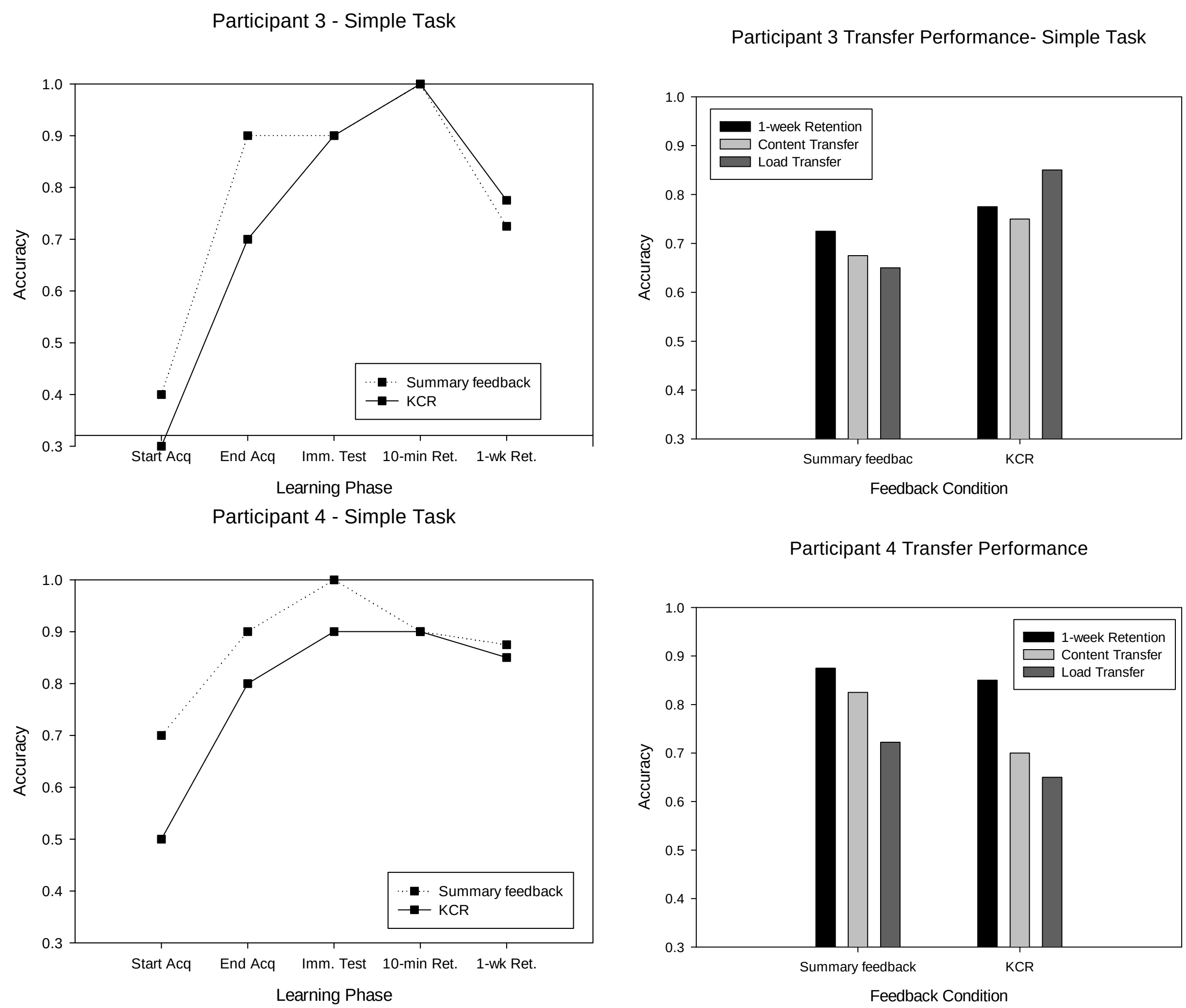


Figure 16. Individual participant performances, Experiment 2.

Individual Participants

See Table 7 for responses. In general, even participants who specifically noted that summary feedback was more helpful to them did not demonstrate higher performance under that condition. This was true in acquisition performance as well as retention and on transfer tests.

Discussion

In general, the four participants in this study were highly motivated and reported putting forth more effort to learn gates supplied with Summary Feedback. However, neither this motivation nor effort resulted in higher accuracies for those gates, either at retention or on transfer tests. The overall pattern of results was similar to the high WMC/simple task data from Experiment 1. For these four highly motivated high capacity participants the summary feedback was still not superior to KCR. One possible explanation was concluded that Summary Feedback did not provide enough information for learning. This question will resurface in Experiment 4.

CHAPTER 5: EXPERIMENT 3 - PRACTICE EFFECTS

Overview

Thus far data from two experiments suggest that the summary feedback condition was insufficient for learning, particularly for low WMC participants. In Experiment 2, high WMC learners demonstrated similar learning via the summary feedback, but even with strategies in place they did not exceed their KCR performance. The low WMC participant results from Experiment 1 suggested the summary feedback produced no learning at all: their performance was often no different than chance on summary feedback gates.

Remaining Questions

It was possible that low WMC participants were not given enough exposure in the first experiment; it was not that they were unable to learn gates via summary feedback, they just needed longer practice than the high WMC group. This would mirror results found in the aging literature where older adults improved task performance equal to that of younger adults, given more practice time (Fisk, McGee, & Giambra, 1988). The main question for the third experiment was whether low WMC participants were unable to learn the complex task due to the feedback provided or due to lack of sufficient exposure.

Hypotheses

It was hypothesized that if the chance performance exhibited by low WMC participants in experiment 1 were due to a deficiency in the summary feedback, the low WMC participants in this study would demonstrate no learning of those gates even with extensive practice. If, however, the chance performance of the low WMC participants in

Experiment 1 were due insufficient acquisition time, they would demonstrate performance and retention of the summary feedback gates in the current experiment.

Method

Participants

All eighteen of the low WMC participants from Experiment 1 were invited to participate in the current experiment. Ten low WMC participants agreed to return to acquire the logic gate task. One of these ten participants declined to return for the 1-week Retention session due to lack of available time. Each participant completed Experiment 1 in its entirety and thus had prior exposure to the LGT. Because of the time commitment required for this experiment, participants were compensated at the rate of $15.00/hour. There was no known systematic difference between those who returned and those who declined to participate based on their demographics information and ability test scores (Table 9).

Task and Feedback Conditions

The task and feedback conditions were identical to Experiment 1.

Procedure

Participants completed 10 sessions of acquisition, each the same length as the acquisition session from Experiment 1. At the end of each session they completed the 10-minute retention block. As in Experiment 1, the interval was filled (with questionnaires not directly concerning the operation of the logic gates). Participants completed two sessions per day with a mandatory break of at least thirty minutes between sessions. The two sessions and thirty minute break comprised about two hours each day. Participants returned for five days.

Table 9

Participant Characteristics for Experiment 3

	n = 10				
	M	SD		M	SD
General demographics			LGT-specific demographics		
Gender	6 males	4 females	Number of times to pass matching test[8]	2.2	0.79
Highest level of education[1]	5	1.5811	Are you familiar with any of the logic gates?	0.2	0.42
Age	25.13	4.1897	Have you ever taken a computer science course that required programming in a a computer language?	0.1	0.32
Handedness[2]	1.8	0.4216			
Near Vision[3]	20	0	If yes, have you ever used operators such as AND, OR, XOR, or NAND in programming?	n = 1 1	0.00
Far Vision[3]	14.1	2.2336			
Shipley Vocabulary Score[4]	28.4	4.9261	Have you ever used operators such as AND, OR, XOR, or NAND in a library or web search?	0.3	0.48
Simple Reaction Time[5]	319.4	69.83	I now feel confident about my ability to solve logic gates. (Start)	3.2	1.14
Choice Reaction Time[5]	344.4	61.437	I now feel confident about my ability to solve logic gates. (End)	1.3	0.48
Digit Symbol Substitution[6]	61.1	20.262	I am motivated to do my best (Start)	2.2	2.04
Reverse Digit Span[6]	7.2	2.2509	I was motivated to do my best. (End)	1.9	1.60
Ao-span absolute[7]	19.6	7.8202	MIA Strategy Subscale Score	2.21	0.49
Ao-span Total[7]	39.7	11.026			

Note . Education rated on an ordinal scale; available in Appendix A. Handedness scored as 1 for left, 2 for right. Vision scored as Snellen acuity. [4]Shipley, 1986 [5]reaction time in milliseconds [6]Wechsler, 1997a [7]Unsworth, et al., 2005 [8]Criterion test of logic gate rule memorization [9]Dixon, et al., 1988; available in Appendix A.

Results

Acquisition

Beginning and End of Acquisition

Contrasts and means for all learning phases are available in Table 10. The first comparison was between performance on Block 1 and Block 120 (the beginning and end of acquisition). It was anticipated that there would be a Block x Feedback interaction, as the feedback conditions should start off similarly and then KCR would produce significantly more accurate performance at the end of the last acquisition session on the fifth day (Figure 17).

Feedback effects and interactions. There was a main effect of Feedback in the expected direction, where KCR produced significantly more accurate performance than did the Summary Feedback, $F(9,1) = 18.57$, $p = .002$, $np^2 = .67$. However, this did not interact with Block.

Other effects. There was a significant difference in accuracy between the beginning and end of acquisition, as would be expected, $F(9,1) = 15.74$, $p = .003$, $np^2 = .64$.

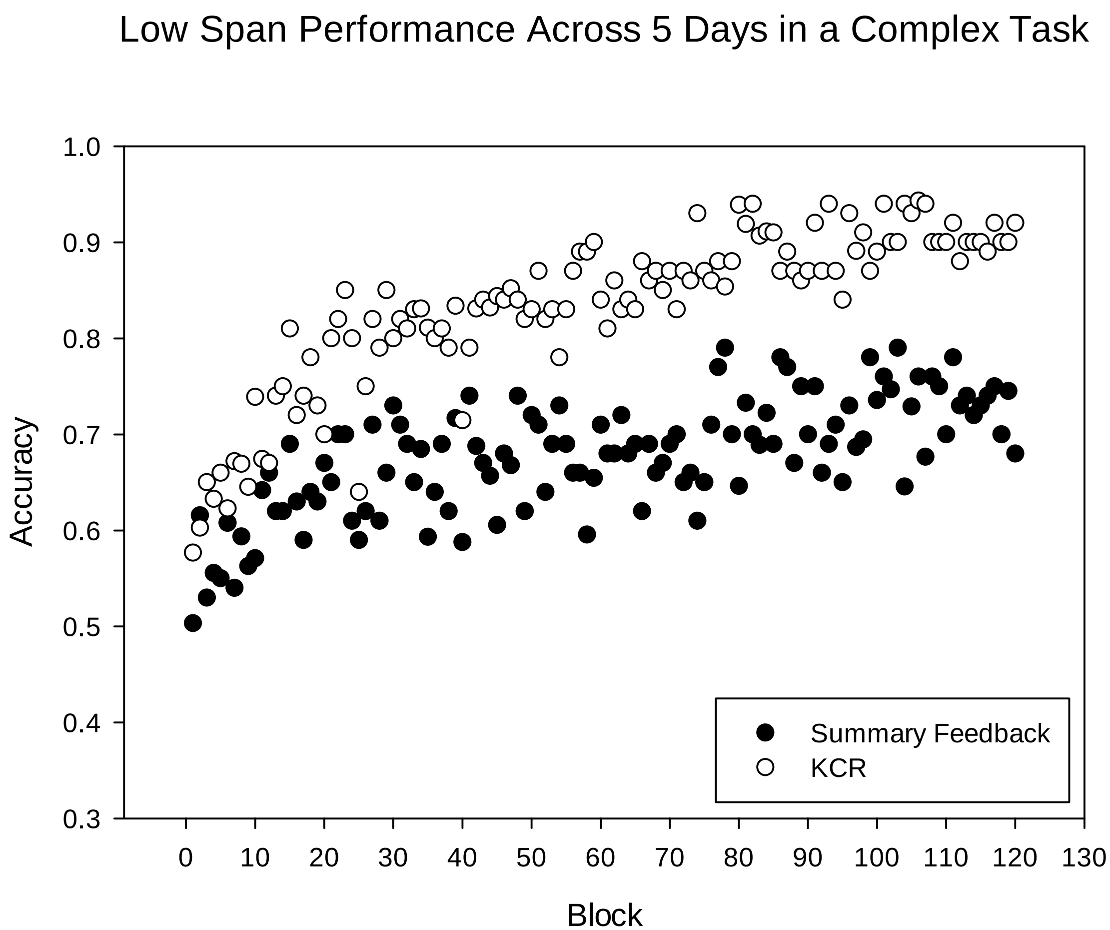


Figure 17. Experiment 3 – Scatterplot of individual performance across 10 sessions for 10 participants.

Table 10

Accuracy Levels for Experiment 3 Divided by Feedback Condition

Beginning of Acquisition					End of Acquisition					10-minute retention				
Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*
End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial	
0.50	0.58	0.20	0.26	-1.14	0.57	0.74	0.20	0.19	-4.40*	0.68	0.92	0.26	0.11	-4.79*

1-week retention					Content transfer					Load transfer				
Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*
End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial		End of block	Every trial	End of block	Every trial	
0.68	0.91	0.29	0.09	-2.87*	0.66	0.84	0.23	0.11	-3.41*	0.72	0.85	0.23	0.13	-1.68

Note. * indicates significant differences between groups ($p < .05$).

Retention

10-minute Retention

Feedback effects and interactions. When comparing the 10-minute retention test from the first session to the 10-minute retention test on the last session, 5 days later, there were effects of Block, $F(9,1) = 6.78$, $p = .029$, $np^2 = .43$, where scores were higher on Day 5, effects of feedback, $F(9,1) = 6.21$, $p = .034$, .41, where KCR was more accurate than Summary Feedback, and an interaction of Block x feedback, where KCR started off higher on Day 1 and improved more by end of acquisition than did Summary Feedback, $F(9,1) = 10.45$, $p = .010$, $np^2 = .54$. These results were similar to the findings from Experiment 1 for the high WMC participants: performance in both feedback conditions was significantly different from chance (Figure 18).

1-week Retention Post Score

After a 1-week Retention interval, there was still an effect of Feedback where KCR gates were more accurate than Summary Feedback gates, $F(8,1) = 8.23$, $p = .021$, $np^2 = .51$.

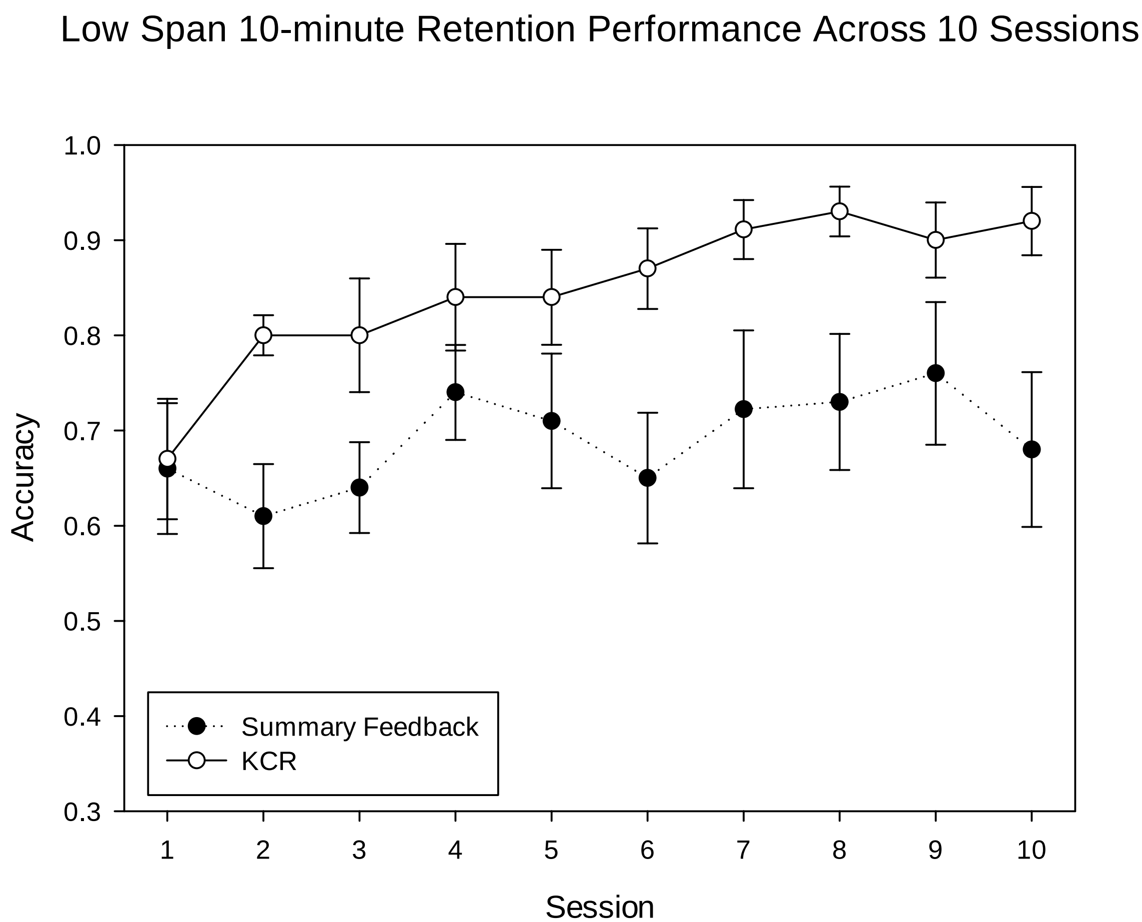


Figure 18. Experiment 3 – Mean performance at 10-minute retention interval across 10 sessions.

1-week Retention Change Score

There were no significant effects in the percent change score from 10-minute retention on Day 5 to 1-week Retention. The lack of effect comes from the impressive retention demonstrated by these groups after 5 days of practice. Their performance did not change more or less with either feedback condition and was not significantly different from zero (Figure 19).

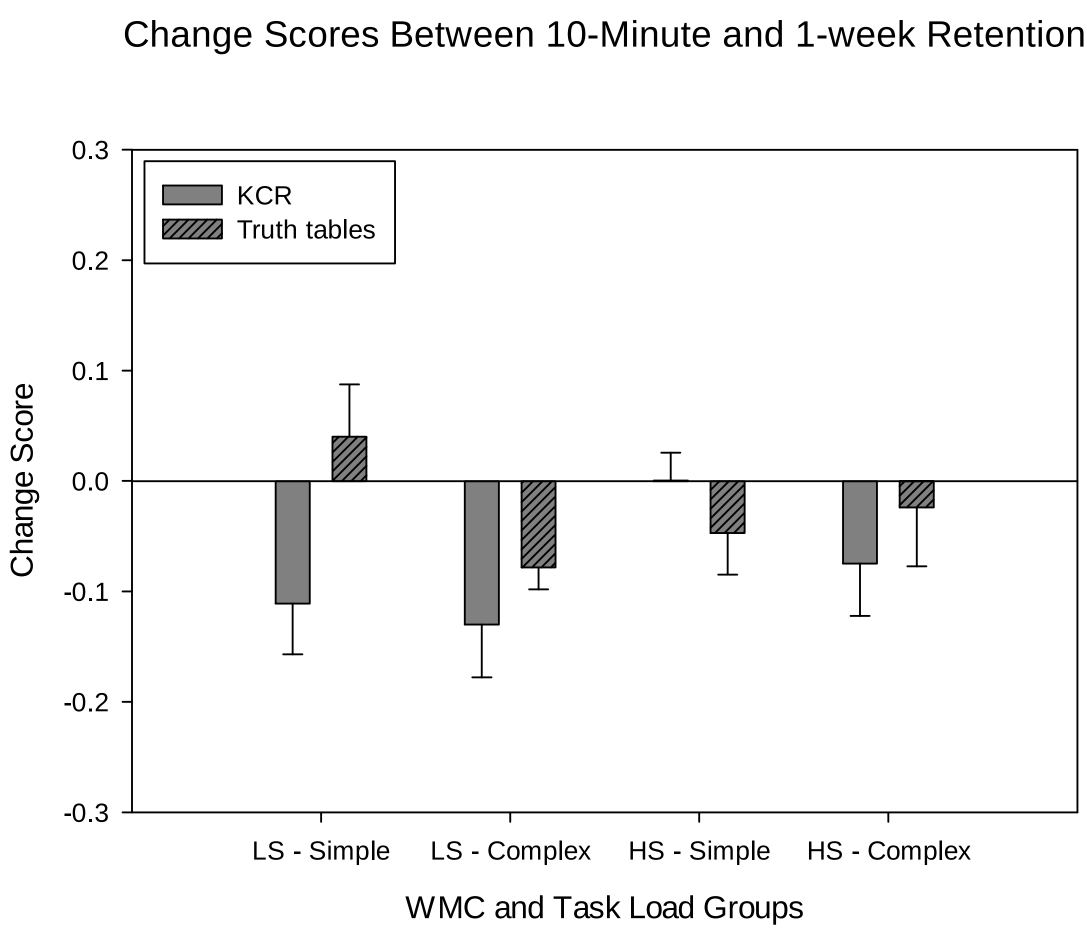


Figure 19. Experiment 3 Change Scores between 10-minute Retention and 1-week Retention.

Transfer

Content Transfer Post Score

Feedback effects and interactions. There was a main effect of Feedback on Content Transfer where participants were more accurate with the KCR gates than on those which had received Summary Feedback, $F(8,1) = 11.67$, $p = .009$, $np^2 = .59$, (Figure 20).

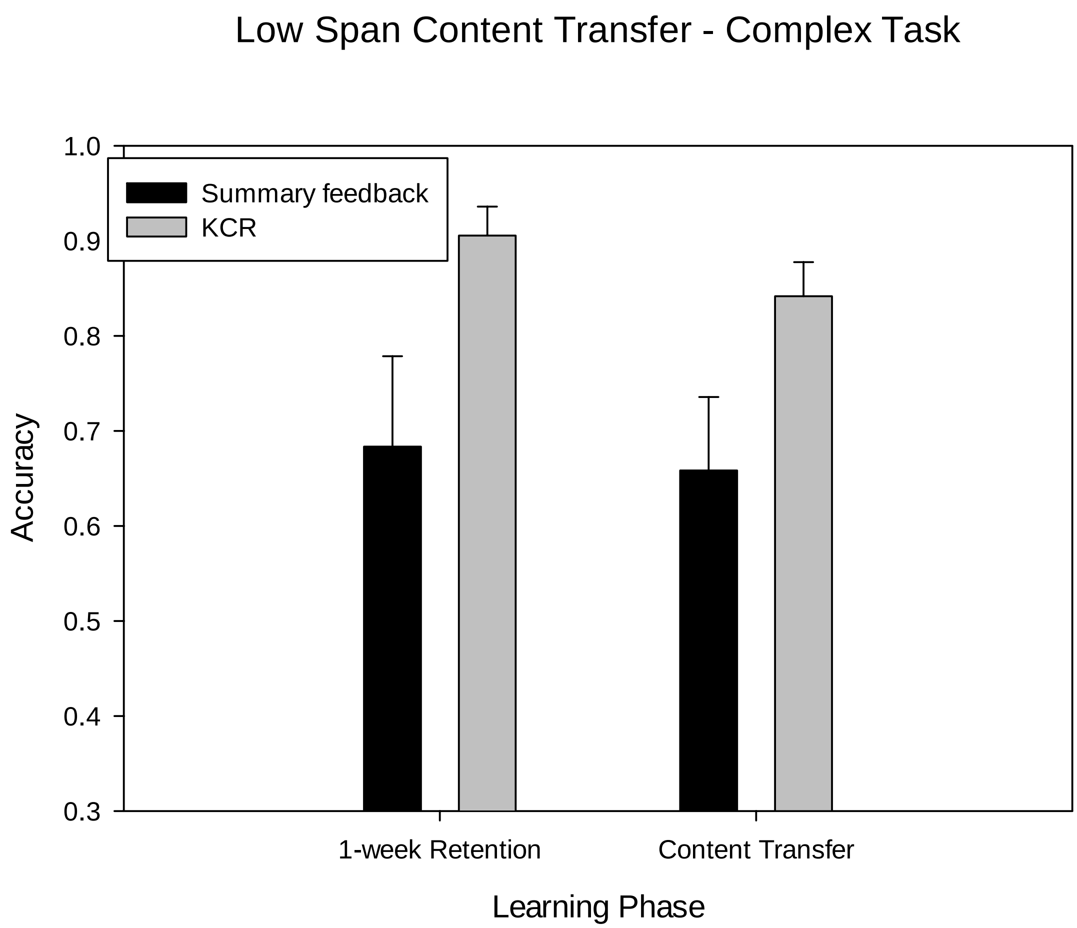


Figure 20. Experiment 3 1-week Retention and Content Transfer Test performance.

Content Change Score

As with the other change scores, there were no effects of feedback (Figure 21).

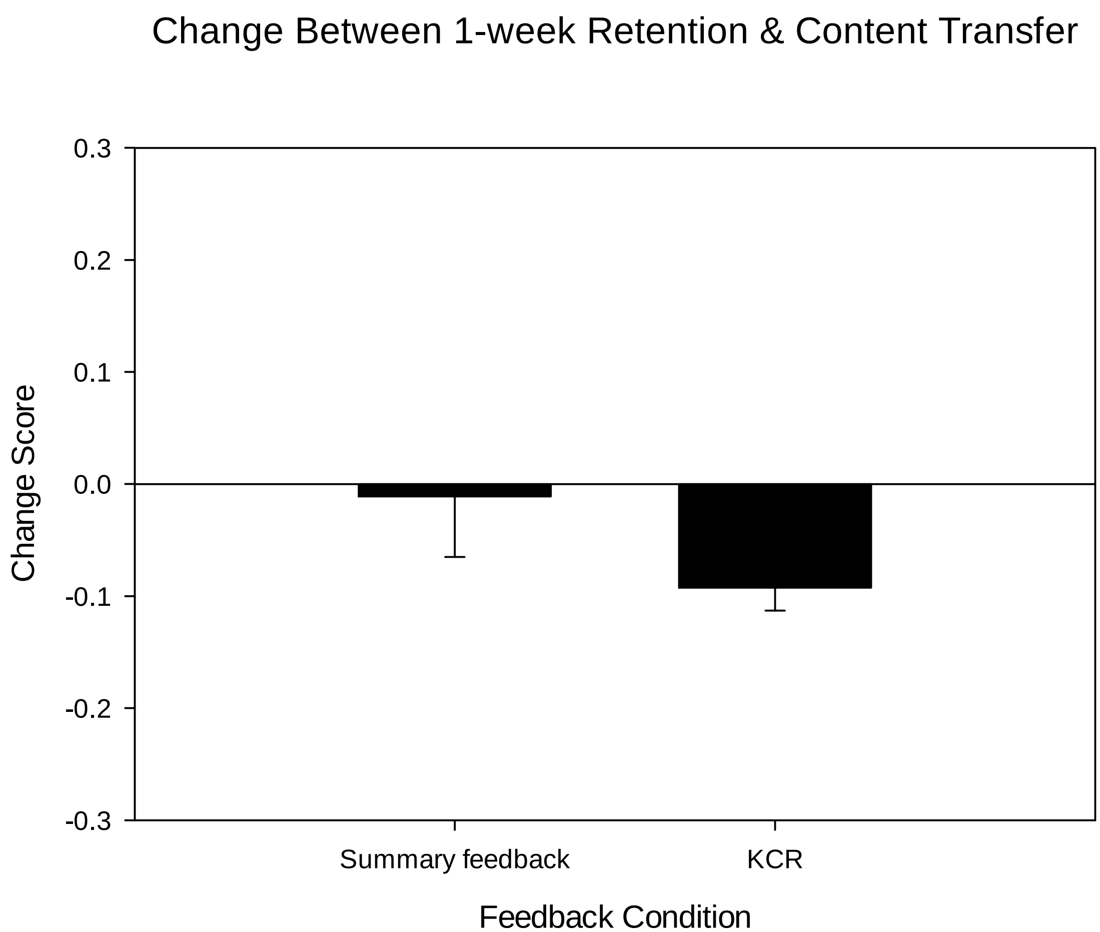


Figure 21. Experiment 3 Change Scores between 1-week Retention and Content Transfer Test.

Load Transfer Post Score

There were no effects found between feedback conditions for the Load Transfer test (Figure 22).

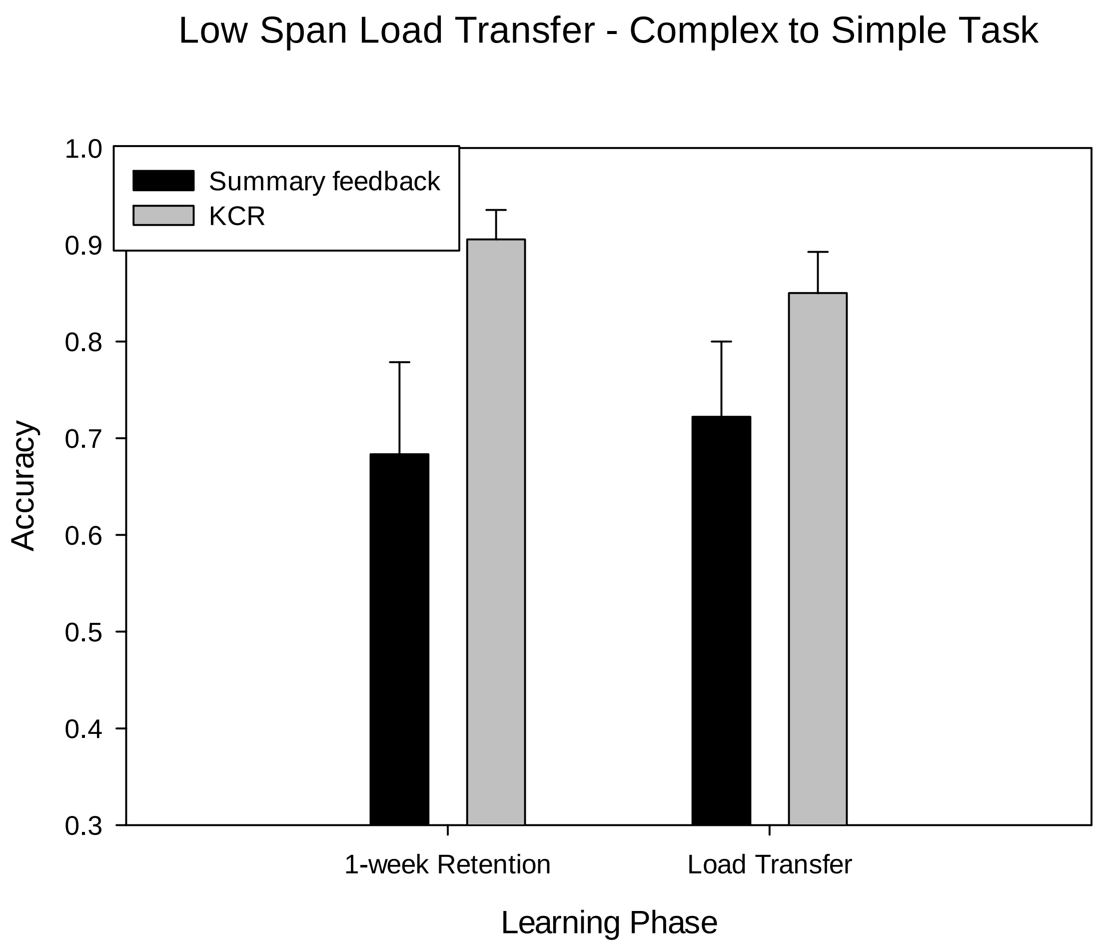


Figure 22. Experiment 3 1-week Retention and Load Transfer Test score.

Load Transfer Change Score

There was no change in performance for the either the Summary Feedback gates or the KCR gates upon switching the task from complex to simple (Figure 23).

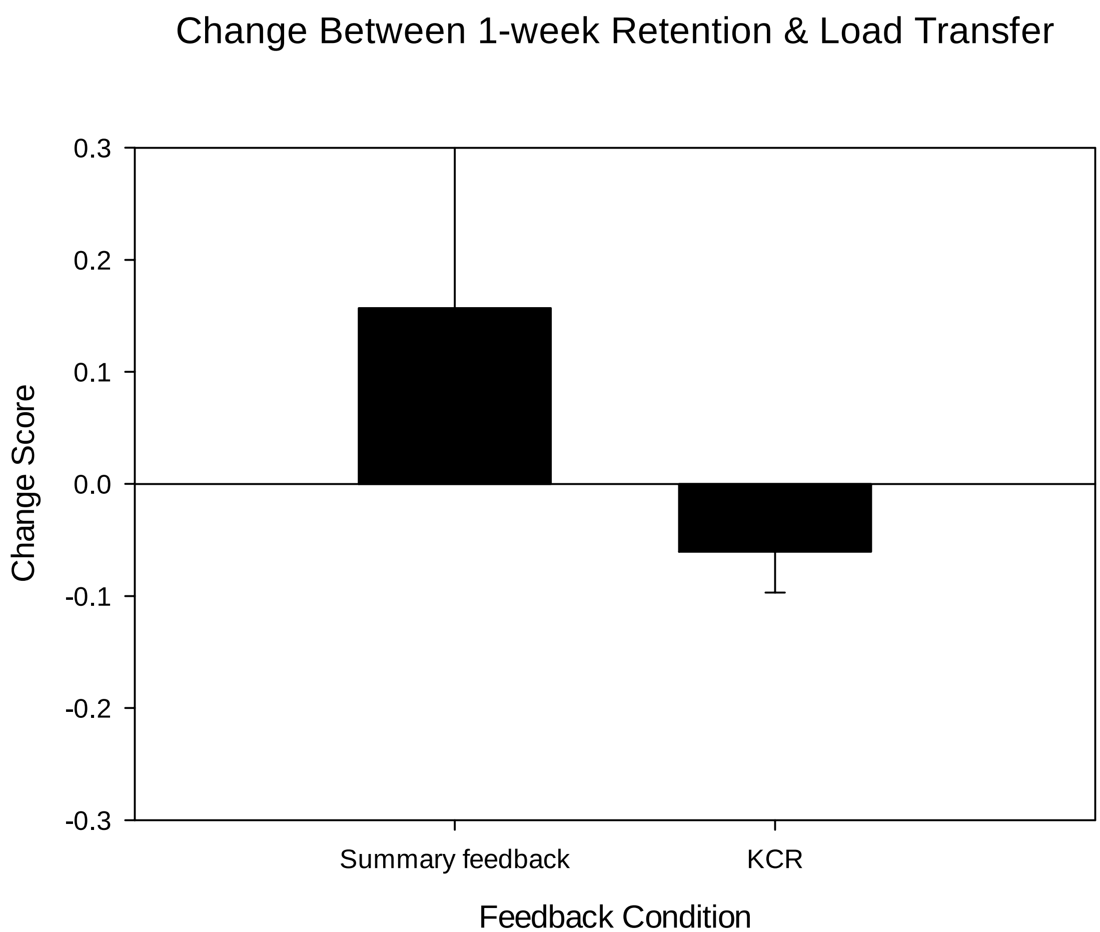


Figure 23. Experiment 3 Change Scores between 1-week Retention and Load Transfer Test.

Subjective Data

Participants were asked about their strategy use for the summary feedback gates during the exit interview. Though most reported having a strategy for learning the gates, there was little indication of in-depth or effective strategy use (Table 11). However, participants did retain the KCR gates. Thus, the increased support offered by KCR produced learning even in the absence of well-developed strategies.

Table 11

Low WMC Participant Responses After 10 Acquisition Sessions

What strategy did you use for learning the no-feedback gates? Please describe.	By looking at the shape of the gate and refering back to a "jingle" I used to help me memorize the gate rules that were given to me at the beginning of experiment by remembering the gates, and what they stood for I didn't have a strategy, if I got 'em wrong then that was it. But I think I got most of them right. I keyed in on my scores per each area upon playing w/ the gates I was unsure about, I use the strategy of I tried to do the opposite of what I would do for the ones that received feedback. I would quickly re-do it in my mind and see whether it was correct or not. Remembering what the last correct answer that was given thruogh the feedback gates. trial and error and also comparing one I forgot how to solve with its companion gate XNOR was the one that did not give any feedback, so I could no do anything. What I try to do, was to use the
How did you use the end of block feedback, with the summary of each gate percent correct to learn the gates?	by looking over what I got wrong I focus on the ones that I got wrong, and try to put a different number the next time I was working on those ones. I tried to remember wether or not the overall gate result was. If I was absolute about a given gate, I merely used I tried to remember what I put for the previous questions that made it incorrect. I also tried to remember the rule exactly, and if I could not, I would think about the rules for its opposite and do the opposite of that for the next I work harder on the ones with the lowest percent. It would show me where I stood and either motivate me or otherwise. I looked at the ones I needed to be careful Made sure to picture + remember the gate + what it looked like. Then I would remember the rule for that gate and find a way in the definition to link it to the answer That however I memorized the rule of the gate was wrong so when I received the feedback, I altered my answer and was that the percentages start to go up the gates that did not receive 100% accuracy got more attention to figure out which one(s) were wrong

Discussion

Experiment 3 was designed to examine the effects of feedback on extended acquisition for low WMC learners. Even for a complex task, low WMC participants improved their performance and demonstrated retention after one week when given supportive feedback. Eleven sessions of acquisition resulted in ~10% increase for low WMC participant retention accuracy for a complex task when compared to that same group in Experiment 1. Therefore, given enough practice, low WMC participants could perform the complex logic gates task. A general conclusion was that learning may be supported through extended feedback as well as through more supportive feedback. What remained to be discovered was whether this learning would continue to increase with higher feedback support.

CHAPTER 6: EXPERIMENT 4 - INCREASED FEEDBACK SUPPORT

Overview

Experiment 2 made strategies for using low feedback support explicit and yet participants still did not appear able to use the summary information to learn the LGT. In Experiment 3, ten sessions of acquisition were adequate to promote learning of the KCR gates, but even after such extensive practice performance did not greatly improve on the summary feedback gates. The current experiment was designed to determine performance differences due to two informative feedback conditions, on the assumption that summary feedback in the previous experiments provided too little information to the learner.

Remaining Questions

It was possible that the operational definitions of feedback "support" in Experiments 1, 2, and 3 were inaccurate. What was called "high" feedback support may have been low and "low" feedback support entirely insufficient for learning.

It is useful at this point to revisit one of the original reasons less feedback support might promote retention and transfer: transfer appropriate processing. Consider how the feedback from Experiment 1, 2, and 3 might have been used by the learner. Both summary feedback and KCR forced a recall of the gate rule prior to receiving feedback. However, KCR forced recall of the gate rule and answer and then allowed inference of the correct answer. This is also required at retention and for transfer.

However, each KCR data point only provided one instance of gate inputs and the correct answer. This allowed the participant to infer the other three conditions and recall the answer for other input-output combinations. The next step was to design feedback that supplied all of the information that the participant was forced to infer from KCR feedback. Thus, in this fourth experiment a feedback support condition was designed to provide higher support than KCR. This feedback condition consisted of a truth table with all possible inputs and outputs for a gate.

Hypotheses

It was hypothesized that if two feedback conditions with informative feedback were supplied, that which gave the highest support (truth table) for performance would result in the most learning by low WMC participants under any task load. High WMC participants should also benefit from additional support for a complex task, however "desirable difficulties" would predict their retention performance should benefit the most from low support (KCR) on a simple task.

Method

Participants

Twenty-eight high WMC young adults (18-35 years of age) and twenty-eight low WMC young adults were recruited from the community, the Georgia Tech participant pool, and from a database collected by the Attention and Working Memory Laboratory. Working memory capacity was determined as described for Experiment 1. Participant demographics and characteristics are presented in Table 12. No participants had experience with logic gates (assessed by a screening survey and prior experience questionnaire.) One low WMC participant was excluded due to inability to complete the task and was replaced to have equal numbers of participants in each cell. All participants were compensated for their time at the rate of $10/hour.

Table 12

Participant Characteristics for Experiment 4

	Low span n = 28					High span n = 28					
	Simple Task		Complex Task			Simple Task		Complex Task			
	M	SD	M	SD	*t*	M	SD	M	SD	*t*	*t*
						General demographics					
Gender	5 males 9 females		5 males 9 females			7 males 7 females		7 males 7 females			
Highest level of education[1]	5.29	0.83	5.38	0.77	-0.32	5.50	0.94	5.07	1.07	1.13	0.19
Age	26.21	4.89	23.71	4.70	1.38	24.93	3.56	22.64	4.81	1.43	0.96
Handedness[2]	1.92	0.28	1.85	0.38	0.59	1.85	0.38	1.79	0.43	0.39	0.70
Near Vision[3]	22.31	5.99	21.54	3.76	0.40	23.57	4.97	20.71	2.67	1.89	-0.18
Far Vision[3]	19.54	7.47	19.93	7.32	0.89	23.07	7.43	16.93	3.81	2.75*	-0.14
Shipley Vocabulary Score[4]	27.43	3.39	26.86	3.23	0.46	29.50	4.00	28.71	8.43	0.32	-1.43
Simple Reaction Time[5]	295.07	60.85	289.43	57.62	0.25	273.50	54.39	260.79	32.55	0.75	1.81
Choice Reaction Time[5]	328.21	47.79	337.14	50.70	-0.48	319.93	32.63	305.57	33.40	1.15	1.79
Digit Symbol Substitution[6]	65.93	16.89	66.00	12.31	-0.01	71.00	8.12	71.93	9.08	-0.29	-1.73
Reverse Digit Span[6]	6.50	1.99	7.21	2.29	-0.88	9.29	2.20	9.71	2.61	-0.47	-4.37*
Ao-span absolute[7]	12.86	8.38	14.43	8.60	-0.49	59.57	6.38	59.21	4.76	0.17	-24.13*
Ao-span Total[7]	34.29	16.41	34.00	14.04	0.05	68.79	4.69	68.50	3.88	0.18	-11.73*
						LGT-specific demographics					
Number of times to pass matching test[8]	2.43	1.16	2.14	1.03	0.69	1.57	0.65	1.36	0.63	0.89	3.46*
Are you familiar with any of the logic gates?	0.14	0.36	0.07	0.27	0.59	0.00	0.00	0.14	0.36	-1.47	0.46
Have you ever taken a computer science course that required programming in a a computer language?	0.07	0.27	0.07	0.27	0.00	0.21	0.43	0.50	0.52	-1.59	-2.73*
	n = 1		n = 1			n = 3		n = 8			
If yes, have you ever used operators such as AND, OR, XOR, or NAND in programming?	1.00	0.00	1.00	0.00	--	0.67	0.58	0.75	0.46	-0.25	0.78
Have you ever used operators such as AND, OR, XOR, or NAND in a library or web search?	0.43	0.51	0.36	0.50	0.37	0.50	0.52	0.86	0.36	-2.11*	-2.20*
I now feel confident about my ability to solve logic gates. (Start)	4.29	1.27	4.36	1.74	-0.12	5.57	1.50	4.93	1.49	1.14	-2.32*
I now feel confident about my ability to solve logic gates. (End)	4.00	1.41	3.57	1.83	0.69	3.71	1.49	4.29	2.02	-0.85	-0.47
I am motivated to do my best (Start)	6.14	1.61	6.07	0.92	0.14	6.64	5.93	1.14	2.02	2.05*	-0.59
I was motivated to do my best. (End)	4.14	2.63	5.50	1.70	-1.62	1.71	0.91	4.50	2.14	-4.48*	2.90*

Materials

Ability Tests

Participants completed the same ability tests as Experiment 2 (Appendix A).

Experimental Task

The experimental task was identical to Experiments 1 and 2 with changes to the feedback support conditions.

Feedback Conditions

What had been "high support" in Experiments 1 and 2 was maintained and an additional "extremely high support" condition was introduced. This condition consisted of a truth-table presented at the same time as the gate (Appendix B). This conformed to the concept of increased support by providing additional information beyond KCR. Every possible input and output for a gate was represented in the truth-table feedback.

Instructions

Participants were given the same instructions as Experiment 1, but were told that there would be different types of feedback given for the gates. When the experimenter walked through the practice gates with the participant, the differences between feedback types were pointed out (but no strategy was offered to the participant).

Procedure

Participants conformed to the same procedure outlined in Experiment 1.

Questionnaires

Several questions were added to the exit interview for this study specifically concerning participant use of feedback. These questions are detailed in the results section. Participants were also asked to solve one of every possible type of logic gate on paper during the exit interview as well as write in words the rule they used for each gate. They were instructed to write both the correct rule if they remembered it, and the rule they used when solving the gate.

Design

The study was a 2 (WMC group: low WMC, high WMC) × 2 (task load: simple, complex) × 2 (feedback condition: Summary Feedback [low support], KCR [high support]) factorial with Feedback Condition as a within-participant factor, Task Load a between-participant factor, and WMC Group a between participants grouping variable. The dependent variable of interest was accuracy of performance, however response time was also measured.

Hypotheses

It was expected that learning would vary according to the support of the feedback provided, but this effect would be different for each of the WMC groups depending on task load. Specifically, in a simple task, high WMC individuals should learn more with less supportive feedback whereas the opposite should be true for low WMC individuals. However, in a complex task, high WMC individuals should also benefit from increased feedback support.

Results

Means and contrasts between feedback conditions are available in Table 13 for the learning phases discussed in this section.

Table 13

Accuracy Levels for Experiment 4 Divided by Feedback Condition

		Beginning of Acquisition					End of Acquisition					10-minute retention				
		Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*
WMC	Task Load	KCR	Truth table	KCR	Truth table		KCR	Truth table	KCR	Truth table		KCR	Truth table	KCR	Truth table	
Low	Simple	0.51	0.55	0.20	0.07	-0.86	0.65	0.71	0.17	0.14	-1.46	0.71	0.60	0.18	0.18	3.62*
	Complex	0.36	0.41	0.16	0.11	-1.62	0.61	0.66	0.18	0.14	-1.68	0.71	0.68	0.13	0.07	1.09
High	Simple	0.55	0.53	0.18	0.17	0.41	0.75	0.75	0.20	0.16	0.31	0.71	0.76	0.16	0.16	-1.43
	Complex	0.61	0.59	0.19	0.21	0.72	0.81	0.84	0.22	0.18	-0.36	0.84	0.88	0.18	0.15	-0.50
		1-week retention					Content transfer					Load transfer				
		Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*	Mean accuracy		Standard deviation		*t*
		KCR	Truth table	KCR	Truth table		KCR	Truth table	KCR	Truth table		KCR	Truth table	KCR	Truth table	
Low	Simple	0.61	0.62	0.13	0.12	-0.43	0.58	0.60	0.18	0.20	-0.95	0.65	0.59	0.12	0.12	1.16
	Complex	0.56	0.60	0.11	0.05	-1.83	0.52	0.60	0.08	0.09	-3.78*	0.62	0.68	0.13	0.10	-1.47
High	Simple	0.72	0.73	0.14	0.15	-0.54	0.66	0.68	0.15	0.14	-0.61	0.74	0.64	0.14	0.18	1.47
	Complex	0.79	0.86	0.17	0.15	-1.72	0.78	0.84	0.20	0.17	-1.26	0.83	0.88	0.16	0.13	-1.30

Note. * indicates significant differences between groups ($p < .05$).

Acquisition

Beginning and End of Acquisition

Feedback effects and interactions. As in previous studies, the beginning and end of acquisition was compared in a repeated measures ANOVA. There was an interaction of feedback condition and WMC where low WMC participants benefited more from the truth table feedback than the KCR, while high WMC participants performed similarly with both, $F(52,1) = 3.80$, 1,52, .057(m), $np^2 = .07$, (Figure 24).

Other effects. Performance improved from beginning to end of acquisition, $F(52,1) = 60.78$, $p < .001$, $np^2 = .54$. High WMC participants were more accurate than low WMC participants, $F(52,1) = 18.26$, $p < .001$, $np^2 = .26$. WMC interacted with task load in that high WMC participants were more accurate on a complex task compared to the simple task while the opposite was true for low WMC participants, $F(52,1) = 6.97$, 1,52, .011, $np^2 = .12$. It is thought that this strange pattern might be due to the time limits placed on each trial, making the simple task more difficult than the complex task (see post hoc caveats in Overview of the Experiments).

Immediate Test

Feedback effects and interactions. There was a main effect of Feedback on the immediate test block, $F(52,1) = 5.331$, 1, 52, .025, $np^2 = .09$, where participants were more accurate on gates where they had received truth tables.

Other effects. High WMC participants continued to be more accurate than low WMC participants, $F(52,1) = 23.30$, $p < .001$, $np^2 = .31$. The interaction of WMC and Task Load was maintained with the same pattern as in acquisition, $F(52,1) = 4.08$, 1,52, .048, $np^2 = .07$.

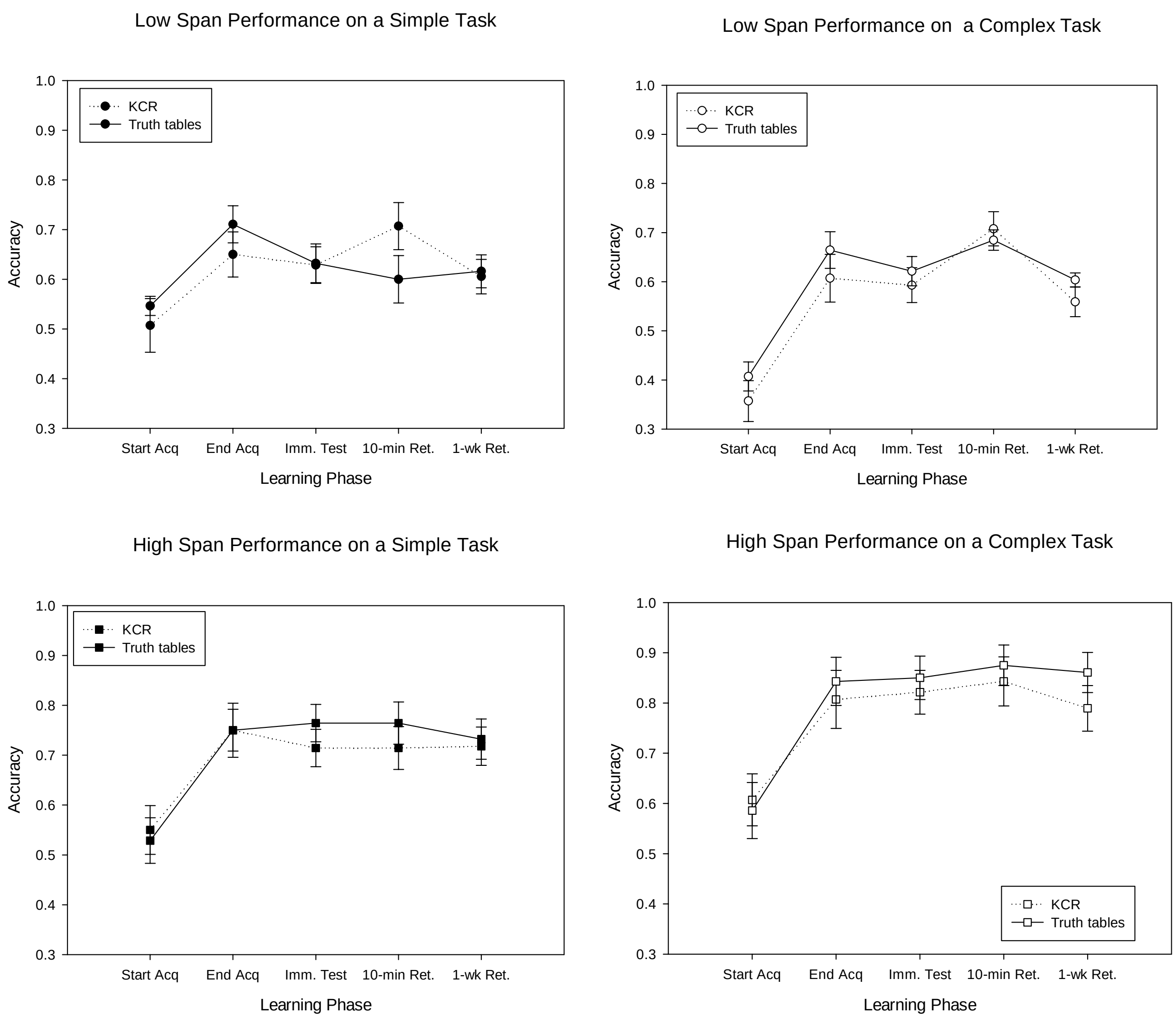


Figure 24. Experiment 4 performance across learning stages, divided by working memory capacity and task load.

Retention

10-minute Retention

Feedback effects and interactions. There was an interaction of Feedback and WMC at the 10-minute retention interval, $F(51,1) = 8.85$, $p = .004$, $np^2 = .15$, where low WMC participants' performance on truth table gates was lower than their performance with KCR, $t(26,1) = 3.36$, $p = .002$, while high WMC participants performed more accurately in the truth table condition. There were no significant differences in high WMC participant performance between feedback conditions ($p > .05$). Further, there were only marginal differences between WMC groups for KCR feedback, $t(53,1) = -1.80$, $p = .078$, however high WMC participants were significantly more accurate than low WMC participants when provided with truth table feedback, $t(53,1) = -4.53$, $p < .001$). As there was a marginal interaction of WMC and Feedback with Task Load, $F(51,1) = 2.30$, $p = .09$(m), $np^2 = .06$ (Figure 25), it appears that most of the interaction of feedback and working memory capacity is due to the simple task being more accurate with the truth tables for the low WMC participants, while the complex task is similar between feedback conditions and shows a larger difference in the KCR condition for the high WMC participants (Figure 26).

Other effects. High WMC participants were more accurate than low WMC participants, $F(51,1) = 11.98$, $p = .001$, $np^2 = .19$, and those in the complex task were more accurate than those in the simple task condition, $F(51,1) = 3.89$, $p = .054$, $np^2 = .07$.

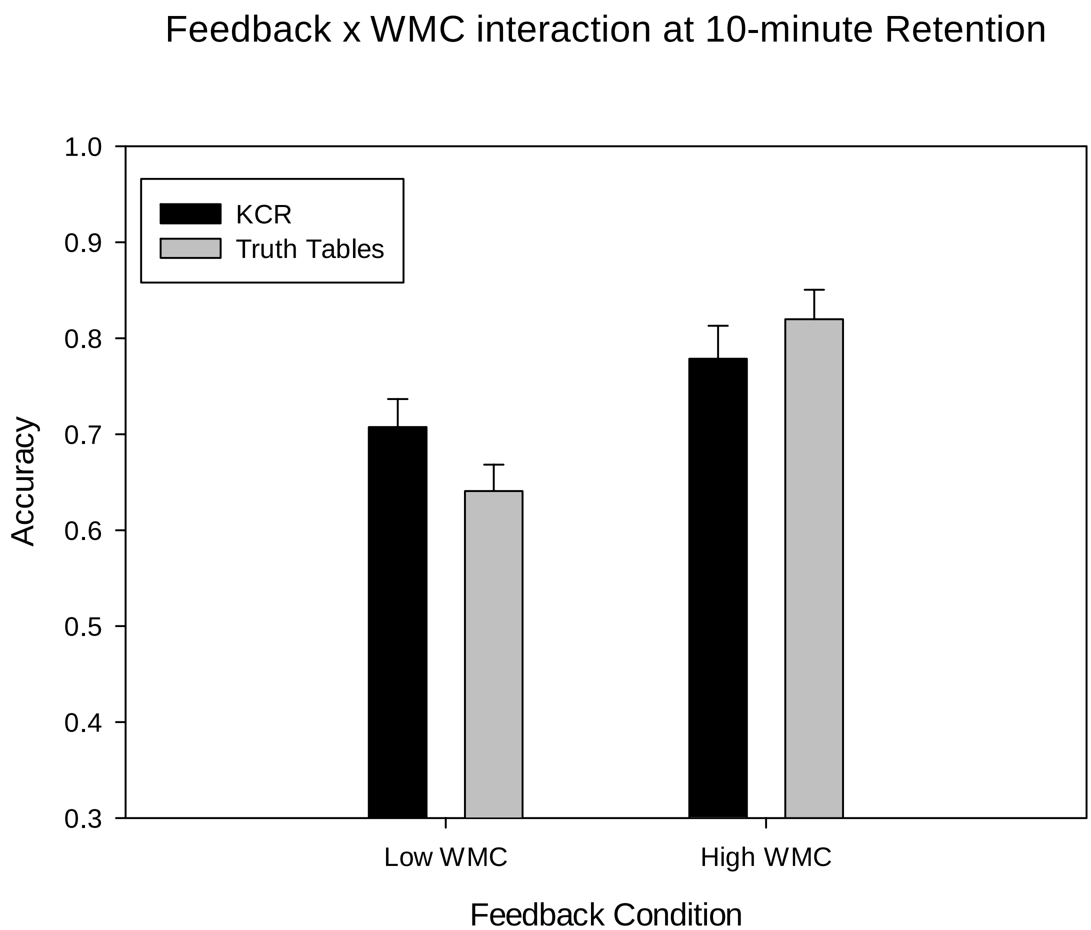


Figure 25. Interaction of Feedback and Working Memory Capacity at the 10-minute Retention test.

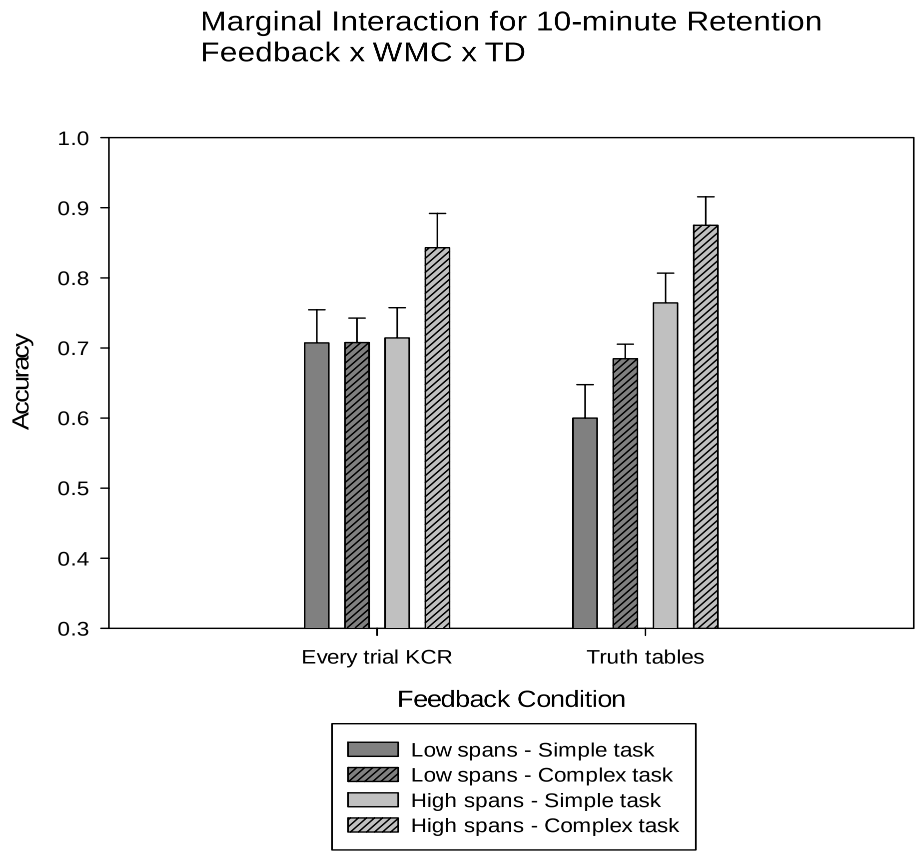


Figure 26. Experiment 4 data showing a marginal interaction of Feedback, Working Memory Capacity, and Task Demand at the 10-minute Retention test.

10-minute retention summary. For the initial retention measure, high WMC participants appeared able to use the additional information contained in the truth table feedback to increase their learning. Low WMC participants did not show this same benefit, and were more accurate with the less supportive feedback condition. It is possible the low WMC participants were overwhelmed with the information available in the truth table feedback.

1-week Retention Post Score

Feedback effects and interactions. At 1-week Retention, a feedback main effect was found in favor of the truth table condition, $F(52,1) = 4.90$, $p = .031$, $np^2 = .09$. The low WMC participant benefit for KCR at the 10-minute retention interval was not evident (Figure 27).

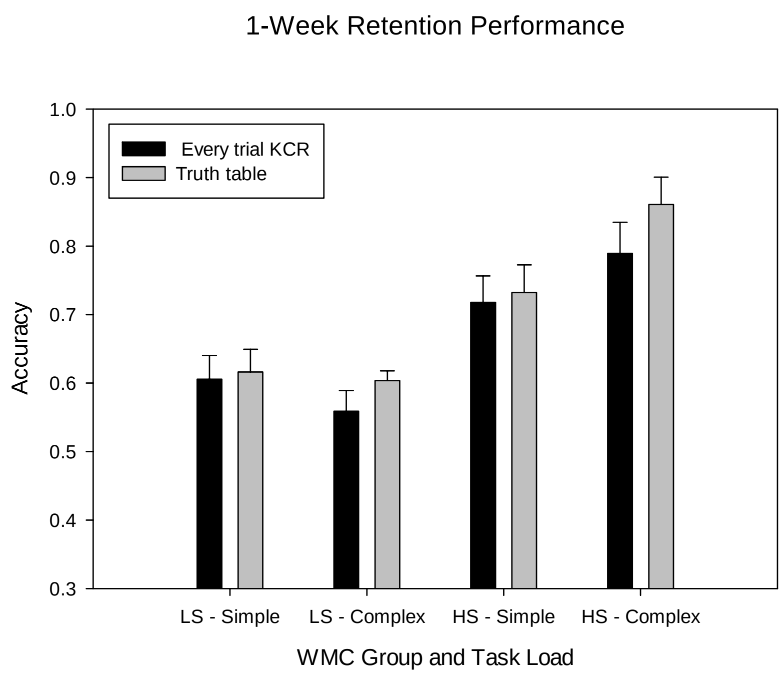


Figure 27. Experiment 4 1-week Retention performance divided by Working Memory Capacity and Task Load.

Other effects. High WMC participants were more accurate than low WMC participants, $F(52,1) = 31.67$, $p < .001$, $np^2 = .38$, and those in the complex task were more accurate than those in the simple task condition, $F(52,1) = 5.23$, $p = .026$, $np^2 = .09$. High WMC participant performance was significantly different from chance under all conditions of task load and feedback. Low WMC participant performance was significantly higher than chance in the simple task condition for both feedback types, but was no different in the complex task condition with KCR feedback, $t = 1.93$, $p = .076$.

Retention Change Score

A change score was computed to further examine 1-week Retention using the following formula:

(1-week accuracy – 10-minute accuracy)/10-minute accuracy

Feedback effects and interactions. A significant Feedback x WMC x Task Load interaction indicated differences in retention change scores, $F(51,1) = 5.27$, $p = .026$, $np^2 = .09$, (Figure 28). All feedback conditions for all WMC groups showed decline or no change. In the simple task, low WMC participant performance did not change across the retention interval for truth table feedback ($t = .85$, $p = .411$) However, there was a significant loss for KCR performance ($t = -2.42$, $p = .031$). Both feedback conditions showed loss for the low WMC participants in a complex task (p’s > .05). High WMC participants showed no change with either task load or feedback type (all p’s > .05). The Feedback x WMC interaction was also significant for this reason, $F(51,1) = 5.36$, $p = .025$, $np^2 = .10$. Contrasts indicated a difference between feedback conditions for the Low WMC participants on a simple task, $t(13,1) = -2.91$, $p = .012$. There was a main effect of Feedback where performance changed more negatively for the KCR condition, $F(51,1) =$

5.70, p = .021, np^2 = .10

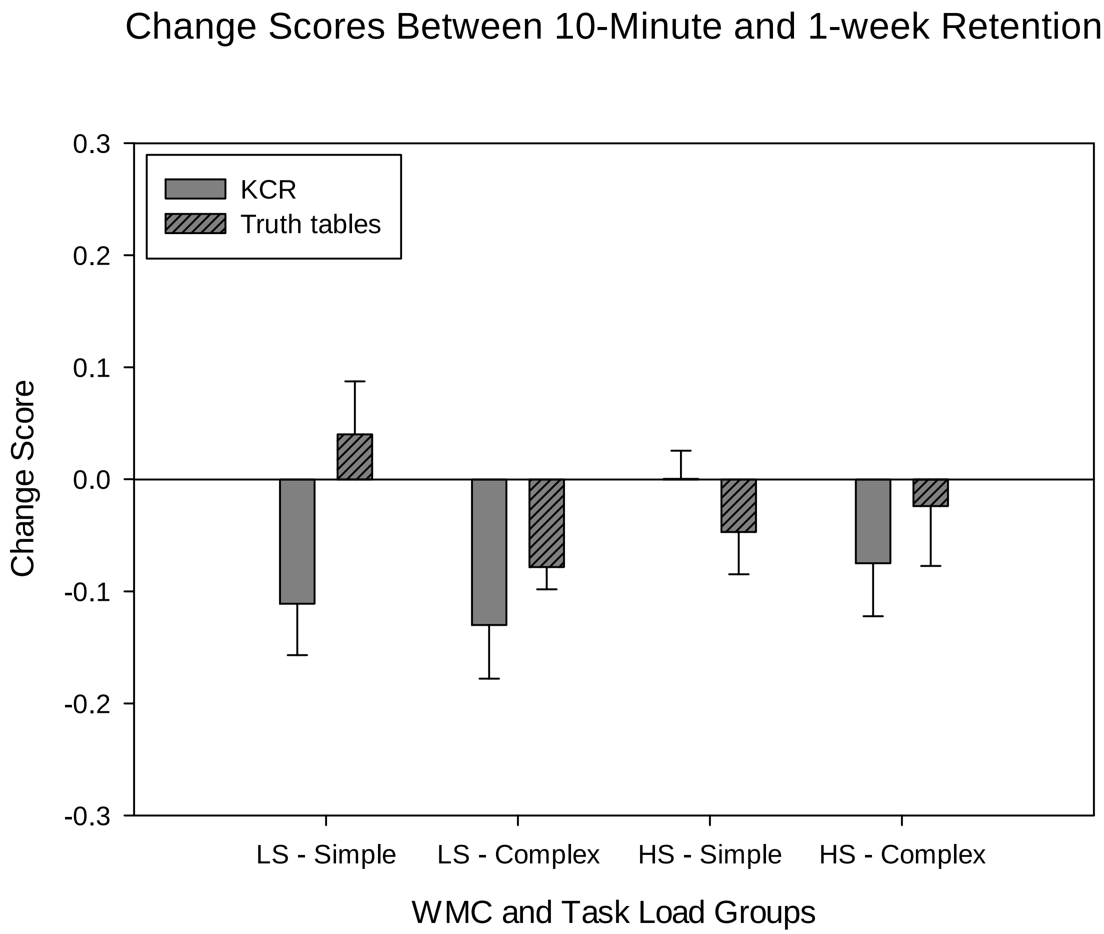


Figure 28. Experiment 4 Change Scores between 10-minute and 1-week Retention.

Summary for retention. In general participants performed more accurately in the truth table condition. High WMC participants retained the task over the week interval, demonstrating no change in performance on the 1-week retention test. Low WMC participants benefited most from KCR at the 10-minute retention interval, however over the course of a week gate accuracy in the KCR condition declined while performance on truth table gates did not change.

Transfer

Content Transfer Post Score

Feedback effects and interactions. There was a main effect of Feedback present for performance on the Content Transfer test where the truth table gates were more accurate than the KCR gates $F(52,1) = 7.84$, $p = .007$, $np^2 = .13$. There were no interactions involving feedback. However, contrasts revealed low WMC participants were more accurate on truth table gates on the complex task than for KCR gates, $t(13,1) = -3.78$, $p = .002$. The increased information present in the truth tables compared to KCR made for more accurate performance in retention and on the Content Transfer test.

Other effects. High WMC participants were more accurate than low WMC participants, $F(52,1) = 17.47$, $p < .001$, $np^2 = .25$. WMC interacted with task load where high WMC participants were more accurate on the complex task while low WMC participants showed no differences between Task Load groups, $F(52,1) = 4.13$, $p = .047$, $np^2 = .07$.

Content Transfer Change Score

There were no effects of performance change between 1-week Retention and the Content Transfer task. No WMC/Task Load group showed any significant change from

1-week Retention to the Content Transfer test.

Load Transfer Post Score

Feedback effects and interactions. The Load Transfer task showed an interaction of feedback and Task Load, $F(52,1) = 7.16$, $p = .010$, $np^2 = .12$. Accuracy was higher for the complex task than for the simple task, $F(52,1) = 8.78$, $p = .005$, $np^2 = .14$. Contrasts between feedback types for each task indicated only marginal differences between feedback types when Task Load conditions were considered separately (p = .068, .061 respectively) with truth tables producing the higher accuracies.

Load Transfer Change Score

Feedback effects and interactions. When Load Transfer was compared to 1-week Retention via a percent change score, there was a marginal effect of Feedback, $F(52,1) = 3.57$, $p = .064$, $np^2 = .06$, on accuracy where KCR showed a larger positive change than did the truth tables. As discussed in the Load Transfer section, truth tables were already at a higher accuracy.

Other effects. There was a main effect of task load where those in the complex condition showed a higher positive change when moving to the simple task than those who learned the gates in the simple task and transferred to complex, $F(52,1) = 4.40$, $p = .041$, $np^2 = .08$.

Summary of Transfer Effects

In general, when there were effects of feedback, there was a benefit for the more supportive truth tables over KCR.

Subjective Data

Participants in Experiment 4 answered exit interview questions about the feedback they received. Specifically, they were asked:

1. Please tell us which one made you learn the gates more thoroughly and why.
2. Which feedback type made you put forth more effort to learn the gates?(Explain)

Participant answers may be found in Tables 14-15, divided by working memory capacity group. In general, though both preferred the truth table feedback, they mentioned different reasons for doing so. Low WMC participants frequently mentioned trying to memorize outcomes from the tables. High WMC participants seemed to mention using the tables to discover patterns, re-discover a forgotten rule, or develop a personal rule from the tables.

Table 14

Low WMC Participant Responses

	Simple Task		Complex Task	
	2 = truth table,3 = KCR, 9 = same for both		**5 = truth table, 4 = KCR, 5 = same for both**	
Which feedback type made you put forth more effort to learn the gates? (explain)	both	I need to know if I'm answering correctly and if not, how I can do so in the future	table	I understand the verbal rule but when it gave me presentations of what input equaled, I had to reapply the verbal rule to the visual input.
	both	Both sp it tells me straight out if I was correct then gives me the possibilities to keep for future answers.	KCR	If I got it wrong being that I only had two options the next time I would know not to pick the same answer
	KCR	I tried to recall based on CPU feedback	KCR	I had to think about what the answer was supposed to be and why as opposed to the computer just giving me the info
	both	To me it went hand in hand. You needed to know if you got it right, and if you didn't what were your possibilities, so you can get it right next time	both	the presentation of all the possibilities helped me find my issues and know how to fix it. Being told correct or incorrect helped my timing
	both	I had a problem memorizing the gates	table	Helps you memorize the gates and keep them in your mind
	table	Obviously, if you make a mistake, any intelligent person with the intent of doing their best would rather have a presentation of possibilities. Only so they can learn along the way with positive reinforcement	KCR	I tend to need black or white, yes or no, correct or incorrect feedback to inspire me to continue forward
	both	the first option gives you a chance to find an easier way to learn the gates. Therefore, it is easier to understand the gates. The second option makes you think harder which makes it more difficult to understand. However, it also forces you to concentrate and think harder which helps in the long run	KCR	Just the kind of person who like to do well, so seeing higher percentages gives me a sense of learning or accomplishment
	both	Both were helpful	both	I say both for me because I fell as though when I got a problem wrong and I saw the feedback with all possibilities that made me want to study the chart and when I saw incorrect that made me want to put more effort to get it right next time
	both	Presentation of all gate possibilities feels hpeless/overwhelming because there are so many; the correct/incorrect states more direct problem.	KCR	Basically the overall understanding of what was a gate
	table	Helped me memorize them	both	The presentation of all gate possibilities made me try to memorize the possibilities. The correct/incorrect feedback made me wonder why my answer was incorrect if so, and the I tried to work out the possibilities in my mind for the next time it comes up
	both	Both helped out. Showing an example of the gate helped out for the one's that followed.	both	The presentation for reference and being told correct/incorrect helps the progress of learning the possibilities
	both	(blank)	table	I like knowing all the posibilities b/c ithelped me remember them throughout the task
			table	Correct or incorrect did not help me learn why I got the answer wrong.
			table	I tried to remember all possible gates when I received that feedback.

Table 15

High WMC Participant Responses

	Simple Task		Complex Task	
Which feedback type made you put forth more effort to learn the gates?(explain)		**4 = truth table,5 = KCR, 5 = same for both**		**9 = truth table, 2 = KCR, 3 = same for both**
	KCR	Because it made me feel bad for not remembering	table	It let me make my own rule
	both	Even though it was very helpful to see the presentation of all gates- knowing if I was right or wrong helped out a lot. Since it was all going so fast- the correct or incorrect feedback was a quick way for me to see the ones I was having difficulty with.	table	I could try to learn the solutions/patterns better
	KCR	When I see incorrect too many times I try harder to be right	KCR	It was easier to learn it looking at the numbers than looking at the words
	table	With correct or incorrect there was no mistake to learn from. With all answers, I could see my answers as well as memorize other scenarios.	table	Correct or incorrect was not helpful to me because I could not see how to fix my mistakes
	KCR	With a presentation of all possibilities I just look and say "ok I got it" but being told correct/incorrect made me think a little harder.	table	I could develop my own rules for gates based on the possibility set.
	both	Correct/incorrect was just as helpful or more helpful for either/or answers such as INV and BUF gates	table	Because if I saw all possibilities it refreshed my memory for when that gate would appear again, so I studied those screens longer.
	table	Because it gave me a chance to make a mental note of future gates and what the correct answer would be.	table	I had information in front of me to put a theory together
	table	Was able to figure out which number pairs would be correct for certain rules, giving me a better understanding of rule	table	By examining all gate possibilities I was able to remember which combination went with that gate
	table	Being able to see all possibilities assured me that I could become more proficient and accurate when attempting to solve each logic gate	KCR	knowing that I get something wrong the first time makes me work harder at it so I don't do it the next time around
	both	By seeing the 2 types of feedback helped me to reinforce the task at hand.	both	even with presentation of possibilities I still got some wrong so knowing it's correct/incorrect helped me see what I did right or wrong
	both	Both feedbacks are very helpful, think I gain motivation to solve each gate from seeing and being told "correct" or "incorrect," than just looking at a formula.	table	being told correct/incorrect didn't show what was wrong
	KCR	I hate being wrong, so I took it as a challenge to get the gates right	both	They both helped me try to keep gaining accuracy.
	KCR	I think this method was better because it made me really think hard in order to regurgitate the information looked over, although it was for a very short time.	table	I spent longer interpreting the first type of feedback
	both	(blank)	both	(blank)

Discussion

As in the previous experiments, the results of this study support the idea that a reducing task load via increased feedback support results in more learning. When there were differences between WMC groups, it tended to be that high WMC participants were better able than low WMC participants to use the increased feedback support to improve their performance in acquisition, retention, and transfer. However, the low WMC participants also benefited from the increased support of the truth tables, if not as much as high WMC participants. As stated previously, feedback was not only a support for performance, it was a tool to be used by those able to take advantage of it.

CHAPTER 7: GENERAL DISCUSSION

A synthesis of the feedback literature revealed differences in studies finding a benefit for increased feedback support during training versus those finding a benefit for reduced support. Those differences tended to be in the populations studied and the tasks that were learned. Studies recommending increased feedback tended to include low ability populations and/or complex or difficult tasks. Studies advocating decreased feedback tended to include highly able learners and/or simple or easy tasks. Situations where learners had high availability of cognitive resources (either because of their internal capacity and/or the simplicity of the task,) were optimal for reduced feedback in training. Training situations with limited available resources required more support from the feedback during training.

Because of these systematic methodological differences in previous studies, a resource-based explanation was proposed and tested in the current series of experiments. The results of the current experiments tested and tuned this resource-based hypothesis by adding knowledge of the many roles feedback may play during training.

The current experiments focused on how feedback is not an entity unto itself, but *part of the task* as it is learned. Feedback support may actually change the task in the way that part-task training changes a whole-task (Wightman & Lintern, 1985). Feedback can remove the step of self-assessment from the task, it can provide more information than would be available in the task alone, and it can guide the learner to a standard of performance.

A series of four experiments demonstrated the complexity of the effects of feedback on learning. Available learner resources were controlled via two variables: learner working memory capacity and cognitive task load. Feedback condition was manipulated by the amount of support it provided for performance. If the utility of feedback is due to available learner resources, it was expected that the high and low WMC learners would flourish under different levels of feedback support.

Both WMC groups benefited from additional feedback support under all levels of task load, with the high WMC participants learning differentially more with additional feedback support. Other important differences were that the high WMC participants were still able to learn when provided with little feedback support while the low WMC participants tended to perform at chance under those conditions.

That the addition of feedback support resulted in learning for both high and low working memory capacity groups may be explained through their different *use* of the same feedback. There is not a one-size-fits-all design for training or even for specific populations. It is most useful to examine what feedback is doing or allowing under conditions of low or high support. The following is an analysis of how different learners utilized the feedback support in the current experiments.

Experiments 1, 2, and 4 specifically included highly capable learners. Though these participants learned the logic gate task under all feedback support levels, they usually learned better via *more* supportive feedback. There were sometimes no differences between feedback conditions, but there was no case in which they learned more from the *less* supportive feedback. These results may be explained in their ability to use the feedback. Subjective measures indicated that these learners looked for patterns or developed concepts as to how the logic gates worked. They were then able to test these theories and re-evaluate upon further feedback receipt. Participants did not use the additional information as much for memorization of answers as they did for hypothesis testing and model development.

All of the functions mentioned thus far involve some amount of meta-cognition, self-analysis, and problem solving. Thus, few of these benefits would necessarily be available to learners with low working memory capacity. Low WMC participants in the current studies also showed more learning under conditions of increased feedback, so the question remained that if they were not also developing and testing hypotheses, what were they doing?

Feedback provided guidance through the task as well as a clear standard of performance. It may also simplify the task by removing the component of self-assessment. Feedback in the current studies also provided more information than would be contained by the task alone which removed the need for the learner to generate that information. There was no time limit on looking at the feedback, so the additional processing load may have been minimal and the benefit of not having to generate the information large. (Indeed, the low WMC participants did not appear able to self-generate this information in Experiments 1 nor even after 10 sessions of practice in Experiment 3).

Thus, even without the meta-cognitive analysis or complex problem solving strategies that could be used by high ability learners, the low ability learners could still benefit from increased feedback support for two reasons: When more feedback support was available, the reduced load allowed low ability learners to maintain some performance while learning the task. Second, it allowed the learner to experience what a correct response looked like. Cognitive load theory supports the first assumption that a reduction in task load should result in more learning (Sweller, 1988) and the ACT-R framework suggests that increasing correct performance in acquisition would mean the learner is held to the correct model of performance, and future performance would also correspond more highly with that model (Anderson, 1982, 1996). Both of these mechanisms seemed to operate on the low WMC participants in the current experiments. Additional feedback reduced task load by providing correct answers. Providing that information held learners to more of an expert model of performance than did less feedback. Of course, this may also explain the better performance of the high WMC learners under additional feedback support, but it seems likely their use of strategies upon evaluating their own performance played a large role as well.

Perhaps the most interesting aspect to this series of experiments is the lack of superiority for *any* relatively less feedback condition. Many articles from the last two decades are adamant about the superiority of “less” feedback compared to more, and cognitive resources provided an obvious source for this superiority. Once explained, it appears to make sense that those with the resources to do so might learn best with little external aid. The current studies controlled for cognitive resources, and yet there was no case in which less outperformed more, despite numerous predictions to the contrary.

For example, transfer appropriate processing theorists would note the summary feedback in Experiments 1-3 was most similar to the retention test. Yet it produced the worst performance at retention. Second, if too much feedback becomes a crutch, or inextricably part of the task, then it would be predicted KCR would produce better retention than truth tables, however this was not the case for any condition or group. Third, if feedback acts as guidance to keep learners from internalizing their own feedback, one would expect superior learning from the summary information, or at least KCR when compared to truth tables. Again, this was not the case. Fourth, if feedback may block other information-processing activities that could result in the capability to perform when feedback is withdrawn, less information should result in better retention. However, the information in the feedback for the current studies only increased retention performance. Fifth, discovery learning theory predicts KCR should have outperformed the truth tables in Experiment 4; KCR gives the needed information but requires the learner to “discover” the other cases for a gate. However, providing those cases explicitly through a truth table resulted in better retention. Last, there are many proponents of desirable difficulties in instruction, because “responding to them (successfully) engages processes that support learning, comprehension, and remembering” (Bjork, 2006). However, in all groups and conditions, performance increased as difficulties were *removed*. Even high WMC participants learning a simple task benefited from a decrease in difficulty during acquisition.

In conclusion, learners benefit from increased feedback support in multiple scenarios. First, users who are very capable may think about their own performance, the feedback, and how the two match or do not match and why. Because they can see patterns in their performance (due to the above) they may benefit from any additional information provided. Of course, this is most likely to happen when the task is not overwhelming, so there is time and attention remaining for considering their performance. This scenario fits with the objective and subjective data obtained on learners with high WMC in the current experiments. Second, learners can benefit from increased feedback support because they are held more closely to a model of expert performance and the task load is reduced by the feedback. This scenario fits with all the WMC groups included in the current studies, but fits especially with the low WMC groups.

Limitations of the Current Experiments

Perhaps the largest limitation of these experiments is the lack of generalization to many learning situations. The feedback conditions for these studies were modeled closely on motor learning experiments that found a benefit for lessened feedback (Schmidt, Lang, & Young, 1990; Schmidt & Wulf, 1997; Schmidt et al., 1989), but did not replicate those findings. Though many equivocate learning domains (Burton et al., 2004; Newell, 1991; VanLehn, 1996), there may be differences worth investigating that were not covered by these studies.

Second, difficulty introduced by time constraints may have differentially changed the simple and complex tasks by making the simple task fairly difficult. Feedback condition did not often interact with task load in these experiments, and the question remains as to whether task load does not produce much effect or whether these tasks were not differentiated enough in difficulty *and* complexity.

Last, although pre- and post-test motivation was measured, it was not specifically manipulated or measured as it pertained to each feedback condition. The subjective measures of “frustration” or other similar language in the exit interviews indicate that motivation may still have had a large effect on the effectiveness of feedback.

Future Research

Future research in this area is promising, not only to eliminate the above limitations, but to explore interesting possibilities raised by the current experiments. First, the difference in the current findings versus previous results from the motor learning literature are intriguing. Future studies should examine what differences in learning requirements or domain could explain these mixed findings. It could be, for example, that although motor learning tasks are also composed of trials, the action is the same for each trial. In cognitive domains (and in the current experiments), the learning required is usually more various per trial. This could create a confound of practice schedule (Schmidt & Bjork, 1992) between domains.

Second, the idea of feedback as both a load and as having the ability to reduce task load should be further explored. Is there a categorical difference between these two actions of feedback? Are they additive? How may they be measured? It may be that the difference between the resources required to understand the feedback and the resources freed by having the feedback present may explain when feedback will increase or decrease learning. Feedback may have both actions simultaneously which will make it difficult to tease apart their functions.

Third, despite the initial foray into strategy use by Experiment 2, there are many questions concerning strategy still to be answered. In experiments of associative learning, participants have been instructed in strategy use (Hertzog & Dunlosky, 2005). Investigating similar questions concerning feedback would go far in explaining why technically useful feedback may not end up being used by a learner. Effective strategies may not be naturally adopted.

Last, there is the question of whether effects from a low WMC younger population generalize to other populations with reduced resources. These populations notably include older adults, who will differ in ways other than WMC to the groups included in the current studies. It is not yet known if increased feedback support would universally help older adults learn a rule-based logic task or whether other supports would be necessary.

Application

Feedback based on the interaction between a learner's working memory capacity and the task load would aid system designers, human tutors, coaches, and teachers. In online coursework, feedback should aid learner retention, not simply correct all errors or provide a grade. Online courses are popular as "learning support" classes for remedial students (Williams, 2007). The current studies indicate the feedback presented in these courses would be particularly important to remedial students. Similarly, automated training could be tailored to the learner rather than using one program to fit all, based on the amount of informative feedback a certain learner could utilize. Finally, human tutors are notoriously bad at giving feedback as the "common sense" approach to feedback (Merrill et al., 1992). Human teachers may be educated in these principles and taught how to give feedback based on the learner and task.

Conclusion

"Everything should be made as simple as possible, but not simpler."

-Albert Einstein

A finding common to the series of experiments performed was that increased feedback support produced higher performance in acquisition, retention, and transfer. Feedback support was increased by providing feedback more frequently, immediately, with higher specificity, and with more units of information. The low and high WMC learners used this information differently, but both benefited from the increased support.

It may be asked, then, why it was not concluded that increased feedback was unilaterally better.

Parsimony seems to suggest we conclude feedback promotes learning by *added* support or information, and the more the better. However, Occam's razor demands the simplest explanation that explains *all* the data, not only the data from this series of experiments. These results demonstrating the superiority of increased feedback break down the reasons why high information content can aid learners with different levels of working memory capacity, while preserving the differences between their use of this information.

APPENDIX A: PARTICIPANT MATERIALS

Experiment Protocols

Session 1: (~3 hrs)

Informed consent

Near and far vision tests

Metamemory in Adulthood questionnaire (Exp. 2 & 5)

Digit symbol substitution

Introduction to logic gates

Prior Experience questionnaire

Memorize gate actions (5 min)

Criterion test(s) of logic gate rules

Pre-acquisition LGT questionnaire

Preferred learning strategy survey (Exp. 2 & 5)

Explicit instruction on feedback conditions (Exp. 2)

Pre-acquisition practice (10 trials with "cheat sheet")

Training (10 blocks of 20 trials each)

Immediate test (1 block)

Strategy survey (Exp. 2 only, ~5 min)

Shipley vocabulary test (~7 min)

2nd Preferred learning strategy survey (Exp. 2 & 5)

10-minute retention (1 block)

Session 2: (~2 hrs)

Simple reaction time test

Choice reaction time test

Logic gate retention test (4 blocks)

Content Transfer test (4 blocks)

Load Transfer test (2 blocks)

Exit Interview

Post-experiment LGT questionnaire

Declarative knowledge test

Reverse digit span test

Debriefing

Payment

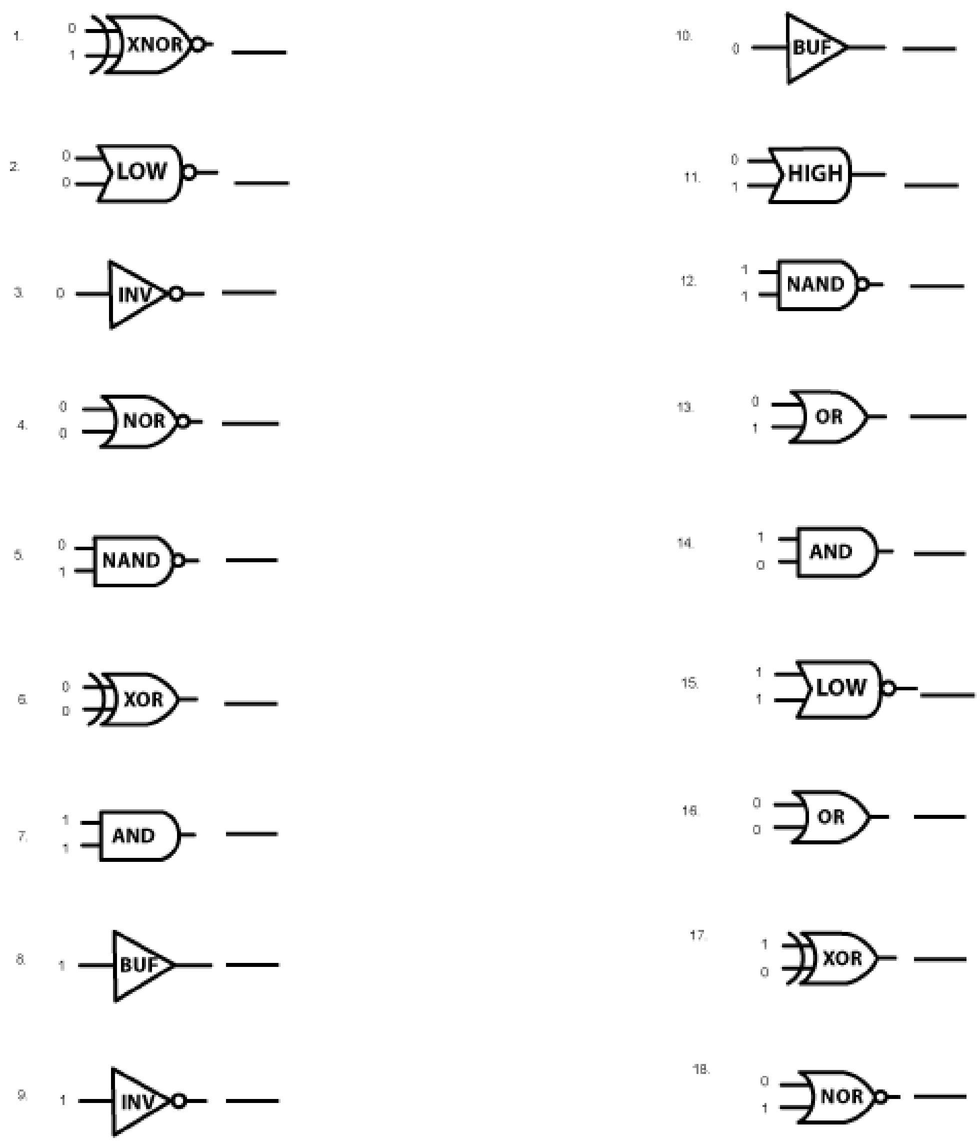
Name: ____________________________

Phone #: __________________________

Please call me if I'm eligible:
YES NO

Write the answer to each problem in the blank space provided.
Guess the answer if you are not sure how to respond. If you have no idea, that's fine!

Put an X in the blank if you have no idea how to respond:________



Consent form

Participant ID ___________

Georgia Institute of Technology

Project Title: Learning a new task through feedback

Investigators: *Anne McLaughlin, Dr. Wendy A. Rogers & Dr. Arthur D. Fisk*

Research Consent Form

You are being asked to be a volunteer in a research study.

Purpose:

The purpose of this form is to tell you about the tests you will be asked to complete today and to inform you about your rights as a research volunteer. Feel free to ask any questions that you may have about the study, what you will be asked to do, and so on.

Thank you for your interest in participating in the study. Our work could not be completed without the help of volunteers. The purpose of this experiment is to understand how people learn to perform a new task and what is the best way to teach them. You will be answering questions presented on a computer screen about the task you are learning.

Procedures

If you decide to be in this study, you will take part in a five-day study. You will answer logic gate problems on a computer for each of these days. You will receive different levels of feedback for different gates.

It is important that you understand the instructions and examples before beginning the study. If anything is unclear at any time, please do not hesitate to ask questions. There is no deception in this study and you can ask any question at any time.

You will be given numerous breaks throughout your participation in this study, including a 30 minute break halfway through each session. We do not wish you to feel fatigued; please let us know if you need to take longer breaks.

1

APPROVED *Consent Form Approved by Georgia Tech IRB: April 19, 2007 - June 13, 2007*

Participant ID ___________

Risks/Discomforts

Participation in this study involves minimal risk or discomfort to you. Risks are minimal and do not exceed those of normal office work. If you experience eyestrain during a session, we recommend that you look around the room for about thirty seconds so that your eyes focus at different distances. The experimenter will explain when breaks are scheduled. However, if you feel that you are getting bored or tired, please feel free to take a break. Please tell us if you are having trouble with any task.

Benefits

There are no direct benefits to participating in the study. After the end of the study, we will share the results with you by mailing a newsletter to all the participants.

Compensation to You

The experiment will last approximately 3 hours each day across five days. At the completion of this study, you will receive $225.00. If you do not complete the study, you will be paid an hourly wage based on the time that you were involved in the study.

Confidentiality

The following procedures will be followed to keep your personal information confidential in this study: The data that are collected about you will be kept private to the extent allowed by law. To protect your privacy, your records will be kept under a code number rather than by name. Your records will be kept in locked files and only study staff will be allowed to look at them. Your name and any other fact that might point to you will not appear when results of this study are presented or published.

Confidentiality cannot be guaranteed; your personal information may be disclosed if required by law. This means that there may be rare situations that require us to release personal information about you, for example, in case a judge requires such release in a lawsuit.

2

APPROVED *Consent Form Approved by Georgia Tech IRB: April 19, 2007 - June 13, 2007*

Participant ID ___________

To make sure that this research is being carried out in the proper way, the Georgia Institute of Technology IRB will review study records. The Office of Human Research Protections may also look at study records.

Because each individual's data and test scores are completely confidential, we cannot mail your individual results. We will mail the group results and a summary of the conclusions once the project is completed.

In Case of Injury/Harm:

Reports of injury or reaction should be made to:
Dr. Wendy Rogers at (404) 894-6775 or
Dr. Arthur Fisk at (404) 894-6066

Neither the Georgia Institute of Technology nor the principal investigator has made provision for payment of costs associated with any injury resulting from participation in this study.

Research Participant Rights

- Your participation in this study is voluntary. You do not have to be in this study if you do not want to be.
- You have the right to change your mind and leave the study at any time without giving any reason, and without penalty.
- Any new information that may make you change your mind about being in this study will be given to you.
- You will be given a copy of this consent form to keep.
- You do not waive any of your legal rights by signing this consent form.

Participant ID ___________

Questions about the Study or Your Rights as a Research Participant

- If you have any questions about the study, you may contact the investigator (Anne) at (404) 894-8344.
- If you have any questions about your rights as a research participant, you may contact Ms. Melanie Clark, Georgia Institute of Technology at (404) 894-6942.

If you sign below, it means that you have read (or have had read to you) the information given in this consent form, and you would like to be a volunteer in this study.

____

Participant Name

____

Participant Signature **Date**

____

Signature of Person Obtaining Consent **Date**

If you must cancel a scheduled time to come to the lab, please call: (404) 894-8344.

Memory Questionnaire

This survey is about how you remember information. There are no right or wrong answers. Circle a number between 1 and 7 that best reflects your judgment about your memory. Think carefully about your responses and try to be as realistic as possible when you make them. Please answer all questions.

1. How would you rate your memory in terms of the kinds of problems that you have

major problems			some minor problems			no problems
1	2	3	4	5	6	7

2. How often do these present a **memory problem** for you?

a. names	1	2	3	4	5	6	7
b. faces	1	2	3	4	5	6	7
c. appointments	1	2	3	4	5	6	7
d. where you put things	1	2	3	4	5	6	7
e. performing household chores	1	2	3	4	5	6	7
f. directions to places	1	2	3	4	5	6	7
g. phone numbers you've just checked	1	2	3	4	5	6	7
h. phone numbers you use frequently	1	2	3	4	5	6	7
i. things people tell you	1	2	3	4	5	6	7
j. keeping up correspondence	1	2	3	4	5	6	7
k. personal dates (e.g. birthdays)	1	2	3	4	5	6	7
l. words	1	2	3	4	5	6	7
m. going to the store and forgetting what you wanted to buy.	1	2	3	4	5	6	7
n. taking a test	1	2	3	4	5	6	7
o. beginning to do something, then forgetting what you were doing	1	2	3	4	5	6	7
p. losing the thread of thought in conversation	1	2	3	4	5	6	7
q. losing the thread of thought in public speaking	1	2	3	4	5	6	7
r. knowing whether you've already told someone something	1	2	3	4	5	6	7

For this next part, we would like you to think about how different people use their memory in different ways in their everyday lives. For example, some people make shopping lists while others do not. Some people are good at remembering names while others are not.

In this survey, we would like you to tell us how you use your memory and how you feel about it. There are no right or wrong answers to these questions because people are different. Please take your time and answer each of these questions to the best of your ability.

Each question is followed by five choices. Draw a circle around the letter corresponding to your choice. Mark ONLY one letter for each statement. Choose the one that comes closest to what you usually do. Don't worry if the time estimate is not exact, or if there are some exceptions.

Keep these points in mind:
Answer every question, even if it doesn't seem to apply to you very well
Answer as honestly as you can what is true for you. Please do not mark something because it seems like the "right thing to say." ☺

For most people, facts that are interesting are easier to remember than facts that are not.
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

I am good at remembering names.
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you keep a list or otherwise note important dates, such as birthdays and anniversaries?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

When you are looking for something you have recently misplaced, do you try to retrace your steps in order to locate it?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

When you have not finished reading a book or magazine, do you somehow note the place where you stopped?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you think about the day's activities at the beginning of the day so you can remember what you are supposed to do?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you post reminders of things you need to do in a prominent place, such as bulletin boards or note boards?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you routinely keep things in a familiar spot so you won't forget them when you need to locate them?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

When you want to take something with you, do you leave it in an obvious, prominent place, such as putting your suitcase in front of the door?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

When you try to remember people you have met, do you associate names and faces?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

When you have trouble remembering something, do you try to remember something similar in order to help you remember?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you consciously attempt to reconstruct the day's events in order to remember something?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you use mental images or pictures to help you remember?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you write yourself reminder notes?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you write appointments on a calendar to help you remember them?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Do you write shopping lists?
1. Agree strongly
2. Agree
3. Undecided
4. Disagree
5. Disagree strongly

Instructions

For this study, you will learn how to use Logic Gates and then solve Logic Gate problems on a screen.

Logic Gates are symbols that represent some transformation to an input. An input goes in the left side of the gate and the gate produces an output.

For example, this first gate is the "AND" gate.

On the left side of the gate are two numbers. These are the inputs that enter the gate.

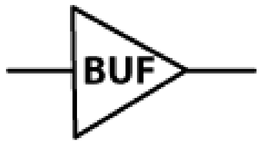
On the right side is the output based on those two numbers. Below is the rule for the AND gate.

> "If *all* inputs are 1, then the output is a 1."
>
> For this gate, the only way for the output to be a 1 is for *both inputs* to be a 1. If they are both 0, then the output is 0. If either input is a 0, then the output is 0.

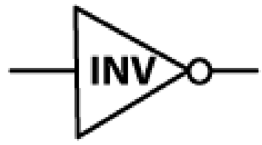
There is a unique rule such as this for every gate.

- First, you will first learn to recognize the gates and their names.
- Then you will learn their functions.
- Only then will you be asked to solve the outputs for multiple gates.
- On the opposite side of this page are the rules for all gates. You may consult these as you learn the gates.

Logic Gate Rules



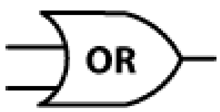
If the input is 1,
then the output will be 1.



If the input is 1,
then the output will be 0.

If *all* inputs are 1,
then the output will be 1.

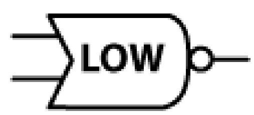
If *all* inputs are 1,
then the output will be 0.



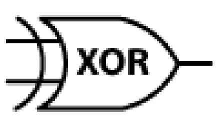
If *one or more* inputs are 1,
then the output will be 1.

If *one or more* inputs are 1,
then the output will be 0.

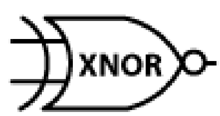
If the inputs are 1 or 0,
then the output will be 1.



If the inputs are 1 or 0,
then the output will be 0.



If *exactly* one input is 1,
then the output will be 1.



If not *exactly* one input is 1,
then the output will be 1.

Prior Experience Survey

Please circle your answers to these questions:
After reviewing the instructions on Logic Gates for this study:

1. Have you ever been exposed to this task? **Yes** **No**

2. Are you familiar with any of the logic gates? **Yes** **No**

   2a. If so, which ones? (please circle)

   AND OR NOR NAND XOR BUF INV

3. Have you ever taken a computer science course that required programming in a computer language? (Examples include BASIC, C++, Visual Basic, and LISP).

   Yes **No**

   3a. If Yes, have you ever used operators such as "AND" "OR" "XOR" or "NAND" in programming?

   Yes **No**

4. Have you ever used operators such as "AND" "OR" "XOR" or "NAND" in a library or web search? (Example would be to search for "country AND fair AND Atlanta OR Marietta")

   Yes **No**

5. Have you ever taken an electrical engineering class? **Yes** **No**

6. Have you ever read books on re-wiring a house? **Yes** **No**

7. Have you ever taken a philosophy or logic class? **Yes** **No**

Pre-acquisition Questionnaire of Logic Gate Confidence

Pre-Test Questionnaire

1. I now feel confident about my ability to solve logic gates.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

2. Solving logic gates will make me physically tired.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

3. The logic gate task will be boring.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

4. The logic gate task will be:

1	2	3	4	5	6	7
Too easy			Neither easy nor hard			Too hard

5. I am motivated to do my best.

1	2	3	4	5	6	7
Strongly disagree			Neither agree nor disagree			Strongly agree

Pre/Post Acquisition Strategy Questionnaire

Preferred Learning Strategies

In the following experimental task you will solve logic gates according to the rules you just memorized. You will need to make a connection between the picture of the gate and the rule so that you will be able to solve each logic gate problem. There are many ways to learn these pairs. Please read the following strategy choices and the brief descriptions of the strategies and rate how effective each strategy is for learning the logic gate rules.

1	2	3	4	5	6	7	8	9	10
least effective				moderately effective					most effective

1. Rote repetition. Say the gate and rule over and over.
   1 2 3 4 5 6 7 8 9 10

2. Attentive reading. Read over or say the gate and rule once in your mind.
   1 2 3 4 5 6 7 8 9 10

3. Semantic reference. Relate the gate and rule to something of meaning in your life.
   1 2 3 4 5 6 7 8 9 10

4. Focal attention. Focus on the gate and rule by looking at them until you can see the pair clearly in your mind.
   1 2 3 4 5 6 7 8 9 10

5. Imagery. Imagine a scene using the two words in the image.
   1 2 3 4 5 6 7 8 9 10

6. Sentence generation. Construct a sentence using both the gate and its rule.
   1 2 3 4 5 6 7 8 9 10

7. Rhyme. Form a pair of words that rhymes with the gate name and its rule.
   1 2 3 4 5 6 7 8 9 10

8. Other strategy. (Please explain) ______________________________

   ____

   ____

   ____

9. Other strategy. (Please explain) ______________________________

   ____

   ____

   ____

10. If you had to pick one and only one strategy, please check the strategy you would prefer to use to learn the logic gates.

______rote repetition
______attentive reading
______semantic reference
______focal attention
______imagery
______sentence generation
______rhyme
______Other ____________
______Other____________

11. Would you use any of the strategies in combination with another strategy to study a single pairing of logic gate and rule? (please circle) YES or NO

If yes, which strategies? ____

____

Preferred Learning Strategies – After session

You have just completed the learning stage of this experiment. You were asked to use a particular strategy to learn the logic gates and their rules. Again, there were many ways to learn these gates and rules. We would like you to think about strategies you might use in the future.

Imagine trying to learn new gates and rules similar to the ones you just learned. Please rate how effective you think each strategy is for learning the gates and rules.

1	2	3	4	5	6	7	8	9	10
least effective				moderately effective					most effective

10. Rote repetition. Say the gate and rule over and over.
    1 2 3 4 5 6 7 8 9 10

11. Attentive reading. Read over or say the gate and rule once in your mind.
    1 2 3 4 5 6 7 8 9 10

12. Semantic reference. Relate the gate and rule to something of meaning in your life.
    1 2 3 4 5 6 7 8 9 10

13. Focal attention. Focus on the gate and rule by looking at them until you can see the pair clearly in your mind.
    1 2 3 4 5 6 7 8 9 10

14. Imagery. Imagine a scene using the two words in the image.
    1 2 3 4 5 6 7 8 9 10

15. Sentence generation. Construct a sentence using both the gate and its rule.
    1 2 3 4 5 6 7 8 9 10

16. Rhyme. Form a pair of words that rhymes with the gate name and its rule.
    1 2 3 4 5 6 7 8 9 10

17. Other strategy. (Please explain) ______________________________

____

____

____

18. Other strategy. (Please explain) ______________________________

____

____

____

10. If you had to pick one and only one strategy, please check the strategy you would prefer to use to learn the logic gates.

______rote repetition
______attentive reading
______semantic reference
______focal attention
______imagery
______sentence generation
______rhyme
______Other ____________
______Other____________

11. Would you use any of the strategies in combination with another strategy to study a single pairing of logic gate and role? (please circle) YES or NO

If yes, which strategies? ____

____

End of Session 1 Questionnaire – Give during 10 minute break

Please circle the answer to each question. Remember we are most interested in what you DID do, not what you were instructed to do. Please be honest, we really want to know what strategies you used. ☺

1. How well did you follow the instructions you were given regarding feedback?

1	2	3	4	5	6	7
Did not follow at all						Followed perfectly

2. When I received feedback that I was INCORRECT after answering a gate, I:
    1. Ignored the feedback
    2. Looked at the feedback briefly
    3. Tried to use the feedback to evaluate the rule I was using for that gate
    4. Looked at the feedback at length and developed a plan for how to answer that gate in the future.
    5. Other strategy____

3. When I received feedback that I was CORRECT right after answering a gate, I:
    1. Ignored the feedback
    2. Looked at the feedback briefly
    3. Tried to use the feedback to evaluate the rule I was using for that gate
    4. Looked at the feedback at length and developed a plan for how to answer that gate in the future.
    5. Other strategy____

4. When I received feedback at the end of a block, I:
    1. Ignored the feedback
    2. Looked at the feedback briefly and moved on to the break
    3. Tried to memorize the gates I did poorly on
    4. Tried to memorize the gates I did poorly on and developed a plan for how to answer those gates in the next block
    5. Other strategy _____________________________________

Preferred Learning Strategies – After session

You have just completed the learning stage of this experiment. You were asked to use a particular strategy to learn the logic gates and their rules. Again, there were many ways to learn these gates and rules. We would like you to think about strategies you might use in the future.

Imagine trying to learn new gates and rules similar to the ones you just learned. Please rate how effective you think each strategy is for learning the gates and rules.

1	2	3	4	5	6	7	8	9	10
least effective				moderately effective					most effective

10. Rote repetition. Say the gate and rule over and over.
    1 2 3 4 5 6 7 8 9 10

11. Attentive reading. Read over or say the gate and rule once in your mind.
    1 2 3 4 5 6 7 8 9 10

12. Semantic reference. Relate the gate and rule to something of meaning in your life.
    1 2 3 4 5 6 7 8 9 10

13. Focal attention. Focus on the gate and rule by looking at them until you can see the pair clearly in your mind.
    1 2 3 4 5 6 7 8 9 10

14. Imagery. Imagine a scene using the two words in the image.
    1 2 3 4 5 6 7 8 9 10

15. Sentence generation. Construct a sentence using both the gate and its rule.
    1 2 3 4 5 6 7 8 9 10

16. Rhyme. Form a pair of words that rhymes with the gate name and its rule.
    1 2 3 4 5 6 7 8 9 10

17. Other strategy. (Please explain) ______________________________

____

____

____

18. Other strategy. (Please explain) ______________________________

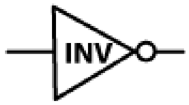
____

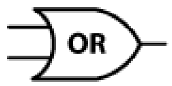
____

____

Post-test of logic gate declarative knowledge (Exp. 1)

Please write the rules you used for each Logic Gate.

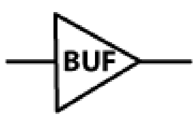
1. 

2. 

3. 

4. 

5. 

6. 

7. 

Exit Interview

Now that you have completed our experiment, we would like you to answer a few questions about your experience in the study. There are no right or wrong answers, please just provide your opinion. Please circle the number that best corresponds to your answer, or, for open-ended questions, write in your response.

1. I now feel confident about my ability to solve logic gates.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

2. Solving logic gates made my arm tired.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

3. The feedback I received helped me memorize the actions of the logic gates.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

4. The feedback I received confused me.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

5. I wanted feedback more often.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

6. By the end of the first session, I no longer needed any feedback.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

7. The two minute break between blocks was too short.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

8. It was easy to stay between 85-95% accuracy.

1	2	3	4	5	6	7
Very easy			Neither easy nor difficult			Very difficult

9. The logic gate task was boring.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

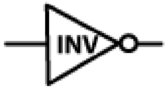
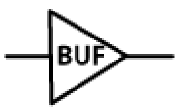
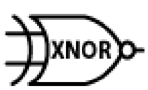
10. The logic gate task was too easy.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

11. The logic gate task was too hard.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

Please rate each Logic Gates in terms of difficulty from 1 (easiest) to 7 (hardest).

		Very easy			Neither easy nor hard		Very hard	
1.		1	2	3	4	5	6	7
2.		1	2	3	4	5	6	7
3.		1	2	3	4	5	6	7
4.		1	2	3	4	5	6	7
5.		1	2	3	4	5	6	7
6.		1	2	3	4	5	6	7
7.		1	2	3	4	5	6	7
5.		1	2	3	4	5	6	7
6.		1	2	3	4	5	6	7
7.		1	2	3	4	5	6	7

13. I was motivated to do my best.

1	2	3	4	5	6	7
Strongly agree			Neither agree nor disagree			Strongly disagree

14. Did you have any physical difficulty or discomfort while using the computer?

☐ Yes

☐ No

If Yes, describe ____

15. Did you feel like you were doing anything you think you "weren't supposed to" during the task, such as daydreaming or taking longer than you needed? (It's perfectly ok if you did, just let us know! ☺)

16. Did you feel rushed through any part of the study?

☐ Yes

☐ No

If Yes, when did you feel rushed?

____

____

17. Were you comfortable in the lab?

1	2	3	4	5	6	7
Not at all comfortable			Somewhat comfortable			Extremely comfortabl

16. Was the experimenter clear in telling you what you were supposed to do?

1	2	3	4	5	6	7
Not at all clear			Somewhat clear			Extremely clear

17. Do you have any general suggestions about how we could improve our study?

____

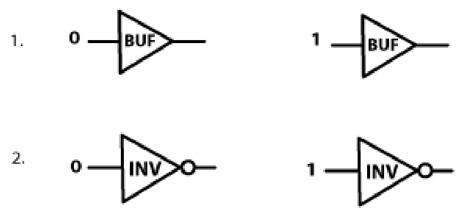
____

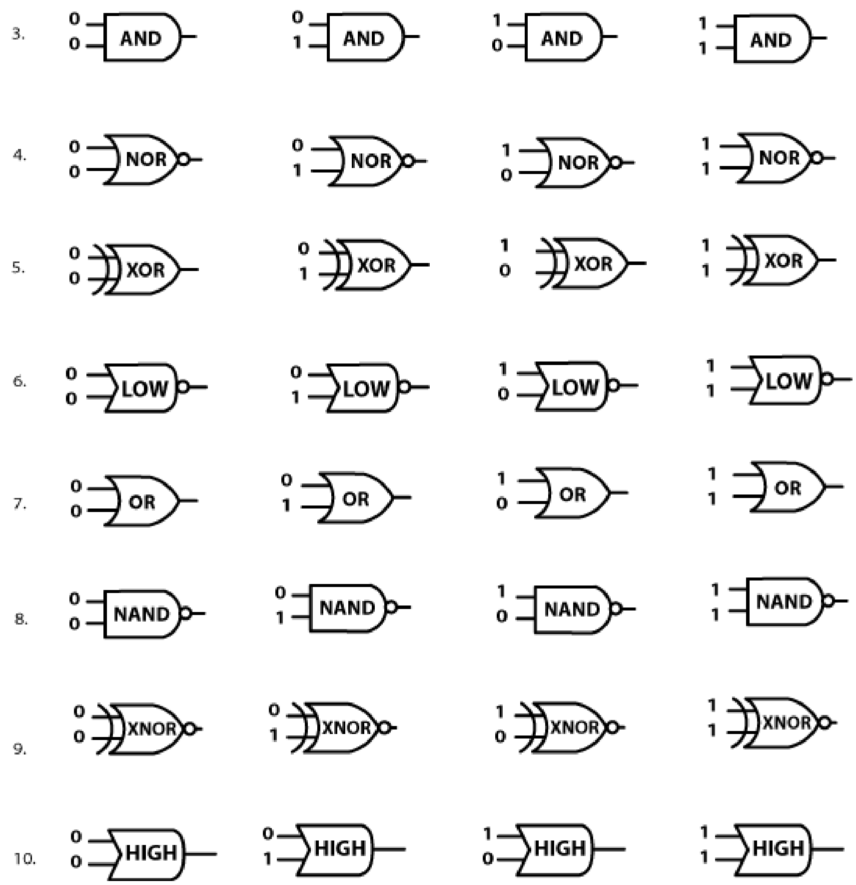
____

____

This page contains every possible logic gate.

Please provide the answer to each logic gate. If you do not know, put an X as your answer.





Addition to Exit Interview – Study 2

On the first day, which gates gave you feedback immediately after answering? Please guess if you don't know for sure (circle).

AND OR XOR NAND BUF INV XNOR HIGH LOW NOR

What strategy did you use for learning the no-feedback gates? Please describe.

____

____

____

____

____

____

How did you use the end of block feedback, with the summary of each gate percent correct, to learnt he gates?

____

____

____

____

____

____

Did you feel **able** to use the end of block feedback? In other words, was it of any use in improving your accuracy? Please give us your thoughts!

____

____

____

We thought that giving only end of block feedback might make you **work "harder"** to learn those gates (because you didn't get much help from the system.)

Did less feedback make you put forth more effort for those gates? Please be honest and explain what you did.

☐ No

☐ Yes

Explain:

____

____

____

____

Addition to Exit Interview – Study 3 & 4

Please describe the two types of feedback you received on the first day.

Now, please tell us which one made you learn the gates more thoroughly and why.

Which feedback type made you put forth more effort to learn the gates?

☐ 1 Presentation of all gate possibilities. (ex: $\begin{matrix}1\\0\end{matrix} = 0, \ldots$)

☐ 2 Just being told "correct" or "incorrect"

☐ 3 I needed the same effort for both types

Please explain:

____

____

____

____

When you were given the table of all possibilities, did you try to create a verbal rule for that gate or did you try to memorize the examples?

☐ 1 Create a verbal rule

☐ 2 Memorized the possibilities for that gate

☐ 3 Both equally

☐ 4 Both, but created a verbal rule more often

☐ 5 Both, but memorized the examples more often

☐ 6 Other ____

When you were told "correct" or "incorrect" only, did you try to think of all the possibilities in the form of a table or did you try to create a verbal rule?

☐ 1 Created a table of possibilities

☐ 2 Create a verbal rule for that gate

☐ 3 Both equally

☐ 4 Both, but created a table more often

☐ 5 Both, but created a verbal rule more often

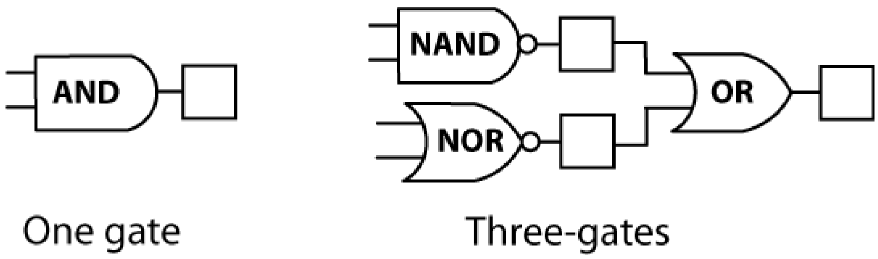
☐ 6 Other ____

Debriefing Information

Thank you very much for participating in this experiment. We could not conduct our research without your help. This study was designed to examine what type of feedback is best suited for the tasks you did. Some of the people in the study received feedback after every logic gate problem they solved. Others only got feedback after a block of 21 logic gate problems.

When you returned for the second session, you solved logic gates with no feedback. This was to measure how the feedback you received in the first session helped you learn to solve the gates. Remembering the gates after a week was an indicator of your learning.

Previous research has shown that people tend to learn a task more thoroughly when they *do not* get feedback after every answer. However, we think that may depend on the difficulty of the task you are trying to learn. Therefore, some participants learned the task with only one gate each time while others learned to solve many gates at once (see below).



Having three gates made the task more difficult, and we predicted that this group of people would learn better with feedback after every gate.

You also completed several different ability tests, such as the vocabulary test. We are interested in determining whether individuals need different amounts of feedback while learning. For example, some people may learn better when given very little feedback on how they are doing while others need feedback more often.

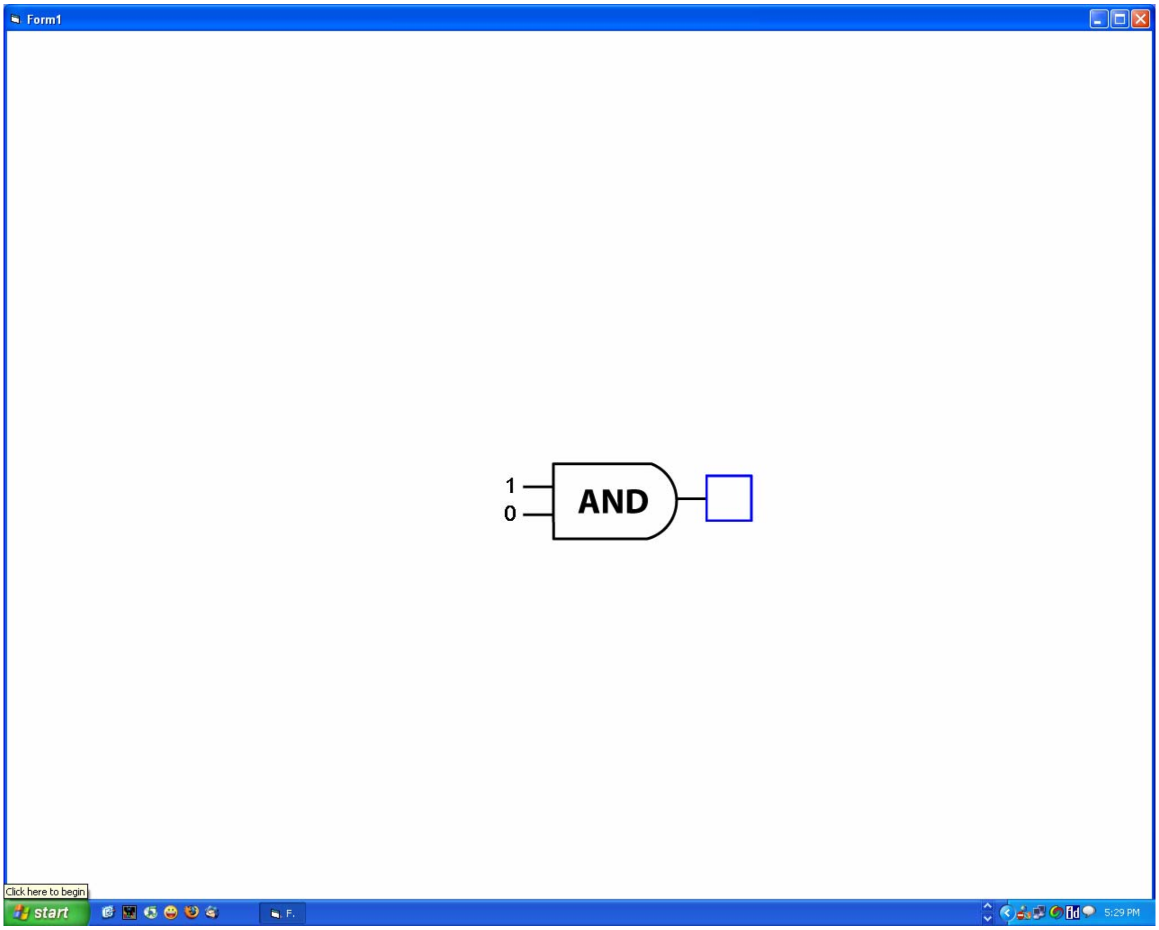
We will use the results of this study to recommend the best type of feedback for groups of people on particular tasks. Again, we would like to thank you for your participation. If you have any questions about the study or any suggestions, please do not hesitate to contact one of the directors of the project:

Anne McLaughlin:404-894-8344, **Dr. Wendy A. Rogers**:404-894-6775, **Dr. Arthur D. Fisk**:404-894-6066

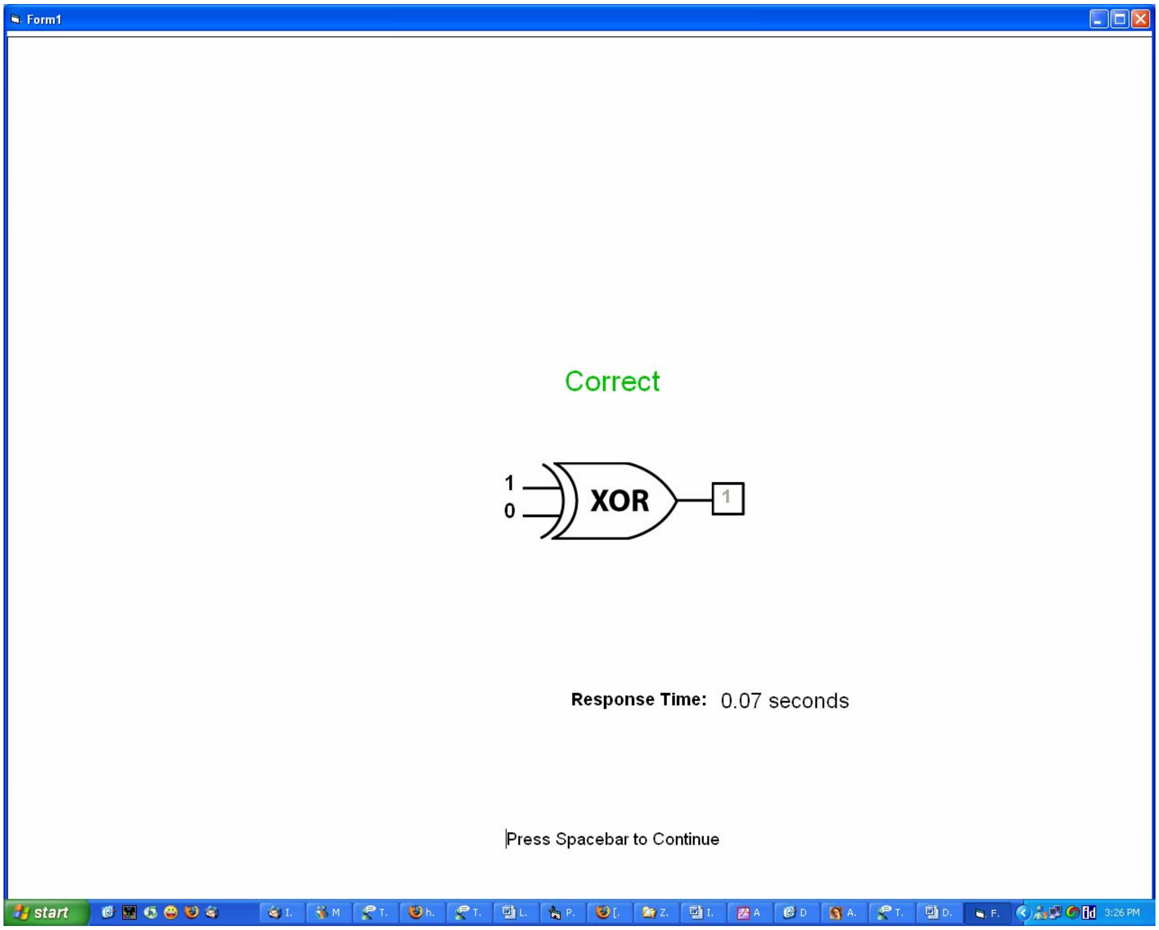
You will either receive a report of the final results of this study via our yearly Human Factors & Aging Lab newsletter or may check on the progress of the study via our website: **www.hfaging.org**

APPENDIX B: EXPERIMENTAL STIMULI

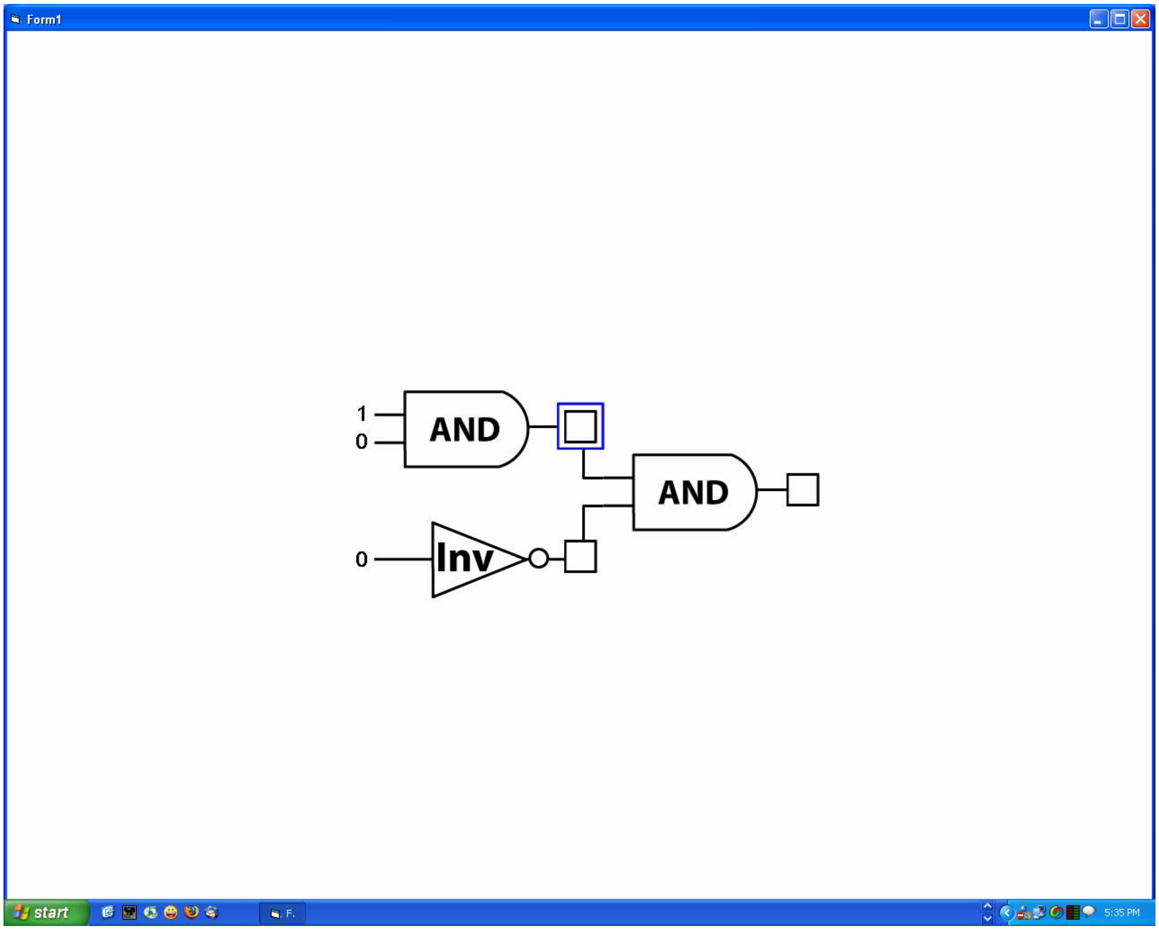
Low-load ("Simple") logic gates screen



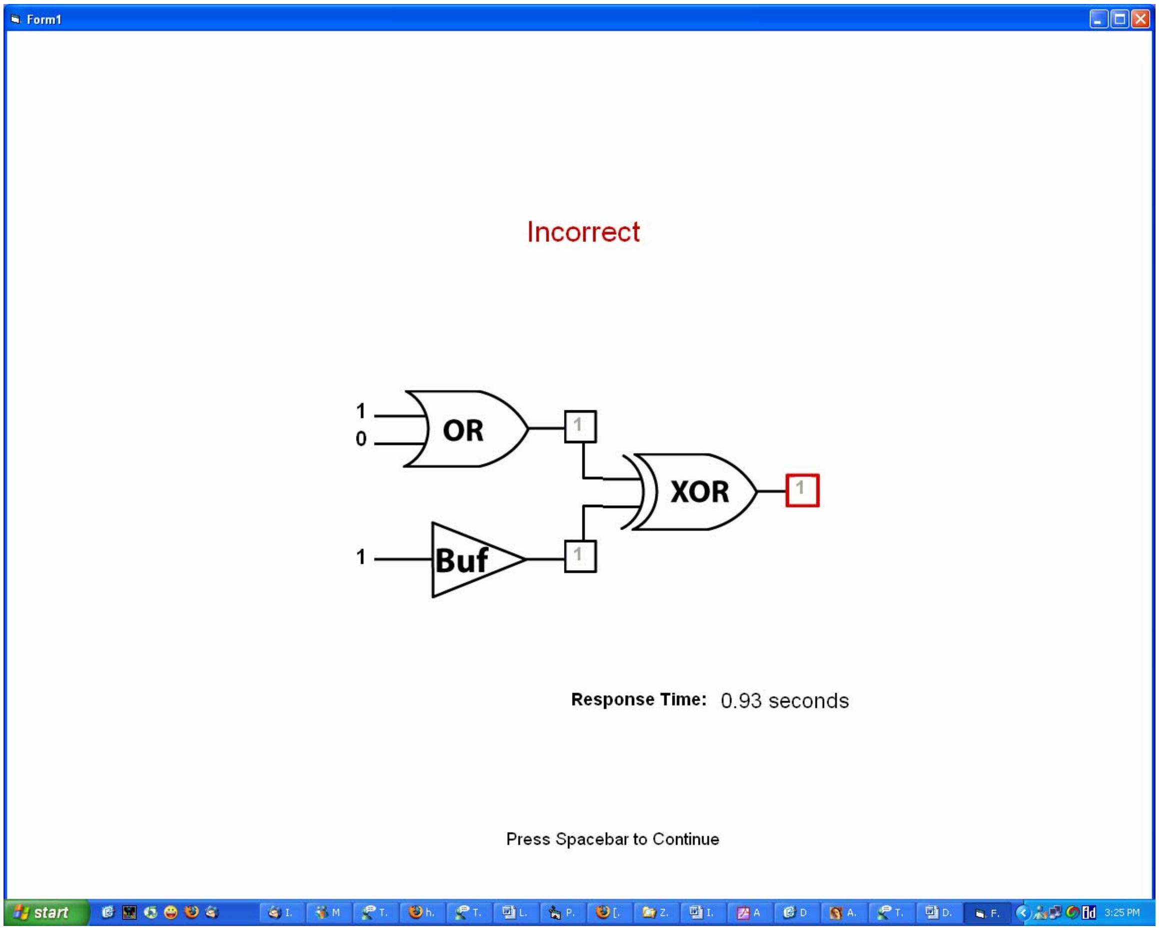
KCR Feedback Support: Every trial, Low-load ("Simple")



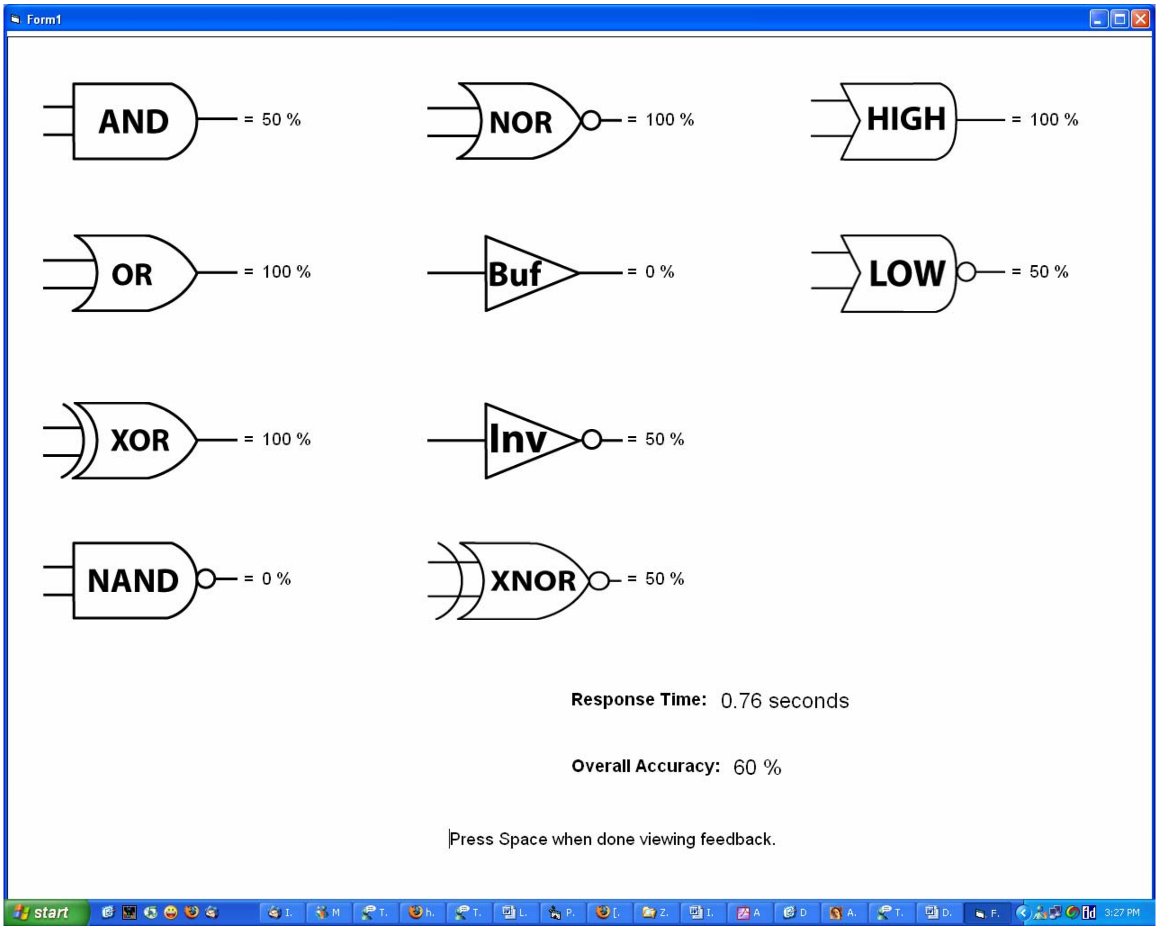
High-load ("Complex") logic gates screens



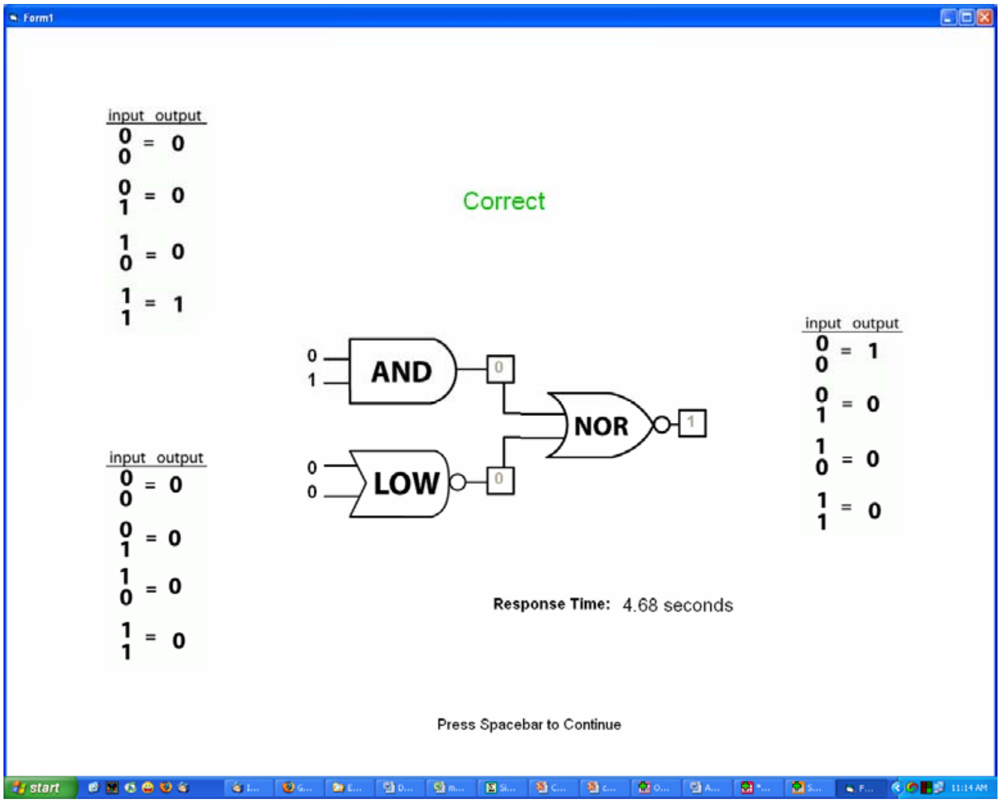
KCR feedback Support: Every trial, High-load("Complex")



Summary feedback Support: End of block feedback



Truth Table feedback Support: Complex task -Every Trial



APPENDIX C: CREATE II BATTERY OF TESTS

Ability Tests - CREATE Group Testing

Test	Ability	Administration time
Number Comparison Test	Perceptual Speed	3 minutes
California Verbal Learning Test	Memory Long-Term	25 minutes
Meaningful Memory	Memory Long-Term	5 minutes
Shipley Vocabulary	Verbal Ability	10 minutes
Alphabet Span	Working Memory	25 minutes
Letter Sets (ETS)	Reasoning/Induction	14 minutes
Information (WAIS-III)	Crystallized Intelligence	7 minutes

Ability Tests - CREATE Individual Testing

Test	Ability	Administration time
Snellen Vision	Far and near vision	5 minutes
Mini Mental State Examination (MMSE)	Cognition	5 minutes
Earscan Audiometer	Hearing	10 minutes
Digit Symbol Substitution	Attention/Concentration	5 minutes
Trailmaking Tests	Attention/Concentration	10 minutes
Choice Reaction Time Task	Psycho-motor speed	5 minutes
Simple Reaction Time Task	Psycho-motor speed	5 minutes

APPENDIX D: PILOT STUDIES

Pilot Study 1

Overview

The purpose of this pilot study was to determine whether feedback could (and should) be manipulated as a within-participant variable. If feedback support were to be a within participant variable, the gates would need to be counterbalanced. It was likely that some gates were more difficult to learn than others. If high feedback support were given only to difficult or easy gates this would confound experiments. Data from this study served as a starting point for designing counterbalance groups.

Method

Participants

Participants in the pilot study were Georgia Tech student volunteers. Participants were given extra credit in their psychology courses at the rate of 1 credit per hour of participation.

Participant working memory capacity was measured via the Ao-span. Most participants in this study were mid to high working memory capacity.

Variables

Independent variables included task load (simple, complex) and feedback support condition (Summary feedback, KCR). All manipulations were between participants

Dependent measures included motivation level, accuracy on each gate, and response time for each gate. As with the other experiments, accuracy was the primary dependent measure.

Procedure

Participants followed the same procedure as Experiment 1 with one exception: there was no trial time limit for this study. There was no time limit on answering gates.

Results

Counterbalance information. Certain gates produced similar accuracies, response times, and subjective ratings of difficulty (Table 3). The matches were as follows: AND/OR, INV/BUF, and NAND/NOR. Objective measures such as accuracy and response time tended to match perceived difficulty. From these data, the groups BUF, OR, XOR, NAND (counterbalance 1) and NOR, INV, AND, LOW (counterbalance 2) were created.

Table 3

Pilot Data for Logic Gates

	Accuracy		Response Time		Perceived Difficulty		
Gate	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Sample size
OR	0.93	0.25	1934.37	1971.37	2.09	1.3	11
AND	0.95	0.22	1823.26	1790.24	2.09	1.64	11
XOR	0.88	0.32	2143.32	2430.95	2.36	1.43	11
INV	0.97	0.17	1718.87	1604.31	1.18	0.4	11
BUF	0.97	0.18	1561.74	1295.01	1	0	11
NOR	0.87	0.33	3418.55	2864.24	4.18	1.78	11
NAND	0.89	0.31	2917.78	2656.04	4.18	1.6	11

Summary

The main purpose of this pilot study was to divide the gates into counterbalance groups for the experiments to follow. Difficulty-balanced counterbalance groups were created from these data.

However, a concern with manipulating feedback support within participants was that enough trials in each condition would not fit into an acquisition session. Thus, it needed to be discovered whether participants would demonstrate learning of the gates in the number of trials provided and how long to present each trial. This was addressed by the second pilot study.

Pilot Study 2

Overview

The main purpose of this pilot study was to determine whether accuracy could be made the main variable of interest by limiting the allowed response time per trial. An initial time limit of 4500ms was chosen for those in the simple task condition and 15000ms for the complex condition. Participant accuracies and mean trial times were recorded to generate a time window for future experiments. This final window was determined for each task load condition by taking the mean response time across Pilot Study 2 and adding one standard deviation. This was to produce time pressure during the study but also to allow time to answer accurately. The time windows generated for Experiment 1 by this pilot study were 3000ms for the simple task and 11000ms for the complex task.

Method

Participants

Eleven young adults took part in this pilot study. Participants in the pilot study were Georgia Tech student volunteers. Participants were given extra credit in their psychology courses at the rate of 1 credit per hour of participation. Most participants were in the 2^{nd} and 3^{rd} quartiles for WMC. Participant demographics and ability test data are available in Table 3.

Variables

Independent variables in included Task Load (simple, complex) and Feedback condition (summary feedback, KCR). Working memory capacity was measured but not manipulated.

Procedure

Procedure was the same as Experiment 1.

Results

Mean response time in acquisition was 1916ms (SD = 974) for the simple task and 7379 (SD = 2971) for the complex task. Mean response time across the study was 2054(1496) for the simple task and 6421(3187) for the complex task.

Summary

A response window of 3 seconds was allotted for the simple task in future experiments. A response window of 11 seconds was allotted for the complex task in future experiments.

Appendix E – ANOVA Tables for all experiments

ANOVA table 1

Exp. 1 - Analysis of variance for performance at beginning and end of acquisition

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	10.47*	0.52	2	5.49*
Task Load(TL)	1	0.713	123.24*	2	59.70*
WMC x TL	1	0.43	0.75	2	0.61
	32			31	
			Within participants		
Block (B)	1	65.76*	140.27*	2	104.24*
B x WMC	1	5.87*	0.2	2	2.94
B x TL	1	2.87	47.54*	2	24.98*
B x WMC x TL	1	0.62	0.179	2	0.40
Feedback (FB)	1	3.84*	0.51	2	2.63
FB x WMC	1	0.168	0.232	2	0.15
FB x TL	1	1.43	0.007	2	0.77
FB x WMC x TL	1	1.26	0.036	2	0.62
FB x B	1	6.97*	0.102	2	3.94*
FB x WMC x B	1	0.19	7.78*	2	5.75*
FB x TL x B	1	1.36	0.019	2	0.77
FB x WMC x TL x B	1	2.28	9.98*	2	4.84*
	32			31	

ANOVA table 2

Exp. 1 - Analysis of variance for performance at 10-minute retention

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	16.17*	1.00	2	7.83*
Task Load(TL)	1	0.42	52.36*	2	28.45*
WMC x TL	1	0.87	0.47	2	0.53
	32			31	
			Within participants		
Feedback (FB)	1	7.66*	1.40	2	3.74*
FB x WMC	1	0.20	0.00	2	0.10
FB x TL	1	0.12	0.66	2	0.03
FB x WMC x TL	1	0.53	0.17	2	0.03
error	32			31	

ANOVA table 3

Exp. 1 - Analysis of variance for performance at 1-week retention

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	7.38*	0.89	2	3.64*
Task Load(TL)	1	0.06	52.60*	2	27.08*
WMC x TL	1	0.12	0.68	2	0.35
	32			31	
			Within participants		
Feedback (FB)	1	13.68*	0.80	2	9.21*
FB x WMC	1	0.29	1.03	2	0.52
FB x TL	1	0.98	1.25	2	0.81
FB x WMC x TL	1	0.00	1.75	2	0.95
error	32			31	

ANOVA table 4

Exp. 1 - Analysis of variance for Retention Change Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	11.43*	1.09	2	7.14*
Task Load(TL)	1	1.30	1.91	2	1.98
WMC x TL	1	2.05	0.01	2	1.02
	32			31	
			Within participants		
Feedback (FB)	1	0.27	3.23*	2	1.76
FB x WMC	1	0.00	0.45	2	0.22
FB x TL	1	0.01	0.94	2	0.47
FB x WMC x TL	1	0.18	0.05	2	0.12
error	32			31	

ANOVA table 5

Exp. 1 - Analysis of variance for Content Change Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	2.08	2.56	2	2.65
Task Load(TL)	1	0.45	3.64	2	1.83
WMC x TL	1	0.01	0.19	2	0.10
	32			31	
			Within participants		
Feedback (FB)	1	5.27*	0.42	2	3.80*
FB x WMC	1	0.01	0.83	2	0.43
FB x TL	1	0.92	0.90	2	1.40
FB x WMC x TL	1	0.16	0.78	2	0.38
error	32			31	

ANOVA table 6

Exp. 1 - Analysis of variance for Content Post Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	14.64*	1.55	2	7.63*
Task Load(TL)	1	0.00	36.13*	2	17.55*
WMC x TL	1	0.24	1.02	2	0.59
	32			31	
			Within participants		
Feedback (FB)	1	1.05	0.04	2	0.56
FB x WMC	1	0.44	0.45	2	0.39
FB x TL	1	3.33*	0.17	2	1.64
FB x WMC x TL	1	0.18	1.08	2	0.66
error	32			31	

ANOVA table 7

Exp. 1 - Analysis of variance for Load Change Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
Between participants					
WMC	1	1.22	0.35	2	0.81
Task Load(TL)	1	6.95*	267.75*	2	130.82*
WMC x TL	1	0.17	0.39	2	0.29
	32			31	
Within participants					
Feedback (FB)	1	3.35	1.70	2	3.28*
FB x WMC	1	0.36	0.24	2	0.23
FB x TL	1	0.53	1.74	2	1.45
FB x WMC x TL	1	4.17*	0.11	2	2.04
error	32			31	

ANOVA table 8

Exp. 1 - Analysis of variance for Load Post Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	13.76*	1.12	2	6.85*
Task Load(TL)	1	2.75	118.79*	2	61.85*
WMC x TL	1	0.00	0.29	2	0.14
	32			31	
			Within participants		
Feedback (FB)	1	2.33	1.71	2	1.49
FB x WMC	1	0.00	1.64	2	0.91
FB x TL	1	0.05	5.57*	2	2.85
FB x WMC x TL	1	4.32*	3.16	2	2.75
error	32			31	

Experiment 3 – Practice Effects for Low WMC Participants on a Complex task

ANOVA table 9

Exp. 3 - Analysis of variance for performance at 1-week retention

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
Feedback (FB)	1	8.23*	0.93	1	4.10*
error	8			7	

ANOVA table 10

Exp. 3 - Analysis of variance for Retention Change Score

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
Feedback (FB)	1	0.08	4.49*	1	4.09*
error	8			7	

ANOVA table 11

Exp. 3 - Analysis of variance for Content Change Score

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
Feedback (FB)	1	1.96	1.49	1	2.37
error	8			7	

ANOVA table 12

Exp. 3 - Analysis of variance for Load Change Score

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
Feedback (FB)	1	1.40	0.58	1	0.62
error	8			7	

ANOVA table 14

Exp. 3 - Analysis of variance for performance on Block 1 Session 1 to Block 10 Session 10

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
			Within participants		
Block (B)	1	15.74*	27.45*	1	15.72*
Feedback (FB)	1	18.57*	0.39	1	9.15*
B X FB	1	2.45	0.48	1	1.16
error	9			8	

ANOVA table 14

Exp. 3 - Analysis of variance for performance at Session 1 10-minute retention to Session 10 10-minute retention

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
		Within participants			
Block (B)	1	10.77*	8.67*	1	15.72*
Feedback (FB)	1	17.84*	0.48	1	9.15*
B X FB	1	6.99*	0.79	1	1.16
error	9			8	

Experiment 4 – KCR versus Truth Tables for Low and High WMC Participants on a Simple or Complex Task

ANOVA table 15

Exp. 4 - Analysis of variance for performance at beginning and end of acquisition

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	18.26*	0.24	2	9.04*
Task Load(TL)	1	0.00	346.32*	2	177.72*
WMC x TL	1	6.97*	0.16	2	3.43*
error	52			51	
			Within participants		
Block (B)	1	60.78*	198.67*	2	146.49
B x WMC	1	1.82	0.59	2	1.36
B x TL	1	1.73	92.30*	2	49.03
B x WMC x TL	1	0.00	0.01	2	0.00
Feedback (FB)	1	1.48	0.10	2	0.73
FB x WMC	1	3.79m	2.63	2	2.57
FB x TL	1	0.08	0.02	2	0.04
FB x WMC x TL	1	0.10	0.87	2	0.57
FB x B	1	0.36	3.11	2	1.63
FB x WMC x B	1	0.00	1.58	2	0.78
FB x TL x B	1	0.00	0.49	2	0.24
FB x WMC x TL x B	1	0.20	0.17	2	0.20
error	52			51	

ANOVA table 16

Exp. 4 - Analysis of variance for performance at 10-minute retention

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	11.98*	1.43	2	6.35*
Task Load(TL)	1	3.89*	224.62*	2	114.27*
WMC x TL	1	2.36	0.53	2	1.35
error	52			51	
			Within participants		
Feedback (FB)	1	0.43	0.32	2	0.54
FB x WMC	1	8.85*	0.06	2	4.59*
FB x TL	1	0.36	0.17	2	0.20
FB x WMC x TL	1	2.99	0.47	2	2.29
error	52			51	

ANOVA table 17

Exp. 4 - Analysis of variance for performance at 1-week retention

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	31.67*	0.01	2	15.87*
Task Load(TL)	1	2.15	160.28*	2	78.77*
WMC x TL	1	5.23*	0.05	2	2.58
error	52			51	
			Within participants		
Feedback (FB)	1	4.89*	1.55	2	3.43*
FB x WMC	1	0.54	0.13	2	0.31
FB x TL	1	1.72	0.93	2	1.42
FB x WMC x TL	1	0.17	0.34	2	0.27
error	52			51	

ANOVA table 18

Exp. 4 - Analysis of variance for Retention Change Score

	F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	0.84	0.90	2	1.00
Task Load(TL)	1	1.69	3.31	2	2.16
WMC x TL	1	0.34	0.07	2	0.23
error	52			51	
			Within participants		
Feedback (FB)	1	5.70*	0.47	2	2.89*
FB x WMC	1	5.36*	0.45	2	2.71
FB x TL	1	0.00	0.28	2	0.14
FB x WMC x TL	1	5.27*	0.06	2	2.59
error	52			51	

ANOVA table 19

Exp. 4 - Analysis of variance for Content Change Score

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	0.01	0.00	2	0.01
Task Load(TL)	1	0.80	2.73	2	1.46
WMC x TL	1	0.01	0.49	2	0.63
error	52			51	
			Within participants		
Feedback (FB)	1	0.16	1.60	2	0.80
FB x WMC	1	0.65	0.08	2	0.33
FB x TL	1	0.01	1.90	2	0.98
FB x WMC x TL	1	0.50	1.41	2	1.11
error	52			51	

ANOVA table 20

Exp. 4 - Analysis of variance for Content Post Score

	F			F	
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	17.47*	0.24	2	8.61*
Task Load(TL)	1	3.42	103.26*	2	51.57*
WMC x TL	1	4.14*	0.49	2	2.21
error	52			51	
			Within participants		
Feedback (FB)	1	7.83*	0.39	2	3.95*
FB x WMC	1	0.24	1.40	2	0.85
FB x TL	1	1.90	1.01	2	1.35
FB x WMC x TL	1	0.09	0.63	2	0.37
error	52			51	

ANOVA table 21

Exp. 4 - Analysis of variance for Load Change Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	1.49	0.10	2	0.82
Task Load(TL)	1	4.40*	669.09*	2	328.19*
WMC x TL	1	0.09	0.00	2	0.04
error	52			51	
			Within participants		
Feedback (FB)	1	3.57	0.19	2	2.12
FB x WMC	1	0.28	0.96	2	0.75
FB x TL	1	2.88	0.33	2	1.87
FB x WMC x TL	1	0.01	0.80	2	0.43
error	52			51	

ANOVA table 22

Exp. 4 - Analysis of variance for Load Post Score

		F			F
	df	Accuracy	Response Time	df	Multivariate Effects
			Between participants		
WMC	1	22.08*	0.37	2	11.52*
Task Load(TL)	1	8.78*	129.94*	2	73.15*
WMC x TL	1	4.33*	0.06	2	2.12
error	52			51	
			Within participants		
Feedback (FB)	1	0.01	1.10	2	0.62
FB x WMC	1	0.05	0.02	2	0.03
FB x TL	1	7.16*	0.55	2	4.77*
FB x WMC x TL	1	0.18	0.29	2	0.18
error	52			51	

REFERENCES

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288-318.

Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review, 89*, 369-406.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1996). A simple theory of complex cognition. *American Psychologist, 51*, 355-365.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036-1060.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: lessons learned. *The Journal of the Learning Sciences, 4*, 167-207.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about Knowing* (pp. 185-205). Cambridge, MA: MIT Press.

Bjork, R. A. (June, 2006). Making things hard on yourself: Desirable difficulties in theory and practice. Psychology Learning and Teaching Conference, York, United Kingdom.

Blackwell, J. R., & Newell, K. M. (1996). The informational role of knowledge of results in motor learning. *Acta Psychologica, 92*, 119-129.

Brackbill, Y., Bravos, A., & Starr, R. H. (1962). Delay-improved retention of a difficult task. *Journal of Comparative and Physiological Psychology, 55*, 947-952.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: a simple proposal with multiple implications. *Review of Research in Education, 24*, 61-100.

Brosvic, G. M., Dihoff, R. E., Epstein, M. L., & Cook, M. L. (2006). Feedback facilitates the acquisition and retention of numerical fact series by elementary school students with mathematics learning disabilities. *Psychological Record, 56*, 35-54.

Burton, J. K., Moore, D. M., & Magliaro, S. (2004). Behaviorism and instructional technology. In D. H. Jonassen (Ed.), *Handbook of research on educational* *communications and technology* (2nd ed., pp. 46-67). Mahwah: Lawrence Erlbaum Associates.

Cabanac, M., & Russek, M. (2000). Regulated biological systems. *Journal of Biological systems, 8*, 141.

Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 517-526.

Carlson, R. A., & Yaure, R. G. (1990). Practice schedules and the use of component skills in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 484-496.

Clariana, R. B. (1990). A comparison of answer until correct feedback and knowledge of correct response feedback under two conditions of contextualization. *Journal of Computer-Based Instruction, 17*, 125-129.

Cope, P., & Simmons, M. (1994). Some effects of limited feedback on performance and problem-solving strategy in a logo microworld. *Journal of Educational Psychology, 86*, 368-379.

Czaja, S. J., Sharit, J., Charness, N., Rogers, W. A., & Fisk, A. D. (2002). CREATE: Center for research and education on aging and technology enhancement. In R. Pieper, M. Vaarama & J. L. Fozard (Eds.), *Gerontechnology: Technology and Aging - Starting Into the Third millennium* (pp. 54-61). Aachen: Shaker Verlag.

Davis, W. D., Carson, C. M., Ammeter, A. P., & Treadway, D. C. (2005). The interactive effects of goal orientation and feedback specificity on task performance. *Human Performance, 18*, 409-426.

Dixon, R. A., Hultsch, D. F., & Hertzog, C. (1988). The Metamemory in Adulthood (MIA) Questionnaire. *Psychopharmacology Bulletin, 24*, 671-688.

Droit-Volet, S., & Izaute, M. (2005). The effect of feedback on timing in children and adults: the temporal generalization task. *Quarterly Journal of Experimental Psychology: A, Human Experimental Psychology, 58*, 507-520.

Elshout, J. J. (2006). Discussion of the zone of tolerable problemacity. In A. C. McLaughlin (Ed.) (pp. Email from Jan Elshout corroborating that the term ZTP was not adopted, but that Sweller's cognitive load theory was a similar idea and the eventual follow-up). Atlanta.

Farquhar, J. D., & Regian, J. W. (1994). *Effects of knowledge representation during computer-based training of console operation skill*.

Feldman-Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological*

Bulletin, 130, 553-573.

Fisk, A. D., Hertzog, C., Lee, M. D., Rogers, W. A., & Anderson-Garlach, M. (1994). Long-term retention of skilled visual search: do young adults retain more than old adults? *Psychology and Aging, 9*, 206-215.

Fisk, A. D., McGee, N. D., & Giambra, L. M. (1988). The influence of age on consistent and varied semantic-category search performance. *Psychology and Aging, 3*, 323-333.

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

Fox, B. A. (1991). Cognitive and interactional aspects of correction in tutoring. In P. Goodyear (Ed.), *Intelligent Tutoring System: Evolutions in Design, Teaching, Knowledge, and Intelligent Tutoring* (pp. 149-172). Norwood, NJ: Albex Publishing Corporation.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612.

Goodman, J. S., & Wood, R. E. (2004). Feedback specificity, learning opportunities, and learning. *Journal of Applied Psychology, 89*, 809-821.

Guadagnoli, M. A. (Writer) (1999). An assimilation of selected research in motor learning: Academic Cafe Television.

Guadagnoli, M. A., & Dornier, L. A. T., R. D. (1996). Optimal length for summary knowledge of results: the influence of task-related experience and complexity. *Research Quarterly for Exercise and Sport, 67*, 239-248.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior, 36*, 212-224.

Hancock, P. A., & Ganey, H. C. N. (2003). From the inverted-U to the extended-U: the evolution of a law of psychology. *Journal of Human Performance in Extreme Environments, 7*, 5-14.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors, 31*, 519-537.

Hertzog, C., & Dixon, R. A. (1994). Metacognition development in adulthood and old age. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition*: MIT Press.

Hertzog, C., & Dunlosky, J. (2005). Aging, metacognition, and cognitive control. In B. H. ross (Ed.), *Psychology of Learning and Motivation*. San Diego, CA: Academic Press.

Hertzog, C., & Robinson, A. E. (2005). Metacognition and intelligence. In O. Wilhelm &

R. W. Engle (Eds.), *Handbook of Understanding and Measuring Intelligence* (pp. 101-119): Sage Publications.

Ho, L., & Shea, J. B. (1978). Effects of relative frequency of knowledge of results on retention of a motor skill. *Perceptual and Motor Skills, 46*, 859-866.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657-690.

Kluger, A. N., & Denisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254-284.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*, 279-307.

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*, 79-97.

Kyllonen, P. C., & Woltz, D. J. (1989). The role of cognitive factors in cognitive skill acquisition. In P. L. A. R. Kanfer, & R. Cudeck (Ed.), *Abilities, Motivation, and Methodology*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lee, T. D., & Carnahan, H. (1990). When to provide knowledge of results during motor learning: scheduling effects. *Human Performance, 3*, 87-105.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information integration category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 650-662.

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human-Computer Interaction, 5*, 381-413.

Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided learning by doing. *Cognition and Instruction, 13*, 315-372.

Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences, 2*, 277-305.

Meyer, B. (1998). *Toward and understanding of age0related use of incidental consistency*. Paper presented at the Human Factors and Ergonomics Society Annual Meeting.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MS: Harvard University Press.

Newell, K. M. (1991). Motor skill acquisition. *Annual Review of Psychology, 42*, 213-237.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2003). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, 1-5.

Pashler, H., Zarow, G., & Triplett, B. (2003). Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 1051-1057.

Ponds, R. W. H. M., & Jolles, J. (1996). The metamemory in adulthood (MIA) questionnaire: structure and effects of age, sex, and education. *Psychology and Aging, 11*, 324-332.

Rogers, W. A. (1996). Assessing age-related differences in the long-term retention of skills. In A. D. F. W. A. Rogers, & N. Walker (Ed.), *Aging and skilled performance: advances in theory and applications* (pp. 185-200). Mahwah, NJ: Lawrence Eerlbaum Associates.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2000). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology, 52*, 453-470.

Rothrock, L., & Kirlik, A. (2003). Inferring rule-based strategies in dynamic judgment tasks: toward a noncompensatory formulation of the lens model. *IEEE Transactions on Systems, Man, and Cybernetics, 33*, 58-72.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95*, 355-386.

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

Schmidt, R. A., Lang, C. A., & Young, D. E. (1990). Optimizing summary knowledge of results for skill learning. *Human Movement Science, 9*, 325-348.

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

Schmidt, R. A., Young, D. F., Swinnen, S., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 352-359.

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

Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors, 27*, 285-300.

Schooler, L. J., & Anderson, J. R. (1990). *The disruptive potential of immediate feedback.* Paper presented at the Proceedings of the Twelfth Annual Conference of the Cognitive Science Society.

Schroth, M. L. (1992). The effects of delay of feedback on a delayed concept formation transfer task. *Contemporary Educational Psychology, 17*, 78-82.

Schroth, M. L. (1997). Effects of frequency of feedback on transfer in concept identification. *American Journal of Psychology, 110*, 71-79.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *5*, 179-187.

Smith, P. L., & Ragan, T. J. (1993). Designing instructional feedback for different learning outcomes. In J. V. Dempsy & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 75-103). Englewood Cliffs, NJ: Educational Technology Publications.

Snow, R. E. (1989). Aptitude-treatment interaction as a framework for reserach on individual differences in learning. In R. J. S. P. L. Ackerman, R. Glaser (Ed.), *Learning and individual differences: advances in theory and research* (pp. 13-60). New York: W. H. Freeman and Company.

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology, 43*, 583-626.

Stuart, A., & Ord, K. (1994). Classical inference and the linear model. In *Kendall's Advanced Theory of Statistics* (Vol. 2a, pp. 753-759). London: Hodder Arnold.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*, 351-362.

Swindell, L. K., & Walls, W. F. (1993). Response confidence and the delay retention effect. *Contemporary Educational Psychology, 18*, 363-375.

Szalma, J. L., Hancock, P. A., Dember, W. N., & Warm, J. S. (2006). Training for vigilance: The effect of knowledge of results format and dispositional optimism and pessimism on performance and stress. *British Journal of Psychology, 97*, 115-135.

Taylor, A., & Noble, C. E. (1962). Acquisition and extinction phenomena in human trial-and-error learning under different schedules of reinforcing feedback. *Perceptual and Motor Skills, 15*, 31-44.

Thorndike, E. L. (1911). Laws and hypotheses for behavior. In *Animal intelligence: Experimental studies* (pp. 241-280).

Trow, W. C. (1928). Recall vs. repetition in the learning of rote and meaningful material. *The American Journal of Psychology, 40*, 112-116.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language, 28*, 127-154.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505.

van Merrienboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38*, 5-13.

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

Vollmeyer, R., & Rheinberg, F. (2005). A surprising effect of feedback on learning. *Learning and Instruction, 15*, 589-602.

Wightman, D. C., & Lintern, G. (1985). Part-task training for tracking and manual control. *Human Factors, 27*, 267-283.

Williams, A. (2007). Use of online classrooms for remedial students at community colleges. Instructor: Georgia Perimeter College, Atlanta, GA.

Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 677-691.

Wishart, L. R., Lee, T. D., Cunningham, S. J., & Murdoch, J. E. (2002). Age-related differences and the role of augmented visual feedback in learning a bimanual coordination pattern. *Acta Psychologica, 110*, 247-263.

Wrisberg, C. A., & Wulf, G. (1997). Diminishing the effects of reduced frequency of knowledge of results on generalized motor program learning. *Journal of Motor Behavior, 29*, 17-26.

Wulf, G., & Schmidt, R. A. (1994). Contextual interference effects in motor learning: evaluating a KR-usefulness hypothesis. In J. R. N. a. R. Seiler (Ed.), *Movement and sport: psychological foundations and effects* (Vol. 2, pp. 304-309). Sankt Augustin, Germany: Academia Verlag.

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review, 9*, 185-211.

Wulf, G., Shea, C. H., & Matschiner, S. (1998). Frequent feedback enhances complex motor skill learning. *Journal of Motor Behavior, 30*, 180-192.