EVALUATIVE PROCESSES AS THE COGNITIVE BASIS FOR THE CONTEXTUAL INTERFERENCE EFFECT: IMPLICATIONS FOR A UNIFIED THEORY OF SKILL ACQUISITION

by

Leroy Darren Kruisselbrink

B.A., University of Alberta, 1989
M.A., University of Victoria, 1993

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the School of Physical Education

We accept this dissertation as conforming to the required standard

Dr. G. H. Van Gyn, Supervisor (School of Physical Education)

Dr. B. L. Howe, Departmental Member, (School of Physical Education)

Dr. J. O. Anderson, Outside Member (Department of Educational Psychology and Leadership Studies)

Dr. M. E. J. Masson, Additional Member (Department of Psychology)

Dr. J. M. Deakin, External Examiner (School of Physical and Health Education, Queen's University)

Supervisor: Dr. G. H. Van Gyn

ABSTRACT

Cognitive effort has been identified as the basis of the contextual interference (CI) effect (Lee, Swinnen, & Serrien, 1994). It has been argued that higher levels of cognitive activity related to either the evaluation of movement information (encoding) or the retrieval of movement plans are demanded by the conditions of random rather than blocked practice. Current theories of skill acquisition appear to more heavily emphasize evaluative/encoding than retrieval processes. Furthermore, a review of evidence from research on the knowledge of results (KR) and observational learning implicates the critical role of evaluative processes as well. A series of three experiments was designed to (a) test the isomorphism of evaluative processes and cognitive effort within the contextual interference paradigm, and (b) use the CI phenomenon as a way to explore the more general role of evaluative processes in motor skill acquisition.

The typical CI effect was replicated in Experiment 1 using three spatial variations of a multi-segment arm movement task. However, this experiment featured the co-occurrence of differential demands for both encoding variability and retrieval practice. In Experiment 2, one of the variations from Experiment 1 was practiced within the context of two unrelated video games. The results showed that no acquisition or retention performance differences emerged between blocked and random practice groups. These results suggest that the role of retrieval practice as the basis of the CI effect should be questioned.

Experiment 3A replicated Experiment 1 with pairs of blocked and random groups. In Experiment 3B, using a second set of three spatial variations, an attempt was made to

reduce differential encoding variability while keeping differential retrieval practice intact between one pair of blocked and random groups (verbalize groups). The blocked group was required to evaluate and associate the features of each pattern variation during the acquisition phase, and to verbalize their thoughts. A random group was also required to verbalize the cognitive strategies they used to learn the patterns. The co-occurrence of differential encoding variability and retrieval was maintained for the remaining pair of blocked and random groups (control groups). The results of Experiment 1 were replicated in Experiment 3A, and by the control groups in Experiment 3B. In Experiment 3B, relative retention and retention performance improved to a greater extent for the blocked verbalize than the blocked control group. However, relative retention and retention performance were not similar between the blocked verbalize and random groups, indicating that the evaluation of pattern variations in isolation does not appear to be an effective intervention with which to reduce the demands for differential encoding variability between blocked and random groups.

Analysis of qualitative data obtained in Experiment 3B indicated differences between blocked and random groups in the degree to which the features of the spatial patterns were compared, suggesting that information derived from single task evaluation may not be equivalent to the information derived from multiple task comparison. Results are discussed within Glenberg's (1979) component levels theory.

Insight into the nature of the cognitive processes underlying the CI effect may have implications for a general explanation of motor skill acquisition. The relationship between cognitive effort, the development of knowledge, and skill acquisition is outlined in a preliminary framework for a unified theory of skill acquisition. The ability of the proposed framework to incorporate a range of experimental data and theoretical views is discussed.

Examiners:

Dr. G. H. Van Gyn, Supervisor (School of Physical Education)

Dr. B. L. Howe, Departmental Member, (School of Physical Education)

Dr. J. O. Anderson, Outside Member (Department of Educational Psychology and Leadership Studies)

Dr. M. E. J. Masson, Additional Member (Department of Psychology)

Dr. J. M. Deakin, External Examiner (School of Physical and Health Education, Queen's University)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF DIAGRAMS

DEDICATION

To Geri,
who cultivated my interest in sport skills
into an academic profession

INTRODUCTION

It is well accepted that repetition plays a critical role in the acquisition of skills, both motor and cognitive. However, many other variables interact with repetition to impact skill acquisition as well. One of these is the sequencing of skills within a practice session. Research on the acquisition of novel motor skills has shown that the sequencing of practice trials can affect both the immediate and longer term performance of the skills to be learned (Lee & Magill, 1983; Shea & Morgan, 1979).

Practicing multiple tasks or skills in a random order (e.g. ACB CAB CBA) typically produces poorer acquisition performance. However, learners are better able to retain what they learned. Conversely, practicing tasks in a blocked order (e.g. AAA BBB CCC) typically produces superior acquisition performance, but learners are unable to retain these levels of performance even over the short term. The effect of blocked and random practice on the acquisition and retention of motor performance is called the contextual interference effect, where contextual interference (CI) refers to the degree of intertask interference produced by the context of practice (Battig, 1972, 1979).

In the motor domain, the CI effect has been shown in numerous laboratory investigations (Carnahan, Van Eerd & Allard, 1990; Lee & Magill, 1983; Lee, Wulf & Schmidt, 1992; Sekiya, Magill, Sidaway & Anderson, 1994; C.H. Shea, Kohl & Indermill, 1990; Shea & Morgan, 1979; Shea & Zimny, 1983, 1988; Wood & Ging, 1991; Wright, 1991; Wright, Li & Whitacre, 1992). This has led to the conclusion that the CI effect is a robust effect.

Theoretical explanations of the contextual interference effect

Two theories have been advanced to explain the contextual interference effect. Lee and Magill (1983, 1985) have hypothesized an action plan reconstruction theory, while Shea and his associates (Shea & Morgan, 1979; Shea & Zimny, 1983, 1985; Shea & Wright, 1991) have proposed a theory of processing elaboration and distinctiveness.

Theory of action plan reconstruction

Lee and Magill's (1983, 1985) theory of action plan reconstruction (TAPR) is an adaptation of Jacoby's (1978; Cuddy & Jacoby, 1982) account for the spacing effect in the verbal behavior domain. The TAPR is based on the notion that "forgetting between repetitions of goal-directed actions will depress acquisition performance yet promote retention" (Lee & Magill, 1985, p. 8). Organizing and implementing the action plans of different tasks between task repetitions, as occurs during random practice, interferes with the maintenance of the to-be-repeated-task in working memory. Thus, prior to implementing the action plan solution for a task, it must be regenerated in working memory. Although more difficult in the short term, the need to continually reformulate action plans benefits learners in the long run because the cognitive processes associated with action plan regeneration receive regular practice.

During blocked practice, however, the lack of intertrial interference during the acquisition phase allows an action plan to be retained in working memory. As a result, the cognitive processes involved in the regeneration of action plans can be largely bypassed during the acquisition period. Since these processes receive little practice during the acquisition phase, the ability to regenerate action plans during retention and transfer tests suffers. Lee and Magill (1983) suggested that the frequency with which action plan regeneration occurs during the acquisition period accounts for the differences in acquisition, retention and transfer performance demonstrated by blocked and random practice subjects.

Theory of processing elaboration and distinctiveness

The theory of processing elaboration and distinctiveness (TPED) proposed by Shea and his associates (Shea & Morgan, 1979; Shea & Zimny, 1983, 1988; Shea & Wright, 1991) is an adaptation of Battig's (1979) account for the contextual interference effect, which was also generated from within the verbal behavior domain. The use of multiple and variable processing strategies during the acquisition period was central to Battig's explanation. Shea and Zimny (1983) suggested that the use of multiple processing strategies to encode task features, combined with the variability with which these strategies were used permitted random practice subjects to encode task relevant information more elaborately and distinctively than blocked practice subjects. According to Shea and his associates, random practice forces subjects to regularly compare, differentiate and store the features of the tasks to be learned in order to maintain a reasonable level of acquisition performance. Retention and transfer performance is hypothesized to benefit from this activity because over time learners generate a greater number of retrieval paths to the appropriate response for given environmental cues. Conversely, subjects' need to actively differentiate and store task features is reduced during blocked practice, making it more challenging for them to easily access an appropriate response.

Cognitive effort in the CI effect

Each theory suggests that the demands for cognitive effort evokes greater engagement in the scope and depth of cognitive processes during random than blocked practice (Lee, Swinnen & Serrien, 1994; Young, Cohen & Husak, 1993). Both theories also suggest that the differences in the demand for cognitive effort, brought about as a function of blocked and random practice contexts, is responsible for the differences in the extent to which acquisition performance is retained. However, the specific cognitive operations supporting cognitive effort differ in each theory.

Emphasis in the TAPR is given to the preparation and formation of movement. The TAPR assumes that working memory can only maintain the action plan representation of one task at a time. Random practice causes tasks to interfere with each other, whereas this is not the case during blocked practice. Repeated demands for preparatory cognitive operations during the acquisition of motor skills, which is characteristic of random practice, amounts to retrieval practice (Schmidt & Bjork, 1992). Retrieval practice is believed to benefit retention performance by enhancing the learner's ability to retrieve and regenerate an action plan when called for (Lee, 1988). Thus, the TAPR emphasizes cognitive operations associated with the retrieval, regeneration and reorganization of action plan information in working memory. The implication from TAPR is that retention performance will benefit to the extent that the context of practice provides opportunities to actively regenerate action plans.

Central to the TPED is the need to differentiate and encode those features that distinguish one task from another so that tasks can be more easily recalled on future trials. Cognitive operations involved in this process include the comparison, evaluation, and differentiation of tasks along any number of meaningful dimensions, combined with the integration of newly discovered distinctions with existing knowledge. Thus, in the TPED, emphasis is given to cognitive operations that become engaged following the execution of movement. Repeatedly engaging these cognitive operations during the acquisition phase amounts to practice evaluating and encoding task relevant information. Evaluative practice is believed to benefit retention performance by creating multiple access routes to the appropriate response when cued, which implies that skills will be better retained when the context of practice that provides learners with opportunities to evaluate, differentiate and integrate task relevant information.

To date, it is unclear as to whether preparatory or evaluative cognitive operations underlie the CI effect. A number of theoretical accounts of skill acquisition provide support for evaluative cognitive operations (Anderson, 1983, 1993; Bereiter & Scardamalia, 1993; Ericsson, Krampe & Tesch-Romer, 1993; Ericsson & Charness, 1994; Higgins, 1991; Ohlsson, 1996). Specific to the CI effect, Wright (1991) and Wright et al. (1992) have found that requiring blocked practice participants to assess the similarity in the features of to-be-learned tasks facilitates the recall of the tasks on retention trials. However, the nature of the relationship between evaluation and cognitive effort in the acquisition of motor skills has not been tested directly.

Purpose and hypotheses

The purpose of this research project was to determine if evaluative processes constitute the cognitive operations that demand cognitive effort and produce the CI effect. To accomplish this, a series of three experiments were designed. The first experiment replicated the traditional CI effect using a multi-segment arm movement task. In the second experiment, task similarity was manipulated in order to separate the theoretical predictions of the TAPR and the TPED and, subsequently, the relative contribution of preparatory and evaluative processes. Part of the problem in determining the nature of cognitive effort required by each type of cognitive operation lies within the CI paradigm itself. The use of task variations performed within an environment that does not change from trial to trial or task to task produces low demands for both preparatory and evaluative cognitive operations during blocked practice, but high demands for both during random practice. Therefore, preparatory and evaluative cognitive operations must be dissociated in order to determine the relative contribution of each. Using dissimilar tasks, the TPED predicts that no CI effect should emerge because the tasks do not share common dimensions that can be compared, evaluated, and differentiated. However, the TAPR predicts that the CI effect should emerge because, regardless of the degree of dissimilarity between tasks, retrieval practice is still a characteristic of random practice.

In the third experiment, subjects learned a set three task variations in each of the two parts of the experiment. As a manipulation in the second part of the experiment, half of the blocked practice participants purposefully evaluated the features of the tasks during the acquisition phase and generated meaningful associations with existing knowledge in order to mimic the evaluative processes carried out by random practice participants. In addition, these participants were required to verbalize this information. Half of the random practice participants were required simply to verbalize the strategies they were using to remember the tasks during the acquisition phase. If evaluation and differentiation underlie the CI effect, then requiring these processes of blocked practice participants should produce retention performance similar to that of random practice participants. Thus, by manipulating different variables within the contextual interference paradigm, the theoretical predictions generated by the TPED and the TAPR can be tested.

REVIEW OF LITERATURE

Motor skill acquisition is a general learning phenomenon in which learners' capacity to perform motor tasks is enhanced through practice. Preparatory and evaluative cognitive processes have been implicated as being involved in the skill acquisition process (Lee, Swanson, and Hall, 1991; Lee, Swinnen, et al., 1994; Schmidt & Bjork, 1992). Contextual interference is one of a number of practice variables that have been shown to influence the acquisition of motor skills. Other variables include, but are not limited to, knowledge of results, and observational learning.

With this in mind, the purpose of this review of literature is two fold; first the theoretical bases of skill acquisition and their implications will be explored in order to gain a greater understanding of how cognitive processes are involved in skill acquisition, this will be followed by an examination of the relationship between these theoretical implications and empirical evidence from the literatures of observational learning (OL), knowledge of results (KR), and CI. The primary objectives of this review are to use these analyses to suggest a unified view of the cognitive processes involved in motor skill acquisition, and to argue that using the CI paradigm is an ideal way to test the predictions of this view.

Information, knowledge, and motor performance

Adaptive behavior involves an actor interacting with his or her physical and social environment with the intention of accomplishing goals (Allard & Burnett, 1985; Gentile, 1987). At a molar level of analysis, this interaction occurs through action, which simply refers to a change in the state of the actor-environment relationship. Action represents how the actor has chosen to solve a problem generated by a goal, and can vary in effectiveness (Higgins, 1991). At a finer grain level of analysis, movement is the means by which actions unfold (Allard & Burnett, 1985; Gentile, 1987; Higgins, 1991) and involves coordinating and implementing a pattern of movement for an action to be realized. Motor skill implies that actions, and the movements of which they are comprised, successfully solve movement problems posed by goals. Thus, skillful behavior is the consistent, effective, and economical application of available resources in service of solving a motor problem.

Higgins (1991) argues that actors' level of skill is constrained by a number of factors including: (a) their capacity to abstract relevant information from the environment, from which action goals can be defined (environmental analysis), (b) their capacity to determine the most appropriate course of action in relation to the goal and relative to current abilities, which depends on knowledge of possible action alternatives coupled with an understanding of current morphological and biomechanical constraints (task analysis and self-knowledge), and (c) their capacity to organize, implement, and monitor movement, which depends on an understanding of the external forces involved, the internal forces needed, and how these two sets of forces interact (the force problem). Each of these capacities is driven by information. Since the information to which actors are attuned represents some proportion of the total amount of available information (Gibson, 1979), the information driving each capacity can only operate within the confines of the evolutionary state of the actor's cognitive system.

Higgins (1991) argues further that four basic sources of information interact to drive these capacities. These sources are (a) the performance environment (exteroceptive information), (b) the position of the actor's body relative to itself (proprioceptive information), (c) the position of the actor's body relative to the environment (expproprioceptive information) and (d) the actor's cognitive capacities (episodic, semantic, and procedural memory). Knowledge, as the term will be used in this paper, is synonymous with actors' control over access to information from these four sources (Allard, 1993; Magill, 1998). Thus, knowledge necessarily defines the boundaries within which action and its underlying movements emerge. In this sense, the efficacy of motor skill performance is dependent on the status of actors' knowledge.

Expert - novice differences

Empirical evidence supporting a knowledge based view of motor skill performance can be found in the literature on expert-novice differences. Because of large variations in the criteria used in defining "experts" and "novices" between studies in this area, the terms "more skilled" and "less skilled" will be used in describing this research in order to retain the relative distinction in skill level as intended by the researcher without reference to absolute expertise. What follows is a list of the principles that have been generated from data gathered from expert-novice research. When considered together, the results of expert-novice research shows that the state of knowledge of more skilled performers distinguishes them from less skilled performers.

1. Expertise is context sensitive (Abernethy, 1993a). In other words, expertise is a situation based, rather than an individual based, phenomena.

2. More skilled performers are better able to abstract relevant regulatory information from the performance environment (e.g., Abernethy, 1993b; Goulet, Bard & Fleury, 1989; Magill, 1998a; Ripoll, 1991; Williams & Davids, 1998). Some studies have shown differences in the visual search characteristics of more and less skilled performers (e.g., Ripoll, Kerlirzin, Stein, & Reine, 1995; Williams & Davids, 1998; see Abernethy, 1993b and Williams, Davids, Burwitz & Williams, 1993 for reviews), while others have not (e.g., Abernethy, 1990; Abernethy & Russell, 1987b; Shank & Haywood, 1987). Williams & Davids (1998) suggest that skill related differences in visual search strategy are likely more a function of context demands rather than any ubiquitous difference related to level of skill. Regardless of visual search strategy, more skilled performers have been shown to consistently make better use of available information.

3. More skilled performers make use of earlier occurring information to anticipate upcoming events (e.g. Abernethy, 1987; Abernethy & Russell, 1984, 1987a; Buckolz, Prapavesis, & Fairs, 1988; Isaacs & Finch, 1983; Jones & Miles, 1978; Williams & Davids, 1998).

4. More skilled performers are better able to recognize meaningful patterns within their domain. (e.g., Allard, Graham, & Paarsulu, 1980; Borgeaud & Abernethy, 1987; Helsen & Pauwels, 1993; Starkes, 1987; Starkes & Deakin, 1984).

5. More skilled performers represent patterns of information at more abstract, principled levels (Allard & Burnett, 1985; Allard, Deakin, Parker, & Rodgers, 1993).

6. More skilled performers make better decisions under time stress (Helsen & Pauwels, 1993).

7. More skilled performers consider more action alternatives (Abernethy, 1993a; Abernethy, Neal, Engstrom & Koning, 1993).

8. More skilled performers possess greater declarative knowledge of their domain (Allard et al., 1993; Williams & Davids, 1995).

9. Knowledge of situational probabilities affects the performance of skilled performers (Abernethy, 1993a; Allard, 1993)

10. More skilled performers possess superior self-monitoring skills (Abernethy, 1993a).

Taken together, the results of expert-novice research suggest that knowledge, in the multidimensional sense of the word, plays a critical role in motor performance (Allard, 1993; Magill, 1998a; see also Bereiter & Scardamalia, 1993).

Acquiring skill

As knowledge appears to be a central component of motor performance, there remains the issue of skill acquisition. If motor performance operates within the boundaries of existing knowledge, then expanding these boundaries should allow skill to develop. This specific proposition is supported by current theories of skill acquisition.

Anderson's ACT* production system

One of the most comprehensive theories of skill acquisition is Anderson's (1983) Adaptive Control of Thought, version star (ACT*). In this theory, performance behavior is goal oriented, which means that actors are viewed as operating in a world of problems in which current states are separated from goal states in both time and space. Performance behavior, whether cognitive or motor, is adaptive to the extent that it reduces the conflict between current and the goal states. Resolution of this conflict requires the application of a problem solving procedure. To find an appropriate procedure, a search is conducted to match the information about current conditions residing in working memory with existing knowledge. The procedure that is eventually selected can be guided through the interpretive application of declarative knowledge or the direct application of procedural knowledge. Interpretive application of procedures occurs when specific production rules are not available. In the absence of specific production rules, problems must be solved through the use of general problem solving operators or analogy. Direct application of procedural knowledge occurs when information contained in the condition side of specific production rules match the information in working memory about current conditions. Production rules are condition-action pairs housed in procedural memory that link specific environmental conditions to specific action procedures.

Anderson (1983) suggests that in the process of acquiring new skills, knowledge will initially be in declarative form and applied interpretively. Once it has been determined that a procedure or set of procedures is functionally adaptive, the compilation of production rules begins. The process of knowledge compilation involves a pair of processes. Proceduralization refers to the shift in the control of performance behavior from declarative to procedural knowledge bases. It is a gradual process in which specific condition-action productions replace interpretive application as the dominant mode of procedural control. Composition is the process by which a sequence of productions are collapsed into a single new one.

Anderson (1983) maintains that production rules are also subject to tuning. Production tuning involves generalization, where production rules are made broader in their range of applicability, or discrimination, where production rules are made narrower in their range of applicability. Generalization involves encoding only general features on the condition side of the rule so that when certain general features characterize the performance environment a general procedure is implemented. Discrimination involves encoding more specific features in a production rule in response to encoding deficiencies on either the condition or action side of the rule. The process of discrimination occurs when the application of a production rule was deemed inappropriate (condition deficiency), or when the production rule was deemed appropriate but it was misapplied (procedure deficiency). In response to error information, further constraints are added to the production rule to more narrowly control the situations in which it becomes activated, or to more narrowly constrain how the procedure is implemented. The process of discrimination is completely reliant on feedback, as it can only occur in response to error information. This information may be provided by internal or external sources. Thus, in Anderson's ACT* production system, the essence of procedural learning involves learning by doing and the assessment of error.

In summary, Anderson's (1983) ACT* production system is a knowledge based system in which skill develops through the acquisition and gradual proceduralization of declarative knowledge, followed by the more gradual process of tuning production rules. Since productions have been proposed as the basis of action in ACT*, this system clearly advances the notion that knowledge development underlies skill acquisition.

Ohlsson's production system

Although Ohlsson's (1996) production system is similar to ACT*, it has some unique aspects. Action again is viewed as goal oriented, the purpose of which is to change aspects of a particular performance situation in the direction of goal achievement. This is accomplished through perceive-decide-act cycles, where the actor perceives the current situation, decides on what action to take (consciously or otherwise) and implements it. Action alters the performance situation, thereby initiating another perceive-decide-act situation.

The decision phase of the cycle is dependent on knowledge, which Ohlsson (1996) has labeled as practical knowledge. Practical knowledge links actions, goals, and performance context. A three way relationship is needed because actions are related to goals, but the relationship is context dependent. The same goal in a different context may require a different action. The relationship between goals, context and action are encoded in production rules such that $G, C \rightarrow A$, when striving for goal G in context C, action A will be considered. If all the features perceived in a particular performance context are present in the condition side of a production rule, the rule becomes activated. All relevant production rules are activated in parallel. A successful method for completing a task may involve anywhere from a few to thousands of production rules. Methods can be task specific or general in nature depending on the generality or specificity with which the production rules define the task environment. The most effective methods for solving action problems define the task environment narrowly, and thus, by definition, will only be applied in restricted situations.

Learning in Ohlsson's (1996) framework involves the successive elimination of performance errors. Central to this process is the need for actors to compare the intended and perceived outcomes of actions to detect error. It is assumed that discrepancies between intended and perceived outcomes contain information that can be used to correct error.

Error detection requires actors to evaluate the outcomes of action to determine the discrepancy between the actual and expected outcome. Expected outcomes are based on declarative knowledge which prescribes the situation that ought to be the case upon goal completion. Thus, declarative knowledge provides a system of constraints that operate as a self-monitoring device with which to judge the efficacy of action. The violation of any constraint constitutes error, whereas action will be considered successful when no constraints have been broken.

Since performance behavior is generated by practical knowledge, performance errors indicate a fault with the practical knowledge that generated the action. The problem is thought to lie with a production rule that is overly general. Rule revision involves incorporating more specific information about the goal, the context or the action into the faulty production rule, thus preventing it from displacing a production rule specifying a more appropriate action. The process by which rules are revised occurs as a function of blame assignment and error attribution. Blame assignment involves locating the faulty rule, while error attribution involves identifying a feature from the context in which the error occurred that will exclude it from being activated in similar contexts in the future. Each time an actor engages in blame assignment and error attribution represents a single learning event. Ohlsson (1996) attributes the lengthy period of skill acquisition to the number of learning events that are required to successively eliminate performance errors since the method for goal achievement may contain many production rules, and each rule may contain multiple faults. He states that "as learning continues, the condition side of each rule becomes more and more specialized and the rule will consequently become active in fewer and fewer situations [until] finally, the rule is activated only in those situations in which the action it evokes is the correct thing to do" (p. 248).

In summary, Ohlsson's (1996) production system is also a knowledge based system where skill develops through the gradual refinement of action alternatives. Since action is generated through the application of practical knowledge, refinement in action must be preceded by refinement of practical knowledge.

Parallel Distributed Processing and connectionist models

A neural network model based on the principles of connectionism and parallel distributed processing (PDP) contains layers of processing units connected together to form a network. The purpose of these layers of processing units is to transform a pattern of input into a desired pattern of output (Horak, 1992; Masson, 1991; McCloskey & Cohen, 1989; Shea & Graf, 1994). In the context of a connectionist model, knowledge is represented as a distributed pattern of weighted connections between the processing units of a network. Learning involves adjusting the connection weights between processing units, which subsequently alters the pattern of activation across the network and the corresponding output pattern.

A typical network consists of an input layer, a number of hidden layers, and an output layer. Units of the input layer are activated by receiving input from outside the network. Input units transmit this pattern of activation to units in an adjacent layer along a weighted set of connections. The value of these weights signify the strength of connection between units. For example, if an input unit received an input activation of 1 and was connected to a hidden unit in the next layer by a strength of .4, the input received by the unit in the second layer from the unit in the first layer would be .4. The level of activation for a single processing unit is calculated as the sum of all weighted inputs received from connections with units of a previous layer. A pattern of activation travels through the network until it finally reaches the output layer. The pattern of output is then compared to the desired pattern of output. If the two are incongruent, the difference is calculated and connection weights in the network are adjusted by a backwards propagation algorithm in order to make the output pattern more closely correspond to the target output pattern on the next trial.

The effects of blocked and random training schedules has been simulated a number of times by connectionist networks (Horak, 1992; Masson, 1991; McCloskey & Cohen, 1989; Shea & Graf, 1994). Similar to the CI literature, the acquisition performance of a network is superior under blocked than random (or alternating) conditions. However, when the network is tested for its ability to generate target outputs for each input pattern that was provided during training, it produces greater errors when it had been trained under blocked than random practice conditions. This has been explained by the fact that the pattern of connection weights in a network are adjusted each time the output pattern does not match the target output (Horak, 1992; Masson, 1991; McCloskey & Cohen, 1989; Shea & Graf, 1994). Since the same connections are used in response to each input pattern, when the weights are adjusted in response to output errors for one input pattern, the network's response to other input patterns will be altered as well. Blocked training conditions allow the pattern of connection weights in a network to stabilize more quickly thereby producing output patterns that are more consistent and accurate during training. However, because the pattern of connection weights in a network change with different patterns of input, training the network with multiple input patterns presented in blocks of trials produces connection weights that are appropriate for the last input pattern, but not for previous ones. When the network is tested on each input pattern it received during training, the connection weights of the network produce large amounts of error for all but the last input pattern received during training (Horak, 1992; McCloskey & Cohen, 1989). Training conditions in which input patterns are presented randomly, or alternately in serial fashion, invoke greater initial changes to the pattern of connection weights in the network, however once they stabilize, the pattern of connection weights is better able to produce accurate output patterns for all of the input patterns provided during training.

The network designed by Shea and Graf (1994) represents an advance in the use of connectionist networks to model motor behavior. Their model includes multiple hidden layers similar to the models reported previously. However, unlike the previous multiple layer models in which all hidden layers contribute to the solution every time (single level processing), Shea and Graf created a multiple layer, multiple level model. A multiple level model means that activation can make its way from the input to output layer by means of one of three levels of processing, which they arbitrarily call planning, programming and execution. In the transformation of a pattern of activation across input units to a desired pattern of activation across output units, the network can utilize either one, two, or all three of the hidden layers between the input and output layers. All three hidden layers are utilized by the planning level, the second and third hidden layers are utilized by the programming level, and the third hidden layer is utilized by the execution level. Thus, the activation received by the output layer from the units of the third hidden layer can be modified if the second layer is involved, and the effect of the second and third hidden layers can be modified if the first hidden layer is involved.

In the Shea and Graf (1994) model, once the units of the input layer are activated, the execution level is provided with the first opportunity to produce the target output. An internal criterion is set to determine whether the solution is satisfactory and output should be generated. If the solution falls outside the criterion range, "the program level of processing is brought to bear on the task and works in parallel with the execution level to achieve a satisfactory solution. If a satisfactory solution is achieved, output is generated. Otherwise, the planning level contributes its influence in parallel with the first two levels and output is generated" (pp. 78-79). The results of their simulation of the CI effect replicated those customarily reported in CI experiments; less error was produced during training under blocked than random practice conditions, while more error was produced at retention. The results also showed that more layers of processing units were used during random than blocked training conditions, which Shea and Graf suggest provides "a measure for the level of cognitive processing brought to bear on the learning of a motor task" (p. 85).

In summary, according to PDP connectionist models, knowledge is represented by the pattern of connection weights among the processing units of a network at a given point in time. Reducing the discrepancy between target and produced output occurs as a function of adjusting the connection weights among processing units, thereby altering the state of the system. Thus, the state of the network determines the accuracy of the output it can produce.

Logan's Instance Theory of Automatization

Logan's (1988) instance theory of automatization is also a knowledge based theory of skill acquisition, but of a different kind than the production systems of Anderson and Ohlsson. Logan's theory deals mainly with the mechanisms involved in generating a response, and how they change with experience.

Performance behavior in Logan's theory is viewed as the product of either algorithm based or memory based processes. Solutions to novel problems are initially produced as a result of algorithms. With experience, specific solutions are learned for specific problems. Each encounter with a problem is encoded along with its solution such that over time, an accumulation of stimulus-response instances develops. Response generation gradually shifts from reliance on algorithm based processing to the direct retrieval of instances from memory as the speed of memory retrieval rivals and surpasses that of the algorithm.

A number of assumptions, each of which Logan (1988) maintains is supported by empirical data, make this system work. It is assumed that encoding into memory and retrieval from memory are obligatory, unavoidable consequences of attention. That is, attending to a stimulus is sufficient to commit it to memory and to retrieve from memory whatever it has been associated with in the past. In addition, each encounter with a stimulus is individually encoded, stored, and retrieved. In this sense, each instance in which a stimulus is encountered is viewed as a processing episode that produces an episodic memory trace. Each trace contains information about the goal that was sought, the stimulus that was encountered in pursuing the goal, how the stimulus was interpreted relative to the goal, and the response that was made. Upon future encounters with the stimulus within the context of the same goal, instances containing these stimulus and goal features are retrieved as an unavoidable consequence of attention. At the same time, the use of the algorithm begins to generate a solution. Whichever process is completed first controls the response. As the sample of instances becomes larger, the likelihood of retrieval processing producing a response ahead of algorithm processing is enhanced until eventually memory retrieval finishes first every time. Logan suggests that memory retrieval itself is also a race with each instance competing against all others. In accordance with statistical principles regarding sampling, over time, the increase in the size of instance samples produces the speed up and negative acceleration effect characteristic of the ubiquitous power law of practice (Newell & Rosenbloom, 1981).

In summary, according to Logan's instance theory, skill develops as the basis for generating solutions shifts from a reliance on algorithm based to memory based processing. The accumulation of instances containing specific responses for specific situations produced by experience is what makes instance theory a knowledge based theory. Instance theory assumes that algorithm and memory mechanisms do not change with practice, the only thing that changes is the size of the data base from which memory retrieval occurs. Thus, in the words of Logan (1988), "automatization reflects the development of a domain-specific knowledge base; nonautomatic performance is limited by a lack of knowledge rather than by the scarcity of resources" (p. 501).

Higgins framework for motor behavior and skill acquisition

Although they most certainly contain implications for the acquisition of motor skill, the theories of Anderson (1983), Ohlsson (1996) and Logan (1988) were developed primarily for cognitive skill acquisition. By way of contrast, the framework provided by Higgins (1991) was developed specifically to explain motor skill performance and acquisition.

Higgins (1991) offers a holistic perspective of motor behavior in which the individual is viewed as an actor in the environment with movement as a means by which goals are achieved. The actor is viewed as a problem solver since the concept of goal achievement occupies a central role in the framework. Problems stem from the pursuit of goals, which arise as a function of the actor interacting with his or her environment. Problems define the task at hand. Action represents a class of movements operating on the environment to effect change in the direction of goal achievement. They represent modes of available interaction between the actor and the environment.

The nature of the actions involved in the solution of problems is dependent on the effectiveness with which actors analyze the task, generate an appropriate coordinative relationship between their body and the environment, and control the spatial and temporal allocation of internal and external forces underlying its production. The degree to which actors are effective in these areas represents their level of skill. Thus, skillful behavior reflects the actor's understanding of the demands of the task posed by the problem as well as reflecting how available resources are perceived in relation to these demands. Resources include the "morphological and psychological characteristics of the individual, the structural characteristics of objects and events in the external environment, and the characteristics of the field of external forces inherent in the environment" (Higgins, 1991, pp 127-8). Solutions are only as effective as the knowledge from which they are generated, therefore action choice and movement topology reflect the evolutionary state of the actor's knowledge.

Motor skill acquisition involves the interaction of discovery and mastery over the knowledge bases that support behavior. This results in a greater understanding of the demands of the task, and greater understanding and control over the relationship between body segments, over the relationship between the body and the environment, and over the nature of the forces that underlie the production of movement.

In summary, motor skill performance in Higgins' (1991) framework involves the interaction of knowledge with information derived from many sources. Thus, the development of motor skill occurs in relation to the development of knowledge.

Commonalties among theories

The common feature underlying these theories is that the structure of knowledge must be altered in order to produce an associated change in skill. In this sense, each theory can be considered as a knowledge based theory. A central concept encapsulating these knowledge based theories is that actors are continuously involved in solving problems since they are constantly interacting with the environment in the pursuit of goals. Anderson (1983) and Ohlsson (1996) suggest that knowledge constrains action through the application of production rules that link what performers perceive (environmental conditions) with performance procedures (action). The acquisition of skill requires the addition of information on either side of a production rule to restrict the range of conditions in which it will be considered. Logan (1988) suggests that knowledge can be represented by the data base of instances from which responses are selected. Over time, as this data base grows, it results in performance behavior that is faster, more accurate, and less variable. Higgins (1991) suggests that knowledge drives task analysis which, in turn, constrains response selection, and that knowledge also constrains the organization and control of movement. Knowledge is also implicated as a central concept in connectionist models in that the state of the network, characterized by the pattern of connection weights between units and layers of units, determines the output that is produced. In order to modify the output of a connectionist network, the distribution of connection weights, which represent knowledge, must be altered.

Cognitive activities considered to be critical for skill acquisition

It has been argued that knowledge plays a central role in motor performance, and that knowledge development is the basis of the acquisition of skill. The major question to be answered, then, is how this knowledge develops. What types of cognitive activity are critical to enable the acquisition of skill? A variety of cognitive activities considered to be critical for learning have been proposed within the theories that have previously been reviewed (Anderson, 1983; Higgins, 1991; Logan, 1988; Ohlsson, 1996; PDP models), as well as elsewhere (Bereiter & Scardamalia, 1993; Ericsson & Charness, 1994; Ericsson, et al., 1993) to account for the process by which skill is acquired and developed.

Anderson (1983, 1993) has proposed five types of cognitive activities to account for learning within the ACT* system; two types are involved in knowledge compilation--proceduralization and composition, and three types are involved in production tuning--generalization, discrimination, and strengthening. Proceduralization involves a shift from declarative to procedural knowledge bases as production rules are created and specialized.

Composition refers to the amalgamation of a set of adjacent production rules into a single production rule. Generalization and discrimination perform opposite operations on the range of the applicability of production rules such that rules are widened through generalization and narrowed through discrimination. Strengthening refers to the process by which better rules are strengthened and poorer rules are weakened.

In Ohlsson's (1996) production system, three critical cognitive activities have been proposed for learning to take place: Error detection, blame assignment, error attribution. Error detection is experienced subjectively as a discrepancy between the outcome that the actor perceived should have occurred and the actual outcome. It is a function of the prescriptive use of declarative knowledge, which serves to bring discrepancies to the actor's attention. Blame assignment refers to the cognitive activity responsible for determining which production rule caused the error, and error attribution refers to the cognitive activity that allows the feature in the production rule that led the system astray to be identified and corrected.

In PDP models, error detection and back propagation are the activities by which the state of the network is modified. Error detection involves an internal comparison of the output produced by the network with target output through a difference computation. Following the detection of error, the obtained difference is used to adjust the connection weights throughout the network so that error can be minimized on the next trial.

Logan (1988) makes reference to only one type of cognitive activity, that being accumulation (of episodic traces). Accumulation is the process of encoding similar processing episodes or instances in memory. Increasing the sample size of applicable instances in memory facilitates the likelihood that retrieval processing will be completed more quickly than algorithm processing and that selected instances will gradually be drawn from a more restricted range of the distribution of all similar instances.

While Higgins (1991) does not make mention of specific types of cognitive activity, the general principles of discovery and mastery are the impetus for skill acquisition in her framework. Discovery involves cognitive activity directed towards searching for sources of exteroceptive, proprioceptive, and / or expropriocepive information with which to alter future performance behavior. This search is stimulated by the comparison of intended with actual performance outcomes. Mastery refers to the process of gaining control over these sources of information.

Recently, Ericsson and his colleagues (Ericsson, 1996b; Ericsson, et al., 1993; Ericsson & Charness, 1994; Ericsson & Lehmann, 1996) have proposed that deliberate practice is a critical variable in the skill acquisition process. Two features distinguish deliberate practice from other types of experience; these include the goal of improvement as the primary motive for practice, and the need for repeated opportunities to execute specific skills. Since the goal of improvement is a critical feature of deliberate practice, the willingness of actors to direct their full attention towards monitoring their practice behavior for the purpose of generating feedback is critical to skill development according to Ericsson et al. (1993). In fact, Ericsson and Charness (1994) suggest that "if the regular activities in a domain did not offer accurate and preferably immediate feedback or opportunities for corrected repetitions, improvements in performance with further experience would not be expected" (p. 739). Ericsson, et al. (1993) and Ericsson (1996a) suggest that two types of cognitive activity are critical to acquire skill; careful monitoring of performance is needed to generate feedback about the efficacy of performance, and necessary adjustments need to be implemented on the basis of this feedback. In this way, new strategies and performance methods can be developed which alter performance in the direction of improvement. These activities are best supported with full concentration under the watchful eye of a knowledgeable mentor (Ericsson et al., 1993). Ericsson and colleagues (see Ericsson, 1996a; Ericsson & Charness, 1994; and Ericsson & Smith, 1991a) argue that accumulated hours of deliberate practice can reliably account for individual differences in attained level of expertise. Recently, however, there has been some debate as to whether deliberate practice is a critical learning activity in its own right, or the consequence of separate but related activities and variables (Sternberg, 1996, 1998).

Bereiter & Scardamalia (1993) have suggested 'reconceptualization' as the cognitive activity critical for learning. They argue that it is more helpful to view expertise as a process or way of approaching problems than as a characteristic of an individual. Expertise as a process is characterized by reinvesting mental resources into the progressive reconceptualization of domain related problems in order to expand existing knowledge to fit problems. Progressive reconceptualization involves the pursuit of more difficult problems, or the pursuit of representing more of the complexities of existing problems as the attentional demands of problem solving subside during the normal course of learning. In opposition are those individuals who disengage from the process of expertise. These individuals reduce domain related problems to fit existing problem solving routines. Thus, the essence of learning in the framework of Bereiter and Scardamalia is the ongoing process during which knowledge is used in solving problems, transformed through experience, and enhanced and attuned to situations as a function of trying to explain why some solutions failed while others did not.

The basis of critical cognitive activities

An overview of the cognitive activities considered to be critical for skill acquisition reveals that, with the exception of the cognitive activity proposed by Logan (1988), evaluation and analysis are the cognitive processes that are central to these cognitive activities.

In ACT*, a prerequisite for discrimination and generalization is feedback. Anderson (1983) indicates that "there are two basic ways in which ACT can identify that a production application is in error and determine the correct action. One is through external feedback, and the other is through internal computation" (p. 248). In the case of internal computation, the action of the production provides indirect evidence about the correctness of the production rule itself. Unsuccessful production procedures are identified through goal structures. Anderson claims that goal structures direct discrimination provided they are available in working memory at the time of discrimination. Since successful completion of an action goal requires successful completion of a series of subgoals, evaluation of subgoals and their execution relative to overall goal is required to identify error.

Error detection is the dominant theme within Ohlsson's (1996) production system. While constraint violations, the process by which errors are detected in Ohlsson's theory, may not be an effortful cognitive venture, blame assignment and error attribution are. Error correction requires evaluation to determine which production rule caused action to drift from successfully achieving the goal, and more specifically which part of the rule is in need of correction.

Like Ohlsson's (1996) production system, connectionist networks are modified on the basis of performance error. Thus, in PDP models, error detection also represents the activity upon which networks depend to learn. The determination of error requires evaluation of the output generated by the network in relation to target output.

Discovery within Higgins (1991) framework requires individuals to use "existing observational and analytical tools that support task analysis, self-analysis, and environmental analysis" (p. 135). Discovery, by definition, cannot occur in the absence of evaluation. Thus, the comparison of intended and actual performance drives evaluation of the demands of the environment, the demands of tasks within the environment, and the pool of resources available to meet these demands, all of which promote the discovery of more effective or more efficient action strategies.

The two types of cognitive activity upon which learning is dependent according to Ericsson and his colleagues include monitoring actions to generate feedback, and formulating a corrected response. Indeed, Ericsson (1996a) suggests that effective learning requires actors to "monitor their processes and performance to determine necessary adjustments and corrections" (p. 33) thereby implicating performance evaluation and analysis as critical processes.

Bereiter and Scardamalia (1993) acknowledge that reinvestment of mental resources into progressive problem solving (reconceptualization) is an effortful cognitive venture that involves paying attention to one's surroundings, thinking about alternate courses of action, and trying to achieve goals that are just out of reach. Analysis within the framework they propose involves evaluation of the situations in which problems occur or of the strategies currently being used for the purpose of searching for ways to do things better, or better ways of doing things.

In summary, the preceding overview of skill acquisition theories and frameworks indicate a convergence on the view that evaluation of the processes and products of action, for the purpose of generating feedback regarding their efficacy, appears to be a critical foundation upon which knowledge and skill develop.

Motor skill learning and the cognitive effort hypothesis

In the domain of motor behavior, Bernstein (1967) has suggested that motor learning is a gradual process in which optimal solutions are sought for movement related problems. Thus cognitive activity associated with process of solving movement problems is critical to the learning of motor skills. Indeed, Lee, Swanson, et al. (1991) and Lee, Swinnen, et al. (1994) suggest that when solving movement problems during practice demands high levels of cognitive effort, learning is enhanced. The cognitive activities that have been identified as demanding cognitive effort include "the evaluation of feedback for a previous action as well as the formulation of a new plan of action for a forthcoming movement" (Lee, Swanson, et al., 1991, p. 152).

The following review of motor learning literature related to the major variables that impact on the rate and efficacy of learning processes supports the view that cognitive effort affects motor skill learning. Furthermore, research evidence from studies on observational learning, knowledge of results, and contextual interference show the isomorphism between the demand for cognitive effort and the demand for evaluation; the conditions of practice that demand a greater cognitive investment in evaluation, enhance learning to a greater degree than practice conditions that minimize this demand.

Observational Learning

The method by which movement information is communicated to learners can affect the skill acquisition process. The problem with which learners are faced while learning a new skill or a new task involves knowing what to do and how to do it (Martens, Burwitz & Zuckerman, 1976). Knowing what to do involves selecting a performance strategy that solves the problem posed by the task, while knowing how to do it involves organizing and controlling a movement pattern that successfully applies the selected strategy. Demonstration is commonly used to convey movement information regarding what to do and how to do it. It involves having individuals watch a live or videotaped model perform a skill to-be-learned, after which they can modify aspects of their performance based on an analysis of what they have seen.

How does observing a model facilitate learning? Scully and Newell (1985) have suggested that the visual system of the learner picks up information about the relative motion of the model's activity. That is, the visual system registers information about the motion of key body parts relative to each other. This notion is supported by Bandura (1986) who has also argued that the benefits of modeling are likely more a function of learning movement rules than merely mimicking the observed response. Moreover, like Martens et al. (1976), Bandura has also suggested that the problem for learners once movement rules have been discovered is to coordinate the body and limbs to reproduce the relative motion that was observed and to scale these movements appropriately. The movement that learners ultimately produce reflects their understanding of what to do in response to the demands of the task as well as their understanding of the reactive forces involved in successfully executing the strategy that has been chosen (Higgins, 1991).

Although performers are capable of acquiring this knowledge through self discovery, the process can be accelerated by observing a model. Compared to self-discovery, observing a model would seem to encourage the discovery of more task relevant information since learners have the opportunity to evaluate the model's performance in addition to monitoring and evaluating the effectiveness of their own performance. The following review supports this conclusion.

Observing a model perform before attempting a skill can produce superior acquisition performance in observers as compared to the performance of individuals left to learn the skill on their. For example, Adams (1986) compared the acquisition performance of a control group, who served as unskilled models, to two observer groups. Observers either did or did not receive the KR of the control group model who was learning a skill. Participants from the observer groups watched individuals from the control group learn a three segment timing task before initiating any practice trials. One observer group received KR for the model's performance while the other observer group did not. All groups received KR for their own performance. Adams reported that watching an unskilled model learn the task--with or without their KR--produced initial performance benefits for both observers groups. Initial performance for observers approximated the level of performance that control group models had reached following 20 trials of practice. In later trials the performance of the observer group that had not received the model's KR tended to become more similar to the performance of the control group, whereas the performance of the observer group that had received the model's KR remained superior throughout the acquisition period. Adams concluded that watching someone learn the task engaged observers in problem solving activities similar to those experienced by the model they were watching, and that these benefits were enhanced by also receiving the model's KR.

Subsequent research has provided support for Adams' findings. In an experiment involving a timed barrier knock-down task, Lee and White (1990) reported that observers who watched models learn the task one day demonstrated superior acquisition performance when they practiced the task themselves the next day. Lee and White (1990), and Pollock and Lee (1992) also found that watching a model play a computer game with KR prior to any game play produced and maintained better acquisition performance in observers compared to individuals who had to learn the game on their own. Observers in the study of Whiting, Bijlard, and den Brinker (1987) watched videotape of expert slalom ski performance on a ski simulator before each of their own practice trials. A second group of participants were left to learn the task on their own. After five days of training, the smoothness and speed of movement was more fluent and consistent for the group of observers than for the group learning on their own. Using the same task Magill (1993b) also reported that observers who watched expert models before every practice trial acquired movement skill more quickly over a three day period and also showed superior coordination characteristics compared to individuals left to learn the task on their own. These results suggest that additional problem solving activity demanded of individuals who observe a model resulted in greater knowledge gains, and may be the basis of the superior acquisition rates of observers.

Following this line of reasoning, Hand and Sidaway (1992) investigated whether increasing the frequency of observing a model would benefit learning the golf chip shot. Three modeling groups received a modeling to trials ratio of 1:1, 1:5, and 1:10 while practicing 150 golf chip shots; a control group practiced on their own. Hand and Sidaway reported that better acquisition performance was associated with groups receiving higher modeling frequencies, with the 1:1 group outperforming the other three groups.

In summary, a review of representative literature from observational learning research supports the idea that analysis and evaluation benefits skill acquisition. Furthermore, observing a model has also been shown to facilitate learning as enhanced levels of performance have been found to persist over time (Adams, 1986; Lee & White, 1990; McCullagh & Caird, 1990; Magill, 1993b; Whiting et al. 1987). The evidence suggests that watching a model allows observers to assess the performance strategies, as well as the organizational and control parameters of movement more frequently than learners who are not exposed to modeling, and that the additional cognitive effort demanded by assessing the performance errors and correction strategies of a model is associated with elevated levels of acquisition performance. Indeed, Lee, Swanson, et al. (1991) conclude that watching an unskilled model learning a task allows observers to learn about the problem-solving process associated with correcting errors by evaluating how errors occur, how the model attempts to correct them, and how successful the corrections were.

Feedback

Feedback is information about the process or outcome of movement that has the potential to modify future performance. Feedback can be obtained through internal mechanisms or it can be augmented by an external source. In motor behavior research, the majority of investigations involving feedback have focused on verbalizable information about the outcome of performance provided by an external source, also referred to as knowledge of results (KR; Magill, 1993a). Early research on feedback (e.g., Bilodeau & Bilodeau, 1958) led to the proposal that "any variation of KR that provides more information more precisely, more accurately, or more often is predicted to have a positive effect on movement learning" (Schmidt, Young, Swinnen, & Shapiro, 1989, p. 353). In a review of KR research, Salmoni, Schmidt and Walter (1984) refuted this proposal noting that, because performance data were obtained by observing practice performance only, much of the evidence upon which the proposal was based included both the temporary and permanent effects of feedback. Salmoni et al. found that on a transfer test, where the influence of these temporary effects were controlled, exposure to higher levels of KR during practice produced better performance but typically also translated into poorer retention performance. Conversely, exposure to lower levels of KR during practice may depress acquisition performance to some extent, but led to superior retention performance (Magill, 1993a). The following review of the frequency of presenting KR and the scheduling of KR suggests that when learners are forced to rely more heavily on internal than external sources for KR, acquisition performance may or may not be negatively affected, but retention performance is superior.

Frequency of KR

The relative frequency of presenting KR refers to the percentage of trials after which KR is given. In three experiments, Winstein and Schmidt (1990) showed that reducing the relative frequency of KR as individuals practiced a multiple flexion-extension movement about the elbow was not detrimental to acquisition performance. However, practice under reduced KR conditions produced more accurate performance on both no-KR and KR-present retention conditions than practice under a high (100%) KR condition.

Similarly, Lee, White and Carnahan (1990) found that acquisition performance on a reciprocal tapping task with fixed time parameters was not significantly affected by giving augmented feedback on either 100% or 50% of practice trials. On no-KR retention tests, however, the results of all three of their experiments in concert revealed that giving augmented feedback on 50% of the practice trials produced more accurate and consistent performance.

The frequency of presenting KR can also be reduced by withholding KR for a specified number of trials, even though the KR includes information from all trials. Lavery (1962) reported that although receiving KR following every trial for three ball accuracy tasks produced more accurate acquisition performance than receiving summary of KR following 20 trials, 20-trial summary KR benefited performance during a no-KR retention period. Schmidt, Young, et al. (1989) presented summary KR to groups after 1, 5, 10, and 15 trials during the acquisition of a timed double-reversal arm movement task and found that longer summary KR conditions produced acquisition performance that was systematically less accurate. Performance accuracy on a delayed no-KR retention test, however, showed a complete reversal of group order with longer summary KR conditions producing systematically more accurate performance. Using a more complex coincident timing task, Schmidt, Lange and Young (1990) reported an inverted "U" effect of summary KR length with the 5-trial summary KR condition performing most effectively on both a no-KR and every-trial-KR retention test.

When considered together, evidence from research on the frequency of presenting KR has led Schmidt and his associates (Salmoni, et al., 1984; Schmidt, 1988; Schmidt, Young, et al., 1989; Schmidt, Lange, et al., 1990; Winstein & Schmidt, 1990) to propose the guidance hypothesis as an explanation for the effects of KR on performance and learning. They suggest that KR has a strong informational capacity to guide learners toward the correct movement pattern. A high relative frequency of KR benefits initial acquisition performance by allowing performance to be maintained at or near the criterion. However, maintaining a high relative frequency of KR after a relatively correct movement pattern can be produced may be detrimental to learning because it may reduce the learner's motivation to process other task-relevant information. That is, augmented feedback can provide learners with more easily attainable, precise, and relevant information about performance than can be derived intrinsically, which reduces their need or desire to actively process other task-relevant information. Consequently, performers tend to rely on the guiding properties of KR rather than expending the effort to discover relevant KR on their own. Over time, reliance on externally provided KR can interfere with the development of error-detection capabilities.

Lavery (1962) provided additional support for the benefit of relying on internal sources to generate KR during practice. He warned individuals receiving KR after every trial during practice to pay attention to task-relevant cues while they practiced, explaining that perceiving task-relevant cues would determine their success during an no-KR retention period that was to follow the acquisition period. Presentation of KR was delayed for 10 seconds to provide participants with the opportunity to process task relevant information. Results showed that compared to a 20-trial summary KR group, acquisition performance of the every-trial-KR group was superior, providing support for the benefits of the guiding properties of KR when it is present. However, no differences in performance emerged between the two groups during a retention period where the guiding effects of KR were removed, thereby providing support for the benefits of processing information during practice on the learning of motor skills.

Scheduling of KR

The importance of developing the capability to generate KR internally on motor skill learning is further supported by research on the timing of KR after a trial has been completed. Swinnen, Schmidt, Nicholson and Shapiro (1990) hypothesized that providing KR immediately following a trial for a timed double-reversal arm movement task should interfere with the processing of task relevant information, while forcing subjects to estimate their performance prior to receiving KR should enhance the processing of this information. Their study involved three groups, (a) a that received KR following an eight second delay (control group), (b) a group that was provided with KR immediately following each trial (interference group), and (c) a group that was required to estimate their performance prior to receiving KR (estimate group). All groups demonstrated similar performance during acquisition, but on a delayed no-KR retention test, the group that received immediate KR performed less accurately than the group that estimated their performance prior to receiving KR. That no significant performance differences emerged between KR-delayed and estimate groups led Swinnen et al. to suggest that individuals spontaneously estimate their performance errors during the KR delay interval, with the immediate presentation of KR interrupting this process. The detrimental effects of interfering with the processing of task-relevant information was further supported in a follow-up experiment using a more complex coincident timing task. Compared to delaying KR for a mere three seconds during practice, providing immediate KR produced poorer performance later in acquisition as well as on immediate and delayed no-KR retention tests. Based on their two experiments, Swinnen et al. concluded that instantaneously presented KR blocked or interfered with spontaneous error estimation activity and thus interfered with the normal development of error detection capabilities.

In a related series of experiments, Swinnen (1990) included a secondary task that required attention during the KR delay interval of a timed double reversal arm movement task. He hypothesized that diverting attention to another task during the KR delay interval would be detrimental to retention performance by interrupting the development of error detection capabilities. Similar to Swinnen et al. (1990), no performance differences emerged between the estimation group and KR-delay group during either acquisition trials with KR or retention trials without KR, thereby providing support for the notion that error estimation occurs spontaneously during the KR delay interval. The secondary activity, which involved estimating the error of a model, negatively affected the accuracy of acquisition performance, especially when the movement to be estimated was slower than the criterion movement. On a delayed no-KR retention test, the interference caused by estimating the error of a second observed movement during the KR delay interval, whether faster or slower than the criterion movement, was detrimental to performance accuracy. This effect was replicated in all three experiments. When secondary activity was introduced during the post-KR interval, Swinnen (1990) found that retention performance was negatively affected, but to a smaller extent.

In summary, the evidence from this review of KR research supports a cognitive effort hypothesis. Experimental conditions that demanded higher levels of cognitive effort by constraining the environment so that learners must process task-relevant information to maintain a reasonable standard of performance, produced better performance on no-KR retention tests. Learning benefits when developing a sensitivity to task-relevant information becomes a necessary component of the practice context. When KR is made regularly and immediately available to the learner, the need to actively process error information becomes redundant and the effortful process of directing attention towards task-relevant cues to guide performance can be averted.

Scheduling of practice

A relatively recent practice variable that has received attention in motor behavior research is the scheduling or ordering of practice. This phenomenon is known as the contextual interference effect and refers to the differences in acquisition and delayed retention performance produced by practicing multiple tasks within the context of a blocked (AAA BBB CCC) or random (ACB CAB CBA) practice schedule. The paradox of the contextual interference (CI) effect is that when the order in which tasks are practiced is random (high CI), acquisition performance is maintained on tests of retention and transfer. However, a practice context in which tasks are practiced in blocks (low CI) produces superior performance during the practice period, but poorer performance on tests of retention and transfer. The effect of practice context on immediate and later performance was introduced in the verbal learning domain by Battig (1966, 1972, 1979) who reported that memory tasks that were made more difficult at original encoding tended to produce better delayed recall.

Shea and Morgan (1979) applied Battig's discovery to motor skills research and reported similar findings. In what Chamberlin & Lee (1993) call the prototypical contextual interference experiment, Shea and Morgan used a barrier knock-down task in which participants practiced each of three spatial movement patterns 18 times. Each pattern was illustrated on a diagram card that remained in full view for all acquisition and retention trials. Illuminating one of three colored stimulus lights indicated which spatial pattern was to be produced, after which individuals released a start button and executed the pattern as quickly as possible. One group practiced each movement pattern in blocks of 18 trials while the other group practiced the three patterns randomly for 54 trials. During the acquisition period, both the reaction time and movement time of the blocked practice group was faster than that of the random practice group for each movement pattern. Furthermore, the blocked group produced significantly fewer sequencing errors as the random group. Based only on the results of practice performance, the context of blocked practice would be identified as the most effective practice condition for learning. However, retention and transfer tests following a ten minute delay revealed a reversal in the efficacy of these conditions. On both blocked and random retention tests, the reaction time and movement time of the random group was faster than that of the blocked group. Moreover, the random group made fewer sequencing errors during the retention test. Similar results were found on retention and transfer tests given ten days later. These results demonstrate that the context in which motor skills are practiced can affect how well they are learned.

Similar results have been reported in subsequent laboratory investigations in which a variety of spatial tasks (Shea & Zimny, 1983, 1988; Wood & Ging, 1991), timing tasks (Carnahan, et al., 1990; Lee & Magill, 1983; Lee, Wulf, et al., 1992; Sekiya, et al., 1994) and force production tasks (C. H. Shea, Kohl, & Indermill, 1990) have been used (see Magill & Hall [1990] for a review). The contextual interference effect has had limited support in field studies involving badminton (Goode & Magill, 1986; Wrisberg, 1991;

Wrisberg & Liu, 1991), baseball (Hall, Domingues & Cavazos, 1994), kayaking (Smith & Davies, 1995), and rifle shooting (Boyce & Del Rey, 1990). To date the CI effect has not emerged in studies involving tennis (Hebert, Landin, & Solmon, 1996) and volleyball (Bortoli, Robazza, Durigon & Carra, 1992; French, Rink & Werner, 1990).

Theoretical accounts of the CI effect

Theory of processing elaboration and distinctiveness

Shea and his associates (Shea & Morgan, 1979; Shea & Zimny, 1983, 1988; Shea & Wright, 1991) have proposed a theory of processing elaboration and distinctiveness (TPED) for the contextual interference effect. They adapted Battig's (1979) theoretical explanation for contextual interference in which it is suggested that contextual variety is the primary factor influencing the use of multiple and variable processing strategies during learning. The use of multiple processing strategies to encode task solutions along with the variable use of these strategies demands that task relevant information be processed more elaborately and distinctively (Shea & Zimny, 1983). More elaborate processing involves encoding more features of a stimulus, while more distinctive processing refers to the degree to which the similar features of multiple tasks can be contrasted. In this theory, processing elaboration and distinctiveness is enhanced by task similarity and random practice. Similarity between tasks requires the learner to look for dimensions along which tasks can be differentiated. This process is most effectively accomplished through comparison. Conversely, processing elaboration and distinctiveness is restricted in blocked practice, because only one task occupies working memory at a time thereby providing little opportunity for comparison; and by dissimilar tasks which are, by their nature, distinct and do not require further differentiation. Shea and Zimny suggest that increased contextual interference during learning leads to better retention and transfer performance because it stimulates processing elaboration and distinctiveness, which in turn enhances the learner's ability to recall the appropriate response given the stimulus. Thus, in TPED, the increase in cognitive effort associated with random practice conditions is a function of higher demands for evaluative processes.

Theory of action plan reconstruction

Lee and Magill (1983, 1985) adapted Jacoby's (1978; Cuddy & Jacoby, 1982) theoretical explanation for the spacing effect to motor skills, formulating the theory of action plan reconstruction (TAPR) or forgetting theory. This theory is based on the notion that "forgetting between repetitions of goal-directed actions will depress acquisition performance yet promote retention" (Lee & Magill, 1985, p. 8). Lee and Magill suggest that when different tasks must be performed between repeated trials of a task, the action plan for the to-be-repeated task cannot be retained in working memory due to the interference created by the intervening tasks. Thus, each time a task is repeated within the context of random practice, the appropriate action plan must be reconstructed in working memory. In effect, Lee and Magill claim that the context of random practice forces the learner to engage in problem solving activities to re-create the task solution since the interference generated by intervening tasks prevents the learner from simply remembering it. The benefit of this process is that the learner's ability to recall appropriate solutions becomes strengthened through practice, which benefits learning as evidenced by superior retention performance. The cost is that the process of refining action plan solutions becomes relatively less important than being able to reconstruct them properly, and therefore acquisition rates are slowed. Indeed, Lee and Magill argue that it is the repetition of problem solving activity associated with reconstructing action plans that facilitates retention and transfer. Conversely, the appropriate action plans for tasks practiced under a blocked schedule need only be recalled once, after which they can simply be remembered and successively refined. Acquisition performance benefits during blocked practice because a solution can be more accurately refined on successive repetitions. However, because the need to continuously reconstruct it is largely absent, much processing activity can be bypassed. While successively refining movement parameters benefits immediate performance, the infrequent use of processing strategies to reconstruct the appropriate action plans results in reduced retention and transfer performance. Thus, in TAPR, the increase in cognitive effort associated with random practice conditions is a function of higher demands for retrieval processes.

In summary, the evidence provided by research on the CI effect also supports the cognitive effort hypothesis. Theoretical orientation aside, it appears that cognitive effort or engagement is the major determinant of learning within the contextual interference paradigm. Furthermore, when the demand for cognitive effort is reduced as it is under blocked practice conditions where one solution can be applied repeatedly over a number of trials, learning appears to progress more slowly. However, each theoretical account for the CI effect differs as to the problem solving activity demanding cognitive effort. The TAPR posits that cognitive effort is required to generate the appropriate solution for the task to be performed while the TPED predicts that cognitive effort is required to produce a more elaborate encoding of the appropriate solution.

Summary of evidence for the cognitive effort hypothesis

When considered together, evidence provided by motor learning research on KR, OL, and CI suggests that the development of knowledge related to action selection and the organization, and control of subsequent movement, can be enhanced or minimized by manipulating cognitive effort. Furthermore, the manipulation of cognitive effort may be a function of manipulating opportunities to evaluate various sources of regulatory information.

Experiments involving knowledge of results (KR) suggest that motor learning benefits to the extent that learners develop a sensitivity to critical sources of regulatory information involved in movement production. Diverting learners attention away from discovering and processing this information during the acquisition of motor skills (Swinnen, 1990) by providing feedback often (Lavery, 1962; Lee, White, et al., 1990; Schmidt, Young, et al., 1989; Winstein & Schmidt, 1990) or immediately (Swinnen, et al., 1990) has been shown to interfere with the normal development of error detection capabilities; capabilities needed for accurate retention performance.

Observational learning (OL) research has shown that regulatory information pertaining to the selection, organization and control of movement can be acquired more quickly by observing a model (Adams, 1986; Hand & Sidaway, 1992; Lee & White, 1990; Magill, 1993b; Pollock & Lee, 1992; Scully & Newell, 1985; Whiting, et al., 1987). Furthermore, the enhanced levels of performance achieved by subjects who were provided with demonstrations were found to persist over time (Adams, 1986; Lee & White, 1990; McCullagh & Caird, 1990; Magill, 1993b; Whiting, et al., 1987).

CI research indicates that regulatory information pertaining to the selection, organization and control of movement can be better retained under the conditions of random than blocked practice (Hall et al., 1994; Lee & Magill, 1983; Lee, Wulf, et al., 1992; C. H. Shea et al., 1990; Shea & Morgan, 1979; Shea & Zimny, 1983; Smith & Davies, 1995).

Rationale for using the CI paradigm to test the evaluation hypothesis

In their reviews of the motor learning literature pertaining to KR, OL, and CI, Lee, Swanson, et al. (1991) and Lee, Swinnen, et al. (1994) have suggested that cognitive effort may be required as a result of problem solving activity associated with either the evaluation of feedback or the formulation of a motor plan that is not currently being held in working memory. However, the present review of literature has indicated that evaluation may be the fundamental cognitive activity influencing motor skill learning. To date, however, the link between evaluative processing and cognitive effort has not been tested directly. Therefore, the specific purpose of this research project was to test the validity of evaluative processing as the cognitive activity primarily responsible for producing the CI effect. The more general purpose of this research project was to explore whether evaluative processing might be considered as a common cognitive process underlying a number of phenomena (i.e. KR, OL, and CI) related to the acquisition of motor skill.

The CI paradigm appears to be an ideal way to test this hypothesis. Of the two theories that have been proposed to explain the CI effect, the TPED (Shea & Morgan, 1979) defers to evaluative processing and encoding variability while the TAPR (Lee & Magill, 1983) defers to memory retrieval and working memory reconstruction. One confound that exists in most CI research is the simultaneous manipulation of both evaluative and retrieval processes as a function of practice scheduling due to the characteristics of the experimental tasks that have been used. As a result, the data arising from much of this research can be equally explained by either an evaluative/encoding based or a retrieval based theory. Experiment 2 is designed to dissociate the simultaneous manipulation of both evaluative/encoding and retrieval processes by altering the characteristics of the experimental tasks. When considered in conjunction with Experiment 1, whose purpose is to (a) establish that the CI effect will emerge using a multisegment arm movement task, and (b) to maintain the simultaneous manipulation of both evaluative/encoding and retrieval processes, the results should indicate whether the cognitive effort producing the CI effect is a function of evaluative/encoding or retrieval processes. Experiment 3 is designed to directly test whether increasing cognitive effort under blocked practice conditions via purposeful task evaluation will eliminate the CI effect.

EXPERIMENT 1

The purpose of this experiment was to replicate the findings of Shea and Morgan (1979) and Lee and Magill (1983, experiment 1) using a multisegment arm movement task, and to ensure that the CI effect would emerge with this task. Experiment 1 also served as a control condition for the experimental manipulations to be made in Experiment 2.

Methods

Participants

Participants were 24 students (16 female, 8 male, mean age 20.5, sd=1.8 years) enrolled in a course offered by the School of Physical Education at the University of Victoria. Informed consent was obtained for all participants. Course credit was given in exchange for voluntary participation in the experiment.

Apparatus and Task

Sixteen spring loaded electronic response keys are arranged in a 4 x 4 grid and mounted onto the face of a square board. Each response key is 3 cm in diameter and can be depressed to a maximum of 7 mm. The distance between the columns and rows of keys is 15 cm. The face of the push button board is 60 cm wide by 62 cm high and is attached on a 45 degree angle to a 44 cm high by 44 cm long support frame. This frame sits on a table top 69 cm from the ground. Participants sit in front of the apparatus on an adjustable seat so that their mid-section is approximately level with the top of the table.

A computer monitor was mounted 2 cm above the board. It was used to provide participants with information about the spatial pattern to be executed, and to provide feedback. Three color coded pattern cards, arranged vertically, were located 3 cm to the left of the push button board. The bottom of the lowest card was level with the top of the push button board. Each pattern card was 12 cm square and contained 16 circles arranged in a 4 x 4 grid corresponding to the 16 response keys on the push button board. On a pattern card, the start key was labeled "S", and four response keys were labeled A, B, C, and D. A diagram of the experimental environment is provided in Figure 1.

The Push Button Challenge (PBC) task involved depressing response keys A - D following the release of the start key. The position of the start key and the sequence of response keys to be depressed were specified on colored pattern cards. To begin a trial, a participant was cued by the researcher to depress and hold down the start key with their dominant hand. After approximately five seconds participants heard a 10 millisecond (msec) low pitched tone. The low pitched tone served as a warning signal and indicated the beginning of the stimulus foreperiod, which varied randomly from 500 to 2500 msecs. Following the foreperiod participants were presented with a 10 msec high pitched stimulus signal. The name of the color of the pattern card containing the pattern to be executed appeared on the computer screen simultaneous with the onset of the auditory stimulus. The computer program recorded the time (msecs, starting at time = 0) at which each response key was pressed and released following the 10 msec auditory stimulus signal.

The main objective of PBC was for participants to learn how to minimize total response time following presentation of the stimulus. Therefore, participants were instructed to release the start key and depress the appropriate response keys as quickly as possible with their dominant hand following the high pitched tone. However, participants were also instructed not to release the start key until they were prepared to hit the four response keys at full speed. This instruction was given to discourage cognitive activity related to movement planning during the execution of the movement and encourage planning the movement as completely as possible in advance of movement execution.

Participants received feedback about their total response time (msecs) five seconds following the release of the last response key. Total response time (TRT) was the amount of time from the onset of the stimulus to the release of the last response key. Faster response times were considered as the index of skilled performance, in accordance with Shea and Morgan (1979).

Hitting response keys other than those specified in the pattern, hitting specified response keys out of sequence, or waiting for longer than two seconds between hitting specified response keys were designated as decision making errors since these errors were considered to be the result of faulty movement planning operations (i.e. the participant did not know what to do at that moment). Missing or miss-hitting a specified response key (making contact with a response key but it fails to register) were designated as execution errors since these errors were considered to be the result of faulty movement execution operations (i.e. the participant knew what to do, but did not execute the movement properly). Participants were asked not to go back and correct an execution error since this would artificially inflate TRT. Rather they were instructed to hit an unspecified response key so that the trial could be repeated. For any error, a message appeared on the computer monitor indicating where an error occurred, and the trial was repeated immediately until executed properly.

Procedures

Each participant (N=24) attended an orientation session where informed consent was obtained. Course credit was given in exchange for voluntary participation in the experiment. Following verbal instructions about the objectives and procedures of Push Button Challenge, participants received ten orientation trials. The purpose of these trials was to allow participants to become familiar with experimental task and procedures, and to provide data with which to match pair participants into experimental conditions on the basis of average speed of response. TRT for the last five orientation trials were blocked, and block means were used to rank participants from fastest to slowest. Participants were then paired by rank and, within each pair, randomly assigned to either a blocked or random practice condition using a random numbers table.

All participants practiced three spatial patterns, which were specified on colored pattern cards (see Figure 2). Each pattern was practiced until it had been completed successfully 18 times, for a total of 54 successful acquisition trials. Feedback on TRT was provided five seconds following each practice trial. This delay was scheduled to provide participants with an opportunity to spontaneously assess their performance (Swinnen, 1990). Ten seconds following the presentation of feedback, participants were cued for the next trial. Approximately five seconds later the computer generated the warning signal. Thus, the intertrial interval was approximately 20 seconds. After each set of 18 successful acquisition trials, participants were given a one minute rest.

Participants assigned to the blocked practice condition completed all 18 successful trials of one pattern before practicing another (AAAA... BBBB... CCCC...). Order effects were counter balanced across participants. Participants were told in advance the order in which patterns would appear.

Participants assigned to the random practice condition performed the three patterns in a random order until each pattern had been performed successfully 18 times. The order in which patterns appeared was determined by a computer program with the stipulation that each pattern be presented once in each set of three trials (ABC BAC ACB...). This stipulation ensured that no pattern would be practiced more than twice in a row. Participants were not made aware of this stipulation.

Following successful completion of all acquisition trials, participants were provided with a filled 10 minute retention interval during which they performed a cognitively demanding distracter task. Superfection™ is a spatial orientation task that involves locating two plastic pieces in an array of 90 different pieces and fitting them together to form a solid cube. The cubes had been divided into two parts in a variety of unique ways and each part had a different color. There were 90 pieces in all that, when fitted correctly with its mate, formed 45 cubes. At the beginning of the retention interval, the pieces were arranged randomly in a pile. Participants' task was to locate and combine each pair of cube pieces and complete as many cubes as possible. They were instructed to separate the pieces and continue if all 45 cubes were solved in less than ten minutes.

After the ten minute filled retention interval, participants performed retention trials on each task until each task had been executed successfully three times (9 successful trials total). Retention trials for all participants were presented in a random order as determined by the computer program. Feedback about TRT was not provided following successful trials. Prior to the first retention trial, participants were reminded that after the stimulus signal they were to be prepared to execute the pattern at full speed, but that the goal was still to minimize TRT. Thirty seconds following these instructions, participants were cued for the first retention trial. If a decision making error was made, the researcher would demonstrate the pattern and the participant would immediately repeat the trial. One week later, all participants returned and repeated these retention procedures.

Data Analysis

Total response time (TRT), reaction time (RT), and movement time (MT) data were collapsed into blocks of three trials for statistical analysis, resulting in six blocks of acquisition data and two blocks of retention data for each dependent measure. Separate analyses, using SPSS 9.0 GLM Repeated Measures, were conducted to assess acquisition and retention performance. For the acquisition period, the factors of Acquisition Condition (blocked and random), Task (red, yellow, & blue patterns) and Trial Blocks (1 - 6) were combined in a 2 x 3 x 6, respectively, multivariate analysis of variance (MANOVA) with repeated measures on all factors. Acquisition condition was treated as a within subjects factor because participants were match paired according to TRT prior to being randomly assigned to groups. Retention performance was assessed by a 2 (Acquisition Conditions) x 3 (Tasks) x 3 (Trial Blocks: acquisition block 6, 10 minute retention, 1 week retention) MANOVA with repeated measures on all factors. The last acquisition trial block was included in the analysis to capture information about any changes in performance across the first retention interval.

Results

Acquisition

Mean performance for TRT, RT and MT are provided by Figure 3. There were no reliable multivariate main effects for Task, no reliable two way interactions of Task with Acquisition Condition or Trial Block, and no reliable three way interaction of Acquisition Condition, Trial Block and Task. The lack of any reliable effect related to task indicated that the pattern of results within each experimental condition was similar across tasks during the acquisition phase. Therefore, the data were collapsed across tasks and re-entered into a 2 (Acquisition Conditions) x 6 (Trial Blocks) repeated measures MANOVA with TRT, RT and MT as dependent measures.

A reliable multivariate effect was found for Acquisition Condition (AC), Wilks' exact $F(2, 10) = 19.54$, $p < .001$, $\eta^2 = .796$. Univariate analyses showed that, overall, TRT, RT and MT performance was reliably different between blocked and random practice conditions, TRT $F(1, 11) = 40.06$, $p < .001$, $\eta^2 = .785$, RT $F(1, 11) = 29.76$, $p < .001$, $\eta^2 = .730$, and MT $F(1, 11) = 13.37$, $p < .01$, $\eta^2 = .549$. The size of these effects are shown on Table 1. The data show that participants who practiced under blocked conditions performed the tasks as a whole (TRT), and each component of the tasks (RT & MT) more quickly than participants who practiced under random conditions.

The pattern of performance change over the acquisition phase differed between acquisition conditions as indicated by a reliable AC x TB interaction, Wilks' exact $F(10, 108) = 8.36$, $p < .001$, $\eta^2 = .436$. Univariate tests showed reliable effects for each of TRT, $F(5, 55) = 22.84$, $p < .001$, $\eta^2 = .675$, RT, $F(5, 55) = 6.86$, $p < .01$, $\eta^2 = .384$, and MT $F(5, 55) = 9.25$, $p < .01$, $\eta^2 = .457$. TRT performance improved to a greater extent under random than blocked practice conditions between each trial block. Partialing TRT into its component parts, RT improved to a greater extent under random than blocked practice conditions between trial blocks 3, 4, and 5, while the speed of pattern execution (MT) improved more substantially under random than blocked acquisition conditions between trial blocks 1-2, 2-3 and 5-6. Values of Fs, ps and η^2s for each of these reliable AC x TB contrasts are shown on Table 2.

Mean number of recall errors made by blocked and random practice participants during the acquisition, immediate and delayed retention phases of the experiment are shown on Figure 4. Analysis of the number of recall errors made during the acquisition period revealed that participants practicing under random conditions made substantially more recall errors during the acquisition phase (M=4.42, sd=2.78) than participants practicing under blocked conditions (M=1.17, sd=1.27), $t(22) = 3.69$, $p<.001$.

Retention

A 2 (Acquisition Conditions) x 3 (Tasks) x 3 (Trial Blocks) repeated measures MANOVA revealed no reliable multivariate effects involving Task, therefore, data were collapsed across tasks and reanalyzed in a 2 (Acquisition Conditions) x 3 (Trial Blocks) repeated measures MANOVA.

A reliable multivariate effect was found for Trial Block, Wilks' exact $F(4, 38) = 5.15$, $p <.01$, $\eta^2 = .351$. Univariate analyses showed that the TB main effect was reliable for each of TRT, $F(2, 22) = 13.02$, $p <.001$, $\eta^2 = .566$, RT, $F(2, 22) = 7.22$, $p <.05$, $\eta^2 = .419$, and MT, $F(2, 22) = 6.65$, $p <.01$, $\eta^2 = .399$. Univariate contrasts for each of TRT, RT and MT showed reliable overall performance differences between the last trial block of the acquisition phase and the trial block of the immediate (10 minute) retention phase, TRT, $F(1, 10) = 14.85$, $p <.01$, $\eta^2 = .598$, RT, $F(1, 10) = 12.18$, $p <.01$, $\eta^2 = .549$, and MT $F(1, 10) = 10.12$, $p <.01$, $\eta^2 = .503$, but no overall performance differences between the trial blocks of the immediate and delayed (1 week) retention phases.

However, the overall performance effect across trial blocks was overshadowed by a reliable AC x TB interaction, Wilks' exact $F(4, 38) = 10.58$, $p<.001$, $\eta^2 = .527$. Univariate analysis revealed that the interaction was reliable for each dependent measure, TRT $F(2, 20) = 27.15$, $p<.001$, $\eta^2 = .731$, RT $F(2, 20) = 20.60$, $p<.001$, $\eta^2 = .673$, MT $F(2, 20) = 8.13$, $p<.01$, $\eta^2 = .448$. For each dependent measure, performance of participants who had practiced under blocked conditions slowed considerably from the end of acquisition to immediate retention, while the performance of participants who had practiced under random conditions remained at levels achieved by the end of the acquisition phase, TRT $\underline{F}(1, 10) = 14.85$, $\underline{p}<.01$, $\eta^2 = .598$, RT $\underline{F}(1, 10) = 12.18$, $\underline{p}<.01$, $\eta^2 = .549$, MT $\underline{F}(1, 10) = 10.12$, $\underline{p}<.01$, $\eta^2 = .503$ (see Figure 3). The size of these interaction effects are shown on Table 3. Immediate and delayed retention performance was similar across acquisition conditions.

Recall errors were analyzed by a 2 (Acquisition Conditions) x 3 (Experimental Phases: acquisition, immediate retention, delayed retention) factorial ANOVA with repeated measures on the last factor. Neither the main effect for Acquisition Condition nor Experimental Phase were found to be reliable, however a reliable effect was found for their interaction, $\underline{F}(2, 44) = 15.69$, $\underline{p}<.001$, $\eta^2 = .416$. Figure 4 illustrates that, from the acquisition to the immediate retention phase, recall error frequency under random practice conditions decreased while they increased under blocked practice conditions, $\underline{F}(1, 22) = 30.22$, $\underline{p} <.001$, $\eta^2 = .579$. These trends reversed from the immediate to delayed retention phases with an increase in recall error frequency under random practice conditions coupled with a slight decrease under blocked practice conditions, $\underline{F}(1, 22) = 10.60$, $\underline{p} <.01$, $\eta^2 = .325$. However, it should be noted that mean recall error frequency for random practice participants during the delayed retention phase was significantly less than it was during the acquisition phase, $\underline{t}(11) = 1.94$, one tailed $\underline{p}<.05$, while mean recall error frequency for blocked practice participants was significantly greater during the delayed than the acquisition phase, $\underline{t}(11) = 2.09$, one tailed $\underline{p}<.05$.

Discussion

The results of this experiment replicate the findings of previous studies on the contextual interference effect (e.g. Lee & Magill, 1983; Shea & Morgan, 1979), providing further evidence for the robustness of the effect in controlled settings. In typical fashion, the conditions of practice made it more difficult for random than blocked participants to perform during the acquisition period as evidenced by their performance times. The criterion level of performance established at the end of the acquisition phase was maintained by random participants during the immediate and delayed retention phases but not by blocked participants.

The finding that RT differed between acquisition conditions, and increased at retention under the blocked practice condition should not be surprising. During the acquisition period blocked participants performed each task within a simple RT paradigm. They knew ahead of time which response would be specified and could organize it prior to the onset of the stimulus. However, participants in the random condition performed within a choice RT paradigm where information about the upcoming response was specified simultaneous with the onset of the stimulus. At retention all participants performed within a choice RT paradigm, meaning that the RT paradigm for blocked participants changed from simple to choice RT. Choice RT is expected to be longer than simple RT due to the added demands on cognitive processes involved in identifying the stimulus (reading the name of a color) and selecting the appropriate response following onset of the stimulus (Hick, 1953; Hyman, 1953). Indeed, Lee and Magill (1983, experiment 1) found that the RT differences between blocked and random conditions during acquisition were caused by simple and choice RT paradigms and vanished during immediate retention trials when all participants performed within a choice RT paradigm.

The similarity of RT data at retention suggests that participants from both acquisition conditions were able to recall the appropriate response to the presented pattern color word within a similar time frame. However, the error data suggests that participants in each acquisition condition were not able to recall this information equally well. Relative to acquisition, blocked participants made significantly more decision making errors at retention while random participants made significantly less. These results replicate those found by Shea and Morgan (1979) and suggest that the ability to retrieve pertinent information regarding a movement sequence that had been paired with a color was considerably more difficult on retention trials for blocked than random participants when the color name was provided.

As was found by Shea and Morgan (1979) and Lee and Magill (1983), the speed with which patterns were executed was considerably slower for blocked participants on retention than acquisition trials. In addition, not only was the movement speed for blocked participants slower at retention relative to the end of acquisition in the present experiment, it was also slower relative to initial performance. It would appear that processes involved in the organization and execution of movement changed for blocked participants from acquisition to retention trials. As was mentioned by Shea and Morgan (1979), an explanation for this result may be found in the mode of movement control used in each phase of the experiment. Since blocked participants performed under simple RT conditions during acquisition and could organize a response in advance of stimulus onset, its execution may have been more than less ballistic. That is, the movement may have been preprogrammed and run off as a unit in the absence of substantial on-line control (van Donkelaar & Franks, 1991). The evidence regarding the difficulties blocked participants experienced recalling movement sequences, and the considerable slowing of movement execution on retention trials suggests that execution of movement sequences at retention may have been less than more ballistic and, therefore, may have involved more on-line control. In terms of movement planning, this suggests that whereas movement sequences may have been planned, organized and executed as a unit during acquisition, their organization and execution on retention trials may have been planned to a greater degree following movement initiation.

Unlike previous experiments (e.g. Lee & Magill, 1983; Shea & Morgan, 1979) there were no reliable differences between blocked and random practice conditions at retention on either of TRT, RT and MT. In experiments by both Lee and Magill (experiment 1), and Shea and Morgan, performance of blocked and random conditions were similar at the end of acquisition and became different at retention, producing a reliable Acquisition Condition x Trial Block interaction. The Acquisition Condition x Trial Block interaction was also reliable in this experiment. However, at the end of acquisition, performance of blocked and random conditions were reliably different in favor of the blocked condition, (TRT $t(11) = 4.33$, two tailed $p<.001$, RT $t(11) = 4.69$, two tailed $p<.001$, MT $t(11) = 2.39$, two tailed $p<.05$) and became similar at retention.

From this experiment it can be concluded that the capability to recall, organize and execute a multisegment arm movement sequence in response to the color to which it had been paired was disrupted for blocked participants on retention trials but remained intact for random participants, demonstrating the classic CI effect. Either CI theory is equally capable of explaining these results. The TAPR suggests that the effect is due to differences in retrieval practice provided by the conditions of blocked and random practice. The TPED suggests that it is due to differences in the strategies used to encode the S-R mapping which developed as a function of each practice condition. To dissociate the predictions made by each theory, Experiment 1 was replicated in Experiment 2 with the exception that the tasks to be learned were not variations of a task within a constant task environment (i.e. different movement sequences on the same apparatus), rather they were completely different tasks.

EXPERIMENT 2

In this experiment, the blue spatial pattern used in Experiment 1 was practiced along with two video games. Maze of Doom is a video game that involves negotiating a game player through a three dimensional labyrinth, and Ultris Puzzle is a video game in which the actor must align and place falling bricks along the bottom of the playing screen to form a solid wall. When the CI paradigm is used but the characteristics of the multiple tasks being learned are dissimilar, different predictions are made by the TAPR and the TPED.

Recall that the TAPR maintains that retrieval practice underlies the CI effect. Even though the tasks-to-be-learned in this experiment are distinctly different from each other, the TAPR predicts that the CI effect should once again emerge because random practice participants must still reformulate the action plan for each task on virtually every trial.

The TPED predicts that differentiating and encoding the features of the tasks underlies the CI effect. In this experiment, however, the features of the tasks-to-be-learned are clearly distinct from each other from the outset. This being the case, the TPED predicts that performance under blocked and random practice conditions should not be reliably different because the need to differentiate task features in the formation of distinct memory structures will have been eliminated under random practice conditions. Experiment 2 was designed to test these predictions.

Methods

Participants

Participants were 24 students (16 female, 8 male, mean age 20.3 years, sd=2.6) enrolled in a course offered by the School of Physical Education at the University of Victoria. Informed consent was obtained for all participants. Course credit was given in exchange for voluntary participation in the experiment.

Apparatus and Tasks

The experimental tasks in this experiment included the blue spatial pattern of PBC, a maze pattern from the computer game Maze of Doom, and a puzzle pattern from the computer game Ultris Puzzle. The blue PBC pattern was chosen for this experiment because it showed the largest contextual interference effect on immediate retention trials in Experiment 1. All task requirements for PBC were identical to Experiment 1.

Maze of Doom (MD) is a three dimensional computer game, readily available on the Internet. It was modified for the purposes of this experiment. The participants' objective was to negotiate a game player through the maze as quickly as possible. The maze was seen from the first person perspective of the game player so that, visually, it appeared to participants as if they were walking through the hallways and rooms of a three dimensional building. Participants negotiated the game player around the maze with a joystick. Moving the joystick along the saggital plane moved the game player forwards and backwards, while moving the joystick along the horizontal plane allowed the game player to rotate to the left and right.

For the game player to reach the end of the maze, it was necessary to successfully complete a number of tasks and successfully maneuver around a number of obstacles. First, participants needed to open a number of doors through out the maze. One of the doors required a key which could only be obtained at a fixed position in the maze. Pressing a button on the side of the joystick opened and closed doors. Second, the doors to exit one of the rooms in the maze were not visible until participants triggered a switch located on one of the walls. The same joystick button that opened doors, triggered the switch. Third, at the end of the maze was another switch that participants needed to trigger to end the trial. Triggering the switch stopped the game's timer. Initiation of the timer was controlled by the researcher. Fourth, at a number of points in the maze, participants had to decide which way to proceed. Incorrect choices resulted in reaching a dead end or entering a circular path. And fifth, the game player was not alone in the maze. A number of opponents were located at various points through out the maze. The starting position of these opponents were the same for each trial. They remained stationary until activated by the proximity of the game player. Once activated, opponents would track and target the game player. Each hit taken by the game player reduced its operating power. Participants could retaliate by squeezing the trigger on the joystick with their index finger. A series of hits taken by an opponent eliminated it from play. Throughout the maze participants were constantly challenged to decide between self defense and speed of maze completion. Successful completion of the maze did not require the acts of self defense.

The researcher provided participants with a standardized set of verbal instructions which informed them of the tasks and obstacles involved in completing the maze. Following each trial, participants were provided with feedback on the duration of the trial (in minutes and seconds) as well as the percentage of opponents eliminated. Time to complete the maze, in seconds, was used as the dependent measure.

Ultris Puzzle (UP) is a computer puzzle game in which participants arrange descending geometric bricks at the bottom of the playing screen to create a solid wall. Geometric bricks are formed by configuring four small squares into various geometric designs with the stipulation that the squares have to be attached to each other on at least one side. Bricks appear one at a time at the top of the computer screen and begin to descend to the bottom of the screen. As soon as one brick was placed at the bottom of the playing screen the next brick would appear. The identity of the next brick to appear at the top of the screen was hidden from participants. The completed puzzle contained fourteen bricks that together formed four solid lines of cubes.

Participants received points each time a brick was placed at the bottom of the screen. The faster bricks could be moved, rotated and dropped into place, the greater the number of points awarded. Movement of the bricks was controlled by the two clusters of keys on a computer keyboard that lie between the main key pad and the number pad. In the upper cluster, bricks could be rotated counter clockwise by pressing the "home" key, clockwise by pressing the "page up" key, and could be dropped immediately to the bottom of the screen by pressing the "page down" key. In the lower cluster, the left arrow shifted bricks left, the right arrow shifted bricks right, and the down arrow key shifted bricks down. The order in which bricks appeared at the top of the screen was the same for each trial. Participants' objective was to complete the puzzle as quickly as possible to obtain the highest possible score. Time to complete the puzzle, in seconds, served as the dependent measure.

Procedures

As in Experiment 1, each participant (N=24) attended an orientation session in which informed consent was obtained. Participants were then given verbal instructions and an opportunity to become familiar with each of the experimental tasks. Participants were provided with ten orientation trials on PBC, two orientation trials on MD, and five orientation trials on UP. The pattern sequence for PBC and UP were different in the orientation session than in the acquisition period, while the maze for MD was the one to be learned during the experiment. Total response time (TRT) performance data on the last five orientation trials for PBC were blocked, and block means were used to rank order participants from fastest to slowest. Participants were then paired by rank and, within each pair, randomly assigned to either a blocked or random practice experimental group using a random numbers table.

Procedures during the practice and retention phases of this experiment were identical to those in experiment 1. Participants in the random practice group received a rest period after completing 18, 36 and 45 trials. Participants in the blocked practice group received a rest period between tasks and in addition received a rest break following trials 6, 12 and 15 of MD, and following trials 6 and 12 of UP. Additional breaks were provided to blocked practice participants due to the fatigue produced by the length of trials for MD and UP.

Data Analysis

Total response time (TRT; in seconds) for PBC, MD and UP were collapsed into blocks of three trials for statistical analysis, resulting in six blocks of acquisition data and two blocks of retention data. Separate analyses were conducted to assess acquisition and retention performance using SPSS 9.0 GLM Repeated Measures. For the acquisition period, the factors of Acquisition Condition (blocked and random), Trial Blocks (1 - 6) and Task (PBC, MD, UP) were combined in a 2 x 6 x 3, respectively, ANOVA with repeated measures on all factors. The Acquisition Condition factor was treated as a within subjects factor because participants were match paired according to PBC TRT prior to being randomly assigned to groups. Retention performance was assessed by a 2 (Acquisition Conditions) x 3 (Trial Blocks: acquisition block 6, 10 minute retention, 1 week retention) x 3 (Tasks) ANOVA with repeated measures on all factors. The last acquisition trial block was included in the analysis to capture information about immediate retention performance relative to the level of performance achieved during the last acquisition trial block.

Results

Acquisition

Performance summaries for each acquisition condition on each task are shown in Figure 5. Reliable multivariate main effects emerged for Task, Wilks' exact $\underline{F}(2, 10) = 753.67$, $\underline{p}<.001$, $\eta^2 = .993$, and Trial Block, $\underline{F}(5, 7) = 41.52$, $\underline{p}<.001$, $\eta^2 = .967$.

Due to differences in the demands of each task not surprisingly the time to complete each task was reliably different. Completion times were shortest on PBC (M=1.25 secs) followed by UP (M=37.75 secs) and MD (M=73.96 secs).

Contrasts between adjacent trial blocks revealed that, overall, mean performance times improved between all but the last two acquisition trial blocks. Values of $\underline{F}$s, $\underline{p}$s and η^2s for each of these reliable contrasts are shown on Table 4. However, the TB main effect was overshadowed by the interaction of TB with Task, Wilks' exact $\underline{F}(10, 2) =$

46.88, $p<.05$, $\eta^2 = .996$. Post hoc analysis indicated that participants showed more substantial improvement on UP than PBC across all trial blocks. MD performance also improved at a greater rate than PBC performance across all trial blocks, however the rate of improvement was only found to be reliably different between the first four trial blocks. Lastly, the rate of improvement for MD and UP was similar across all trial blocks except between trial blocks 2 and 3 where MD performance times improved to a greater extent than UP performance. Values of Fs, ps and η^2s for each of these reliable Task x TB contrasts are shown on Table 5.

Figure 6 shows the mean frequency of recall errors on PBC for blocked and random practice participants during the acquisition, immediate retention and delayed retention phases. As can be seen, compared to Experiment 1 (see Figure 4), very few recall errors were made by participants in either acquisition condition, and no reliable difference in recall error frequency emerged during the acquisition period as a function of acquisition condition.

Retention

A 2 (Acquisition Conditions) x 3 (Tasks) x 3 (Trial Blocks) repeated measures ANOVA revealed reliable multivariate main effects for Task, Wilks' exact $F(2, 10) = 2101.85$, $p<.001$, $\eta^2 = .998$, and Trial Block, Wilks' exact $F(2, 10) = 9.95$, $p<.01$, $\eta^2 = .666$. Again, task completion times were shortest for PBC (M=1.160 secs), followed by UP (M=30.60 secs) and MD (M=64.45 secs).

The reliable Trial Blocks effect revealed that, overall, completion times continued to decrease between the last trial block of the acquisition phase and the trial block of the immediate retention phase, $\underline{F}(1, 11) = 15.31$, $\underline{p} < .01$, $\eta^2 = .582$, but that they increased between the immediate and delayed retention phases, $\underline{F}(1, 11) = 21.11$, $\underline{p} < .001$, $\eta^2 = .657$.

However, a reliable interaction between TB and Task, Wilks' exact $\underline{F}(4, 8) = 19.68$, $\underline{p} < .001$, $\eta^2 = .908$, indicated that the overall pattern of change across trial blocks was not consistent between tasks. Post hoc analysis of TB x Task contrasts showed that mean completion times for MD and UP continued to decrease from the last acquisition trial block to immediate retention while completion times for PBC stayed the same, although the size of the PBC vs UP contrast was small, PBC vs MD, $\underline{F}(1, 11) = 14.91$, $\underline{p} < .01$, $\eta^2 = .576$, PBC vs UP $\underline{F}(1, 11) = 4.71$, $\underline{p} = .053$, $\eta^2 = .300$. Between the immediate and delayed retention phases, completion times slowed on all tasks, as indicated by the reliable TB contrast, with completion times for UP and MD slowing more substantially than for PBC, but not more slowly than each other, PBC vs MD $\underline{F}(1, 11) = 8.27$, $\underline{p} < .05$, $\eta^2 = .429$, PBC vs UP $\underline{F}(1, 11) = 76.87$, $\underline{p} < .001$, $\eta^2 = .875$.

Analysis of recall errors showed no reliable Experimental Phase, Acquisition Condition or interaction effects. Although the data indicate a trend towards a greater frequency of recall errors made during the delayed retention than immediate retention and acquisition phases, the size of these effects were small and not reliable at acceptable levels.

Discussion

Conspicuously absent in this experiment are any traditional contextual interference effects. Traditionally, participants have a more difficult time under random than blocked conditions while learning the tasks, presumably because the interleaving of tasks during random practice makes it more difficult for participants to remember what they have learned.

According to the TAPR, learning under random practice conditions should have been more difficult during the acquisition phase, and hence performance should have suffered, because the interference caused by interleaving tasks required action plans to be repeatedly reformulated (Lee & Magill, 1983, 1985). Although TRT was higher for random than blocked participants on the first trial block, the difference was not reliable and disappeared by the second trial block. The similarity of acquisition performance between random and blocked practice participants in the present experiment indicated that, virtually from the outset, switching between tasks during the acquisition phase did not produce retrieval difficulties for random practice participants. This suggests that the difficulties associated with random practice conditions is not simply a function the order in which tasks are practiced, as is implicated by the TAPR.

The similarity of performance under blocked and random conditions during the immediate retention phase, coupled with the similarity of performance on the last block of the acquisition phase, indicated that practice conditions did not affect participants' ability to retain the levels of performance that they had achieved during the acquisition phase. The similarity in recall error frequency for the PBC pattern supports this finding. Since random practice still forced participants to retrieve task solutions repeatedly throughout the acquisition phase, the similarity in plan access and implementation under blocked and random practice conditions suggests that retrieval practice alone during the acquisition period is insufficient to produce contextual interference. Taken together, the results of the present experiment fail to support the predictions made by the TAPR. Evidently, factors other than the order in which tasks are practiced are responsible for the contextual interference effect.

GENERAL DISCUSSION: EXPERIMENTS 1 & 2

When the results of the experiments 1 and 2 are considered together, a clearer picture of the cognitive operations underlying the contextual interference effect emerges. The blue PBC pattern was common to both experiments. In Experiment 1 it was learned in the context of two similar tasks (2 patterns of the same length on the same apparatus) and in Experiment 2 it was learned in the context of two dissimilar tasks (2 computer games on a different apparatus). Figure 7 overlays the results from Experiments 1 and 2 on the blue pattern. These figures show that performance under random and blocked practice conditions was substantially different during the acquisition phase of Experiment 1, but similar in Experiment 2. From this comparison it can be concluded that random practice was more difficult when tasks were variations within a constant performance environment than when the tasks and performance environments of each task were unique. Furthermore, the performance of blocked participants slowed substantially over the immediate retention interval in Experiment 1 and remained at that level for the delayed retention phase, whereas this was not the case in Experiment 2. This suggests that the retrieval of task solutions during immediate and delayed retention phases was more difficult for blocked participants for task variations than unique tasks.

Shea and colleagues (Shea & Morgan, 1979; Shea & Zimny 1983, 1988) have long maintained that task similarity plays a prominent role in producing intertask interference. The results from Experiments 1 and 2 support this contention and extend the definition of similarity to include consideration of the features of the environment in which tasks are performed.

A tentative explanation for how the similarity of tasks interacts with practice order includes discussion of the degree of ambiguity associated with the stimulus, the regulatory conditions in which the stimulus is given, and the locus of the cues used by participants to formulate task responses.

In Experiment 1, the stimulus to which each response pattern was paired was arbitrary (a color). Thus, the features of the stimulus alone could not meaningfully specify the response to be made. Furthermore, the regulatory and non regulatory features (Gentile, 1987) of the performance environment prior to initiating a trial did not change from trial to trial. For random practice participants, this meant that information about the upcoming stimulus was generally not readily available from either the regulatory or nonregulatory features of the performance environment prior to stimulus presentation. In this sense, the performance environment was ambiguous. For blocked participants there was no ambiguity regarding the stimulus that was to be presented. However, this was a function of the lack of uncertainty associated with event probabilities, and was not a function of differences in the regulatory or nonregulatory features of the performance environment. Under these circumstances, access and implementation of appropriate action plans for random practice participants, upon presentation of the stimulus, was solely dependent on their ability to recall the response that had been paired with the stimulus, as well as the correct sequence of response keys contained in the pattern. In adapting to these environmental conditions, acquisition performance was maintained for random practice participants.

The performance environment experienced by participants in Experiment 2 contained little ambiguity. Both regulatory and nonregulatory features of the performance environment for each task uniquely specified the response that would be required on the upcoming trial. The responses themselves were also very distinct. This complete lack of ambiguity made it easy for participants in both acquisition conditions to recall and organize the appropriate response for each task. Adaptations under these conditions produced similar retention performance under random and blocked practice conditions.

Since the CI paradigm was used in both experiments, the demand for cognitive processes involved in action plan retrieval were the same in each experiment for random practice participants. According to the TAPR the results for random practice participants should have been similar between experiments. The lack of support for the theoretical predictions of the TAPR when tasks and performance environments were dissimilar suggests that cognitive effort associated with action plan reconstruction may not underlie the contextual interference effect.

Indeed, Wood and Ging (1991) have also suggested that action plan reconstruction may not underlie the CI effect. The tasks in their experiment, rather than being unambiguous due to uniqueness, were unambiguous because they were identical (spatial patterns forming the letter N, in three sizes). They compared performance on multiple tasks with identical spatial patterns with performance on tasks where the spatial patterns did not resemble each other (pattern variations). The results showed that the CI effect emerged with the pattern variations, but did not emerge with the identical patterns. They concluded that the emergence of the CI effect was dependent on the features of the tasks to be learned rather than "reformulations of the movement plans on each trial" (p. 25), and ruled out the TAPR as an acceptable explanation of the CI effect.

Since retrieval practice was ruled out as an explanation for the CI effect in Experiment 2, it would seem that the CI effect may be primarily a function of cognitive processes involved with the encoding of task solutions, as Shea and Morgan (1979) originally hypothesized. Adopting the TPED, it would seem that in Experiment 1, the CI effect was a product of (a) the comparison, evaluation, and differentiation of the features of the tasks, and (b) the formation of a meaningful retrieval structure (by way of association with current knowledge) with which to recall this information (Chase & Ericsson, 1981). These cognitive processes may have been active throughout the acquisition phase of Experiment 1 for random practice participants. Therefore, task solutions became more elaborately and distinctively encoded, allowing the memory structures for each task to become less ambiguous and more distinct. Conversely, evaluative cognitive processes were generally inactive for blocked practice participants. The predictability of the practice situation under blocked conditions likely did not evoke the same urgency in participants to develop meaningful associations for either the S-R pairing or the sequence of response keys. During the retention phases of Experiment 1, retrieving the appropriate response when given its stimulus was challenging for blocked practice participants likely because the S-R pairing, as well as the sequence of response keys, had not been encoded as distinct memory structures. The ambiguity surrounding which response sequence was to be paired with each stimulus can be seen by the increase in the number of recall errors made by blocked practice participants during retention phases, and the slowing of movement speed during pattern execution.

In Experiment 2, it was not necessary for random practice participants to compare, evaluate, differentiate, and associate the features of the tasks to form distinct retrieval structures. Since the demands for cognitive effort were reduced for random participants by the unique regulatory features of the tasks, the memory advantage that they traditionally demonstrate over blocked participants disappeared. Therefore, Experiment 3 was conducted to more closely examine the nature of these cognitive processes. More specifically, the aim of Experiment 3 was to see whether actively engaging blocked participants in evaluation, differentiation, and association would eliminate the CI effect.

EXPERIMENT 3

In the previous experiment, by practicing three tasks with unique task demands and environmental features, blocked and random practice participants were equally capable of forming distinct memory representations of each task. It was concluded that with unique tasks demands for cognitive effort were similar regardless of practice condition thereby eliminating the CI effect. The aim of the present experiment was to increase the demand for cognitive effort in blocked practice participants through purposeful evaluation of task variations, that were to be performed in a constant performance environment. Analysis of task demands for PBC indicated that successful retention performance was dependent on participants' ability to (a) quickly establish the correct S-R pairing upon presentation of a color word, (b) recall the correct sequence of response keys for the spatial pattern requested, and (c) move efficiently through the response key sequence. In Experiment 1, it was suggested that under random practice conditions, the ambiguity of the S-R pairings and response key sequences diminished over the course of the acquisition phase because the interleaving of tasks allowed various features of the tasks to be compared, evaluated, differentiated and meaningfully associated (encoded).

Evaluation can be seen to play a pivotal role in differentiation, since the differences of task features along a particular dimension must be evaluated before tasks can be properly differentiated along that dimension. According to the TPED, the basis of the CI effect lies in the distinctiveness of the memory representations created by the process of differentiation, which arises as a function of random practice. Thus, the present experiment was an extension of Experiment 1 in which a second group of blocked and random practice participants were added. Blocked verbalize (BV) participants were asked to evaluate, differentiate, and meaningfully associate the features of each task during the acquisition phase, and verbalize their thoughts. Random verbalize (RV) participants were asked to simply report the cognitive strategies they used to learn and perform the patterns. Analyzing the features of each task throughout the acquisition phase should increase the demands for cognitive effort in blocked practice participants and allow them to create distinct memory representations for each task, similar to random practice participants. It was hypothesized that, practicing under these conditions, blocked practice participants would retain the level of performance achieved at the end of the acquisition period and perform similar to random practice participants at retention, thereby eliminating the CI effect.

Previous experiments have manipulated intertask and intratask interference during blocked practice. Wright (1991) and Wright, et al. (1992) supplemented two groups of blocked participants with either intratask or intertask interference. Results showed that extra cognitive processing during the acquisition phase did not affect acquisition performance, but did facilitate retention performance such that the retention performance of supplemented blocked groups was not reliably different from a random group. Whereas in the Wright, and Wright et al. experiments evaluative processing was indirectly manipulated through intratask and intertask interference, in the present experiment evaluative processing was manipulated directly by requiring analysis of the features of the tasks.

To observe the effects of evaluation and verbalization, the present experiment was conducted in two parts. In Part A, two blocked groups and two random groups practiced the PBC spatial patterns from Experiment 1 under traditional blocked and random practice conditions. Part A of the experiment served two purposes; to replicate Experiment 1, and to establish a within subjects control condition against which to compare the effects of the intervention provided in part B of the experiment. In Part B, one blocked group served as a control group (BC) and performed under the same conditions as Part A. A second blocked group (BV) was instructed to purposefully evaluate the spatial patterns and to verbalize these evaluations. Similarly, one random group served as a control group (RC) performing under identical conditions in Parts A and B, while the other random group (RV) was instructed to verbalize their thoughts throughout the acquisition period. The purpose of this experiment was to observe the patterns acquisition and retention performance of the groups during Parts A and B of the experiment.

It was expected that the results of Part A would replicate the results obtained in Experiment 1 with pairs of random and blocked groups. For Part B, acquisition performance was hypothesized to be similar to Part A, in accordance with the findings of Wright (1991) and Wright, et al. (1992). Retention performance between random groups was expected to be similar since, although more meta-cognitive activity may have been required by the RV group in the form of introspection, the cognitive demands elicited by their practice schedule during the acquisition period were similar between RC and RV groups. Conversely, the BV group was expected to have engaged in more evaluative processing than the BC group during acquisition, and therefore was expected to retain a greater portion of what they had learned. Thus, it was expected that the change in performance over the first retention interval would be similar among the BV, RV, and RC groups.

Methods

Participants

Participants were 48 students enrolled at the University of Victoria. The data from one participant was removed from the study because the participant showed a lack of interest in learning the tasks. Of the remaining 47 participants, there were 26 females and 21 males (mean age 21.5 years, sd=2.9). Informed consent was obtained for all participants. Course credit was exchanged for voluntary participation in the experiment.

Apparatus and Task

PBC was the experimental task used in this experiment. See Experiment 1 for a detailed description of the apparatus and task.

Procedures

Following the orientation session, the last five orientation trials were blocked, and block means were used to rank order subjects (originally n=36) from fastest to slowest. Triads were formed by rank order and, within each triad, subjects were randomly assigned to either a blocked control, blocked verbalize, or random verbalize experimental condition using a random numbers table. The RC condition (n=12) was added to the experiment later. These experimental groupings became important in the second part of the experiment.

Part A was a replication of Experiment 1 using pairs of blocked and random practice groups. Participants practiced the spatial patterns used in Experiment 1 (red, yellow, and blue; see Figure 2) according to the same procedures.

Part B began following the conclusion of the 1 week retention phase of Part A. In Part B, participants practiced a second set of three spatial patterns labeled white, fuchsia and lime (see Figure 8). Blocked control (BC) and random control (RC) participants practiced under the same conditions as they had in Part A. Blocked verbalize (BV) participants practiced according to a blocked schedule but were prompted to (1) generate strategies that would allow them to quickly associate a color word with its corresponding spatial pattern, (2) encode the sequence of keys in each pattern in a meaningful way, (3) practice estimating their movement speed so it could be estimated in the absence of feedback, and (4) devise an eye focusing strategy upon presentation of the stimulus similar to the strategy they would have to use during the retention phases (transfer appropriate processing). Participants were instructed to evaluate their performance in practice with reference to how effective it would be in the context of retention, and to verbalize these evaluations during intertrial intervals. Random verbalize (RV) participants were instructed to reflect on what kinds of strategies they used to quickly get off the start key following the stimulus signal, how they remembered the sequence of keys in each pattern, and how they knew whether a trial was performed relatively quickly or slowly during intertrial intervals, and to verbalize their thoughts during intertrial intervals.

The think out loud protocol was used to ensure that BV participants engaged in cognitive activity associated with task evaluation, and to examine the cognitive processes and types of task analyses entertained by BV and RV participants.

Following completion of all 54 acquisition trials, 9 no-feedback retention trials were given following intervals of ten minutes and one week according to the same retention procedures that were used in Experiments 1 and 3A.

<u>**Data Analysis**</u>

TRT, RT, and MT data were collapsed into blocks of three trials for statistical analysis, resulting in six blocks of acquisition data and two blocks of retention data for each dependent measure. Separate analyses using SPSS 9.0 GLM Repeated Measures were conducted to assess acquisition performance, the changes in performance over the first and second retention intervals, and average retention performance. For the acquisition period, the factors of Acquisition Condition (BC, BV, RC, RV), Task, and Trial Blocks (1 - 6) were combined in a 4 x 3 x 6, respectively, multivariate analysis of variance (MANOVA) with repeated measures on the last two factors. Acquisition condition was treated as a between subjects factor due to low correlations between groups on the matching factor, and the addition of the RC group.

Retention performance was assessed by a 4 (Acquisition Conditions) x 3 (Tasks) x 3 (Trial Blocks: acquisition block 6, immediate [10 minute] retention, delayed [1 week] retention) MANOVA with repeated measures on the last two factors. The last acquisition trial block was included in the analysis to capture information about relative changes in performance across the first retention interval. In Part A, the random groups (RC, RV) and the blocked groups (BC, BV) were considered together in all between condition comparisons.

Qualitative analysis of data obtained through the think aloud protocol during Part B of the experiment proceeded according to the procedures for interpretive inquiry outlined by Côté, Salmela, Baria, and Russell (1993), and Côté, Salmela, and Russell (1995). Verbalizations for the blocked and random verbalize participants were noted by the experimenter during each testing session. Following the end of the experiment these notes were transcribed and the text was reduced into meaning units; statements in which a single thought or idea was communicated. Meaning units were then grouped together into clusters based on the principles of internal homogeneity and external heterogeneity (Côté, et al., 1993). That is, meaning units within a cluster shared a common theme or idea that was not shared by meaning units of other clusters. Meaning unit clusters were then grouped together and organized into progressively higher order categories according to the same criteria.

Results

Part A: Acquisition

Results from a 4 (Acquisition Conditions) x 3 (Tasks) x 6 (Trial Blocks) factorial MANOVA with repeated measures on the last two factors and TRT, RT and MT serving as dependent measures revealed a reliable main effect for Task, Wilks' exact $F(4, 40) = 2.95$, $p<.05$, $\eta^2 = .228$. Univariate analyses revealed a reliable effect for MT, $F(2, 86) = 6.99$, $p<.01$, $\eta^2 = .140$, which was reliably slower on the blue pattern (M=1119.3 ms) than either the yellow (M=1086.8 ms) or the red (M=1053 ms) patterns. There were no reliable multivariate two way interactions of Task with either Acquisition Condition (AC) or Trial

Blocks (TB), and there was no reliable three way interaction of Task, AC and TB, indicating that, although the blue pattern was executed more slowly than the others, this trend was consistent across acquisition groups for the duration of the acquisition phase. Therefore, the data were collapsed across tasks and re-entered into a 4 (AC) x 6 (TB) factorial MANOVA with repeated measures on the second factor. Mean performance of the four acquisition groups for each of TRT, RT and MT, are shown on Figure 9.

Multivariate analysis of the contrast between the blocked and random groups revealed a reliable AC main effect, Wilks' exact $F(2, 42) = 37.09$, $p<.001$, $\eta^2 = .638$. The size of these effects are shown on Table 6. Overall, TRT, RT and MT acquisition performance of random practice participants was reliably slower than that of blocked practice participants, TRT $F(1, 43) = 41.85$, $p<.001$, $\eta^2 = .493$, RT $F(1, 43) = 73.00$, $p<.001$, $\eta^2 = .629$, MT $F(1, 43) = 10.09$, $p<.01$, $\eta^2 = .190$. Performance of the blocked groups did not differ from each other, $Fs<1$. While MT did not differ between the random groups, RT was reliably different, $F(1, 43) = 11.17$, $p<.01$, $\eta^2 = .206$, with the RC group reacting slower than the RV group. Differences between the random groups on TRT were marginally reliable, $F(1, 43) = 3.99$, $p=.052$, $\eta^2 = .085$.

The reliable AC x TB interaction indicated that the rate of performance change between trial blocks differed among the blocked and random groups during the acquisition period, Wilks' exact $F(10, 34) = 22.25$, $p<.001$, $\eta^2 = .867$. Univariate analyses revealed that these differences were reliable for each of TRT, RT and MT. For each dependent measure, the random groups showed reliably greater improvement than the blocked groups across each of the six acquisition trial blocks, with the exception of RT which did not show reliable improvements between the fourth and fifth trial blocks, and MT which did not show reliable improvements between the fifth and sixth trial blocks. Table 7 provides the values of F, p and η^2 for each reliable AC x TB contrast. The blocked groups did not reliably differ from each other at any time on any dependent measure. The random groups were not reliably different from each other across the first five trial blocks, however, the rate of performance change between the RC and RV groups was reliably different on each dependent measure between the fifth and sixth trial blocks, TRT $F(1, 43) = 9.61$, $p<.01$, $\eta^2 = .183$, RT $F(1, 43) = 8.11$, $p<.01$, $\eta^2 = .159$, MT $F(1, 43) = 4.25$, $p<.05$, $\eta^2 = .090$. As can be seen on Figure 9, the RC group showed a greater rate of improvement between fifth and sixth trial blocks than the RV group.

Finally, as shown on Figure 10, the random groups made significantly more recall errors (RC M=4.57; RV M=4.37) than the blocked groups (BC M=0.17; BV M=0.42) during the 54 trials of the acquisition phase, $F(1, 43) = 57.59$, $p<.001$, $\eta^2 = .573$. Neither of the blocked or random groups were reliably different from each other.

Part A: Retention

Results of a 4 (AC) x 3 (Tasks) x 3 (TB) factorial MANOVA on TRT, RT and MT with repeated measures on the second and third factor revealed a reliable multivariate effect for Task, Wilks' exact $F(4, 40) = 4.61$, $p<.01$, $\eta^2 = .315$, but no reliable two or three way interaction effects involving Task. Once again, MT for the blue pattern (M=1089.9 ms) was reliably slower than either the yellow (M=1042.5 ms) or the red (M=1028.8 ms) pattern. The lack of reliable interactions involving Task indicated a consistent pattern of results among acquisition groups from the last trial block of the acquisition phase, to the trial blocks of each retention phase. Therefore, the data were collapsed across tasks and re-entered into a 4 (AC) x 3 (TB) factorial MANOVA with repeated measures on the TB factor.

A reliable multivariate AC x TB interaction effect was found between the blocked groups and the random groups, Wilks' exact $\underline{F}(4, 40) = 20.45$, $\underline{p} < .001$, $\eta^2 = .672$, indicating that the rate of performance change across the retention intervals was different between the blocked and random groups. Univariate analyses showed that the rate of performance change between the blocked and random groups differed considerably from the end of acquisition to immediate retention on each of TRT, $\underline{F}(1, 43) = 48.85$, $\underline{p} < .001$, $\eta^2 = .532$, RT, $\underline{F}(1, 43) = 31.39$, $\underline{p} < .001$, $\eta^2 = .422$, and MT, $\underline{F}(1, 43) = 34.34$, $\underline{p} < .001$, $\eta^2 = .444$. The rate of performance change was not reliably different among either of the pairs of blocked or random groups. Figure 9 shows that performance on all dependent measures remained relatively stable for the random groups, but slowed substantially for the blocked groups. Effect size calculations for these differences are shown on Table 8. The change in performance between the blocked and random groups did not differ reliably between immediate and delayed retention.

While the blocked groups demonstrated superior performance relative to the random groups during the acquisition period, this trend was almost completely reversed on tests of retention, Wilks' exact $\underline{F}(2, 42) = 5.32$, $\underline{p} < .01$, $\eta^2 = .202$. TRT and RT retention performance was found to be superior for the random groups compared to the blocked groups, TRT $\underline{F}(1, 43) = 9.56$, $\underline{p} < .01$, $\eta^2 = .182$, RT $\underline{F}(1, 43) = 9.88$, $\underline{p} < .01$, $\eta^2 = .187$, while neither of the blocked (BC, BV) nor random (RC, RV) groups were reliably different from each other. Group means for MT among random and blocked groups were not found to differ reliably.

Figure 10 shows mean recall error frequency for each experimental group during the acquisition, immediate, and delayed retention phases of Part A of the experiment. Analysis of recall errors contrasting the random with the blocked groups revealed a reliable AC x Experimental Phase interaction, Wilks' exact $\underline{F}(2, 42) = 38.74$, $\underline{p}<.001$, $\eta^2 = .648$. It was found that, from the acquisition to the immediate retention phase, recall error frequency decreased for the random group while at the same time it increased for the blocked groups, $\underline{F}(1, 43) = 77.81$, $\underline{p}<.001$, $\eta^2 = .644$. The change in error frequency from immediate to delayed retention was not reliably different between random and blocked groups. Furthermore, neither of the blocked or random groups were reliably different from each other between any of the experimental phases.

A between groups analysis of recall error frequency at retention (IR & DR) showed that the random groups made an average of 1.7 fewer recall errors than the blocked groups, $\underline{F}(1, 43) = 34.70$, $\underline{p}<.001$, $\eta^2 = .224$. Once more, neither of the blocked nor random groups were reliably different from each other.

Part B: Acquisition

Analysis of the acquisition data using a 4 (AC) x 3 (Tasks) x 3 (TB) factorial MANOVA on TRT, RT and MT with repeated measures on the second and third factor revealed reliable multivariate effects for task, Wilks' exact $\underline{F}(4, 40) = 6.31$, $\underline{p}<.001$, $\eta^2 = .387$. Reliable univariate effects were found for TRT, $\underline{F}(2, 86) = 6.97$, $\underline{p}<.01$, $\eta^2 = .140$, and RT, $\underline{F}(2, 86) = 12.26$, $\underline{p}<.001$, $\eta^2 = .222$, but not MT. Overall, participants performed the white task reliably faster (1386.5 ms) than either the fuchsia (1465.8 ms) or lime tasks (1452.5 ms). Participants also responded more quickly, when presented with the white than fuchsia or lime tasks, as evidenced by lower RT's, 409.3, 452.7, 462.4 ms respectively.

The acquisition data could not be collapsed across tasks in this part of the experiment because the presence of reliable two way interactions of task with AC, Wilks' exact $F(12, 106) = 3.36$, $p<.001$, $\eta^2 = .247$, and with TB, Wilks' exact $F(20, 24) = 2.04$, $p<.05$, $\eta^2 = .629$, signified that the pattern of performance for the white, fuchsia and lime tasks were not consistent either among groups or over trial blocks.

Univariate analysis of the Task x AC interaction revealed differences in RT performance among groups between tasks, $F(6, 86) = 7.74$, $p<.001$, $\eta^2 = .351$. No reliable differences emerged for MT and TRT performance. Group differences for RT were found between the white and fuchsia tasks, $F(3, 43) = 13.51$, $p<.001$, $\eta^2 = .485$, and white and lime tasks, $F(3, 43) = 10.39$, $p<.001$, $\eta^2 = .420$. Follow up tests showed that while RT performance was consistent for the BV and BC groups between the white and fuchsia tasks, it increased on the fuchsia task for both of the random groups, and increased to a greater extent for the RC than the RV group. In addition, the increase in RT between the white and lime tasks was greater for the RC than each of the RV, BV, and BC groups, whose RT performance on these two tasks were not reliably different from each other. These results are illustrated on Figure 11. Table 9 includes the values of F, p and η^2 for all reliable Task x AC interactions.

Follow up tests of the Task x TB interaction revealed that both TRT and RT improved between trial blocks 1 and 2 to a greater extent on the lime (by 438.5, 260.6 ms respectively) than the white task (by 226.1, 89.2 ms respectively), TRT $F(1, 43) = 14.37$, $p<.001$, $\eta^2 = .250$, RT $F(1, 43) = 10.67$, $p<.01$, $\eta^2 = .199$.

Task differences not withstanding, a reliable multivariate effect emerged for acquisition conditions, Wilks' exact $F(6, 84) = 10.52$, $p<.001$, $\eta^2 = .429$. Analysis of the effect showed that, similar to the results of Part A, overall TRT, RT and MT performance of the blocked groups was faster than that of the random groups, TRT $F(1, 43) = 27.96$, $p<.001$, $\eta^2 = .394$, RT $F(1, 43) = 73.47$, $p<.001$, $\eta^2 = .631$, MT $F(1, 43) = 7.00$, $p<.05$, $\eta^2 = .140$. Mean performance of the four acquisition conditions for each of TRT, RT and MT, are shown on Figure 12. The size of these effects are shown on Table 10. Similar to the results of Part A, performance of the blocked groups did not differ from each other, $Fs<1$. RT performance between the random groups was found to be reliably different, $F(1, 43) = 11.71$, $p<.01$, $\eta^2 = .214$, with the RC group reacting more slowly than the RV group to the presentation of the tasks. Although the RC group executed the tasks more slowly than the RV group, differences in MT were not reliable while differences in TRT were marginally reliable, $F(1, 43) = 3.98$, $p=.053$, $\eta^2 = .085$.

As expected, the rate of performance change across acquisition trial blocks differed between blocked and random groups, Wilks' exact $F(10, 34) = 6.89$, $p<.001$, $\eta^2 = .670$. For each of TRT, RT, and MT the random groups showed greater improvement than the blocked groups between the first two trial blocks, TRT $F(1, 43) = 33.13$, $p<.001$, $\eta^2 = .435$, RT $F(1, 43) = 22.61$, $p<.001$, $\eta^2 = .345$, MT $F(1, 43) = 20.75$, $p<.001$, $\eta^2 = .326$. The same trend emerged for MT between trial blocks 4-5, $F(1, 43) = 10.09$, $p<.01$, $\eta^2 = .190$. This was because the rate of MT improvement for the RV group was reliably greater that of the RC group between trial blocks 4-5, MT $F(1, 43) = 4.86$, $p<.05$, $\eta^2 = .101$. Improvement rates did not differ between the BC and BV groups on any dependent measure at any time during the acquisition phase.

Lastly, as shown on Figure 13, group differences again emerged regarding the frequency of recall errors made during the acquisition phase. A one way ANOVA revealed that, as in Part A, the frequency of recall errors was greater for the random groups (RC M=5.58; RV M=2.09) than for the blocked groups (BC M=0.50; BV M=0.92), $F(1, 43) = 22.27$, $p<.001$, $\eta^2 = .341$. While the error frequency among blocked groups was not reliably different, the RC group was found to have experienced more recall errors than the RV group, $F(1, 43) = 13.57$, $p<.001$, $\eta^2 = .240$ during the acquisition phase.

Part B: Retention

Results of a 4 (AC) x 3 (Tasks) x 3 (TB) factorial MANOVA on TRT, RT and MT with repeated measures on the second and third factor revealed no reliable multivariate main or interaction effects involving Task, indicating a consistent pattern of performance among acquisition conditions over the tasks from the last trial block of acquisition, to the trial blocks of immediate and delayed retention. Therefore, the data were collapsed across tasks and re-entered into a 4 (AC) x 3 (TB) factorial MANOVA with repeated measures on the second factor.

Retention performance for the BV group was expected to be similar to that of RC and RV groups, which together was expected to be different from the BC group. A reliable AC x TB interaction confirmed that performance change for the BC group was different than that of the BV, RC, and RV groups combined, Wilks' exact $F(4, 39) = 17.38$, $p<.001$, $\eta^2 = .641$. Performance was found to have slowed considerably from the end of acquisition to immediate retention for the BC group compared to the other three groups, TRT $F(1, 42) = 41.29$, $p<.001$, $\eta^2 = .496$, RT $F(1, 42) = 25.30$, $p<.001$, $\eta^2 = .376$, MT $F(1, 42) = 43.82$, $p<.001$, $\eta^2 = .511$. Individual comparisons were made between the BV and BC groups to see if the deterioration of performance between the end of acquisition, immediate retention and delayed retention was similar between them, as had been the case in Part A of the experiment. A reliable multivariate AC x TB interaction effect was found, Wilks' exact $F(4, 39) = 2.61$, $p<.05$, $\eta^2 = .211$. Univariate tests showed that TRT and MT change from the end of acquisition to immediate retention had slowed to a greater extent for the BC than the BV group, TRT, $F(1, 42) = 5.29$, $p<.05$, $\eta^2 = .112$, MT, $F(1, 42) = 6.00$, $p<.05$, $\eta^2 = .125$. While RT also slowed to a greater extent for the BC than the BV group, the magnitude of the effect was not large enough to be considered reliable, $F(1, 42) = 3.14$, $p=.09$, $\eta^2 = .069$.

The change in performance between the RC and RV groups was equivalent from the end of acquisition to immediate retention, as can be seen in Figure 12.

However, similar to the results of Part A, a reliable multivariate effect was found between the BV and random groups as well, Wilks' exact $F(4, 39) = 10.41$, $p<.001$, $\eta^2 = .516$. Performance for the BV group was found to have deteriorated to a greater extent than for the random groups between the end of acquisition and immediate retention on each dependent measure, TRT, $F(1, 42) = 25.96$, $p<.001$, $\eta^2 = .382$, RT, $F(1, 42) = 16.31$, $p<.001$, $\eta^2 = .280$, and MT, $F(1, 42) = 26.13$, $p<.001$, $\eta^2 = .383$. Table 11 shows effect sizes for the AC x TB interaction for each of these pairwise comparisons.

Analysis of performance change from immediate to delayed retention revealed no reliable differences between the BV and BC groups. Performance of the RC group was not reliably different from that of the RV group for TRT and RT, however MT was found to have slowed to a greater extent for the RC than for the RV group, MT, $F(1, 42) = 6.05$, $p<.05$, $\eta^2 = .126$. Performance for the BV group was reliably different from the random groups for each of TRT, RT and MT. Compared to the improvements shown by the BV group, performance deteriorated for the random groups for TRT and RT. MT slowed for the random groups but was maintained for the BV group.

A comparison of average retention performance between the BC group and the other three groups combined revealed that the retention performance of the BC group was inferior to that of the other three groups combined, Wilks' exact $\underline{F}(2, 41) = 3.52$, $\underline{p} < .05$, $\eta^2 = .146$. In particular, TRT and RT was found to be reliably slower for the BC group, TRT $\underline{F}(1, 42) = 5.11$, $\underline{p} < .05$, $\eta^2 = .108$, RT $\underline{F}(1, 42) = 6.36$, $\underline{p} < .05$, $\eta^2 = .131$, while MT did not differ reliably. Whereas the retention performance of the BV group and each of the random groups was found to differ reliably in Part A of the experiment, they did not in Part B. However, neither did the retention performance of the BV group differ reliably from that of the BC group. As was the case in Part A, the retention performance of the two random groups was not reliably different.

Figure 13 illustrates mean recall error frequency for the four experimental groups during each phase of Part B of the experiment. The change in error frequency across acquisition, immediate retention, and delayed retention was reliably different between the BC group and the other three groups combined, Wilks' exact $\underline{F}(2, 41) = 9.37$, $\underline{p} < .001$, $\eta^2 = .314$. Univariate analyses localized the effect between the end of acquisition and immediate retention where the frequency of recall errors rose for the BC group and, on average, dropped for the other three groups, $\underline{F}(1, 42) = 18.84$, $\underline{p} < .001$, $\eta^2 = .310$. The change in error frequency across experimental phases was not reliably different between the BC and BV groups. However, the changes in recall error frequency between the RC and RV groups were reliably different across experimental phases, Wilks' exact $\underline{F}(2, 41) =$

6.43, $p<.01$, $\eta^2 = .239$. While error frequency for both random groups decreased from acquisition to immediate retention, the decrease was greater for the RC than the RV group, $F(1, 42) = 9.30$, $p<.01$, $\eta^2 = .181$. The trend was reversed from immediate to delayed retention; while recall error frequency increased for both groups across this retention interval, the increase was greater for the RC than the RV group, $F(1, 42) = 6.56$, $p<.05$, $\eta^2 = .135$. These results indicate that pattern recall was more stable for the RV than the RC group in Part B of the experiment. The pattern of recall error frequency between the BV and random groups was also found to differ reliably, Wilks' exact $F(2, 41) = 8.67$, $p<.001$, $\eta^2 = .297$. Recall error frequency remained stable for the BV group between acquisition and immediate retention whereas it decreased for the random groups, $F(1, 42) = 16.14$, $p<.001$, $\eta^2 = .278$. From immediate to delayed retention error frequency rose to a greater extent for the random groups than the BV group, $F(1, 42) = 4.35$, $p<.05$, $\eta^2 = .094$, likely due to the influence of the RC group.

Analysis of recall errors between control groups (BC & RC) produced the expected results between acquisition and immediate retention; error frequency increased for the BC group at the same time as it decreased for the RC group, $F(1, 42) = 38.45$, $p<.001$.,$\eta^2 = .478$. However, between immediate and delayed retention, error frequency increased to a greater extent for the RC than the BC group, $F(1, 42) = 4.70$, $p<.05$, $\eta^2 = .101$, which was not expected.

Comparison of the verbalization groups (BV & RV) revealed no reliable differences across either retention interval, although the difference in error frequency approached acceptable levels of reliability between the acquisition and immediate retention phases.

A between groups analysis of recall error frequency at retention (IR & DR) paralleled the results found for execution performance. Recall error frequency of the BC group was reliably greater than that of the other three groups combined, $\underline{F}(1, 42) = 5.48$, $\underline{p}<.05$, $\eta^2 = .115$. While the BV group made more recall errors than the RV group in Part A of the experiment, the error frequency of these groups was not reliably different in Part B. The difference in recall performance between the blocked groups was marginal, $\underline{F}(1, 42) = 3.78$, $\underline{p}=.059$, $\eta^2 = .083$, with the BV group tending to make fewer recall errors than the BC group. Unlike the error frequency results of Part A, a reliable difference emerged between the random groups, $\underline{F}(1, 42) = 5.09$, $\underline{p}<.05$, $\eta^2 = .108$, with the RC group making more recall errors than the RV group on retention trials.

Changes between Parts A and B: Acquisition

The acquisition data from Parts A and B were entered into a 4 (AC) x 2 (Parts) x 6 (TB) factorial MANOVA with repeated measures on the second and third factor and with TRT, RT and MT serving as dependent variables. A reliable multivariate main effect emerged for part, Wilks exact $\underline{F}(2, 42) = 19.76$, $\underline{p}<.001$, $\eta^2 = .485$. Univariate effects indicated that acquisition performance was reliably faster for each of TRT (by 141.6 ms), RT (by 48.7 ms), and MT (by 92.8 ms) during Part B than Part A of the experiment, TRT $\underline{F}(1, 43) = 38.06$, $\underline{p}<.001$, $\eta^2 = .469$, RT $\underline{F}(1, 43) = 10.69$, $\underline{p}<.01$, $\eta^2 = .199$, MT $\underline{F}(1, 43) = 32.61$, $\underline{p}<.001$, $\eta^2 = .431$.

However, Part also interacted reliably with TB, Wilks exact $\underline{F}(10, 34) = 3.81$, $\underline{p}<.01$, $\eta^2 = .529$, and AC, Wilks exact $\underline{F}(6, 84) = 2.64$, $\underline{p}<.05$, $\eta^2 = .159$, indicating that the pattern of improvement within Parts A and B were not equivalent either between trial blocks or among acquisition groups. Analysis of the Part x TB interaction revealed reliable effects for each of TRT, $F(5, 215) = 6.67$, $p<.001$, $\eta^2 = .134$, RT, $F(5, 215) = 4.55$, $p<.001$, $\eta^2 = .096$, and MT, $F(5, 215) = 6.56$, $p<.001$, $\eta^2 = .133$. RT was found to have improved to a greater extent between the first and second trial blocks in Part B than Part A, $F(1, 43) = 4.57$, $p<.05$, $\eta^2 = .096$. Conversely, each of TRT, RT, and MT was found to have improved to a greater extent between trial blocks 2 and 3 during Part A than Part B, TRT, $F(1, 43) = 15.70$, $p<.001$, $\eta^2 = .267$, RT, $F(1, 43) = 15.69$, $p<.001$, $\eta^2 = .267$, and MT, $F(1, 43) = 4.11$, $p<.05$, $\eta^2 = .087$. These results are illustrated in Figure 14.

Univariate analysis of the Part x AC interaction showed reliable effects for TRT, $F(3, 43) = 4.48$, $p<.01$, $\eta^2 = .238$, and RT, $F(3, 43) = 4.27$, $p<.01$, $\eta^2 = .230$. The improvement in RT, and subsequently TRT, of the random groups was found to be greater than that of the blocked groups between Parts A and B, TRT $F(1, 43) = 12.96$, $p<.001$, $\eta^2 = .232$, RT $F(1, 43) = 10.17$, $p<.01$, $\eta^2 = .191$. Neither of the random or blocked groups were found to differ reliably from each other; however, it is worthy to note that the RT performance of the BV group slowed from Part A to Part B, as can be seen in Figure 15, whereas the RT performance of the other three groups was faster in Part B than Part A.

Differences in the number of recall errors made between Part A and Part B were computed for each experimental group and are shown on Figure 16. A one way ANOVA showed a difference in the pattern of recall errors made during acquisition between Parts A and B by the groups, $F(3, 43) = 5.90$, $p<.01$, $\eta^2 = .292$. Post hoc tests showed that recall error frequency for the RV group improved to a greater extent than for the other three groups, who did not differ from each other.

Changes between Parts A and B: Retention

Difference scores were computed between Parts A and B on each of TRT, RT, and MT for the last trial block of acquisition, immediate retention and delayed retention, and entered into a 4 (AC) x 3 (TB) repeated measures MANOVA. A reliable multivariate TB main effect was found, Wilks exact $F(4, 39) = 5.61$, $p < .001$, $\eta^2 = .365$. Univariate tests revealed reliable effects for both TRT, $F(2, 84) = 3.82$, $p < .05$, $\eta^2 = .083$, and MT, $F(2, 84) = 4.92$, $p < .01$, $\eta^2 = .105$. Tests of adjacent trial blocks indicated that, overall, TRT, RT, and MT performance differences between Parts A and B were larger at immediate retention that at the last trial block of the acquisition period, TRT, $F(1, 42) = 10.55$, $p < .01$, $\eta^2 = .201$, RT, $F(1, 42) = 7.85$, $p < .01$, $\eta^2 = .158$, and MT, $F(1, 42) = 4.34$, $p < .05$, $\eta^2 = .094$.

Figure 17 shows the difference in performance experienced by each group from the end of acquisition to immediate retention during Parts A and B. When expressed as performance change from trial block 6 to IR between Parts A and B, it was expected that performance would improve to a greater extent for the BV group than for the other three acquisition groups. However, no reliable effect emerged when the change in performance of the BV group was compared against that of the other three groups, although the effect for TRT approached reliability, $F(1, 42) = 3.46$, $p = .07$, $\eta^2 = .076$. The performance change across the first retention interval did not differ reliably between the RC and RV group on any measure. However, the MT between the BC and BV groups was found to be reliably different, $F(1, 42) = 4.10$, $p < .05$, $\eta^2 = .089$ (see Figure 17). Although a very small effect, it indicated that the change in MT performance from the last block of acquisition to IR was reliably greater for the BV group than the BC group between Parts A and B.

Univariate analyses following a reliable multivariate TB x AC interaction, Wilks exact $F(12, 103) = 1.94$, $p<.05$, $\eta^2 = .164$, also revealed an interaction effect for MT, $F(6, 84) = 3.61$, $p<.01$, $\eta^2 = .206$. Performance differences were found between groups over the retention interval between immediate and delayed retention, $F(3, 42) = 4.01$, $p<.05$, $\eta^2 = .223$. The differences in MT performance between Parts A and B became larger from IR to DR for all but the RC group, whose performance differences became smaller.

Analysis of average retention performance between Parts A and B revealed a reliable multivariate main effect, Wilks exact $F(2, 41) = 10.30$, $p<.001$, $\eta^2 = .334$. Average retention performance was found to be reliably faster for MT during Part B than Part A of the experiment by 89.0 ms, $F(1, 42) = 20.21$, $p<.001$, $\eta^2 = .325$. No reliable differences emerged for TRT and RT performance between Parts A and B.

Greater improvement from the last trial block of acquisition to IR between Parts A and B should also produce larger retention performance differences between Parts A and B. Thus, it was expected that the largest gain in retention performance from Part A to Part B would be made by the BV group. Differences in retention performance are shown on Figure 18. Analysis of retention performance differences between Parts A and B showed a reliable multivariate effect between the BV group and the other three groups, Wilks exact $F(2, 41) = 3.64$, $p<.05$, $\eta^2 = .151$. Univariate analyses indicated that the difference in MT retention performance was reliably greater for the BV group than the other three groups combined, $F(1, 42) = 6.14$, $p<.05$, $\eta^2 = .128$. The difference in TRT was found to be marginally greater, $F(1, 42) = 3.69$, $p=.062$, $\eta^2 = .081$, while the difference in RT was not reliably different. When contrasted individually with each of the other groups, the BV group showed greater gains in retention MT performance than the BC group, $F(1, 42) = 5.10$, $p<.05$, $\eta^2 = .108$, and the RC group, $F(1, 42) = 6.40$, $p<.05$, $\eta^2 = .132$. Although the gains in retention MT performance for the BV group were also greater than for the RV group, the difference was not reliable, $F<2$.

A comparison of retention performance differences between verbalize and control conditions failed to produce a reliable multivariate effect.

Differences in retention recall error frequency (IR & DR) between Parts A and B were not reliably different between groups (see Figure 19).

Part B: Qualitative data analysis

Verbalizations of the cognitive activity of each verbalization group were transcribed, and the text was reduced into meaning units. Since the demands of Part B of the experiment were identical to Part A for the random verbalize (RV) group, with the exception of verbalizing cognitive activity, the text of this group was analyzed first. The data structure that emerged from the hierarchical arrangement of RV participant meaning unit clusters was used to guide the arrangement of meaning unit clusters from BV participants. This was done so that the cognitive processes typical of RV participants could serve as a basis from which to compare the cognitive processes of BV participants.

The text of RV participants was reduced into 242 meaning units from which 36 clusters emerged, while the text of BV participants was reduced into 256 meaning units from which 33 clusters emerged. Meaning unit clusters from each group were arranged separately into four tier hierarchies which are shown on Figures 20 and 21. Hierarchical arrangement of meaning unit clusters for each group revealed that during the acquisition phase of Part B, the cognitive processes verbalized by both groups shared the two major themes of memory formation and performance analysis, while an additional theme of environmental analysis was unique to RV participants.

Since the hierarchical arrangement of the meaning unit clusters for both groups is similar, results will be presented together. Where groups share category structures, representative meaning unit statements will be provided from each group, along with experimental group and subject identification codes.

Memory formation

The first major theme of memory formation dealt with the formation of knowledge structures related to the three spatial patterns that were to be learned. Sub-categories contained within this theme included association strategies, performance goals, pattern analysis, and retrieval analysis.

Association strategy.

This sub-category contained meaning unit clusters dealing with the strategies that participants used to commit the spatial patterns to memory. It contained seven meaning unit clusters shared by both groups, and three meaning unit clusters unique to BV participants.

The "association process" cluster contained statements indicating the process by which participants committed the patterns to memory. Statements ranged from rote memorization, "I try memorize them" (RV23) and "the repetition just puts it in" (BV17), to deliberate associations, "now I'm making a more conscious effort to associate color with pattern since last time [the previous week] I didn't know" (BV39).

Statements in the "whole sequence" cluster indicated that participants were remembering response key sequences as a whole rather than fractionating the response sequences into smaller chunks, "it's getting to the point where I see the whole picture in my mind, a connect the dots picture" (RV61), and "I visualize the pattern as a whole" (BV28).

Statements related to the association of a pattern color with its first response key location were included in the "A key" cluster. RV participants reported 14 statements indicating that pattern recall and performance speed could be improved by developing an association between the first response key and the stimulus word, "In my mind I'm just going over the first key, the rest flows from there" (RV27). By way of contrast, BV participants provided only three such statements, i.e. "I'm thinking about the A's for all the patterns and how they make a square with the start button" (BV17).

In the "multiple keys" cluster, participants indicated making additional associations between other response keys in a pattern sequence, such as, "I'm focusing on where B is now. If I can hit both A and B it takes half the keys out of the equation" (RV25), and "I'm focusing on hitting the A and D key because my arm has to move way down and then way back up" (BV20).

A number of association strategies were implemented by participants to make the spatial patterns personally meaningful. BV participants most frequently reported associating the spatial patterns with letters or numbers (18 statements) with statements like, "the shape, with the exception of key 6, looks like F for fuchsia" (BV37) and "I'm looking at it as a P. The buttons kind of form a P" (BV24). This strategy was mentioned by only one RV participant, "lime is an up side down Y!" (RV38). Statements in the "movement direction" cluster referred to remembering response key sequences according to movement direction. Examples include, "fuchsia is the big sweep" (RV81) and "I take it [white] through as start from bottom and go up [AB], and start from bottom and go up [CD]" (BV24). In addition to these two association strategies which both groups shared, BV participants also reported associating patterns with sayings or rhymes like, "lime at the bottom of your drink gets you high" (BV22) and "I had fuchsia hair once, so the association is 'pink stinks go fast'" (BV56), and with objects and other miscellaneous images like, "lime kind of looks like a mountain, Mt. Lima" (BV20) and "fuchsia reminds me of a hippo I had as a kid" (BV36).

Statements in the "multitask rehearsal" cluster indicated that participants would review more than one task during an intertrial interval. For RV participants, this was the most frequently reported association strategy (17 statements) and included comments like, "I'm scrolling through the patterns between beeps. Fuchsia is still a little fuzzy" (RV23) and "I'm just going over the patterns in my head" (RV57). A noteworthy finding was that of the 11 statements provided by BV participants, five of them came from one participant, including "I usually try to review everything before I go" and, during the lime practice block, "every now and then I'll review the white one too" (BV24).

BV participants also mentioned having no association strategy at all with comments like, "I can't really think of anything right now" (BV77) and "I haven't thought much about how to remember it" (BV17). The frequency of these reports (13 statements) coupled with the number of participants that provided them (7 of 12) seems to indicate that establishing meaningful associations with the spatial patterns was a somewhat effortful enterprise.

Between group comparisons of the number, type, and distribution of association strategy meaning units shows that BV participants provided more meaning units (93) along more categories (10) than RV participants (45 and 7, respectively). This indicates that BV participants were making efforts to associate colors with patterns, and to attach meaning to the sequences of response keys, and therefore were following the instructions given at the start of the experiment. Comparisons also reveal that the two most frequently reported association strategies utilized by RV participants involved associating a color with the location of the first response key of the pattern, and multitask rehearsal, which together accounted for 69% of association strategy meaning units. In contrast, these two clusters accounted for only 15% of association strategy meaning units for BV participants. Since both of these association strategies involve means of differentiating the patterns in memory, it suggests that formation of distinct memory structures for each spatial pattern was more important in meeting the demands of their practice condition for RV than BV participants.

Performance goals.

The second main sub-category within the memory formation theme was labeled "performance goals". Meaning unit clusters within this sub-category were related to the goals that participants indicated they were trying to accomplish. Only one cluster of goal statements emerged from BV participants while three clusters of goal statements emerged from RV participants.

The category that RV and BV participants shared involved the goals that participants set related to performance speed. This cluster was included within the memory formation theme because, in this thesis, performance speed is hypothesized to be dependent on the speed of recall and the strength of memory. Thus, goals related to performance speed must also be goals related to committing patterns to memory. Statements reflecting performance speed goals included, "I'm trying to get in the 13's" (RV57) and "I'm just trying to get off the start key faster" (BV37).

Two performance goal categories were unique to RV participants. The first involved RV participants' wish to wean themselves from using the pattern card illustrations as the basis for recalling response key sequences. Statements contained within the "use of pattern cards" cluster included, "I'm trying not to look at the cards" (RV61) and "I'm just trying to learn the patterns so I don't have to look at them. If I can remember them I can go faster" (RV31). RV participants realized that it took less time to execute a pattern if it could be planned on the basis of recall versus glancing up at the illustrations. The second category unique to RV participants was labeled "flow of movement". Statements like, "right off the bat I try get accustomed to the patterns and how my arm feels going through them" (RV26) indicated participants' wish to establish kinesthetic awareness of response produced feedback generated by moving through the pattern sequences.

Although RV and BV participants reported the same frequency of performance speed goals, the additional clusters of "use of pattern cards" and "flow of movement" provided by RV participants, which deal explicitly with memory formation goals, seems to indicate that the demands of random practice made memory formation a more explicit goal for RV than BV participants.

Pattern analysis.

The third main sub-category within the memory formation theme was labeled "pattern analysis" and contained statements about the various features of the patterns. This sub-category was further subdivided into statements in which the features of single tasks were evaluated, and statements in which the features of multiple tasks were compared.

The "single task evaluation" sub-category included four meaning unit clusters containing fourteen statements for BV participants and two meaning unit clusters containing five statements for RV participants. Only one cluster, analysis of the difficulty of performing the patterns, was shared. Statements within this cluster included, "white is easy to remember" (RV32) and "lime may be more difficult because it's counter clockwise" (BV36). The cluster unique to RV participants contained two statements dealing with pattern shape including, "why can't I remember that? Doesn't make a shape" (RV57) and "fuchsia seems like a random pattern" (RV23). Unique to BV participants were three clusters of meaning unit statements related to key location, "seems like the keys are far [away from each other]" (BV56), movement direction, "[fuchsia] feels awkward, goes in a bunch of different directions" (BV77), and preference, "don't like it [fuchsia] at all" (BV22).

Participants also compared features between patterns along a number of dimensions. Three meaning unit clusters were extracted from the text of BV participants which were all shared with RV participants. RV participants reported multitask comparisons along three additional dimensions. The first shared cluster contained comparisons concerning the location of keys among patterns, such as, "lime and fuchsia keys seem to be in tighter to each other, white is more spread out. It's easier to get better scores on lime and fuchsia" (RV25) and "I don't like having to go all the way down and then all the way back up [for white] relative to lime which is nice and easy" (BV56). Both groups also contrasted the level of difficulty between patterns with statements like, "white is the easiest" (RV57 & 81) and "the patterns seem a little more difficult this week" (BV37). The last shared cluster involved preference comparisons including, "I like it [white] better than the last one" (BV22) and "my favorite is lime, least favorite is fuchsia" (RV23). Between pattern comparisons unique to RV participants included comparisons of pattern shapes, "Lime is an A, fuchsia is an arrowhead, that's how I differentiate those two. White is the square one" (RV32), movement direction, "I just realized you have to change direction at every key!" (RV61), and comparing a current pattern with a pattern learned the previous week, "white is like red of last week" (RV23 & 81).

Analysis of the number and type of pattern analysis meaning units shows that RV participants made more than double the number of multitask comparison statements (37) than BV participants (14) while making far fewer single task evaluation statements (5 vs 14 respectively). While multitask comparisons represented 50% of pattern analysis meaning units for BV participants, they accounted for 88% of the meaning units for RV participants. Furthermore, the multitask comparisons of RV participants were made along twice as many dimensions. This indicates that substantially different cognitive processes were at work between groups. The data suggests that RV participants directed substantially more cognitive resources towards making contrasts between tasks than BV participants.

Retrieval analysis.

The last main sub-category within the memory formation theme was labeled "retrieval analysis" and included clusters dealing with participants' evaluation of their ability to recall the sequences of response patterns. For RV participants, this sub-category contained only one cluster labeled "state of learning" which was shared with BV participants. Statements in this cluster involved participants' assessments of how well or poorly task solutions could be recalled, such as, "I definitely have white down, just the other two" (RV81), "white is cemented in, lime 75%, fuchsia 25%. I don't know why it's so hard" (RV31), and "what the heck! I'm busy trying to think of them all and it's screwing me up" (BV74). BV participants also made statements indicating the confidence they felt in their ability to remember the associations and response sequences later on. Statements in this "confidence in memory" cluster included, "hopefully I'll be able to repeat this at the test" (BV77) and "I just hope I don't get confused" (BV24). A noteworthy finding is that of the nine statements contained in this cluster, all but one indicated a lack of confidence, "I'll remember. I'm good at associating things" (BV20).

Comparing the groups, RV participants made nearly six times the number of reflective statements about their state of learning than BV participants, who seem more concerned with their future rather than present state of memory. These results seem to provide yet another example of participants adapting to the demands of their practice schedule. Performance during acquisition was easy for BV participants but because they had just finished Part A of the experiment, they were familiar with the difficulties that they experienced recalling the response patterns during the retention phases. However, BV participants had no way to test the strength of their memory during the acquisition phase because their practice trials were blocked. This left them to hope that what they had learned was adequate. Conversely, the conditions of random practice provided participants with multiple opportunities to test recall, which seems to have grounded them in the present.

<u>**Performance analysis**</u>

The second major theme extracted from the data was performance analysis. All sub-categories within this theme were shared by both groups, and dealt with participants' analyses of a variety of performance aspects including (1) perceptual strategy, (2) blocked trial performance, (3) the performance process, (4) performance state, (5) performance problems, (6) execution analysis, and (7) performance evaluation.

Statements within the "perceptual strategy" cluster centred around the strategies participants used to identify the stimulus color word that appeared on the monitor. Since the timer began with the presentation of the color word, participants discovered that performance times could be reduced by decreasing the time it took to identify the word. Statements in this cluster indicated that some participants experimented with word reading versus word recognition, and included, "I try not to read the word, just see it. The words look different from each other" (RV26), and "I should try to recognize the word instead of reading the word, it's faster with recognition" (BV28).

On blocked trials, the upcoming pattern was known in advance. This was constant for BV participants and occurred on re-trials for RV participants. A number of participants reported that, under these circumstances, they felt like they were cheating. Examples included, "I don't like redoing a trial. It feels like cheating" (RV61) and "it's like cheating on a test" (BV28).

The "performance process" sub-category contained two clusters for RV participants but only one for BV participants. Groups shared the cluster labeled "optimal performance". This cluster included statements in which participants reflected on how cognitive processes affected performance, such as, "thinking takes too long. Gotta do, not think" (RV25) and "I think too much" (BV37). Statements contained in the "benefits of experience" cluster were unique to RV participants who made mention of the fact that performing in Part B of the experiment seemed easier because they knew what to expect from having completed Part A. For instance, "I'm off the cards faster this week because I know the process from last week" (RV16).

The "performance state" cluster included participants' comments about their emotional state and how it might affect performance. Examples include, "I find that when I get excited I blow it, so a little relaxation stage [helps]" (RV27) and "I'm tired and this is as fast as I'm likely to get" (BV17).

Six clusters of performance problems emerged from the text of RV participants, four of which were shared with BV participants. Participants in both groups reported that performance problems arose, or were likely to arise as a function of: (1) not being ready, "my mind went blank" (RV26 & 57), "the beep caught me off guard" (BV33), (2) association problems, "it's pink, not fuchsia, so I see fuchsia but think pink, it takes a bit of time" (RV81), "I'm also not looking at the A key because I might get confused and think L, lime" (BV20), (3) excess speed, "I'm pushing myself a bit so I'm losing accuracy, my hand is getting ahead of my brain" (RV32), "I'm trying to slow down, but just can't" (BV74), and (4) recall errors, "I want to hit C as B. It seems to be a more natural movement" (RV81), "I still want to make D C, then both would be going up and it would be easier to remember" (BV17). RV participants reported additional performance problems due to (1) anticipation, "If I think just before it's white, and it's white I fly. But if I think it too early it's like don't get excited, it could be fuchsia or lime and I'm a little slower" (RV27), and (2) key contact, "I'm rolling off [key] 14, never 13, just 14" (RV16).

The "execution" sub-category was further subdivided into three sub-categories that were common to both groups that dealt with an analysis of various aspects of executing the task.

For BV participants, the "visual attention" sub-category contained two clusters of statements involving analyses of where participants focused their point of gaze. Unique to BV participants were 22 statements indicating the location of their visual attention when the stimulus word was presented on the monitor. Statements included, "I'm waiting until I see the word 'white' instead of just going off the beep so come the test..." (BV77) and "I didn't look at the screen that time. I don't like looking at the screen" (BV56). Groups shared the cluster in which participants indicated the location of their visual attention during movement execution. Statements in this cluster included, "my eyes are focused somewhere in the centre of the board, I'm not following my hand as they hit the keys" (RV16) and "I see 'white' and my hand moves to A while my eyes look at B, then it's just my hand following my eyes" (BV20).

Statements dealing with assessments of how response keys were hit were included in the cluster labeled "key contact strategy". Participants from both groups reported experimenting with a variety of different key contact strategies like, "I'm trying to get my fingers to hit the keys so my hand doesn't have to move so far" (RV81) and "I tried different hand strategies; fingertips, palms, knuckles, pound. I liked pound and stuck with it" (BV33).

The "movement analysis" sub-category was made up of clusters dealing with overall movement analysis and analysis of movement for individual segments of response sequences. Groups shared the cluster related to overall movement analysis, which contained statements like, "I don't understand how I can make my hand go faster" (RV61) and "that was faster, I could tell by the feel of my arm" (BV20). Unique to BV participants were assessments of movement between individual segments of response sequences such as, "C-D is getting better now" (BV17).

The last cluster within the performance analysis theme was labeled "performance evaluation". This cluster was shared by both groups and included statements in which participants' evaluated their performance, such as, "that's a bad score" (RV57) and "time differences are probably more a function of getting off the start key" (BV37).

Comparisons of the number, type and distribution of performance analysis meaning units again reveals some noticeable differences in the cognitive processes at work for RV and BV participants during the acquisition phase. RV participants made four times the number of statements than BV participants concerning analyses of performance problems. Also, a greater percentage of meaning units within the performance analysis theme were related to performance problems for RV (44%) than BV (8%) participants. Furthermore, for RV participants, the largest cluster within the performance problems sub-category was related to anticipation problems. These results seem to confirm the idea that the conditions of random practice create more problems for participants than the conditions of blocked practice.

Execution analyses also showed differences between groups. BV participants made substantially more execution analysis statements (75) than RV participants (25). While only four statements were given by RV participants regarding the location of visual attention, BV participants provided 35. First, the number of execution analysis meaning units provided by BV participants as well as the number of meaning units specifically related to visual attention again confirms that BV participants were following the instructions provided at the beginning of the experiment regarding preparing for the demands of retention. Second, the results again suggest that the conditions of practice placed substantially different demands on participants, who in turn, adapted quite differently. Under the conditions of blocked practice, BV participants had a choice of where they could place visual attention when the stimulus was presented. They could choose whether or not to watch the screen when the stimulus word was presented since knowledge of the upcoming pattern was not in doubt. However, since this strategy produced slower performance times, it likely took more of an effort for BV participants to maintain, as indicated by the number of statements they provided. Due to the conditions of random practice, RV participants were forced to watch the monitor for the stimulus word, which may explain why they provided no statements related to visual attention at the stimulus.

Environmental analysis

The last major theme was labeled "environmental analysis" and was unique to RV participants. This theme contained only one cluster in which all 18 statements revolved around the theme of participants trying to discover regularities in the performance environment so that they could try to anticipate the upcoming pattern. Examples of statements in this cluster included, "it's definitely not white this time, I already had two in a row" (RV81), "I try to decrease hesitation by playing the odds on what was last" (RV23), and "the sequence starts over again so I can't guess now, but whatever comes up now will determine the next one" (RV38).

Environmental analysis was one type of analysis not available to BV participants due to the certainty of presentation for upcoming patterns. However, random practice conditions generated a considerable amount of uncertainty, which RV participants attempted to reduce by playing the odds and trying to assess whether or not task presentation followed a recurring pattern.

Summary

Although the hierarchical organization of the data for each group was similar, the frequency and distribution of statements within categories indicated differences in what each group emphasized throughout the acquisition phase. RV participants appear to have emphasized developing memory structures for task information and refining performance through anticipation. RV participants reported that they would rather depend on internal than external cues to perform the task, indicating that memory development and independent recall were explicit goals. To accomplish this, it appears that they created distinct memory structures for the patterns, which they formed by contrasting the features of response sequences and linking pattern colors with the location of the first response key. This was accompanied by evaluation of their progress, which they did more frequently than BV participants.

In contrast, BV participants seem to have emphasized making sense of response sequence patterns, making refinements to performance, and, to a lesser degree, simulating test conditions during practice. They seem to have spent considerably more time making S-R associations and trying to make response sequence patterns meaningful. Their method relied less on differentiation and more on generating meaningful associations. Compared to RV participants, BV participants appeared less concerned with evaluating their state of memory development and more concerned with whether they would be able to recall task solutions during the retention phases. They also made efforts to practice the performance processes that would be demanded on retention trials, such as looking at the monitor until a color word had been presented. Committing task information to memory appeared more difficult under random than blocked conditions as indicated by the increased frequency of reports given by RV participants involving assessments of problem areas.

Discussion

The hypothesis that purposefully directed task evaluation in Part B of the experiment would permit the relative retention and retention performance of BV participants to rival that of random participants was not supported. Based on theoretical considerations and the results of Wright et al. (1992), it was expected that the level of performance achieved by the end of the acquisition period would be maintained for the RC, RV, and BV groups while relative retention and retention performance for the BC group would worsen considerably. This would indicate that the access to and implementation of movement plans (Wright, 1991) by BV participants was strengthened by effortful evaluation of task related information and purposeful practice during the acquisition period. While access to and implementation of movement plans deteriorated to a greater extent for the BC group than the other three experimental groups combined from the end of acquisition to immediate retention in the Part B of the experiment, the results also showed that BV participants incurred greater losses in plan access and implementation than RV and RC participants. However, the losses in relative retention and retention performance shown by BV participants were less than those experienced by BC participants, although the improvements appear to be limited to MT. When compared with the retention results of Part A of the experiment, where the losses of BV and BC participants were similar, the results of Part B suggest that purposeful evaluation did ease movement plan implementation at retention somewhat for BV participants. This conclusion is further supported by observing the changes in retention performance from Part A to Part B. If relative retention demonstrated by all groups was similar in Parts A and B, the differences in retention performance would be similar as well. However, demonstration of greater relative retention by BV participants, as was hypothesized, would produce greater differences in retention performance from Part A to Part B. That MT retention performance for BV participants showed the greatest improvement from Part A to Part B supports the conclusion that purposeful task evaluation had a positive effect on movement plan implementation.

Changes in reaction time and recall error frequency from Part A to Part B for the BC and BV groups were not reliably different which suggests that access to movement plan information benefited as much from experience as from purposeful practice. However, when considered along side the MT results, the data suggest that movement implementation may have involved more closed loop control for the BC than the BV group. This would seem to indicate that the information regarding the movement patterns to be executed was somewhat more complete for the BV group than the BC group.

Qualitative data analysis indicated that BV participants expended substantial effort associating colors with patterns, and making the sequences of response keys meaningful. The number of statements indicating that no associations had yet been formed suggests that generating these associations was no trivial process for BV participants. Nonetheless, improved relative retention performance for BV compared to BC participants in Part B suggests that the process did somewhat counteract the effects of a low CI practice condition. However, the relative retention for the BV group was not equivalent to the relative retention demonstrated by the random groups. This seems to indicate that although effortful practice and single task evaluation can produce some gains in movement production, these cognitive processes do not rival those produced in response to random practice. The information regarding the movement patterns to be executed appears to develop more completely under conditions of high CI. Two explanations may account for this finding. First, single task evaluation may be inferior to multitask comparison; it may be that multitask comparison simply produces more elaborate retrieval structures than single task evaluation. Second, it may be that the amount of effort needed to sustain evaluative processes during blocked practice isn't worth it. Perhaps if blocked practice participants are not sufficiently motivated to improve task performance, the discipline needed to sustain levels of effort that exceed those inherently demanded by the practice situation is too great.

Closer inspection of the data of Wright et al. (1992) reveals that the relative retention results found in Part B in this experiment may not be that different from the results reported in their paper. Wright et al. did not include the last trial block of the acquisition phase in their analysis, choosing instead to assess differences in retention performance. Observing the data they present on Figure 4 of their paper (p. 35), the relative retention of the random group compared to the two blocked groups supplemented with intra and intertask processing shows that the random group did not experience performance decrements from the end of acquisition to immediate retention, whereas both blocked groups did. The changes in performance for the blocked groups appear even though there are no reliable differences in performance between the groups at either retention or the end of acquisition. Furthermore, the performance decrements of the two blocked groups supplemented with intra and intertask processing appear to be less than those experienced by the blocked group that was not supplemented with extra processing. This seems to support the idea that supplementing blocked practice with extra processing can improve retention, but not to the extent of random practice.

Although many experiments investigating the CI effect have been conducted, most have made inferences regarding the cognitive processes underlying performance. Very few (Shea & Zimny, 1988) have made use of verbal protocols to assess the cognitive processes entertained by participants. The verbal protocol implemented in the present experiment advances previous work done by Shea and Zimny in terms of how the data were analyzed. Shea and Zimny grouped the statements of their participants according to fixed categories, including "1. intra-item comments (reports detailing a singular task); 2. inter-item comments (reports involving comparisons across tasks); 3. extra-experimental reports (reports that mapped the task onto pre-existing knowledge); and 4. general comments and performance evaluations" (p. 306). In the present experiment, statements were grouped according to themes that emerged from the data. Although this makes direct comparisons somewhat problematic, Shea and Zimny's finding that random practice participants made double the number of inter-item reports than blocked practice participants was supported. Even though BV participants were encouraged to evaluate the features of the tasks, they reported less than half the number of multitask comparisons than RV participants. Furthermore, Shea and Zimny note that when the response sequences of the patterns are not similar, "subjects showed a startling ability to map onto knowledge outside the domain as a strategy to aid in performance and recall" (p. 307). This is supported in the present experiment by the number of dimensions along which participants in both groups mapped response sequences. It also confirms Chase and Ericsson's (1981) notion that existing knowledge structures provide the basis for representing novel information in a meaningful way.

Acquisition performance in Part B showed little difference between blocked groups, both of which performed better than the random groups. This finding supports the results reported by Wright (1991) and Wright et al. (1992) who also found that supplementing blocked practice with intra or intertask processing did not affect acquisition performance. From this evidence it appears that any interference experienced by blocked practice participants during acquisition as a function of purposeful single task evaluation could be easily overcome by the repetitive nature of a blocked practice schedule. Wright et al. had suggested that because random and supplemented blocked groups engage in similar processing activity during the acquisition phase, similar performance would be predicted during both acquisition and retention. Although in this experiment additional processing requirements were being made of BV participants, at no point during a block of practice was it necessary for participants to fully replace an existing action plan. Thus, it appears that the cognitive model responsible for guiding action (Shea & Zimny, 1988) likely remained in working memory throughout the block of practice, producing performance in BV participants similar to that of BC participants.

Figure 15 showed that the random groups made greater MT performance gains during the acquisition phase of Part B than the blocked groups. There may be several possible explanations for this. First, having experienced the learning process in Part A, random participants may have been better prepared to handle demands of random practice in Part B. Second, the margin for improvement may have been greater for the random than blocked groups. Since movement speed is the dependent measure in PBC, there will be a point at which performers cannot go faster. Perhaps the conditions of blocked practice allowed the performance of participants to come closer to this limit during the first part of the experiment, thereby reducing the opportunity to make large improvements during the second part of the experiment.

The verbalization requirement made of RV participants appears to have made a difference in their ability to recall movement patterns since they showed a reduction in error frequency during the acquisition phase of Part B while the other groups showed slight increases. Perhaps the demand for a greater degree of introspection created by the verbalization condition prompted a more conscientious use of recall strategies, leading to improved recall.

Finally, the reliable task main effect found in Part A showed that MT was longer for the blue pattern than for the yellow or red patterns. This suggests that participants seemed to have a more difficult time implementing the movement plan for this pattern and confirmed that, indeed, the most difficult pattern from Experiment 1 was selected to be included in Experiment 2.

GENERAL DISCUSSION

The specific purpose of this research project was to test whether encoding operations in general, and evaluative processes in particular, are primarily responsible for the CI effect. The more general purpose of this research project was to use the CI phenomenon as a vehicle to explore whether evaluative processing might be considered as a common cognitive process capable of unifying a number of phenomena related to the acquisition of motor skills.

Related specifically to the CI effect, Experiment 1 was replicated twice in Experiment 3; it was replicated with pairs of blocked and random groups in 3A, and it was replicated by the BC and RC groups in 3B. Each time similar results emerged. Effect sizes of practice condition main effects during the acquisition phases were similar in magnitude between experiments (see Tables 1, 6, & 12).

Effect sizes of the AC x TB interaction were larger in Experiment 3 than Experiment 1 (see Tables 3, 8, & 13). In fact, the largest effect sizes were seen between the BC and RC groups in Experiment 3B (see Table 13), indicating that the CI effect became larger rather than smaller as a function of experience. Apparently, the conditions of blocked practice severely restrict participants ability to form easily accessible knowledge structures, even though the participants had already experienced recall difficulties under identical conditions once before. The emergence of a CI effect by the control groups in Experiment 3B suggests that experience alone is insufficient to compensate for the cognitive demands created by the conditions of practice. Furthermore, in all three instances, MT for the blocked groups was greater at retention than on the first trial block, indicating that the cognitive processes underlying performance for these groups changed substantially from the acquisition to retention phases. These findings suggest that the push button board was a reliable experimental task in producing the CI effect. Whereas Wright et al. (1992) produced the CI effect with similar response sequence requirements on the number pad of a computer keypad, the present research project extends these findings from small finger movements to large arm movements.

In Experiment 3A, unlike the results of Experiment 1, the retention performance of the random groups was superior to that of both blocked groups, at least for TRT and RT. This is more in line with existing CI literature in which spatial tasks are used (e.g. Lee & Magill, 1983; Shea & Morgan, 1979). Again, this can be explained by observing the level of performance achieved by the groups at the end of the acquisition phase. The magnitude of effect sizes from the last block of the acquisition phase to the immediate retention block were similar between Experiments1 and 3A. This indicates that differences in retention performance between groups were coupled with smaller performance differences between groups on the last trial block of the acquisition phase in Experiment 3A than in Experiment 1. This finding suggests that including the last trial block of the acquisition phase in analyses of retention data provides valuable information concerning relative retention, which is equally important evidence of the CI effect as performance differences at retention.

The present series of experiments extends and clarifies the theoretical position that cognitive effort underlies the CI effect. Lee, Swanson, et al. (1991), Lee, Swinnen, et al. (1994), and Schmidt and Bjork (1992) have argued that cognitive effort underlies the CI effect, however the specific cognitive operations responsible for producing this cognitive effort have, to date, not been clearly specified. When considered together, the results of the three experiments in this research project suggest that encoding variability rather than retrieval practice is primarily responsible for producing the cognitive effort that underlies the CI effect. The primary importance of encoding variability is evidenced most clearly by comparing the results of Experiments 1 and 3A, with those of Experiment 2. In Experiment 2, the pattern of TRT, RT, MT, and recall errors was similar between blocked and random groups, however in Experiments 1 and 3A it was different. The data indicate that the CI effect emerged when participants practiced movement pattern variations on the Push Button board, but did not emerge when participants practiced three different tasks.

According to the TPED, the CI effect will emerge when the context of practice produces variability in how task relevant information is stored. It was reasoned that minimizing differences in encoding variability between blocked and random practice groups should eliminate the CI effect. Task similarity has been suggested as a necessary element that stimulates differential encoding variability because it stimulates the processing of information about intertask differences for random but not blocked practice participants. Shea & Zimny (1983) suggest that this information yields a more elaborate and distinct memory representation of each task. Therefore it follows that if the intertask comparisons of random participants cannot generate more information along more dimensions than that generated by blocked participants, then differential encoding variability between random and blocked practice groups would be minimized. Since encoding variability is thought to benefit from intertask comparisons caused by task similarity, it was reasoned that intertask comparisons of dissimilar tasks would minimize differential encoding variability between blocked and random groups and, subsequently, produce equivalent retention performance. The results obtained in Experiment 2 support this prediction.

An attempt was also made to minimize differential encoding variability between blocked and random groups in Experiment 3B, this time by facilitating encoding variability under blocked practice conditions. It was argued that requiring blocked practice participants to purposefully evaluate the features of three PBC pattern variations, evaluate the quality of execution of these patterns, and associate the information obtained by evaluation with existing knowledge structures would increase the variability with which these patterns could be encoded during the acquisition phase of the experiment. It was predicted that minimizing differential encoding variability between BV and random groups in this manner would benefit relative retention for these groups to a similar degree. By magnifying differential encoding variability between BV and BC groups, it was predicted that relative retention of the BV group would be substantially better. The MT data indicate that purposeful evaluation allowed the BV group to maintain their acquisition performance somewhat better than the BC group on immediate and delayed retention tests. However, the results were less conclusive than hypothesized since neither the relative retention nor the frequency of recall errors of the BV and random groups were similar between the end of the acquisition phase and immediate retention. The failure of the hypothesis to fully predict the obtained results will be discussed later.

Although the TPED is an encoding based theory and the TAPR is a retrieval based theory, both have been accepted as explanations for the CI effect. Why have the TPED and the TAPR co-existed for so many years? Analysis of the cognitive demands of the tasks in many CI studies (e.g. Lee & Magill, 1983; Shea & Morgan, 1979) indicates that differential encoding variability and differential retrieval practice co-occur due to the use of task variations. The problem with the co-occurrence of differential encoding variability and retrieval is that there is no way of knowing which of the TPED or the TAPR is the better account for the results. The nature of the tasks interacts with the scheduling of practice to simultaneously require high demands for both encoding variability and retrieval practice under random practice conditions, and low demands for both encoding variability and retrieval practice under blocked practice conditions. Either the TAPR or the TPED are equally capable of explaining the results of experiments in which blocked and random practice conditions produce differential demands for encoding variability and retrieval practice at the same time.

The TAPR has been unable to account for the results of experiments in which the demands for encoding variability and retrieval practice have been dissociated. For instance, the TAPR is capable of explaining the results from Experiment 1 and 3A where differential encoding variability and differential retrieval practice co-occur. In Experiment 2, differential retrieval practice remains intact, but differential encoding variability is minimized through the use of dissimilar tasks. As a retrieval based theory, the TAPR predicts that the pattern of acquisition and retention results in Experiment 2 should have been similar to those produced in Experiments 1 and 3A since the demands for retrieval remained unchanged between the groups in these experiments. However, in Experiment 2, the acquisition, immediate and delayed retention performance between blocked and random groups was similar. The TAPR is incapable of offering an explanation for these results. Neither can the TAPR offer an explanation for the improved relative retention demonstrated in Experiment 3B by the BV compared to the BC group. From a working memory reconstruction perspective, the amount of retrieval practice received by the BC and BV groups is equivalent, since both groups practiced under blocked conditions and could implement the solution from the previous trial. This should have produced similar relative retention performance between the groups, however this was not the case. The lack of empirical support for the predictions of the TAPR in Experiments 2 and 3B indicates that retrieval practice was not primarily responsible for the data that were obtained.

Additional evidence supports the view that blocked and random practice groups perform similarly at retention when differential encoding variability has been reduced even though differential retrieval practice remains intact. The CI effect has not emerged in experiments where the tasks to be learned are dissimilar rather than task variations, even though continuous action plan retrieval is demanded of random practice participants. For instance, the CI effect did not emerge when using the three volleyball skills of forearm pass, overhead pass and spike (Bortoli, et al., 1992; French, et al., 1990), the tennis skills of forehand and backhand ground stroke (Hebert, et al., 1996), or computer games of different winter Olympic events (Shewokis, 1997). Furthermore, when differential encoding variability has been reduced between blocked and random groups because the spatial patterns to be learned were nearly identical (Wood & Ging, 1991), the groups performed similarly on retention trials. If retrieval repetition was the locus of the CI effect, as suggested in the TAPR, a CI effect should have emerged in each of these cases. Clearly, the TAPR is incapable of providing an explanation for these results.

While purposeful evaluation in the context of blocked practice produced noticeable benefits in relative retention for BV participants in Experiment 3B, these benefits were not nearly as great as those demonstrated by random participants. Why might this be the case? One explanation may be that even though BV participants were engaged in evaluating pattern features and were attempting to form associations with existing knowledge that would help them recall the spatial layout of each PBC pattern, the information derived from these evaluations may not have been equivalent to the information obtained under the conditions of random practice. As the qualitative data revealed, there was a greater tendency for BV participants to evaluate the features of one spatial pattern at a time whereas the tendency for RV participants was to compare the features of two or more tasks. That relative retention benefited to a greater extent for RV than BV participants indicates that multitask comparison appears to be more effective than single task evaluation in promoting multiple and variable encoding strategies. In other words, information about the features of a task appear to be encoded more meaningfully when these features can be used to distinguish one task from another rather than being used to elaborate on the features of a single task. Theoretical support for this argument is provided by Glenberg's (1979) component levels theory.

To account for the effects of spacing in the verbal learning domain, component levels theory specifies conditions under which certain classes of information will be variably encoded. It is assumed that a multi-component episodic trace is created following the presentation of a stimulus which represents the stimulus in memory. Repeated presentation of the stimulus is effective to the extent that the second presentation allows information to be stored that is distinct from the information stored at the first presentation. The multi-component episodic trace includes three classes of components which represent information about the stimulus and the context in which it is presented; these are contextual components, structural components, and descriptive components.

Contextual components represent the context in which an item is presented and include such information as the characteristics of a particular individual, in a particular environment, at a particular time. Repetition becomes more effective when different information is encoded at repeated presentations of an event, which is produced by a change in some aspect of either the learner's cognitive or affective state, or the environment in which tasks are performed. Glenberg (1979) suggests that differential storage of contextual components for a repeated item is directly related to the lag between presentations of the item.

The structure that a learner imposes on individual events is encoded as structural components. Associations formed between repeated items provide this structure. Structural components are not stored automatically, instead they are dependent on processes controlled by the learner, which means that he or she must engage in "some sort of structural analysis (e.g., relating TBR stimuli within a story, interactive imagery, etc.)" (Glenberg, 1979, p. 97) to generate relations between items. Since it is assumed that associations between items can only be made while they are being processed concurrently, the specific structural components encoded in the trace will be determined by local context and can be induced by contextual change.

Descriptive components are the most unique components encoded in an episodic trace. These involve information specific to the stimulus, such as the shape, color, movement directions, flow, assigned meaning, etc.

Glenberg (1979) suggests that the trace produced from an item repeated after a time lag will include more contextual, structural, and descriptive components than the trace of an item repeated after a shorter lag, or no lag at all. Furthermore, he states that "differential storage of the components on one or more levels...will, in general, facilitate remembering" (p. 99). When applied to the CI effect in the motor domain, component levels theory suggests that the multi-component episodic traces of random participants will include more of each type of component than those of blocked participants, due to differences in contextual change between the groups.

Although component levels theory was formulated to explain the spacing effect in the verbal learning domain, it appears to account for the data produced by this research project remarkably well. In Experiment 3B, BV participants were instructed to evaluate the spatial patterns being learned in order to discover features that they could associate to something meaningful. It was hypothesized that inducing evaluative processing for BV participants would produce a related increase in cognitive effort, thereby allowing BV participants to perform similarly to random participants at retention. However, since BV participants practiced under the conditions of blocked practice, it could be argued that the contextual components that they encoded may not have been substantially different than those encoded by BC participants. Since there was little by way of contextual demands to induce the concurrent processing of more than one task at a time, it is also unlikely that there would be any substantial differences in the encoding of structural components between BV and BC participants. However, the experimental instructions given to BV participants may have induced differential encoding of descriptive components relative to the BC group. Finding that the frequency of statements related to single task comparison was greater for the BV than the RV group, combined with the finding that, at retention, the BV group moved through the spatial patterns somewhat faster than the BC group and made fewer recall errors, suggests that differential encoding of descriptive components appears to facilitate recall to some extent. However, the relative retention of the BV group was still substantially less than that of either random group, suggesting that the recall and implementation of movement is facilitated to a greater extent by the differential encoding of structural components.

The pattern of recall errors and relative retention between blocked and random groups in Experiments 1 and 3A support the predictions of component levels theory as well. In these two experiments, the conditions of blocked practice did not change substantially across repeated presentations of items during the acquisition period. Thus, it would be reasonable to assume that the information encoded by blocked participants at repeated presentations of a given task would not have been substantially different. Conversely, the conditions of random practice required participants to respond to a different stimulus word on virtually every trial, effectively altering at least one component of the learner's environment each trial. Since the learning environment changed each trial, different information would likely have been encoded at different presentations of a task. Furthermore, since multiple items could be concurrently processed by random but not blocked participants, intertask distinctions could be made by random but not blocked participants, thereby producing differential encoding of structural components. Differential encoding of descriptive components as a function of the conditions of practice is also likely, since the conditions of random practice likely required participants to attend more closely to the descriptive components of each task in order to execute it reasonably well on repeated presentations.

The results obtained in Experiment 2 can also be explained by component levels theory. In this experiment, the contextual components related to the perceptual and response characteristics of the experimental tasks changed substantially from task to task. As a result, the episodic traces that developed in response to repeated presentations of each task likely shared few components regardless of practice condition. However, the rate of change in contextual components was still greater for random participants, since a different task was presented on almost every trial. These changes may have differentially affected the contextual components related to participants cognitive and affective state. Trends in the acquisition data indicate superior acquisition performance for random than blocked participants, trends that may have emerged as a function of the different cognitive and affective states produced by the conditions of practice. It is conceivable that blocked participants may have experienced a loss in motivation as a function of blocked repetition whereas the affect of random participants may have been more resistant to motivational loss due to the increased rate of contextual change. Furthermore, since the contextual changes between tasks were considerable for both blocked and random participants, there simply may not have been sufficient grounds for random participants to form relations among the tasks as they were being concurrently processed. This being the case, the differential encoding of structural components between random and blocked participants was likely minimized. The lack of substantive differences between blocked and random participants regarding the encoding of contextual and structural components would easily account for the results obtained in Experiment 2.

Implications for CI Theory

A shift from emphasizing differential retrieval practice to emphasizing differential encoding variability in CI research affects the utility of the theoretical accounts of the CI effect. Analysis of the encoding demands in CI experiments reveals that differential encoding variability co-occurs with differential retrieval practice in those experiments where the TAPR is capable of explaining the results and that differential encoding variability is minimized in those experiments where the TAPR is incapable of explaining the results. However, being an encoding based theory, the TPED predicts that the CI effect should emerge only when practice conditions produce differential encoding variability between groups. Where the demand for the cognitive processes that lead to encoding variability, such as evaluation, comparison, elaboration and distinction, are minimized under blocked and random practice conditions, similar levels of retention performance are predicted. Therefore, it appears that the TPED accounts for the data obtained in experiments where an equally acceptable explanation is offered by the TAPR, and is also capable of accounting for the data obtained in experiments where the TAPR is incapable of offering an explanation. In light of this evidence, it would appear that the range of experimental data that can be accounted for by the TPED includes and extends beyond that of the TAPR, putting into question the utility of retaining it as a viable theoretical alternative for the CI effect.

Shea and Morgan (1979) as well as Shea and Zimny (1983) have suggested that task similarity is a necessary condition for the emergence of the CI effect. Some researchers (e.g. Wood & Ging, 1991) have taken this to mean similarity of features involved in movement coordination, such that tasks are considered similar when coordination features are similar between tasks, and tasks are considered different when coordination features are different between tasks. However the present series of experiments suggest that the definition of similarity should be broadened. Similarity may be considered along a number of additional dimensions including the perceptual features between tasks, as well as the environment in which the tasks are performed (Glenberg, 1979). Thus, multiple tasks containing similar perceptual features that involve minimal changes in local context between them would be predicted to interfere with each other and should produce the CI effect in response to blocked and random practice conditions due to differences in the variability with which task features are encoded. Conversely, multiple tasks containing different perceptual features that involve changes in local context among them would not be predicted to interfere with each other because the distinctiveness of the conditions in which each task is performed should minimize differential encoding variability under blocked and random practice conditions. By this definition, the majority of CI research has been conducted using similar tasks. To better understand the theoretical basis of the CI effect, more research utilizing dissimilar tasks, as previously defined, is recommended.

An issue associated with task similarity is the role of the generalized motor program (GMP) in CI research. In their review of CI literature, Magill and Hall (1990) assessed the commonalties and differences among studies in which the CI effect did and did not emerge. They suggested that the CI effect was most likely to emerge in studies in which each of the tasks required a different pattern of movement coordination (GMP), and was less likely to emerge when the tasks required parameter adjustments for the same pattern of movement coordination. In the present series of experiments, analysis of the tasks in Experiments 1 and 2 indicates that each task in each experiment required a different pattern of movement coordination, yet the CI effect emerged in Experiment 1 but not in Experiment 2. From these results it appears that the CI effect depends on more than movement coordination differences. Focusing research attention on issues related to movement coordination without consideration of environmental and perceptual features may be unnecessarily limiting theory development regarding the CI effect. The results of this research indicate that the coordination and control of movement cannot be considered in isolation from the environment in which it occurs (Glenberg, 1979; Higgins, 1991).

Christina and Shea (1988) have argued that "parsimonious conclusions concerning learning phenomena are unlikely" (p. 291) since any generalization must consider the characteristics of the subjects under investigation, the instructions they received, the characteristics of the task or tasks that were practiced, as well as the criterion task on which subjects are tested. Christina and Shea argue that generalizing an empirical effect is likely to be an exercise in futility since it is unlikely that an empirical effect can account for all possible interactions between subjects, tasks, instructions, and criterion task. Instead of searching for the generality of a phenomenon, they proposed it might be more useful for researchers to search for the conditions under which the phenomenon does and does not occur. The present series of experiments represents such an endeavor. In the second experiment the similarity of task demands and performance environment were altered. The lack of a CI effect indicates that the relationship between the characteristics of the tasks to be learned affects whether or not the CI effect will emerge. It appears that the CI effect is not characteristic of all multi-skill environments, rather it is characteristic of multi-skill environments in which the regulatory features of the tasks are similar. Thus, the CI effect cannot be generalized to all performance environments.

Implications for practice

A shift from emphasizing differential retrieval practice to emphasizing differential encoding variability also has implications for applied practice. From the point of view of applying the results of CI research in the field, the emphasis in practice design changes from simply the spacing of skill repetitions, to the creation of situations in which learners have the opportunity to compare and differentiate critical regulatory information in order to make distinctions between events. This implication runs contrary to the suggestion made by Schmidt (1991). He states, "It is not essential that the tasks presented in a random practice session be similar to each other. In fact, the research evidence suggests that the benefits of random practice are enhanced by large task differences on successive trials. This fosters forgetting the solutions of each task before resuming its practice on a later attempt." (p. 206). Based on the results of the present experiments, it appears that Schmidt's recommendation is incorrect.

Theories of skill acquisition

The present results support the hypothesis that skill is acquired in accordance with the development of knowledge, as predicted by a number of the knowledge based theories of skill acquisition.

In Anderson's (1983) ACT* theory, the compilation processes of proceduralization and composition, and the tuning processes of generalization and discrimination have been proposed as the cognitive activities that underlie the acquisition of skill. Of particular interest to this project is the tuning process of discrimination since it is involved in the development of knowledge regarding what to do in particular performance environments. According to ACT*, selection of an appropriate problem solving procedure is a function of matching features of the environment with existing knowledge. Activation of a problem solving procedure can be limited by specifying in greater detail the characteristics of the conditions under which it is an appropriate option. This is accomplished by the process of discrimination. A prediction arising from ACT* is that as more specific condition features are associated with problem solving procedures, they become activated in fewer situations, but are more likely to be appropriate in those conditions. Conversely, problem solving procedures associated with general condition features will become activated in more situations, and may be inappropriate. The present series of experiments support this prediction. In Experiments 1 and 3A, the movement patterns could be preprogrammed and executed repeatedly for blocked practice participants. At the tone, they did not need to attend to the color word each time to properly implement the appropriate spatial sequence, they only needed to attend to the tone. Since the tone was present for all three patterns, it can be considered as a general condition feature incapable of discriminating between one pattern and the next. Thus, for blocked practice participants, each spatial sequence was likely activated by the same condition feature for many of the trials. Random practice participants, on the other hand, were required to attend to the tone and the color word on each trial. Thus each movement sequence could be differentiated from the others. The relevant environmental stimuli for random practice participants would be the color word since, unlike the tone, this environmental feature was specific to each spatial sequence required as a response. According to ACT*, retention performance should benefit random participants. Indeed, blocked practice participants experienced more difficulty than random practice participants trying to remember what pattern to execute in response to a color word during retention trials.

In Experiment 2, the environmental features specifying each task were unique. This means that only the problem solving procedure for the task at hand would have been activated for each task regardless of the conditions of practice. Under these circumstances ACT* would predict that random practice conditions should provide no benefit over blocked practice in terms of eliciting discrimination processes, and should produce similar retention performance. The results of Experiment 2 support this prediction.

In Experiment 3B the improvement in the relative retention of movement production by the BV over the BC group seems to indicate that the purposeful evaluation elicited by the experimental intervention may have encouraged discrimination to a small extent. Again, the critical environmental feature to which blocked practice participants must have attended during the acquisition phase is the color word, as this condition feature is capable of discriminating one problem solving procedure from another. In this regard, participants were encouraged to associate the color word to something personally relevant. That the retention performance of the BV group was not similar to either of the random groups suggests that the intervention may not have been sufficient to promote discrimination processes. The qualitative data revealed that while BV participants did try to establish associations to the color words, they found it difficult and bothersome to watch the monitor for the appearance of the color word on each trial, as it reduced the speed of their performance. The lack of continual interaction with the color word as a relevant environmental feature involved in producing a response may be one explanation for the lack of similarity between the BV and random groups.

The explanatory power of Ohlsson's (1996) production system is similar to that of Anderson (1983), even though the mechanics of his system are somewhat different. Since a learning event occurs following the detection of error, which stimulates the need to assign blame to a rule and to attribute error to a condition or action feature in a production rule, in Experiments 1 and 3A, blocked participants are likely to have experienced few learning events since it is likely that constraints were not regularly violated. Conversely, it is likely that blame assignment and error attribution were regularly experienced by random participants. The RV group in Experiment 3B reported trying to differentiate the first response key for each pattern early in the experiment suggesting that the first key response production rule for each task may have been activated, as predicted by Ohlsson's theory. It is not unlikely that similar strategies were used by random participants in Experiments 1 and 3A. Associating the correct first key solution with a color word would require constraints to de-activate the production rules of the other first key solutions. Multiple learning events in which erroneous features on either the condition or action side of production rules were deactivated could easily account for the relative retention demonstrated by random participants at retention. Over time, only the correct production rule would be activated due to the constraints placed on alternate rules.

In Experiment 2, since the context in which movement occurred was unique to each task, only the production rule for the specified task would be activated. Therefore the conditions of practice would not be expected to produce differences in the degree to which blame assignment and error attribution were required, and acquisition and retention performance would not be expected to differ between groups. The results support this prediction.

In Ohlsson's (1996) production system features also exist that can explain the results obtained in Experiment 3B. In this experiment, error detection should be similar during the acquisition period for the BC and BV groups; there should be few learning events since few errors occur. The only advantage to be gained by the BV group might be related to their execution performance, and only if they perceived it to be deficient in some way and sought to lay blame and attribute error to some aspect of the action side of a production rule. It is difficult to conceive that either of the BC or BV groups would engage in any kind of regular blame assignment or error attribution for features on the condition side of the production rule, since after one or two trials it is likely that the only production rule to become activated would be the correct one. Thus, Ohlsson's production system would predict that the only advantage to be gained by the BV over the BC group at retention might be associated with movement production. The results confirm this prediction.

According to the PDP connectionist model of J.B. Shea and Graf (1994), more layers of processing units were utilized more often under random than blocked training conditions. Shea and Graf suggested that this was an indication of the demands for cognitive processing required by each practice schedule. The qualitative results obtained in Experiment 3B confirm that more processing is required to produce a response under the conditions of random than blocked practice even when the blocked group was instructed to purposefully evaluate the strategies that they could use to remember each pattern.

In Experiment 2, it was suggested that the CI effect did not emerge because the unique demands of the multiple tasks did not require additional cognitive processing on the part of random practice participants. Based on the results of Experiment 2, it would be predicted that were Shea and Graf to replicate their simulation with unique tasks rather than task variations, no difference should emerge between blocked and random practice conditions on either performance errors or levels of processing used. It would be interesting to see whether a connectionist network could model these results.

Implications for a unified approach to motor skill acquisition

The results of this research have provided an insight into the nature of the cognitive processes that require cognitive effort within the CI paradigm, and may have implications for a general explanation of motor skill acquisition. Within the CI paradigm, the results suggest that (a) cognitive effort is demanded primarily by the variability with which information is encoded rather than retrieval practice, and (b) that the analysis of individual tasks produces encoding variability to a lesser extent than intertask comparison. This seems to imply that intertask comparison may be the cognitive process responsible for generating the information that produces more elaborate and distinct knowledge structures. For the general phenomenon of motor skill acquisition, the implication is that the process of comparison may play a critical role in providing information to enhance knowledge and, as a consequence, skill.

The reason that evaluation within the context of multi-item comparison rather than evaluation within the context single item analysis may constitute the cognitive process primarily responsible for differential encoding variability may be explained by comparing the information produced by comparison and analysis. Comparison facilitates a search for information about the relationships and distinctions between items whereas analysis facilitates a search for meaningful features pertaining to a single item. Over time more and richer information about the items may be discovered through comparison because items may be compared along a greater number of dimensions relative to analysis of the features of a single item. Differences in the quality of information obtained through comparison and analysis would produce differences in the extent to which each item can be distinctively represented in memory. This explanation is supported by the qualitative data obtained in Experiment 3B related to the number of dimensions along which patterns were compared by RV participants, and the frequency of statements along these dimensions relative to BV participants. Although evaluation within the context of analyzing the features of a single item and evaluation within the context of comparing the features between multiple items would both seem to require cognitive effort, the information obtained as a result of this effort can be seen to differ in terms of the effect on the distinctiveness of the memory representation of each item. Hence, the context in which evaluation occurs appears to be an important factor in the generation of elaborate and distinct knowledge structures in memory. The lack of specificity in defining the term evaluation may have been the reason that the hypothesized relationship between evaluation and cognitive effort was not fully supported in Experiment 3B.

Arising from this research are implications that may provide insights into the relationship between cognitive effort, the development of knowledge, and skill acquisition; insights that may allow preliminary specifications of a framework leading towards a unified theory of skill acquisition. It has been argued that knowledge, defined as gaining control over regulatory information derived from exteroceptive, proprioceptive, expproprioceptive, and cognitive sources, is the basis for skilled action. It has also been argued that comparison, which is a specific cognitive process demanding cognitive effort, may lead to the discovery of distinctions between the regulatory information for two or more items, which enhances knowledge. Accordingly, it can be argued that knowledge mediates the relationship between cognitive effort and the acquisition of skilled action. The results of the present research indicate that cognitive effort is a function of encoding variability rather than retrieval practice, and that comparison is likely the cognitive process responsible for demanding cognitive effort. The information produced as a function of comparison may serve as the basis for the distinctiveness and elaboration with which memory representations can be encoded. Based on these arguments, the relationship between cognitive effort, knowledge development, and skill acquisition can begin to be outlined more specifically. In a provisional theoretical framework that provides a unified view of the cognitive processes involved in motor skill acquisition, a view that will be referred to as comparative evaluation producing knowledge (CEPK), it is proposed that knowledge develops most effectively as a function of investing cognitive effort in the process of comparison. Evaluation within the context of comparison is hypothesized to lead to the discovery and encoding of distinctions between the regulatory features of two or more items. Knowledge is proposed to increase as the memory representations containing this information become more distinct and elaborate, which occurs as a natural consequence of the number of dimensions along which the distinctions between items are encoded. Consequently, it is proposed that as knowledge structures become more elaborate and distinct, the acquisition of skill is advanced. The ongoing acquisition of skill is proposed to be a function of repeated engagement in this process.

The proposed CEPK framework unifies a number of theoretical concepts related to the acquisition of skill. It incorporates the central tenets around which the productions systems of Anderson (1983) and Ohlsson (1996) are constructed, that the acquisition of skill is a function of the development of knowledge. It begins to fill in the framework proposed by Higgins (1991) with more precise theoretical constructs. Recall that in Higgins framework, the process by which skill is acquired is the result of discovery and mastery of regulatory information from a variety of sources, although no mention is made regarding what types of cognitive processes may fulfill these functions. In the present framework it is suggested that the comparison of items or events may lead to the discovery of previously unknown sources of information. It is speculated that control over access to regulatory information could be related to the level of elaboration and distinctiveness of knowledge structures.

The CEPK framework incorporates and expands Ericsson's (1993) view that deliberate practice is an effortful cognitive activity characterized by careful monitoring of performance behavior followed by corrected repetition. CEPK expands on this notion by suggesting what this careful monitoring might entail. Comparing two or more performance outcomes, or perhaps comparing performance outcomes to an ideal, may promote the discovery of new sources of regulatory information that can be incorporated into future performance. Repetition of this process over an extended period of time would lead to increasingly elaborate and distinct knowledge structures, which supports the view that expertise is a function of the accumulation of deliberate practice hours.

The CEPK framework is also compatible with Bereiter and Scardamalia's (1993) view of expertise as the process of reinvesting cognitive resources for the purpose of expanding knowledge. Expertise as Bereiter and Scardamalia conceive of it represents the process of searching out information to reformulate problems at progressively higher or more complex levels, the by-product of which is improved performance. Within the present framework, the reformulation of problems would be comprised of the comparison between the efficacy of existing problem solving routines and the efficacy of modified problem solving routines. Since expertise is viewed as a process rather than as a state, individuals may engage or disengage from the process of expertise depending on their willingness to invest cognitive effort in problem reformulation.

The second major implication arising from the proposed theoretical framework has to do with the relationship between cognitive effort and the research findings from the major learning variables of contextual interference, knowledge of results, and observational learning that were previously reviewed. While most would agree that repetition is important in the acquisition of motor skill, the CEPK framework suggests that the distinctiveness with which information can be encoded and the process of comparison that produces this distinctiveness may form the basis of the cognitive effort that underlies the acquisition of motor skill.

More specifically, the value of repetition within each of the CI, KR, and observational learning paradigms may be beneficial to the extent that each learning variable encourages the comparison of critical regulatory information. In CI research, learning may benefit as a function of comparing important perceptual, cognitive (strategic), and motor regulatory features of motor skills; in KR research, learning may benefit as a function of comparing actual results with target output at the level of movement execution, followed by the backwards propagation of error to the coordination and execution features of movement; and in observational learning research, learning may benefit as a function of comparing actual performance with ideal performance at the level of the cognitive and motor regulatory features of motor skills.

Thus, there are implications arising from the CEPK framework regarding the acquisition of perceptual, cognitive, and motor aspects of motor skills that have yet to be explored in the motor domain. For instance, where perceptual skill is important, as is the case in anticipating an upcoming event characteristic of open sports, if learners are prompted to compare the visual displays of similar events, will they be better able to gain control over the distinguishing features of each display? Where cognitive skill is important, as is the case in performance environments that allow learners to choose between one of a number of performance strategies, some of which are more effective than others, if learners are prompted to repeatedly compare the outcomes for different action strategies during training, will they learn to implement the most effective strategy any better? Where movement coordination and force production are key, if learners are prompted to repeatedly assess and compare the kinesthetic feedback from similar types of movement, will each movement become more distinct? Will learners learn to more accurately control their movement?

Conclusion

The results of this research support the hypothesis that during the acquisition period, greater processing demands are required in response to random practice than blocked practice. This is in agreement with the arguments of Lee, Swanson, et al. (1991), Lee, Swinnen, et al. (1994), and Schmidt and Bjork (1992) that cognitive effort is the underlying cause of the CI effect. However, as Experiment 2 showed, retrieval practice is not responsible for the cognitive effort produced. While Lee, Swanson, et al. (1991), Lee, Swinnen, et al. (1994), and Schmidt and Bjork (1992) claim that the CI effect is the result of either differential retrieval practice or differential encoding variability, the results of this project suggest that the locus of the CI effect lies with differential encoding variability. This is in agreement with the TPED forwarded by Shea and Zimny (1983, 1988). Nonetheless, as discovered in Experiment 3B, encouraging the evaluation of tasks in isolation does not appear to be an effective intervention with which to stimulate encoding variability. From the perspective of Glenberg's (1979) component levels theory, it appears that encoding variability is induced to a greater extent through structural analysis than analysis of descriptive components. The implication generated by this interpretation is that the comparison of multiple tasks provides more, and more different information, which leads to more elaborate and distinctive encoding. It is suggested that comparison be tested directly within the contextual interference, knowledge of results, and observational learning research paradigms.

REFERENCES

Abernethy, B. (1987). Anticipation in sport: A review. Physical Education Review, 10, 5-16.

Abernethy, B. (1990). Expertise, visual search, and information pick-up in squash. Perception, 19, 63-77.

Abernethy, B. (1993a). The nature of expertise in sport. In S. Serpa, J. Alves, & V. Pataco (Eds.), International perspectives on sport and exercise psychology (pp. 57-68). Morgantown, WV: Fitness Information Technology Inc..

Abernethy, B. (1993b). Searching for the minimal essential information for skilled perception and action. Psychological Research, 55, 131-138.

Abernethy, B., & Russell, D. G. (1984). Advance cue utilisation by skilled cricket batsmen. The Australian Journal of Science and Medicine in Sport, 16, 2-10.

Abernethy, B., & Russell, D. G. (1987a). Expert-novice differences in an applied selective attention task. Journal of Sport Psychology, 9, 326-345.

Abernethy, B., & Russell, D. G. (1987b). The relationship between expertise and visual search strategy in a racquet sport. Human Movement Science, 6, 283-319.

Abernethy, B., Neal, R., Engstrom, C., & Koning, P. (1993). What makes the expert sports performer better than the novice? Sports Coach, 16(2), 31-37.

Adams, J. A. (1986). Use of the model's knowledge of results to increase the observer's performance. Journal of Human Movement Studies, 12, 89-98.

Allard, F. (1993). Cognition, expertise, and motor performance. In J. L. Starkes, & F. Allard (Eds.), Cognitive issues in motor expertise (pp. 17-34). Amsterdam: Elsevier.

Allard, F., & Burnett, N. (1985). Skill in sport. Canadian Journal of Psychology, 39, 294-312.

Allard, F., Deakin, J., Parker, S., & Rodgers, W. (1993). Declarative knowledge in skilled motor performance: Byproduct or constituent? In J. L. Starkes & F. Allard (Eds.), Cognitive issues in motor expertise (pp. 95-107). Amsterdam: Elsevier.

Allard, F., Graham, S., & Paarsalu, M. E. (1980). Perception in sport: Basketball. Journal of Sport Psychology, 2, 14-21.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). Problem solving and learning. American Psychologist, 48, 35-44.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Battig, W. F. (1966). Facilitation and interference. In E. A. Bilodeau (Ed.), Acquisition of skill (pp. 215-244). New York: Academic Press.

Battig, W. F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R. F. Thompson & J. F. Voss (Eds.), Topics in learning and performance (pp. 131-159). New York: Academic Press.

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

Bereiter, C., & Scardamalia, M. (1993). Surpassing ourselves: An inquiry into the nature and implications of expertise. Chicago, IL: Open Court.

Bernstein, N. (1967). The co-ordination and regulation of movements. Oxford, England: Pergamon Press.

Bilodeau, E. A., & Bilodeau, I. M. (1958). Variable frequency knowledge of results and the learning of a simple skill. Journal of Experimental Psychology, 55, 379-383.

Borgeaud, P., & Abernethy, B. (1987). Skilled perception in volleyball defense. Journal of Sport Psychology, 9, 400-406.

Bortoli, L., Robazza, C., Durigon, V., & Carra, C. (1992). Effects of contextual interference on learning technical sports skills. Perceptual and Motor Skills, 75, 555-562.

Boyce, B. A., & Del Rey, P. (1990). Designing applied research in a naturalistic setting using a contextual interference paradigm. Journal of Human Movement Studies, 18, 189-200.

Buckolz, E., Prapavesis, H., & Fairs, J. (1988). Advance cues and their use in predicting tennis passing shots. Canadian Journal of Sport Sciences, 13, 20-30.

Carnahan, H., Van Eerd, D. L., & Allard, F. (1990). A note on the relationship between task requirements and the contextual interference effect. Journal of Motor Behavior, 22, 159-169.

Chamberlin, C., & Lee, T. D. (1993). Arranging practice conditions and designing instruction. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.), Handbook of Research on Sport Psychology (pp. 213-241). New York: MacMillan.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), Cognitive skills and their acquisition (pp. 141-189). Hillsdale NJ: Erlbaum.

Christina, R. W., & Shea, J. B. (1988). The limitations of generalization based on restricted information. Research Quarterly for Exercise and Sport, 5, 291-297.

Côté, J., Salmela, J. H., & Russell, S. (1995). The knowledge of high-performance gymnastic coaches: Methodological framework. The Sport Psychologist, 9, 65-75.

Côté, J., Salmela, J. H., Baria, A., & Russell, S. (1993). Organizing and interpreting unstructured qualitative data. The Sport Psychologist, 7, 127-137.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. Journal of Verbal learning and Verbal Behavior, 21, 451-467.

Ericsson, K. A. (1996a). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), The Road to Excellence (pp. 1-50). Mahwah, NJ: Erlbaum.

Ericsson, K. A. (Ed.). (1996b). The Road to Excellence. Mahwah, NJ: Erlbaum.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. American Psychologist, 49, 725-747.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. Annual Review of Psychology, 47, 273-305.

Ericsson, K. A., & Smith, J. (1991a). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), Towards a general theory of expertise: Prospects and limits (pp. 1-38). New York, NY: Cambridge.

Ericsson, K. A., & Smith, J. (Ed.). (1991b). Towards a general theory of expertise: Prospects and limits. New York, NY: Cambridge.

Ericsson, K. A., Krampe, R. Th., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100, 363-406.

French, K. E., Rink, J. E., & Werner, D. H. (1990). Effects of contextual interference on retention of three volleyball skills. Perceptual and Motor Skills, 71, 179-186.

Gentile, A. M. (1987). Skill acquisition: Action, movement, and neuromotor processes. In J. H. Carr, R. B. Shepherd, J. Gordon, et al. (Eds.), Movement Science: Foundations for Physical Therapy in Rehabilitation (pp. 93-154). Rockville, MD: Apen Publishers.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston, MA: Houghton-Mifflin.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. Memory & Cognition, 7, 95-112.

Goode, S., & Magill, R. A. (1986). The contextual interference effects in the learning of three badminton serves. Research Quarterly for Exercise and Sport, 57, 308-314.

Goulet, C., Bard, C., & Fleury, M. (1989). Expertise differences in preparing to return a tennis serve: A visual information processing approach. Journal of Sport and Exercise Psychology, 11, 382-398.

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. Perceptual and Motor Skills, 78, 835-841.

Hand, J., & Sidaway, B. (1992). Relative frequency of modeling effects on performance and retention of a motor skill. Research Quarterly for Exercise and Sport, 63 (Suppl.), A57-58.

Hebert, E. P., Landin, D, & Solmon, M. A. (1996). Practice schedule effects on the performance and learning of low- and high-skilled students: An applied study. Research Quarterly for Exercise and Sport, 67, 52-58.

Helsen, W., & Pauwels, J. M. (1993). The relationship between expertise and visual information processing in sport. In J. L. Starkes & F. Allard (Eds.), Cognitive issues in motor expertise (pp. 109-134). Amsterdam: Elsevier.

Hick, W. E. (1952). On the rate of gain of information. Quarterly Journal of Experimental Psychology, 4, 11-26.

Higgins, S. (1991). Motor skill acquisition. Physical Therapy, 71, 123-139.

Horak, M. (1992). The utility of connectionism for motor learning: A reinterpretation of contextual interference in movement schemas. Journal of Motor Behavior, 24, 58-66.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. Journal of Experimental Psychology, 45, 188-196.

Isaacs, L. D., & Finch, A. E. (1983). Anticipatory timing of beginning and intermediate tennis players. Perceptual and Motor Skills, 57, 451-454.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. Journal of Verbal Learning and Verbal Behavior, 17, 649-667.

Jones, C. M., & Miles, T. R. (1978). Use of advance cues in predicting the flight of a lawn tennis ball. Journal of Human Movement Studies, 4, 231-235.

Lavery, J. J. (1962). Retention of simple motor skills as a function of type of knowledge of results. Canadian Journal of Psychology, 16, 300-311.

Lee, T. D. (1988). Transfer-appropriate processing: A framework for conceptualizing practice effects in motor learning. In O. G. Meijer & K. Roth (Eds.), Complex movement behavior: 'The' motor-action controversy (pp. 201-215). Amsterdam: Elsevier.

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

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), Differing perspectives in motor learning, memory and control (pp. 3-22). Amsterdam: North Holland.

Lee, T. D., & White, M. A. (1990). Influence of an unskilled model's practice schedule on observational motor learning. Human Movement Science, 9, 349-367.

Lee, T. D., Swanson, L. R., & Hall, A. L. (1991). What is repeated in a repetition? Effects of practice conditions on motor skill acquisition. Physical Therapy, 71, 150-156.

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. Quest, 46, 328-344.

Lee, T. D., White, M. A., & Carnahan, H. (1990). On the role of knowledge of results in motor learning: Exploring the guidance hypothesis. Journal of Motor Behavior, 22, 191-208.

Lee, T. D., Wulf, G., & Schmidt, R. A. (1992). Contextual interference in motor learning: Dissociated effects due to the nature of task variations. The Quarterly Journal of Experimental Psychology, 44A, 627-644.

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

Magill, R. A. (1993a). Augmented feedback in skill acquisition. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.), Handbook of Research on Sport Psychology (pp. 193-212). New York: MacMillan.

Magill, R. A. (1993b). Modeling and verbal feedback influences on skill learning. International Journal of Sport Psychology, 24, 358-369.

Magill, R. A. (1998a). 1997 C. H. McCloy Research Lecture: Knowledge is more than we can talk about: Implicit learning in motor skill acquisition. Research Quarterly for Exercise and Sport, 69, 104-110.

Magill, R. A. (1998b). Motor learning: Concepts and applications. Dubuque, IA: WCB McGraw-Hill.

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

Martens, R., Burwitz, L., & Zuckerman, J. (1976). Modeling effects on motor performance. Research Quarterly, 47, 277-291.

Masson, M. E. J. (1990). Cognitive theories of skill acquisition. Human Movement Science, 9, 221-239.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. The Psychology of Learning and Motivation, 24, 109-165.

McCullagh, P., & Caird, J. K. (1990). Correct and learning models and the use of model knowledge of results in the acquisition and retention of a motor skill. Journal of Human Movement Studies, 18, 107-116.

Newell, A, & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), Cognitive skills and their acquisition (pp. 1-55). Hillsdale, NJ: Erlbaum.

Ohlsson, S. (1996). Learning from performance errors. Psychological Review, 103, 241-262.

Pollock, B. J., & Lee, T. D. (1992). Effects of the model's skill level on observational motor learning. Research Quarterly for Exercise and Sport, 63, 25-29.

Ripoll, H. (1991). The understanding-acting process in sport: The relationship between the semantic and the sensorimotor visual function. International Journal of Sport Psychology, 22, 221-243.

Ripoll, H., Kerlirzin, Y., Stein, J. F., & Reine, B. (1995). Analysis of information processing, decision making, and visual strategies in complex problem solving sport situations. Human Movement Science, 14, 325-349.

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. (1988). Motor Control and Learning (2nd ed.). Champaign, IL: Human Kinetics.

Schmidt, R. A. (1991). Motor learning and performance: From principles to practice. Champaign, IL: Human Kinetics.

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

Schmidt, R. A., Young, D. E., Swinnen, S. P., & 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.

Scully, D. M., & Newell, K. M. (1985). Observational learning and the acquisition of motor skills: Toward a visual perception perspective. Journal of Human Movement Studies, 11, 169-186.

Sekiya, H., Magill, R. A., Sidaway, B., & Anderson, D. I. (1994). The contextual interference effect for skill variations from the same and different generalized motor programs. Research Quarterly for Exercise and Sport, 65, 330-338.

Shank, M. D., & Haywood, K. M. (1987). Eye movements while viewing a baseball pitch. Perceptual and Motor Skills, 64, 1191-1197.

Shea, C. H., Kohl, R., & Indermill, C. (1990). Contextual interference: Contributions of practice. Acta Psychologica, 73, 145-157.

Shea, J. B., & Graf, R. C. (1994). A model for contextual interference effects in motor learning. In C. R. Reynolds (Ed.), Cognitive assessment: A multidisciplinary perspective (pp. 73-87). New York: Plenum Press.

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

Shea, J. B., & Wright, D. L. (1991). When forgetting benefits motor retention. Research Quarterly for Exercise and Sport, 62, 293-301.

Shea, J. B., & Zimny, S. T. (1983). Context effects in memory and learning movement information. In R. A. Magill (Ed.), Memory and control of action (pp. 345-366). Amsterdam: North Holland.

Shea, J. B., & Zimny, S. T. (1988). Knowledge incorporation in motor representation. In O. G. Meijer & K. Roth (Eds.), Complex movement behavior: 'The' motor-action controversy (pp. 289-314). Amsterdam: Elsevier.

Shewokis, P. A. (1997). Is the contextual interference effect generalizable to computer games? Perceptual and Motor Skills, 84, 3-15.

Smith, P. J. K., & Davies, M. (1995). Applying contextual interference to the Pawlata roll. Journal of Sports Sciences, 13, 455-462.

Starkes, J. L. (1987). Skill in field hockey: The nature of the cognitive advantage. Journal of Sport Psychology, 9, 146-160.

Starkes, J. L., & Deakin, J. (1984). Perception in sport: A cognitive approach to skilled performance. In W. F. Straub & J. M. Williams (Eds.), Cognitive sport psychology (pp. 115-128). Lansing, NY: Sport Science Association.

Sternberg, R. J. (1996). Costs of expertise. In K. A. Ericsson (Ed.), The Road to Excellence (pp. 347-354). Mahwah, NJ: Erlbaum.

Sternberg, R. J. (1998). Abilities are forms of developing expertise. Educational Researcher, 27(3), 11-20.

Swinnen, S. P. (1990). Interpolated activities during the knowledge-of-results delay and post-knowledge-of-results interval: Effects on performance and learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 16, 692-705.

Swinnen, S. P., Schmidt, R. A., Nicholson, D. E., & Shapiro, D. C. (1990). Information feedback for skill acquisition: Instantaneous knowledge of results degrades learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 16, 706-716.

Van Donkelaar, P., & Franks, I. M. (1991). Preprogramming vs. on-line control in simple movement sequences. Acta Psychologica, 77, 1-19.

Whiting, H., Bijlard, M., & den Brinker, B. (1987). The effect of the availability of a dynamic model on the acquisition of a complex cyclical action. The Quarterly Journal of Experimental Psychology, 39a, 43-59.

Williams, A. M., & Davids, K. (1995). Declarative knowledge in sport: A by product of experience or a characteristic of expertise? Journal of Sport and Exercise Psychology, 17, 259-275.

Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. Research Quarterly for Exercise and Sport, 69, 111-128.

Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1993). Visual search and sports performance. The Australian Journal of Science and Medicine in Sport, 22, 55-65.

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.

Wood C. A., & Ging, C. A. (1991). The role of interference and task similarity on the acquisition, retention, and transfer of simple motor skills. Research Quarterly for Exercise and Sport, 62, 18-26.

Wright, D. L. (1991). The role of intertask and intratask processing in acquisition and retention of motor skills. Journal of Motor Behavior, 23, 139-145.

Wright, D. L., Li, Y., & Whitacre, C. (1992). The contribution of elaborative processing to the contextual interference effect. Research Quarterly for Exercise and Sport, 63, 30-37.

Wrisberg, C. A. (1991). A field test of the effect of contextual variety during skill acquisition. Journal of Teaching in Physical Education, 11, 21-30.

Wrisberg, C. A., & Liu, Z. (1991). The effect of contextual variety on the practice, retention, and transfer of an applied motor skill. Research Quarterly for Exercise and Sport, 62, 406-412.

Young, D. E., Cohen, M. J., & Husak, W. S. (1993). Contextual interference and motor skill acquisition: On the processes that influence retention. Human Movement Science, 12, 577-600.

Table 1

Means (msec), standard deviations, and effect size (ES) calculations for the Acquisition Condition main effect during the acquisition phase of Experiment 1.

	Blocked		Random				
Measure	M_B	sd_B	M_R	sd_R	$\lvert M_B-M_R \rvert$	pooled sd	ES
TRT	1261.9	197.5	2319.2	792.0	1057.3	577.2	1.83***
RT	286.0	42.4	883.4	434.6	597.5	308.8	1.93***
MT	976.0	177.4	1435.8	600.9	459.9	443.0	1.04**

**p<.01 ***p<.001

Table 2

Results of analyses on contrasts of adjacent trial blocks for the AC x TB interaction during the acquisition phase of Experiment 1.

Measure	Contrast	$F(1, 11) =$	$p <$	η^2
TRT	1 - 2	12.66	.01	.535
	2 - 3	5.99	.05	.352
	3 - 4	8.91	.05	.447
	4 - 5	5.15	.05	.319
	5 - 6	9.20	.05	.455
RT	3 - 4	5.00	.05	.312
	4 - 5	6.60	.05	.375
MT	1 - 2	6.75	.05	.380
	2 - 3	7.48	.05	.405
	5 - 6	10.51	.01	.489

Table 3

Means, standard deviations, and effect size (ES) calculations for reliable AC x TB contrasts on performance differences (in msecs) between blocked and random conditions for the last block of acquisition (6) and the immediate retention (IR) during Experiment 1.

Contrast	M_6	sd_6	M_{IR}	sd_{IR}	$\lvert M_6-M_{IR}\rvert$	pooled sd	ES
TRT	-663.3	531.3	123.2	593.1	786.5	563.1	1.40***
RT	-419.5	309.6	35.7	345.2	455.1	327.9	1.39***
MT	-243.9	354.3	87.5	371.7	331.4	363.1	0.91***

***$p<.001$

Note: Mean values represent performance differences between acquisition conditions (blocked performance - random performance). Negative values indicate slower random than blocked performance, positive values indicate faster random than blocked performance.

Table 4

Results of analyses on contrasts of adjacent trial blocks during the acquisition phase of Experiment 2.

Contrast	$F(1, 11) =$	$p <$	η^2
1 -2	39.17	.001	.781
2 -3	49.43	.001	.818
3 -4	18.23	.001	.624
4 - 5	18.55	.001	.628

Table 5

Results of analyses on contrasts of adjacent trial blocks for the Task x TB interaction during the acquisition phase of Experiment 2.

Comparison	Contrast	$F(1, 11) =$	$p <$	η^2
PBC vs UP	1 -2	599.29	.001	.982
	2 -3	101.59	.001	.902
	3 -4	24.56	.001	.691
	4 - 5	39.29	.001	.781
	5 - 6	7.76	.05	.455
PBC vs MD	1 -2	8.92	.05	.448
	2 -3	25.87	.001	.702
	3 -4	6.60	.05	.375
MD vs UP	2 -3	5.42	.05	.330

Table 6

Means (msec), standard deviations, and effect size (ES) calculations for the Acquisition Condition main effect during the acquisition phase of Experiment 3A.

	Blocked		Random				
Measure	M_B	sd_B	M_R	sd_R	$\lvert M_B - M_R \rvert$	pooled sd	ES
TRT	1201.9	246.4	1958.3	702.8	756.4	526.6	1.44***
RT	265.5	43.4	720.3	377.1	454.7	268.4	1.69***
MT	936.4	223.7	1238.0	478.8	301.6	373.7	0.81**

**p<.01 ***p<.001

Note: Blocked and Random acquisition conditions represents the combined performance of control and verbalize groups.

Table 7

Results of analyses on contrasts of adjacent trial blocks for the AC x TB interaction during the acquisition phase of Experiment 3A.

Measure	Contrast	$F\ (1, 43) =$	$p <$	η^2
TRT	1 - 2	46.25	.001	.518
	2 - 3	53.46	.001	.554
	3 - 4	9.11	.01	.175
	4 - 5	5.67	.05	.117
	5 - 6	7.01	.05	.114
RT	1 - 2	15.41	.001	.264
	2 - 3	23.43	.001	.353
	3 - 4	5.28	.05	.109
	5 - 6	14.22	.001	.249
MT	1 - 2	21.57	.001	.334
	2 - 3	16.70	.001	.280
	3 - 4	5.57	.05	.115
	4 - 5	7.71	.01	.152

Table 8

Means, standard deviations, and effect size (ES) calculations for reliable AC x TB contrasts on performance differences (in msecs) from the last block of acquisition to immediate retention for Experiment 3A.

Contrast	M_B	sd_B	M_R	sd_R	$\|M_B-M_R\|$	pooled sd	ES
TRT	-949.9	589.2	-53.2	140.7	896.7	428.4	2.09***
RT	-740.7	594.0	-28.6	83.9	712.1	424.2	1.68***
MT	-209.2	119.2	-24.6	91.5	184.6	106.3	1.74***

***p<.001

Note: Difference values represent mean performance differences across trial blocks. Relative to performance at the end of acquisition, negative values indicate slower performance at immediate retention. Blocked and Random conditions represents the combined performance of control and verbalize groups.

Table 9

Reliable between-group contrasts for the Task x AC interaction during the acquisition phase of Experiment 3B.

Task Contrast	Group Contrast	F (1, 43) =	p <	η^2
White vs Fuchsia	RV vs BC	4.64	.05	.097
	RV vs BV	4.30	.05	.091
	RC vs BC	30.71	.001	.417
	RC vs BV	29.83	.001	.410
	RC vs RV	10.67	.01	.199
White vs Lime	RC vs BC	24.61	.001	.364
	RC vs BV	21.60	.001	.334
	RC vs RV	12.86	.001	.230

Table 10

Means (msec), standard deviations, and effect size (ES) calculations for Acquisition Condition main effects during the acquisition phase of Experiment 3B.

Measure	Blocked M_B	Blocked sd_B	Random M_R	Random sd_R	$\lvert M_B-M_R \rvert$	pooled sd	ES
TRT	1143.0	204.0	1733.8	677.2	590.8	500.2	1.18***
RT	264.3	63.4	623.0	341.4	358.7	245.6	1.46***
MT	878.7	182.0	1110.7	455.4	232.1	346.8	0.67*

*p<.05 ***p<.001

Note: Blocked and Random acquisition conditions represents the combined performance of control and verbalize groups.

Table 11

Means, standard deviations, and effect size (ES) calculations for reliable AC x TB contrasts on performance differences (in msecs) from the last block of acquisition to immediate retention for Experiment 3B.

Measure	M_1	sd_1	M_2	sd_2	$\lvert M_1-M_2 \rvert$	pooled sd	ES
	BV	vs	BC				
TRT	-504.2	333.8	-775.4	418.5	271.2	378.5	0.72*
RT	-366.0	281.7	-546.1	372.8	180.2	330.4	0.55
MT	-138.2	108.7	-229.2	115.4	91.0	112.2	0.81*
	RV	vs	RC				
TRT	-6.7	49.6	51.4	172.4	58.2	126.8	0.46
RT	-17.0	32.5	8.5	133.4	25.5	97.1	0.26
MT	10.3	48.6	43.0	67.4	32.7	58.7	0.56
	BV	vs	R				
TRT	-504.2	333.8	23.6	129.8	527.8	253.2	2.08***
RT	-366.0	281.7	-3.7	97.7	362.3	210.8	1.72***
MT	-138.2	108.7	27.3	60.2	165.5	87.8	1.88***

*$p<.05$ ***$p<.001$

Note: Difference values represent mean performance differences across trial blocks. Relative to the last trial block of acquisition, negative values indicate slower performance at immediate retention, positive values indicate faster performance at immediate retention.

Table 12

Means, standard deviations, and effect size (ES) calculations for reliable AC main effects for BC and RC experimental groups during the acquisition phase of Experiment 3B.

Contrast	M_{BC}	sd_{BC}	M_{RC}	sd_{RC}	$\lvert M_{BC}-M_{RC}\rvert$	pooled sd	ES
TRT	1130.5	208.0	1884.3	802.1	753.8	585.9	1.29***
RT	249.3	32.0	719.8	396.6	470.5	281.4	1.67***
MT	881.2	197.4	1164.5	528.1	283.3	398.6	0.71*

*p<.05 ***p<.001

Table 13

<u>Means, standard deviations, and effect size (ES) calculations for reliable AC x TB contrasts on performance differences (in msecs) from the last block of acquisition to immediate retention for BC and RC experimental groups during Experiment 3B.</u>

Contrast	M_{BC}	sd_{BC}	M_{RC}	sd_{RC}	$\lvert M_{BC}-M_{RC}\rvert$	pooled sd	ES
TRT	-775.4	418.5	51.4	172.4	826.8	320.0	2.58***
RT	-546.1	372.8	8.5	133.4	554.6	280.0	1.98***
MT	-229.2	115.4	43.0	67.4	272.2	94.5	2.88***

***$p<.001$

Note: Difference values represent mean performance differences across trial blocks. Relative to the last trial block of acquisition, negative values indicate slower performance at immediate retention, positive values indicate faster performance at immediate retention.

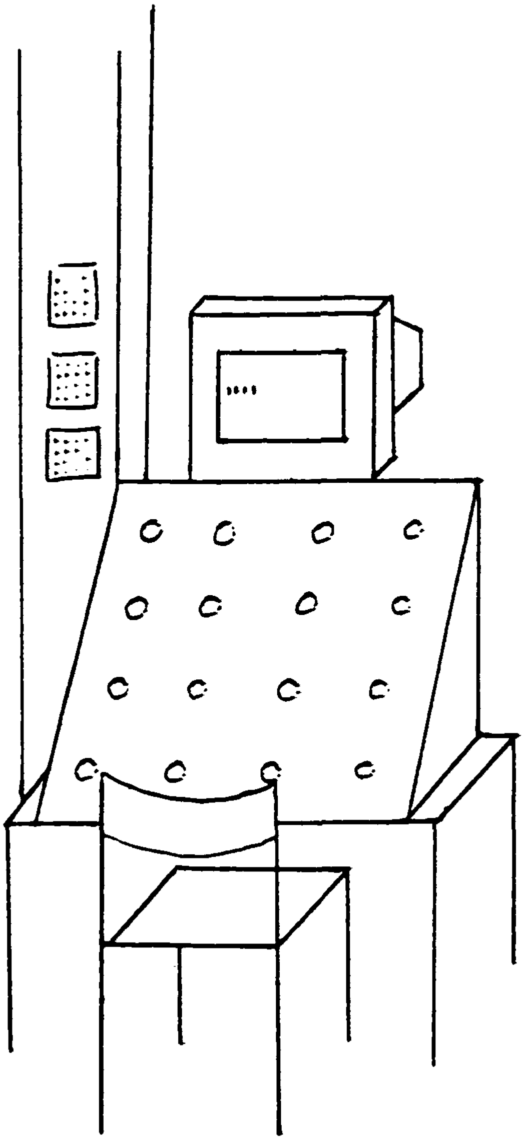
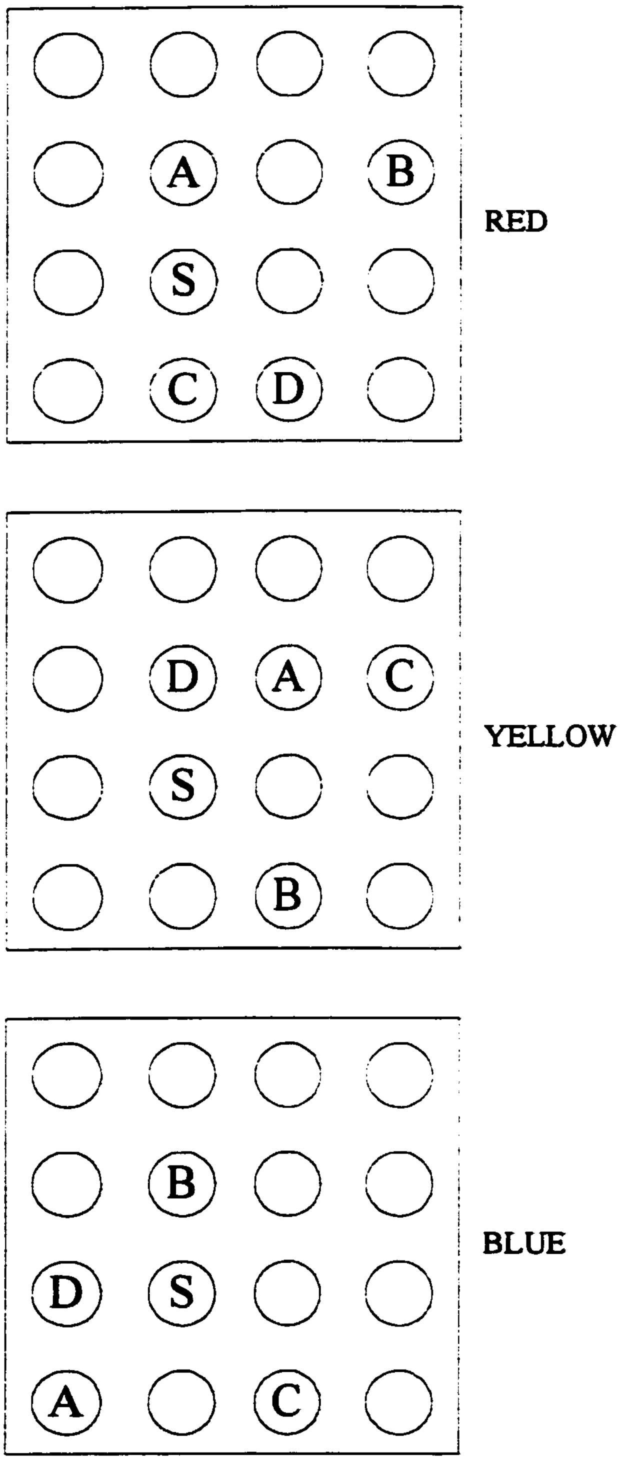


Figure 1. Physical arrangement for the Push Button Challenge task.



<u>Figure 2.</u> The pattern cards indicating the location and order in which response keys were to be depressed for Experiment 1. Each movement pattern card was copied on appropriately colored paper.

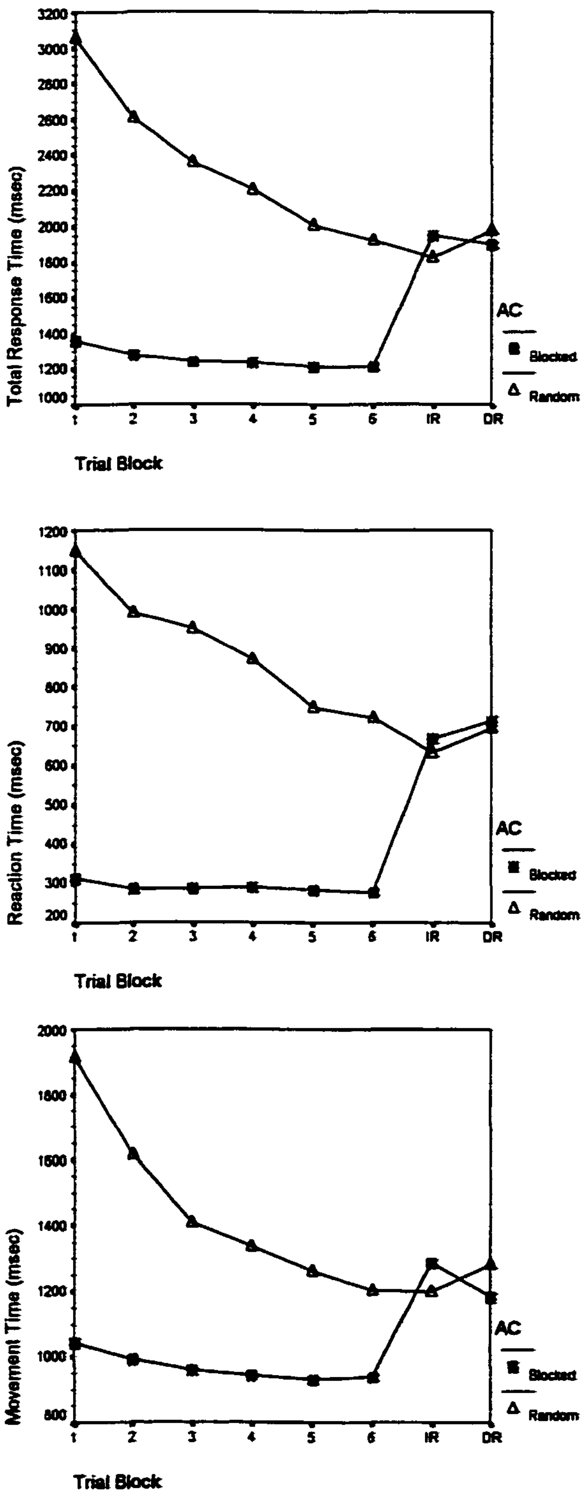


Figure 3. Mean acquisition, immediate (IR) and delayed (DR) retention performance for TRT, RT, and MT under blocked and random practice conditions during Experiment 1.

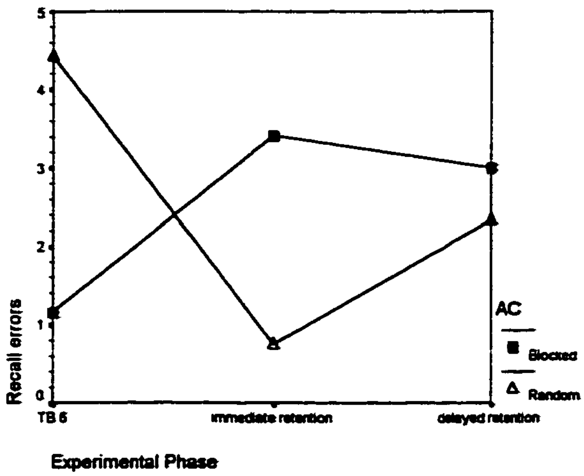


Figure 4. Mean recall error frequency by blocked and random practice conditions through out each phase of Experiment 1.

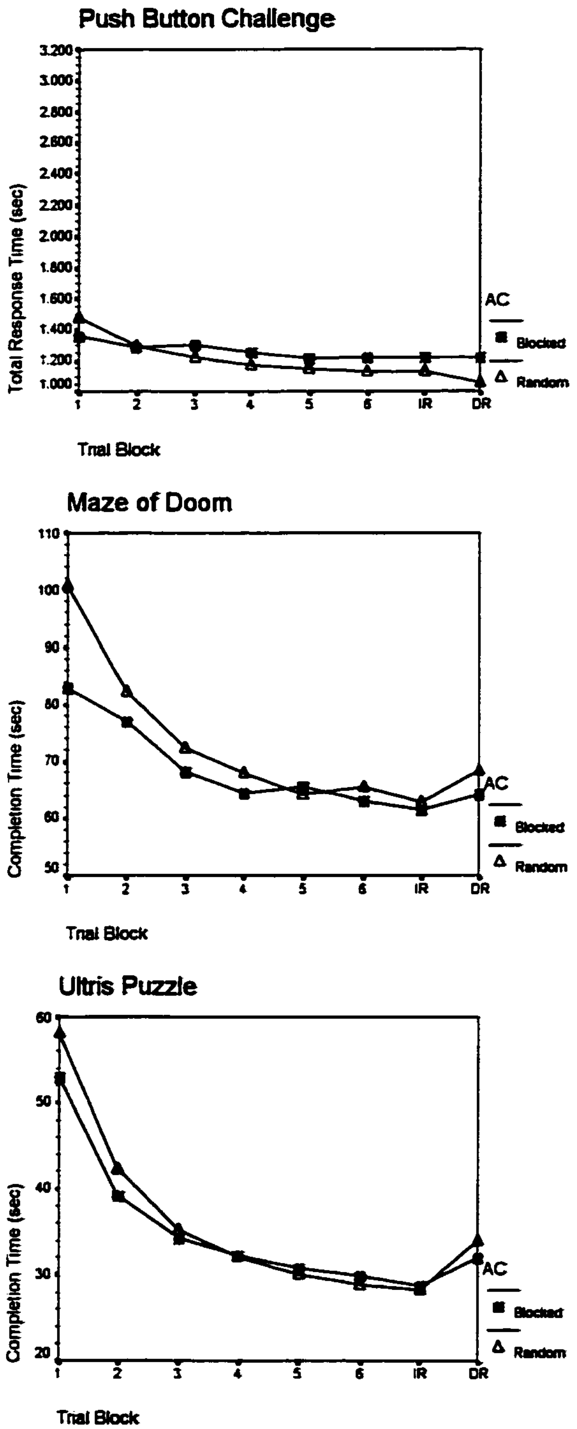


Figure 5. Mean acquisition, immediate (IR) and delayed (DR) retention performance for PBC, Maze of Doom, and Ultris puzzle under blocked and random practice conditions during Experiment 2.

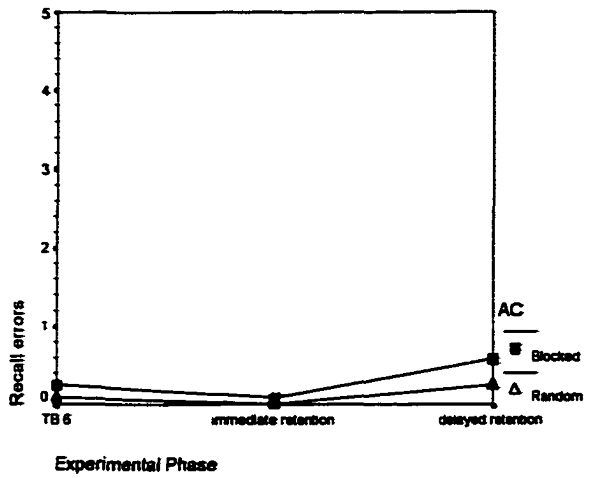


Figure 6. Mean PBC recall error frequency under blocked and random practice conditions throughout each phase of Experiment 2.

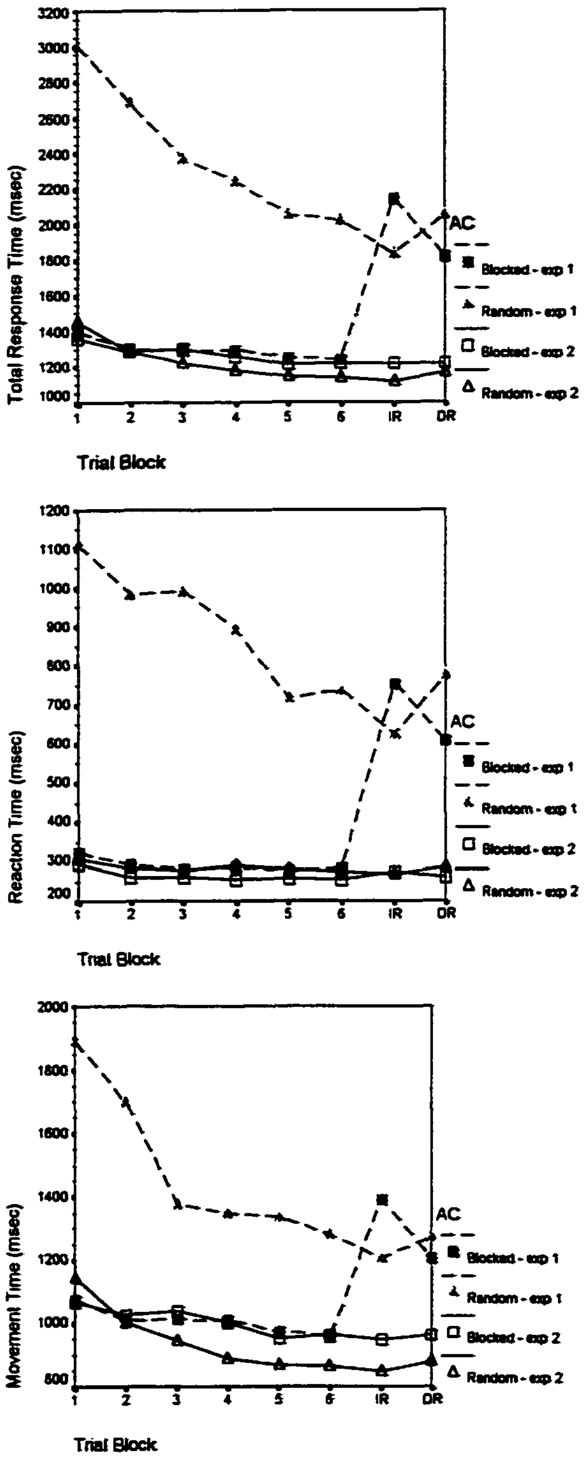
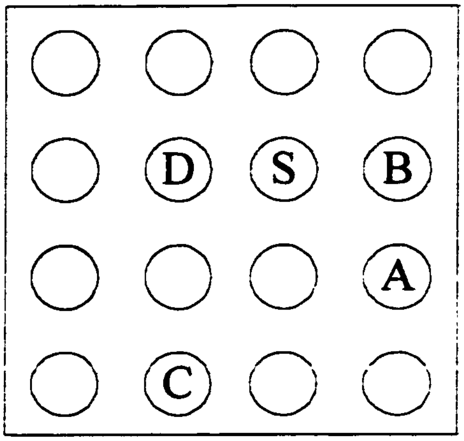
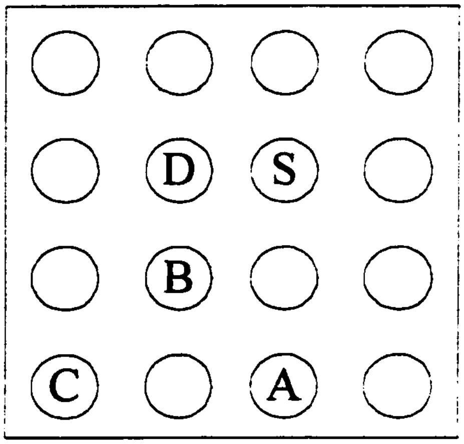


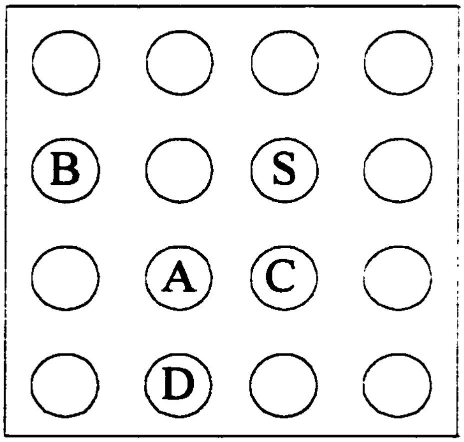
Figure 7. Comparison of mean TRT, RT, and MT for blocked and random acquisition, immediate (IR) and delayed (DR) retention performance between experiments 1 and 2.



WHITE



LIME



FUCHSIA

Figure 8. The pattern cards indicating the location and order in which response keys were to be depressed for Experiment 3B. Each movement pattern card was copied on appropriately colored paper.

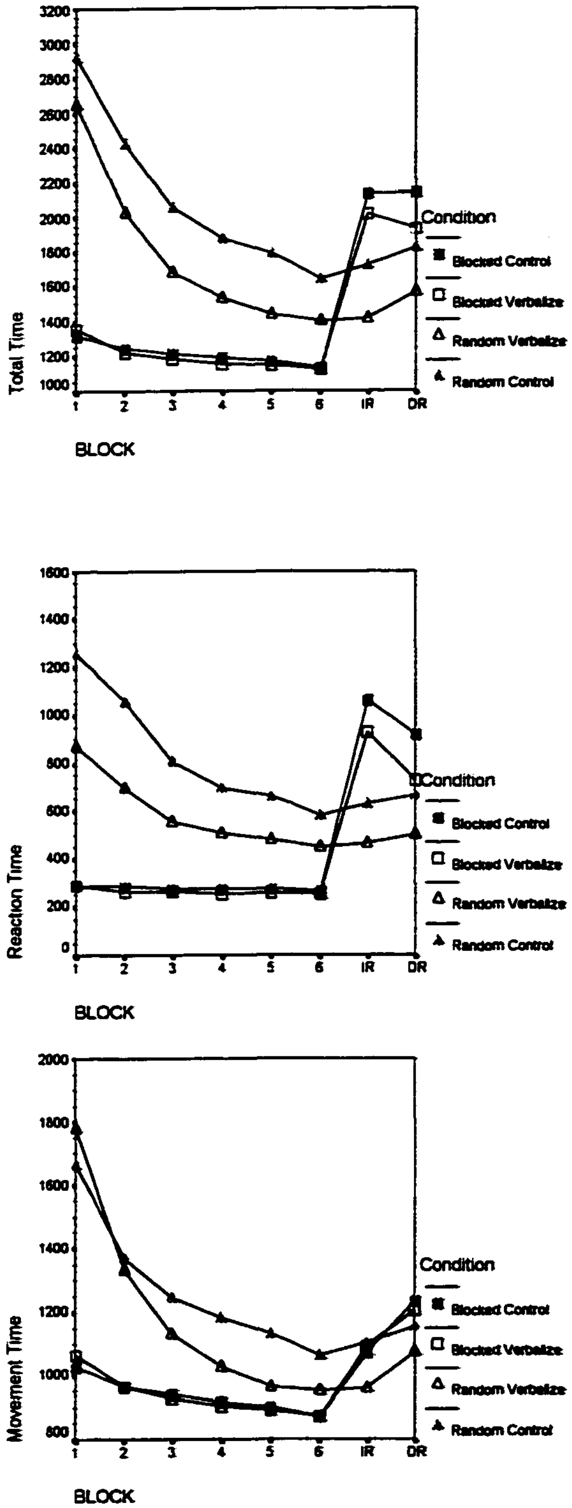


Figure 9. Mean performance (in msecs) for acquisition, immediate retention (IR), and delayed retention (DR) trial blocks for TRT, RT, and MT during Experiment 3A.

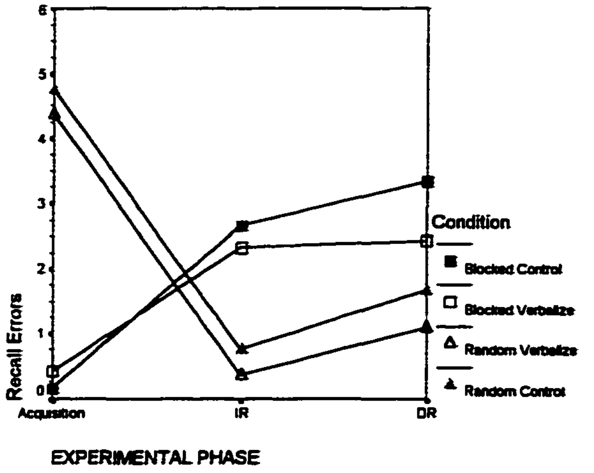


Figure 10. Mean recall error frequency by experimental condition during acquisition, IR, and DR of Experiment 3A.

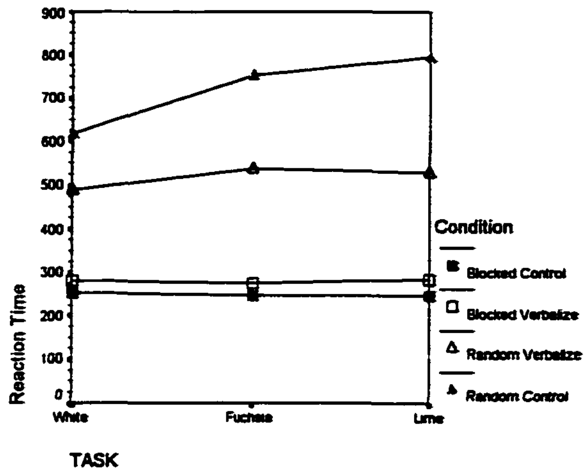


Figure 11. Reaction times (in msecs) on the white, fuchsia, and lime tasks during the acquisition period of Experiment 3B.

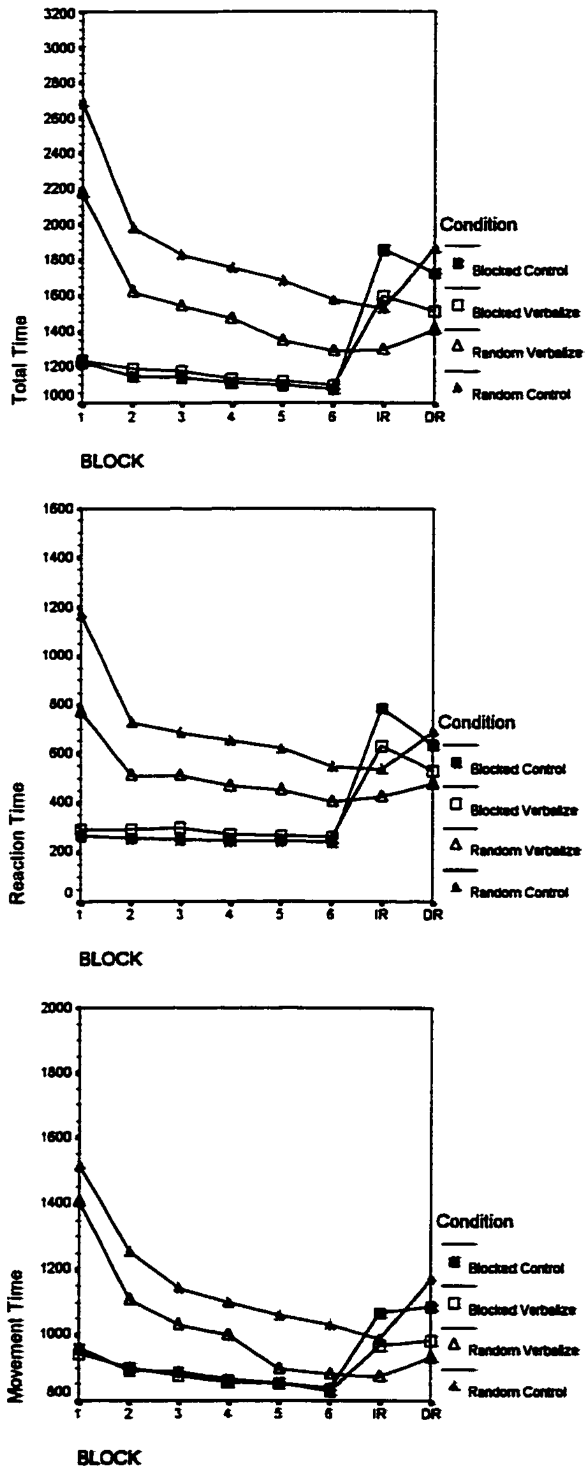


Figure 12. Mean performance (in msecs) for acquisition, immediate retention (IR), and delayed retention (DR) trial blocks for TRT, RT, and MT during Experiment 3B.

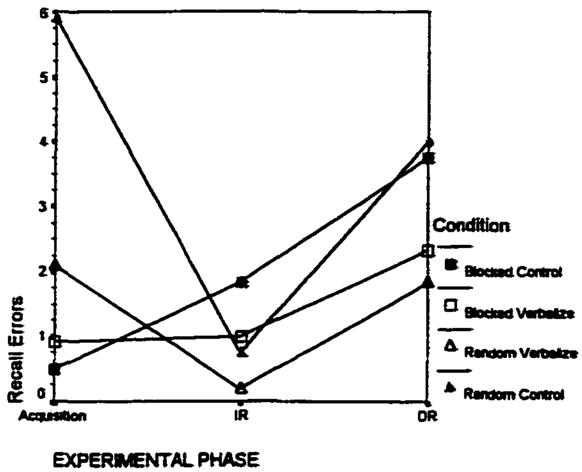


Figure 13. Mean recall error frequency for each experimental condition during acquisition, IR, and DR of Experiment 3B.

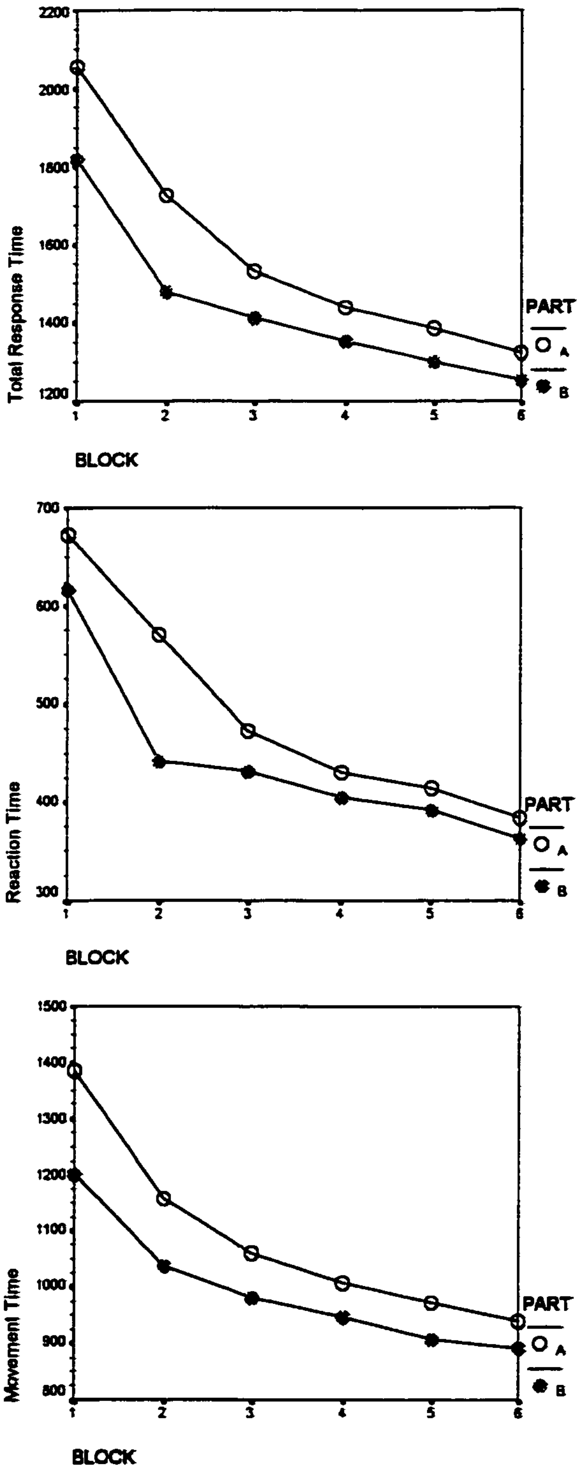
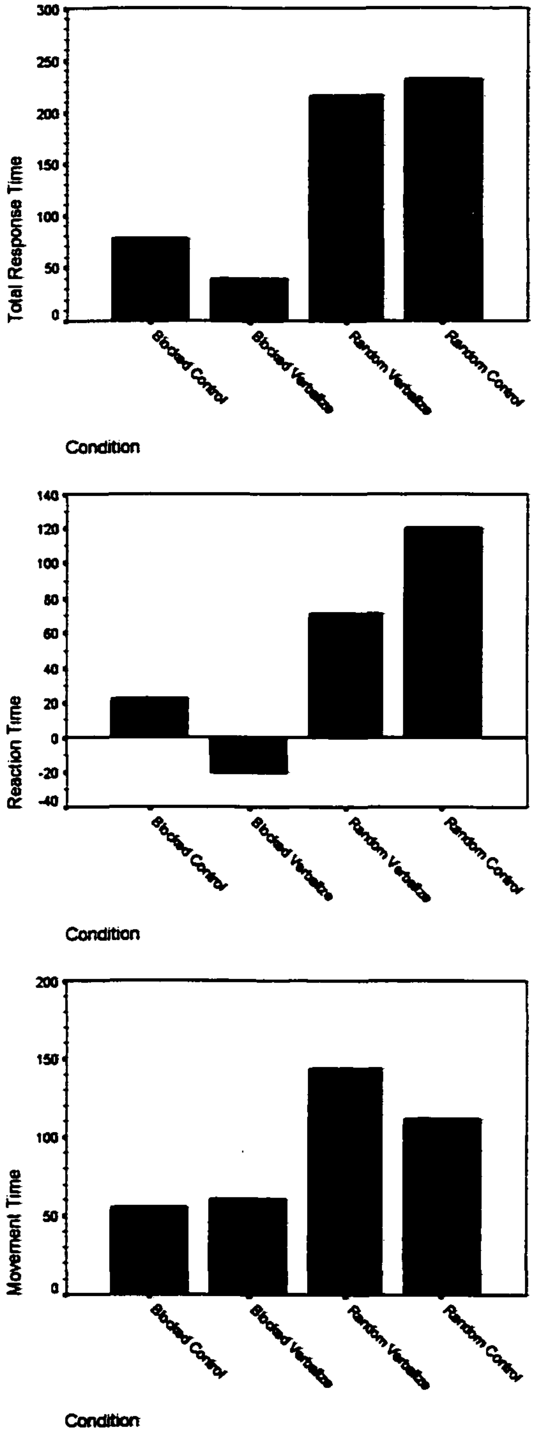


Figure 14. Mean performance (in msecs) for acquisition trial blocks for TRT, RT, and MT during Experiment 3A and 3B.



<u>Figure 15</u>. Differences in acquisition performance (in msecs) for TRT, RT, and MT between 3A & 3B; positive values indicates faster performance during 3B than 3A, negative values indicate slower performance during 3B than 3A.

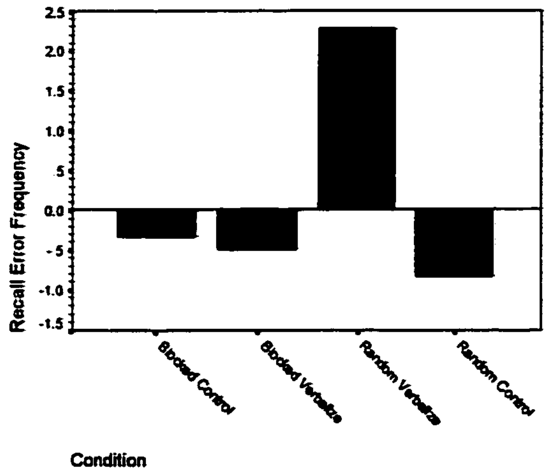
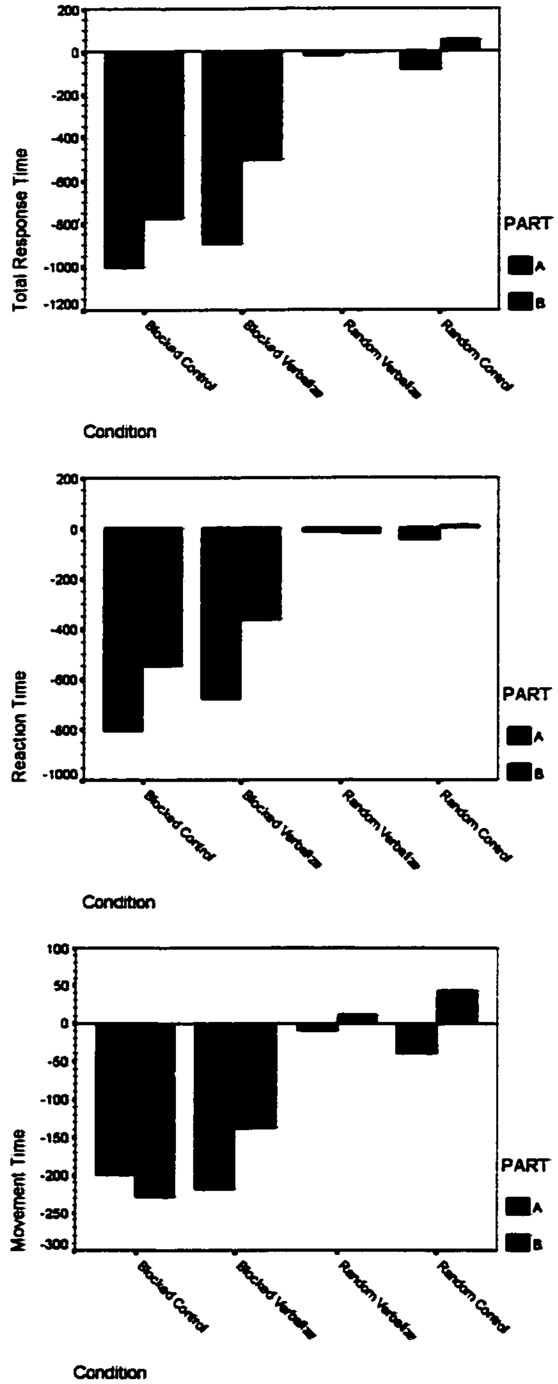


Figure 16. Differences in acquisition recall error frequency between 3A & 3B; positive values indicate fewer recall errors during 3B than 3A, negative values indicate more recall errors during 3B than 3A.



<u>Figure 17</u>. Mean performance change (in msecs) from the last acquisition trial block to immediate retention for TRT, RT, and MT from 3A to 3B. Positive values indicate faster IR than acquisition performance, negative values indicate slower IR than acquisition performance.

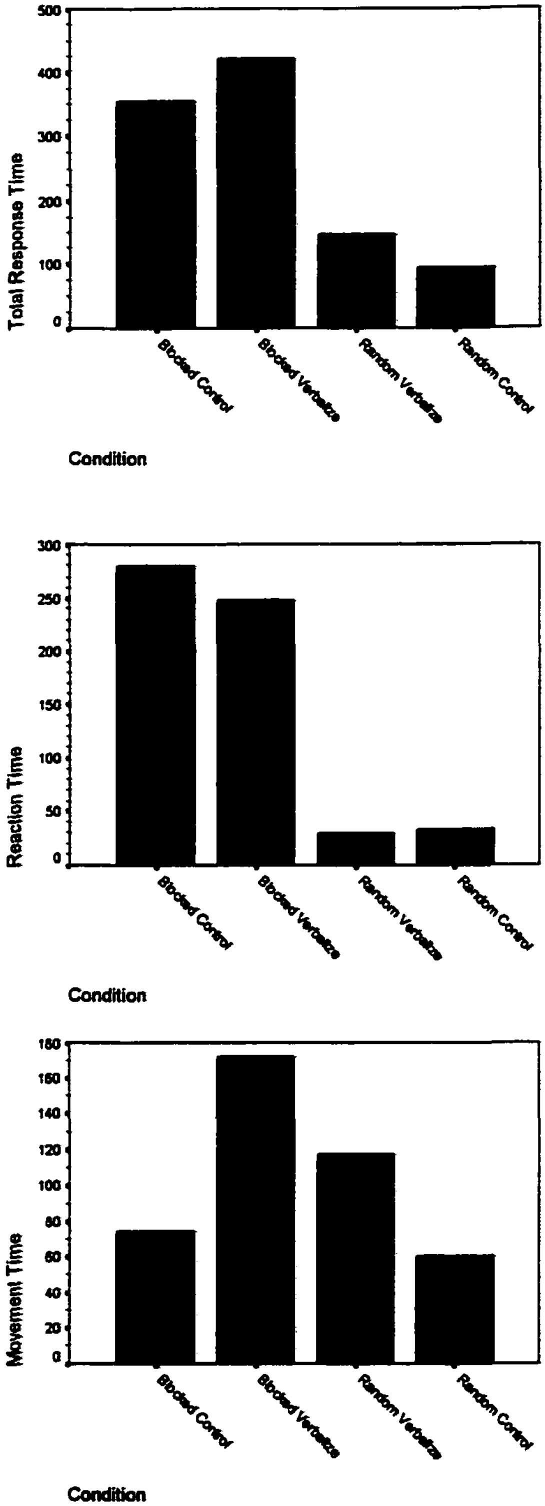


Figure 18. Differences in retention performance (in msecs) for TRT, RT, and MT between 3A & 3B; positive values indicates faster performance during 3B than 3A.

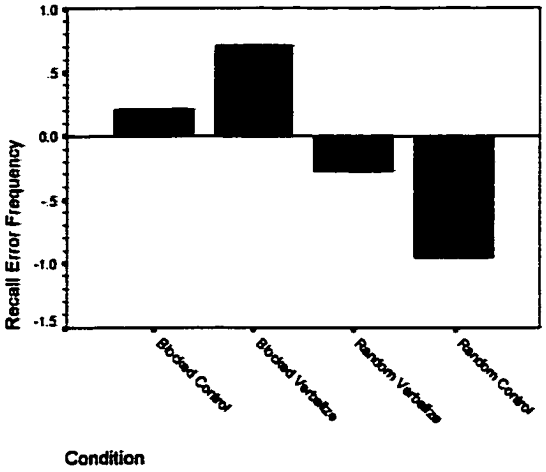


Figure 19. Differences in retention recall error frequency between 3A & 3B; positive values indicate fewer recall errors during 3B than 3A, negative values indicate more recall errors during 3B than 3A.

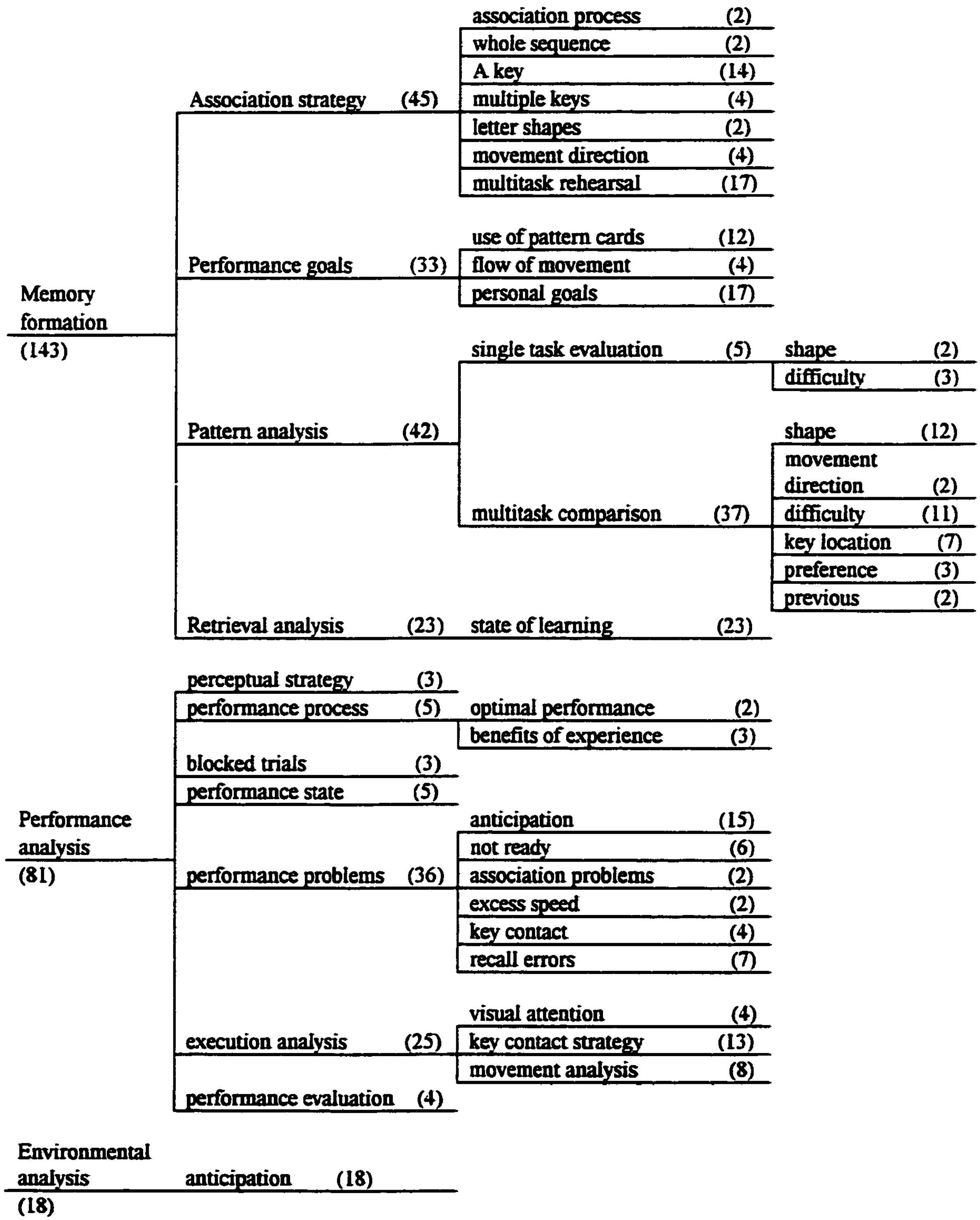


Figure 20. Results of qualitative analysis on verbalizations from the random verbalize group obtained during Experiment 3B.

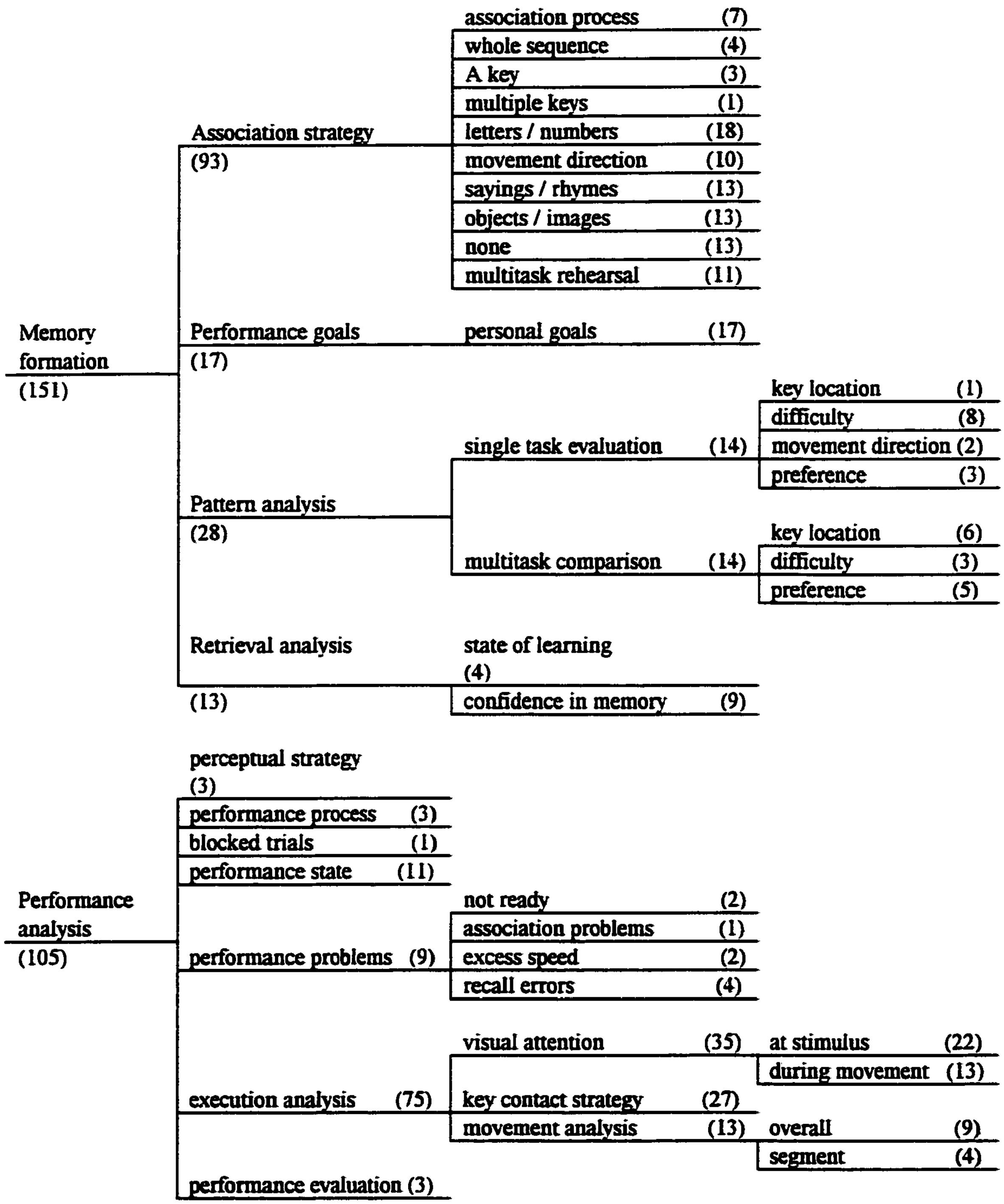


Figure 21. Results of qualitative analysis on verbalizations from the blocked verbalize group obtained during Experiment 3B.

APPENDIX A

Informed Consent form for Experiment 1

Informed consent

for participation in the research project of Darren Kruisselbrink entitled *Motor Skill Acquisition as Improved Problem Solving through Knowledge Specialization*

Purpose of the research project

I am a PhD student under the supervision of Dr. G. H. Van Gyn (local 8381) and have designed a study to look at the mental processes underlying motor skill acquisition, and how different practice conditions affect these processes and your ability to perform.

Requirements of participants

Here's what I'll be asking you to do. Your goal is to learn three spatial patterns on the push-button board. The push button board is a 2' x 2' board on which 16 buttons have been mounted to form a 4 x 4 grid. You will be asked to press a specific sequence of 4 of these buttons with one hand as quickly as you can following the release of a start button.

During the acquisition phase of the experiment, you will be asked to practice each pattern 18 times in the Human Motor Performance Laboratory (McKinnon 193). This will take roughly 25 minutes. Following these practice trials, you'll be tested on what you've learned, once following 10 minutes of working on a puzzle, and once a week later. You will be asked to perform each pattern another 3 times on these occasions, Each of these testing sessions will last about 10 minutes. All in all, during the course of this study your time commitment will total about an hour and a half.

Conditions of participation

I understand that whether or not I choose to participate in this research project, my academic standing in this course will not be negatively influenced what so ever.

I understand that my participation is completely voluntary and that I will receive a bonus mark for it. I also understand that I may withdraw from the study at any time, without explanation, but that withdrawal also nullifies receipt of the bonus mark. Once I withdraw I understand that the data collected from me to that point will be destroyed.

I understand that all data referring to me will remain completely confidential; that it will be kept in a locked cabinet in a locked office. I also understand that once the requirements of participation have been fulfilled, all references associating my name to data generated by me will be destroyed and that any data generated by me will be identifiable only in code form. I understand that the principle researcher may retain this anonymously coded data indefinitely.

Name ______________________________ Age ________

Signature______________________________ Date ________________

Researcher: Darren Kruisselbrink
Phone: 721-8392

APPENDIX B

Task Instructions And Procedures For Push Button Challenge: Experiment 1

The general description
Located near the computer screen you will find any number (1-7) of colored cards. Each card contains 16 circles. In 5 of those circles you'll see the letters S, A, B, C, and D. In Push Button Challenge, you'll start each trial by holding down the start key (marked S on the pattern card) with your dominant hand. After the second of two beeps, release the start key and hit keys A B C and D.

The details
To begin a trial, hold the start key down when the experimenter says "ready". While you're holding down the start key you'll hear two beeps, first a low pitched one and then a high pitched one. The low beep means get ready because the color name (ie BLUE, GREEN, RED...) of one of the pattern cards will flash up on the screen soon. At the high beep, a color name will be displayed on the computer screen. The objective of Push Button Challenge is to, following the high beep, release the start key (the key marked S on the pattern card), and hit the keys marked A B C D as fast as possible using your dominant hand.

BUT, release the start key ONLY when you're prepared to hit keys A - D at full speed. I don't want you to lift the start key as fast as possible following the high beep and then have to spend time thinking about what to do next.

Your score is the time it takes you to hit keys A - D, given in 1000th's of a second. Your score will be shown on the screen 5 seconds after you hit the D key of a pattern. The timer starts at the high pitched beep and ends once you've hit the D key.

As you move through a pattern, you may find that you'll miss hit or skip off a key. If this happens to you, DO NOT go back and correct your error since this will add time and artificially inflate your score. Just bail out by hitting any other key. When you hit a wrong key all that will happen is that an error message will come up on the screen, and you'll have to do that trial over again.

10 seconds after your feedback is given, I'll say ready again to start the next trial.

APPENDIX C

Informed Consent form for Experiment 2

Informed consent

for participation in the research project of Darren Kruisselbrink entitled *Motor Skill Acquisition as Improved Problem Solving through Knowledge Specialization*

Purpose of the research project

I am a PhD student under the supervision of Dr. G. H. Van Gyn (local 8381) and have designed a study to look at the mental processes underlying motor skill acquisition, and how different practice conditions affect these processes and your ability to perform.

Requirements of participants

Here's what I'll be asking you to do. Your goal is to learn three tasks: Push Button Challenge, Doom, and Ultris. Push Button Challenge involves pressing a specific sequence of 4 buttons on the push button board as quickly as you can following the second of two beeps. Doom involves negotiating your way through a three dimensional maze with a joystick. The objective is to finish the maze as quickly as possible while avoiding roving storm troopers. Ultris is a tetris like computer game that involves pressing various keys on a computer keyboard to rotate and place geometric shapes as they descend from the top of the screen.

During the acquisition phase of the experiment, you will be given the opportunity to learn each task by practicing it 18 times. This will take roughly one hour. Following these practice trials, you'll be tested on what you've learned, once following 10 minutes of working on a puzzle, and once a week later. You will be asked to perform each task another 3 times on these occasions, Each of these testing sessions will last about 10 minutes. All in all, during the course of this study, your time commitment will total about an hour and a half. All practice and testing sessions will be held in the Human Motor Performance Laboratory (McKinnon 193).

Conditions of participation

I understand that whether or not I choose to participate in this research project, my academic standing in this course will not be negatively influenced what so ever.

I understand that my participation is completely voluntary and that I will receive a bonus mark for it. I also understand that I may withdraw from the study at any time, without explanation, but that withdrawal also nullifies receipt of the bonus mark. Once I withdraw I understand that the data collected from me to that point will be destroyed.

I understand that all data referring to me will remain completely confidential; that it will be kept in a locked cabinet in a locked office. I also understand that once the requirements of participation have been fulfilled, all references associating my name to data generated by me will be destroyed and that any data generated by me will be identifiable only in code form. I understand that the principle researcher may retain this anonymously coded data indefinitely.

Name ______________________________ Age ________

Signature ___________________________ Date ________________

Researcher: Darren Kruisselbrink

Phone: 721-8392

APPENDIX D

Instructions for Push Button Challenge: Experiment 2

OVERVIEW
Located to the left of the computer screen you will find three colored cards (enter color names here). Each of these cards contains 16 circles. In 5 of those circles you'll see the letters S, A, B, C, and D. In Push Button Challenge, you'll start each trial by holding down the S key (start key) with your dominant hand. After the second of two beeps, release the start key and hit keys A B C and D as fast as you can.

DETAILS
To begin a trial, when the experimenter says "ready" hold down the start key. While you're holding down the start key you'll hear two beeps, the first one is low pitched, the second one is high pitched. At the high beep, name of a color will be displayed on the computer screen. The color name indicates the pattern to be executed. Your main objective is to release the start key and hit keys A B C and D as fast as you can with your dominant hand following the second beep.

5 seconds after you lift your hand off of the last key (D key), your score will be shown on the screen. Your score is the time it takes from the second beep until you lift off the D key, given in 1000^{th}'s of a second. Your main objective is to get this score as low as possible.

BUT, while your objective is to be as fast as possible, I want you to release the start key ONLY when you're prepared to hit keys A B C and D at full speed. I don't want you to be thinking about what to do next while you're moving. I want you to do all that thinking BEFORE you release the start key.

As you move through a pattern, you may find that you'll miss-hit or skip off a key. If this happens to you, DO NOT go back and correct your error since this will add time and artificially inflate your score. Just bail out by hitting any other key. When you hit a wrong key all that will happen is that an error message will come up on the screen, and you'll get to do that trial over again.

10 seconds after your feedback is given, I'll cue you up to start the next trial.

APPENDIX E

Instructions for Ultris Puzzle: Experiment 2

OVERVIEW
14 bricks will appear one at a time at the top of the screen. As soon as they appear they will begin to fall. The objective of this task is to rotate the bricks and get them into their proper position at the bottom of the screen as fast as you can.

DETAILS
Each of the 14 bricks needs to be placed in it's proper position to successfully complete three full lines of cubes. The proper position for each brick is indicated on a map that will be available to you for the first 6 trials.

Following each trial your score will appear on the screen. Your main objective is to get as high a score as you can. To do this you'll need to move and rotate the bricks quickly once they appear at the top of the screen (using the keys marked with arrows), and drop them into place with the key, because you get bonus points for dropping bricks. The closer the bricks are to the top of the screen when you drop them, the more bonus points you get. So increasing your speed is the essence of increasing your score.

A final reminder; the map will only be available to you for the first 6 trials.

APPENDIX F

Instructions for Maze of Doom: Experiment 2

OVERVIEW

Maze of Doom is a 3D game where every thing you see on screen is from your own perspective. You control your movement through the maze by the joystick.

THE OBJECTIVE

The primary objective of this task is to get to the end of the maze as fast as you can. Once you're at the end, you need to throw a switch to end the trial. To accomplish this, there are a few things you should know.

1. There are a number of doors you will need to go through as you move through the maze. To open any door, press the bottom button on the left side of the joystick (the one with the white and pink rings around it).

2. One of the doors is a red security door, which means you can only open it if you have the red key. You'll need to find and pick up the red key first. It's location is shown on the map, which will only be available to you for the first 6 trials.

3. Not all the doors you need to go through are immediately visible. To open one set of doors you'll need to throw a switch first (using the same button that opens doors). Once you throw the switch, you'll see the door. Again, it's location is shown on the map.

4. At one point in the maze it will appear that you have hit a dead end. You haven't, you're just in an elevator. Walk up to the wall and wait, and the elevator will lower you into another room.

5. Roaming around the maze are 12 armed storm troopers trying to hinder your progress. They move in relation to you, so if you stay in one spot for any length of time they will hone in on you and open fire. To protect yourself you have been provided with a pistol and some ammunition. To fire the pistol, press the trigger on the joystick (marked in red). But remember, firing your pistol wastes valuable time. So use it sparingly. As you go through the maze remember to emphasize speed rather than retribution!

Once again, your mission is to get to the end of the maze and throw the switch as fast as possible. We leave it up to you to decide on the strategy that will best serve this goal. Good luck!

APPENDIX G

Informed Consent form for Experiment 3

Informed consent

for participation in the research project of Darren Kruisselbrink entitled *Motor Skill Acquisition as Improved Problem Solving through Knowledge Specialization*

Purpose of the research project

I am a PhD student under the supervision of Dr. G. H. Van Gyn (local 8381) and have designed a study to look at the mental processes underlying motor skill acquisition, and how different practice conditions affect these processes and your ability to perform.

Requirements of participants

Here's what I'll be asking you to do. Your goal is to learn three spatial patterns on the push-button board. The push button board is a 2' x 2' board on which 16 buttons have been mounted to form a 4 x 4 grid. You will be asked to press a specific sequence of 4 of these buttons with one hand as quickly as you can following the release of a start button.

You will be asked to come in to the Human Motor Performance Laboratory (McKinnon 193) for two 45-50 minute practice sessions and a 5-10 minute follow up session, each 1 week apart. This means coming to the lab on the same day of the week for three consecutive weeks. During the practice sessions you will practice each of the three patterns 18 times. The set of three patterns will change from the first to the second practice session. Following practice, you'll be tested on how well you've learned each of the patterns, once following 10 minutes of putting together puzzle cubes, and once a week later. You will be given 3 test trials on each pattern on these occasions. All in all, during the course of this study your time commitment will total about two hours.

Conditions of participation

I understand that whether or not I choose to participate in this research project, my academic standing in this course will not be negatively influenced what so ever.

I understand that my participation is completely voluntary and that I will receive a bonus mark for it. I also understand that I may withdraw from the study at any time, without explanation, but that withdrawal also nullifies receipt of the bonus mark. Once I withdraw I understand that the data collected from me to that point will be destroyed.

I understand that all data referring to me will remain completely confidential; that it will be kept in a locked cabinet in a locked office. I also understand that once the requirements of participation have been fulfilled, all references associating my name to data generated by me will be destroyed and that any data generated by me will be identifiable only in code form. I understand that the principle researcher may retain this anonymously coded data indefinitely.

Name ______________________________ Age ______________

Signature____________________________ Date ______________

Researcher: Darren Kruisselbrink
Phone: 721-8392

APPENDIX H

Instructions for Push Button Challenge: Experiment 2

Pre-test Instructions for Push Button Challenge

OVERVIEW

Located to the left of the computer screen you will find a green colored card. It is a pattern card. The circles on the card correspond to the buttons on the board. In 5 of those circles you'll see the letters S, A, B, C, and D. In Push Button Challenge, you'll start each trial by holding down the S key (start key) with your dominant hand. After the second of two beeps, release the start key and hit keys A B C and D as fast as you can. Your main objective in this game is to learn how to increase your speed.

DETAILS

The experimenter will give you a cue to start each round. When he / she says "ready" that's your cue to hold down the start key. While you're holding down the start key you'll hear two beeps, the first one is low pitched, the second one is high pitched. As soon as you hear the second beep, release the start key as quickly as you can and fly through the pattern.

5 seconds after you lift your hand off of the last key (D key), your score will be shown on the screen. Your score is the time that passes from the second beep until you lift off the D key. The timer starts at the second beep and stops once you release the D key. Your time will be given in 1000^{th}'s of a second. The challenge in Push Button Challenge is to learn how to improve this score, since skill in this game is associated with movement speed.

Practice Instructions for Push Button Challenge

This is a learning study. Learning involves improving and being able to maintain this improvement over time. Therefore, the main objective of Push Button Challenge is to learn the three movement patterns you see to your left, and learn how to increase your speed on each pattern so that by the end of practice you're consistently hitting your top speed. To do this, you'll get 18 rounds of practice on each pattern. The 18 practice rounds are your opportunity to try different things to find out what will help you increase your speed and what won't.

After every 18 rounds of practice we'll take about a 1 minute break, and after you've practiced each pattern 18 times we'll take a 10 minute break. During the 10 minute break you get to put together puzzle cubes. After the break we'll test you on how well you've learned the patterns with 3 test rounds on each pattern. Being tested means that you won't have the cards to look at, you won't get to see your feedback, and the patterns will come up in a random order so you'll have to base your performance on what you've learned, there will be no external aids to help you.

NOW, while movement speed is your objective, I want you to release the start key ONLY when you're prepared to hit keys A B C and D at full speed. I don't want you to be thinking about what button to hit next after you've released the start button. I want you to do all your thinking and movement planning BEFORE you release the start key.

Again, as you move through a pattern, if you don't think you hit a button, or find yourself in the neighborhood of a button but miss it, DO NOT go back and correct yourself since this will add time to your score and won't accurately reflect how quickly you're moving. Just bail out by hitting any other button. When you hit a button out of sequence all that will happen is that an error message will come up on the screen, and you'll get to re-do that round.

Approximately 10 seconds after you see your score, I'll cue you up to start the next trial.