Efficacy and Generalizability of a Memory-Training Paradigm: Application to a Population of Brain-Injured Individuals

by

Rema Andrea Lillie
B.A., Washington University in St. Louis, 1998

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

MASTER OF SCIENCE

in the Department of Psychology

We accept this thesis as conforming
to the required standard

University of Victoria

Supervisor: Dr. Catherine A. Mateer

ABSTRACT

The current project applied a new theoretically-driven training paradigm shown to be effective at improving memory performance in a group of aging adults (Jennings & Jacoby, 2003) to a group of 10 individuals with Acquired Brain Injury (ABI). Training effects were assessed on the paradigm itself and other measures of memory and attention. Performance on cognitive measures was compared to a group of 9 healthy, young adults to control for practice effects. Results showed a replication of previous findings in terms of both frequency and magnitude of improvement in this new population. Some proximal effects of training were found on a similarly-structured task (false fame) but no distal effects of training were seen on other cognitive measures. Limitations of the current project included small sample sizes. Recommendations are provided for future research. Implications for a dual-process model of memory and clinical practice are discussed.

Examiners:

Table of Contents

List of Tables

List of Figures

Acknowledgments

This project never would have been possible without the support of a variety of individuals, some of whom are acknowledged below.

Firstly, I would like to thank my supervisor, Katy Mateer. Without her constant support and guidance, I never would have been able to take on such a large project. You are a true mentor. Thank you to Steve Lindsay who provided valuable insight into the cognitive science literature and guidance in the initial planning stages. Thank you also to my committee member, Roger Graves, who was patient enough to help me tackle some of the methodological issues associated with such an ambitious project.

I thank Larry Jacoby for providing me with a working version of the training program as well as other stimuli used in my project. For providing practical support, I thank my research assistants, especially Tomoko Arimura, without whose time and assistance I would not have been able to make it through this project. Thank you to Kim Kerns for helping me make sense of my data. Not very many people would have had the patience for a three-hour stats session. Thank you to Cindy Bukach for helping me organize counterbalancing issues and to Vincenza Gruppuso for loaning me her stimuli in progress. Special thanks to Jen Michel for last minute help on power analyses.

A big thank you to all my friends and family that have listened to the ups and downs of this project over the years. I never would have made it without you all. I would especially like to thank my friend, Leora Dahl, for always being there for me in whatever way I needed. I could never fit all your contributions onto one page. And lastly, I would like to thank Jeff Prochnow for his ever-constant support and unwavering belief in me. Thanks for being there for me every step of the way.

Dedications

To my mom,

Everything I am is because of you.

I miss you.

Introduction

Memory dysfunction is one of the most common and persistent impairments seen in Acquired Brain Injury (ABI; Butters, Soety, & Glisky, 1998; Evans, Wilson, Needham, & Brentnall, 2003; Lincoln & Brooks, 1992). In a recent survey of community dwelling brain-injured individuals, "Problems remembering new things" was most frequently ranked at the highest levels of severity as compared to other cognitive abilities (Ricker et al., 2002). Sixty-five percent of the sample (46 of 71 respondents) indicated "moderate" to "a lot" of problems in memory functioning.

Memory impairments have been shown to relate to reintegration into the community and return to work (Brooks, McKinlay, Symington, Beattie, & Campsie, 1987; Schacter, Glisky, & McGlynn, 1990) and are of considerable importance to patients, their families, and health-care providers (Evans, 1992; Hellawell, Taylor, & Pentland, 1999; Kinsella et al., 1996; Sohlberg & Mateer, 1989). Schacter et al. (1990) point out that memory impairments "frequently have devastating effects on patients' ability to function in the real world." (p.231)

In addition, as memory dysfunction can mediate adherence to treatment regimens and outcome, it is of critical concern for clinical teams working on the assessment and rehabilitation of brain injury. As Mateer, Sohlberg, and Crinean (1987) noted, prospective memory (e.g., the ability to remember to keep an appointment, to return a phone call, or to take medications) is crucial to optimal outcome. Recognizing, then, that memory impairments are common following ABI and can impact functional outcome, are there intervention strategies available to treat the dysfunction?

Cognitive rehabilitation of memory dysfunction

Various attempts at the cognitive rehabilitation of memory functions have met with mixed success. In the mid-1980's, rehabilitation attempts focused primarily on the application of practice and drills. The repetitive nature of these tasks was intended to stimulate damaged memory and neural processes so as to restore their function (Glisky, 2002). This attempt at remediation has been compared to conceptualizing memory as a muscle that requires exercise. In general, this approach has not been shown to be effective (Tate, 1997). Specifically, although practice can lead to the acquisition of at least some new information (Glisky & Glisky, 2002), general memory benefits are not observed. Modern applications of repeated drills tend to focus on the acquisition of relevant information to the patient at hand, as there appears to be no general benefit associated with practice of meaningless material (Glisky & Schacter, 1989; as cited in Glisky & Glisky, 2002).

In contrast, rehabilitation techniques that focus on the training of internal memory strategies have met with some success. The use of mnemonics and visual imagery, such as that used in the pegword or keyword methods and the method of loci, can improve recall performance both in healthy individuals and those showing memory problems (Baine, 1989; Bellezza & Reddy, 1978; Gruneberg, 1992; Roediger, 1980). Although commonly used to treat learning disabilities (e.g., Scruggs & Mastropieri, 1984) and improve memory in a variety of populations including children and the elderly (e.g., Wood & Pratt, 1987), these techniques have not proven as effective in the ABI population. Generally, most classic mnemonic strategies are useful for remembering unrelated lists of words and, as such, their practical value in influencing the everyday lives of memory impaired people is limited (Tate, 1997). Surveys have shown that most individuals would instead write down lists of words to be remembered (Harris, 1992). Visual imagery techniques, on the other hand, often used to help form associations between unrelated words or objects, have been effectively employed to help patients learn people's names (Wilson, 1987), a common problem in memory-impaired populations. In general, individuals with ABI may need assistance generating imagery and benefit from concrete drawings for rehearsal.

Other strategies that have proven successful for improving performance in ABI involve principles of organization and association, generally used at encoding. This class of intervention would include teaching patients to cluster information into natural groupings, count the number of items on a list, or link a series of simple steps to complete more complex tasks. Individuals can be taught to form categories when learning, or associate new information to already-learned information. Examples include chunking a telephone number into smaller groups (e.g., recalling 555-21-85 instead of as an individual series of numbers) or forming an analogy between new information and previously learned information (e.g., that person has the same name as someone who went to my high school). The technique of "chaining," as described by Glisky and Glisky (2002) involves learning simple steps in a larger sequence and then linking each new step until the sequence as a whole is acquired. Using this technique, Glisky and Schacter (1989a; as cited in Glisky & Glisky, 2002) taught a patient a complex data entry job and Wilson (1996b; as cited in Glisky & Glisky, 2002) taught patients how to transfer from wheelchairs to regular chairs. Another commonly used mnemonic for individuals with memory problems uses the acronym 'PQRST' for organizing and recalling prose or narrative information, such as a newspaper article or book (Glisky & Glisky, 2002). The acronym stands for "preview, question, read, state, and test" and is intended to provide a deeper level of processing and organization at encoding to enhance recall.

In general, the teaching of mnemonic strategies to those showing memory impairments has been conceptualized as a treatment focused on optimizing residual function (Glisky & Glisky, 2002). As such, it is most appropriate for individuals showing only mild or moderate impairments. Similar to other forms of memory treatment it faces certain limitations; generalizability is often an issue as strategies do not tend to be employed spontaneously outside the laboratory or clinical setting. Instead, generalization to everyday activities often has to be rehearsed directly.

External aids are another general category of compensatory devices. Aids are often used to reduce disability associated with memory dysfunction by circumventing memory. They are especially useful for performing tasks that rely on prospective memory, or memory for future intentions (Harris, 1992). Written reminders, including displaying to-be-remembered information in prominent locations, are commonly used both by healthy adults and those showing memory impairment. The difference for individuals with ABI is that these reminders may serve to prompt memory for information most of us take for granted including, for example, how to get to the dining room in a rehabilitation facility or the names of grandchildren. Additionally, research has shown that use of these types of reminders, including signposts in a hospital, requires actively teaching patients to seek out and use them (Moffat, 1992).

Advances in technology over the past decade have revolutionized the accessibility of pagers, timers, and electronic organizers. Unlike past attempts at complex computer-based reminding systems that required bulky equipment, current technology allows for small, portable devices that blend in with those being used by the general population. Computer technology provides assistive devices that can not only sound a tone when a task is required, but can display a message indicating what task needs to be completed (e.g., Neuropage). These types of tools can be invaluable for memory-impaired individuals who are trying to juggle a daily schedule or return to work. Such devices (e.g., PDA's) can also store basic information, such as insurance policy numbers or driver's license information that can be easily accessed when away from home. Problems with these types of aids include the original expense of purchasing such equipment, the fact that they can often be misplaced, and the fact that learning to use them may pose challenges to memory-impaired individuals.

More basic assistive devices include the use of appointment books, calendars, and diaries. Although less sophisticated, these types of aids can provide valuable information for everyday living. They not only can be used to record actions to be performed in the future, but also can provide important cues about past actions. For example, they can provide a cue that a phone call needs to be made or provide a reminder that the task was completed. As with high-tech devices, limitations include misplacing one's agenda. Additionally, use of such a system can be difficult for individuals with more moderately- or severely-impaired memory and generally needs to be taught directly.

Sohlberg and Mateer (1989) provided a three-stage behavioural approach for training the use of compensatory memory books. They provided a model of a comprehensive notebook system that provides information about past and future actions as well as personal information and other useful material. Notebooks can be individually-tailored to the needs of the client but may include sections such as 'orientation' (e.g., phone number, medications, doctor's names, etc.), a 'memory log' to record events of the day (e.g., phone conversations, tasks completed), a calendar and 'things to do' section where the patient can record and update tasks to be completed, transportation information (e.g., bus schedules, phone numbers for services), a 'relevant people' section where the patient can record names, contact information, and identifying information about people he or she commonly encounters, and a work-related information section including names of coworkers, locations, and instructions on how to complete work-related tasks. Guidelines for enhancing generalizability include 1) actively planning for everyday needs from the onset of treatment, 2) identifying reinforcements in the natural environment, 3) providing practice stimuli that are common to both the training environment and the real world, 4) using sufficient examples in training, and 5) measuring generalization directly (Sohlberg & Raskin, 1996).

One last general category of treatments focuses specifically on the acquisition of domain-specific learning. This category of rehabilitative strategies rests on empirical findings showing that although amnesic patients have great difficulty acquiring new memories that can be explicitly retrieved, they are generally able to retain some record of prior experience that can be expressed implicitly (e.g., Warrington & Weiskrantz, 1974).

These two general ways of retrieving memory for prior experiences have been labelled "implicit" and "explicit" memory (e.g., Hintzman, 1990; Schacter, 1987). Implicit memory describes a facilitation or change in test performance that is attributable to information or skills acquired during a prior study episode, even though subjects are not required to, and may even be unable to, recollect the study episode. Generally, patients with organic amnesic syndromes show substantial, and occasionally, normal levels of performance on implicit memory tasks, despite the absence of recollection for having previously performed the tasks (Schacter, Chiu, & Ochsner, 1993). This phenomenon has been most often studied through repetition or direct priming in which exposure to a word or object on a study list facilitates its subsequent identification when degraded perceptual cues are provided (e.g., Tulving & Schacter, 1990; as cited in Schacter, Chiu, & Ochsner, 1993).

Findings of intact implicit memory in amnesics have led to the development of a variety of techniques designed to help memory-impaired individuals learn key bits of critical information. These techniques are applicable to even the most severely-impaired individuals in whom improvement of general memory functioning is unlikely to occur (Butters, Soety, & Glisky, 1998) and include the method of vanishing cues, errorless learning principles, and spaced retrieval.

The method of vanishing cues involves practice at remembering something when provided with partial cues. Initially, as much cue information is provided as is needed for correct responding. Cues are then gradually withdrawn across learning trials. Using this technique, researchers have been able to teach memory-impaired individuals specific information relevant to their everyday lives including computer operations, word processing, and vocation tasks such as computer data entry (Glisky & Glisky, 2002). Possible applications of this technique are by no means limited to computer-based administration and have been proposed to treat naming difficulties, dysgraphia, and apraxia (Moffat, 1992). Fundamental requirements include the possible construction of a graded series of prompts that can be introduced systematically.

Some have argued that the success of the vanishing cues method is at least in part attributable to the fact that cues constrain responses and thus prevent errors (Baddeley & Wilson, 1994; as cited in Glisky & Glisky, 2002). Wilson and colleagues in a series of studies demonstrated that errorless learning, as compared to trial-and-error learning, provides benefits to a range of patients with memory impairment. In errorless learning paradigms, individuals are provided with sufficient information so as to avoid making mistakes and are discouraged from making random guesses. Using this format, errorless learning has been used successfully to teach names, the use of an electronic memory aid, items of orientation and general knowledge, the use of a memory notebook, and word-processing skills (Glisky & Glisky, 2002).

A third technique that takes advantage of intact implicit memory is that of spaced-retrieval. Spaced-retrieval is often used in combination with other rehabilitative techniques noted above. When learning new information, it has been found that distributed practice over multiple sessions can improve both the rate of learning and subsequent retention (Baddeley, 1992). Teaching of rehabilitative strategies, then, is best administered in the context of repeated rehearsal over multiple, relatively brief sessions. Expanding this technique one step further, a spaced-retrieval approach focuses on having individuals learn new information through repetition across increasing delays (Schacter, Rich, & Stamp, 1985). A common schedule is one that doubles the time interval for each subsequent retrieval (e.g., 0, 2, 4, 8, 16, 30 minutes, followed by 1, 2, 4 hours; Moffat, 1992). This series can be modified at the individual level to respond to the needs of each patient. Using this technique, researchers have successfully taught patients critical pieces of information including name-face associations, locations of objects, and items of orientation (Glisky & Glisky, 2002).

A commonality of each of these approaches is a focus on teaching domain-specific knowledge by creating habitual, overlearned responses. By focusing on intact, implicit memory, researchers and clinicians are able to strengthen access to information that may or may not be consciously recalled. Rehabilitative efforts that focus on improving processes involved in the conscious recall of information have in general met with limited success. Instead, impairments in the conscious recollection of material are often treated with the application of compensatory devices – either in the form of teaching strategies to assist the encoding and retrieval of information or through the use of external aids. These techniques are limited by the amount of time required for teaching and a lack of spontaneous generalization outside the rehabilitation setting.

Recent advances in neuroscience: Impact on rehabilitation

Recent advances in neuroscience have shown that the human adult brain is capable of a greater degree of plasticity than once thought. These findings have led to a proliferation of studies meant to capitalize on brain plasticity through rehabilitative efforts. Perhaps the most well-recognized of these have been in the domain of Constraint-Induced Movement Therapy, or CIMT.

CIMT is a popular new technique for treating motor deficits that has come to the forefront of scientific interest and led to a number of case reports and clinical trials. CIMT grew directly out of the animal literature and is based on the principle of "learned non-use" from deafferentiation experiments with monkeys (Taub & Uswatte, 2003). Essentially, researchers found that when afferent motor pathways were lesioned in monkeys, the animals would quickly learn to suppress movement of the affected limb. However, by constricting movement of the unaffected limb, typically through use of a splint or cast, abilities in the affected limb increased dramatically. In the 1990's, this technique was modified for use in humans with much success (Taub et al., 1994). Imaging studies suggest that use of the affected limb leads to cortical reorganization (Schaechter et al., 2002; Wittenberg et al., 2003).

In the domain of cognitive processing, studies of physiological change following cognitive rehabilitation are rare. Limited research is available to suggest that changes in regional cerebral blood flow (rCBF), functional magnetic resonance imaging (fMRI), and SPECT are detectable following intervention (Laatsch, Thulborn, Krisky, Shobat, & Sweeney, 2004; Lindgren, Hagstadius, Abjornsson, & Orbaek, 1997; Penades et al., 2000). These findings support the possibility of detectable changes in processing following cognitive rehabilitation.

The success of CIMT suggests that forced, focal use of an impaired system can lead to improved functioning. Findings of change on physiological measures of brain functioning following cognitive rehabilitation suggest that such interventions can impact brain processing at the physiological level. Put together, these findings raise the question of whether procedures similar to CIMT could be applied to cognition to impact change at the level of memory processing. Interventions of this type would expand on previous research applying repetition and drills as they would focus on a specific aspect of memory performance. A major barrier to exploring interventions of this type is determining how to limit alternate cognitive processing mechanisms to provide forced, focal use of an impaired system.

Jacoby's dual-process model of memory

Historically, implicit and explicit memory have been assessed by using different tasks for each type of memory (e.g., Schacter, 1987). Implicit memory tasks (e.g., stem completion, priming experiments), generally indirect in nature, were thought to measure a fundamentally different type of memory as compared to direct, explicit tasks (e.g., list recall, narrative prose, etc.).

One of the problems with directly comparing results from such investigations is an assumption that each task is a "pure" measure of a specific type of memory. As Jacoby (1991) argued, indirect tests are not a pure measure of implicit memory as they can be influenced by explicit recall that was undetected by the experimenter. For example, one cannot be certain that subjects in an "unconscious perception" condition were unaware of supposedly subliminal stimuli. Similarly, explicit memory tasks may be influenced by automatic processing (e.g., "guessing" on a multiple choice test may be influenced by unconscious memory). Even in studies of amnesics, Jacoby stated, one cannot assume that no explicit memory exists; amnesic subjects may at times intentionally use memory to aid their performance on indirect tests of memory. By exploring ways to assess different types of memory within the same task, namely by tracking recognition performance under differing conditions, researchers can focus on the relative contribution of each memory process while moving away from assumptions regarding the "purity" of the measure.

Jacoby's process dissociation framework presents a model for distinguishing automatic versus intentional uses of memory within the same task. This framework is based on a dual-process approach that conceptualizes memory as consisting of two independent processing streams: one labelled "familiarity" and related to automatic forms of processing, and the other labelled "recollection" and reflective of intentional forms of processing (Jacoby, 1991; Mandler, 1980). To distinguish between the relative influence of each type of memory, Jacoby imposed two distinct demands on recognition performance that are described below.

As both types of memory are assessed using the same task, the general procedures are the same. Typically, a study list (words, names) or task involving working with words (e.g., anagrams) is presented followed by a recognition test. During the test phase, some items are made to seem familiar through various manipulations including presentation during the study phase (usually under conditions of limited attention) or repetition during the test phase. In an inclusion condition, automatic influences of memory lead to the same memory judgement as recollection, thus creating a facilitation effect. Specifically, subjects are instructed that those words that seem familiar, typically due to prior presentation of some sort, were definitely on the study list. In contrast, a second condition is administered using the same general task format where feelings of familiarity are placed in opposition to conscious recollection and are used as a basis to exclude items. In these types of manipulations, subjects are told that if a word seems familiar, then it was definitely not on the study list. One can then compare the probability of calling a word "old" under each of these conditions (inclusion vs. exclusion) to determine the relative contribution of each type of memory processing (for specifics, see Jacoby, 1991).

This framework for dissociating feelings of familiarity from conscious recollection has been investigated in a variety of contexts, with different populations (young adults, aged adults, individuals with ABI), and has been well-supported (e.g., Hay, Moscovitch, & Levine, 2002; Hertel & Milan, 1994; Jacoby, Toth, & Yonelinas, 1993). Critiques of the procedure (e.g., Buchner, Erdfelder, & Vaterrodt-Plunnecke, 1995; Curran & Hintzman, 1995; Russo, Cullis, & Parkin, 1998) have generally focused on assumptions underlying the mathematics involved in determining the relative contribution of each processing stream and are not focused on the theoretical framework for understanding memory processing (for exceptions, see Dodson & Johnson, 1996; Gruppuso, Lindsay, & Kelley, 1997).

Jennings and Jacoby memory training paradigm

Building on the process dissociation framework, a training paradigm created by Jennings and Jacoby (2003) uses an incremented-difficulty approach to improve conscious recollection. The training paradigm consists of an exclusion procedure that places memory processes of recollection and familiarity in opposition. Participants are first presented with a list of words that they are told to remember. They are then presented with a test list and are to indicate if the presented words were part of the original study set. The trick is that some of the new words in the test list are repeated. Participants are correctly informed that if a word in the list is repeated, then it is *not* one of the original study items. The repetition makes these items more familiar and participants are, therefore, more likely to mistakenly identify them as having been on the original study list. To complete the task successfully, individuals must counteract feelings of familiarity by consciously recollecting their prior presentation on the test list. This task is quite difficult for individuals with intact processing of familiarity but impaired recollection, as is the case for many individuals with memory impairments as discussed above.

The training program uses an incremented-difficulty approach, beginning with 1 to 2 words (lags) between repetition and increasing the lag once the participant has met the criteria of performance at the level seen in healthy, young adults. Since the process dissociation framework supports the idea that conscious recollection of previous presentation is required to accurately complete the task, the improvement of performance over time is interpreted as an improvement in the process of conscious recollection. An important distinction is made, then, between this training paradigm and some of the others previously described in that the focus here is on training a *process* of recollection and not on the acquisition of specific pieces of critical information.

In the context of research on CIMT, this paradigm has the potential to distinguish itself from previous attempts at training memory processes through drills or repetition as it attempts to constrain intact processing and forces focal use of an impaired memory processing system. Unlike previous rehabilitative efforts that have capitalized on intact implicit memory to teach amnesics key pieces of critical information (e.g., errorless learning, method of vanishing cues), this paradigm works to have participants suppress intact feelings of familiarity. By more specifically focusing intervention at the point of dysfunction, more targeted practice is provided. Importantly, the goal here is not to improve a general process such as list-learning or recall performance as has been previously attempted, but is instead focused on a mechanism that has been linked to everyday memory errors commonly seen in the aging (Hay & Jacoby, 1999; Jacoby, Jennings, & Hay, 1996) and ABI populations (Arciniegas et al., 1999; Ste-Marie, Jennings, & Finlayson, 1996) such as repeatedly telling the same story. This type of theoretically-driven research is critical to the field of rehabilitation. As Harris (1992, p. 59) pointed out, "Improvements to techniques seemed more likely to come from a better understanding of memory processes, the object of most memory research." By using emerging theories in the cognitive sciences to develop new treatment strategies for rehabilitation, we can capitalize on the strengths of each field to advance intervention techniques.

Current study

Using the Jennings and Jacoby (2003) procedure, aging adults have been shown to improve from meeting criteria at lags of only 1 or 2 to meeting criteria at lags of 28 after seven days of training. While the training paradigm has demonstrated some success in a sample of normally aging individuals, it is unknown whether the paradigm is effective in individuals with other sorts of memory difficulties and/or with neurological conditions. In the current study, I propose to apply this training paradigm to a group of individuals with acquired brain injury (ABI) to assess generalizability to a new population. In addition, though this paradigm has proven effective at training recollection on this specific task, the generalizability of its effects to other memory measures has not been reported. As part of the current study, a variety of tasks were administered both pre- and post-training to assess proximal and distal effects of training, specifically in the areas of memory and attention, to further define the nature of any improvements.

Due to the similarity in the patterns of impairment seen in the aging population and brain-injured individuals, namely intact familiarity in the context of impaired recollection memory (e.g., Ste-Marie, Jennings, & Finlayson, 1996), it is hypothesized that the training paradigm will show similar effects in this new population. Based on previous research (Jennings & Jacoby, 2003), an improvement in performance, defined as an increase in the lag interval at which participants can reach criteria from Day 1 to Day 7, is expected in the majority of trained individuals. Whether training effects will generalize to other measures of memory is unknown, but it is hypothesized that proximal effects will be found on similarly-structured tasks. Distal effects of training on more clinical measures of memory are not anticipated but will be assessed as a means of testing the limits of any improvements.

It has been previously suggested that the gains seen on this training task are due to improvements in the process of recollection at test. However, an alternate hypothesis is that these gains reflect changes in the allocation of attentional resources over time. This idea is supported by previous self-reports that participants were "working to concentrate more on the task overall" (Jennings & Jacoby, 2003, p. 435). The hypothesis that improvements on this task are linked to attention will be assessed by the pre- and post-training administration of various measures of attention and working memory.

In summary, the aims of this study were threefold:

(1) To assess whether training effects found in a previous study (Jennings & Jacoby, 2003) generalize to a new population with a similar pattern of impairment;

(2) To assess whether other measures of memory are affected by training;

(3) To assess whether patterns of change across measures of memory and attention lend insight into mechanisms underlying training or suggest directions for future research.

The primary goal of this project was to assess generalizability of the training paradigm. Items 2 and 3 above were considered exploratory in nature.

Methods

Participants

Eleven adults (6 men, 5 women) between the ages of 18 and 59 with a history of acquired brain injury (ABI) were recruited from the community for participation in this study. Flyers were used to advertise in local rehabilitation centres and community resource centres. Clinicians in the community working with brain-injured individuals were provided with information on the project. Due to the large time commitment required for study completion, brain-injured individuals were compensated $50 for participation. Reimbursement was also given for parking fees and bus fare incurred through study participation.

Ten university students (2 men, 8 women) were recruited from undergraduate psychology courses to serve as controls and twenty-four undergraduate students participated in pilot testing of materials. Undergraduate participants received course credit for participation. Approval was obtained from the Human Research Ethics Committee at the University of Victoria before beginning study procedures. Informed consent was obtained from all participants.

Inclusion and exclusion criteria

Experimental subjects. For participation in this study, individuals must have experienced an acquired brain injury (ABI) resulting in subjective complaints of memory dysfunction. History of ABI was based on personal report and was verified by medical records whenever possible (5 of 11 participants). In our sample, 9 of 11 individuals had experienced a traumatic brain injury (TBI) acquired through motor vehicle accidents. Severity ratings of TBI ranged from mild (3 cases) to severe (6 cases). The two remaining individuals in our sample had experienced brain aneurysms affecting the frontal (1 case) or temporal (1 case) lobes and requiring surgical intervention.

Medical history questionnaires were completed by each participant that included specific items related to the nature of the brain injury. Participants were excluded if they sustained their injury prior to age 15, had a pre-existing history of any neurologic disorders (e.g., seizure disorder, brain tumour), or were within one-year post-injury.

Before commencing study procedures, participants completed a screening to assess intellectual capacity and reading ability. An estimate of Full Scale IQ was calculated by administering the Vocabulary and Block Design subtests of the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III; Weschler, 1997). These two subtests have excellent reliability, correlate highly with the Full Scale IQ over a wide age range, and are thought by many experts to be good measures of a general intelligence factor, *g* (Sattler, 2001). Age-corrected scaled scores on each of these measures were summed and converted to an estimated Full Scale IQ using a table provided by Sattler (2001). Using an adaptation of a procedure from Tellegen and Briggs (1967; as cited in Sattler, 2001), this composite shows satisfactory reliability and validity ($r = .86$, $r_{xx} = .91$, across age groups). Participants with estimated Full Scale IQ scores more than two standard deviations below the mean (i.e., less than 70) were excluded from the study.

Reading ability was assessed using the reading subtest of the Wide Range Achievement Test, version 3 (WRAT3; Wilkinson, 1993), one of the most frequently used measures of academic achievement that has proven useful as a quick, but gross, screening device (Spreen & Strauss, 1998). Performance on this measure was compared to a large normative sample ($N = 4{,}433$) stratified according to age, regional residence (within the United States), gender, ethnicity, and socioeconomic level (Wilkinson, 1993). Since the training paradigm relies on intact reading ability, those individuals whose performance was below one and a half standard deviations from the mean of the age-matched normative sample were excluded from the study. In addition, due to the strong language component of this study, we only included individuals who spoke English as their native language.

For descriptive purposes, participants were screened for memory dysfunction using the California Verbal Learning Test (CVLT; Delis et al., 1987), a well-validated and widely used clinical measure. Scales of interest included the total number of words correctly identified over five learning trials, learning on Trial 1, and recall following a delay. Performance was compared to a large normative sample (N=1,087), matched to a recent U.S. Census in terms of demographic variables (education, race/ethnicity, and geographic region), and was considered impaired if consistently poor (at least 1.5 SD below the mean) across measures. Since the type of memory assessed in this measure (i.e., structured, categorical learning across several trials) is not directly linked to the recollection memory used in training, a lack of impairment was not used as a grounds for exclusion.

To capture the broadest age range possible while taking into account developmental issues, all individuals between the ages of 18 and 59 years who met the above criteria were included in this study. It was anticipated that this age range would capture mature memory processes (Cycowicz, 2000; Kee, 1994) prior to any effects of age-related cognitive decline (Klein, Houx, & Jolles, 1996; Small, 2001).

For our purposes, intact hearing and vision were defined as the ability to participate in standard testing procedures at the screening visit.

In summary, experimental participants met the following criteria:

i. Had a history of ABI;
ii. Resulting in subjective complaints of memory dysfunction;
iii. Aged 18 to 59 years old;
iv. No significant history of any neurologic disorders prior to injury (e.g., seizure disorder, brain tumour);
v. At least 15 years old at the time of injury;
vi. At least one year since injury;
vii. An estimated Full Scale IQ above 70;
viii. No significant hearing or vision problems as evidenced by performance at the screening session;
ix. English as native language;

x. Able to read as evidenced by performance on a subtest of the WRAT3 within one and a half standard deviations from the mean of age-matched neurologically-intact individuals.

Control subjects. Control participants completed all screening measures (Block Design, Vocabulary, WRAT3, CVLT) and provided basic information on age, native language, and neurologic history. Individuals with a history of any neurologic condition (seizure disorder, brain tumour, etc.) were excluded. Two individuals reported closed head-injuries in childhood (1 individual fell from a swing requiring stitches, 1 individual experienced a concussion in a ski accident). These injuries, occurring prior to age 10, resulted in no noticeable long-term cognitive effects and so were not used as a basis for exclusion from study participation.

Procedure

Experimental subjects. Study participants completed a screening visit as outlined above. To control for practice effects, a multiple-baseline design was used. Once study eligibility was ascertained, a preliminary baseline evaluation was completed (Baseline). During this visit, participants completed the full cognitive test battery as described below. Approximately one week later (5-8 days), participants repeated the cognitive test battery for a second baseline evaluation (Pretraining). Training commenced one to two days following the second baseline visit, except in one case where training began 5 days after the second baseline evaluation due to scheduling conflicts. On all study days, breaks were taken as necessary to reduce possible effects of boredom, fatigue, and frustration. When a short break was insufficient for easing discomfort, data gatherers responded with active listening. In no circumstances was discomfort judged to be sufficient to warrant discontinuation of testing or referral to community resources.

Training sessions were administered following the protocol established by Jennings and Jacoby (2003). Four training sessions were administered per day for a total of seven training days. Sessions were self-paced and lasted, on average, between 7 and 15 minutes so that most training days included approximately one-half hour to one hour of training. A weekend break occurred following day 3 (7 participants) or day 4 (4 participants). Four participants experienced breaks between training days due to illness (3 participants) and a missed appointment that could not be rescheduled (1 participant).

Post-training administration of the full cognitive test battery (Posttraining) typically occurred one day after training concluded (day 8). In two cases, post-training testing was administered the same day as the final training sessions (day 7). In these situations, a break was provided between completion of training and commencement of cognitive testing. On the final study visit, following post-training testing, participants answered questions regarding techniques or strategies used to complete the task. Debriefing occurred once all study procedures were completed.

See Table 1 for a summary of study visits.

Control subjects. Control participants completed a screening visit as outlined above. To mirror performance of brain injured individuals, control participants completed the full cognitive battery, described below, once a week for three consecutive weeks. Although testing of brain injured individuals as controls would have been ideal, this procedure provided an estimate of practice effects on each of our cognitive measures as we would see in young, healthy controls.

Jacoby's memory training paradigm

Materials

Materials were provided by Jennings and Jacoby and were thus identical to a previous application of the training program (Jennings & Jacoby, 2003). Specifically, 2100 concrete nouns from The Toronto Word Pool (Friendly et al., 1982; as cited in Jennings & Jacoby, 2003) and the Thorndike and Lorge word norms (1944; as cited in Jennings & Jacoby, 2003) were divided into two sets of lists to be used as study and filler items. Sixteen hundred and eighty of these words were divided into 56 lists of 30 words each that were balanced for frequency of occurrence in the language. Twenty-eight of these lists acted as study items while the other lists comprised the new items that were repeated. The remaining 420 words were divided into 28 lists of 15 words to act as filler items. Six items were randomly selected from each filler list to be presented as new words that were not repeated; this was necessary to create the appropriate spacing for lag intervals. Lists designated as study, new, and filler were identical for each participant. Since a relatively large number of study and distractor lists that are balanced for word frequency, randomly assigned to different conditions (i.e., study, new, filler), and held constant across individuals were used, it is unlikely that a list effect could be mistaken for training (Jennings & Jacoby, 2003). List items were presented on a PC compatible computer in lower case letters centered in the middle of the screen. Letters were white presented on a black screen. The character size of the stimuli was approximately 3 x 5mm.

Procedure

The procedure generally followed the protocol established by Jennings and Jacoby (2003). To reiterate, four training sessions were given per day for a total of seven days with a weekend break after day 3 or 4. Other than the weekend break and instances when a break could not be avoided (illness, missed appointment), training days occurred consecutively.

As devised by Jennings and Jacoby (2003), each training session consisted of a study and test phase, the latter randomly intermixing one study, one new, and one filler list. During the study phase, each word was presented for two seconds and participants were asked to read the word aloud and remember it. The opposition test phase followed in which participants were shown the 30 study words and 30 new words, with the latter repeated at one of two different lag intervals. Participants were asked to identify words read aloud during the study phase by pressing 'yes' (corresponding to '/' on the keyboard) or 'no' (corresponding to 'z' on the keyboard). Prior to the test phase, participants were correctly instructed that some of the test items would be repeated and that if an item was repeated then it *was not* one of the original study items. When a correct response was entered, the computer beeped and the message "CORRECT" appeared across the bottom of the screen; when an error was made, no message appeared. Participants were asked to pay attention to the feedback given and try to learn from their mistakes. Unlimited time was given during the test phase; each study-test cycle lasted approximately 7 to 15 minutes to complete, depending on the individual.

As previously described, an incremented-difficulty technique was used that gradually increased the lag intervals as a function of performance. Specifically, the lag interval was increased if the participant reached the accuracy level shown by young adults (Jennings & Jacoby, 1997; as cited in Jennings & Jacoby, 2003). This criterion was such that a participant could only make one error on the repeated items for lags 1 to 4 and two errors for lags 8 to 40. The lag interval pairs used for training were 1 and 2; 1 and 3; 2 and 4; 2 and 8; 4 and 12; 4 and 16; 8 and 20; 8 and 24; 12 and 28; 12 and 32; 16 and 36; 16 and 40; and proceeded in that order. Pairs were selected so that participants worked on one lag interval they have already mastered and a second new and more difficult interval. Participants progressed to the next lag pair only if criteria were met at both lags. If criteria were not met at both lags, participants continued to work at those intervals for as many cycles as needed to meet criteria. Since this incrementing procedure progressed only if participants met criteria, final performance was evaluated as the difference in interval length between the first and last day of training.

Deviations from previous protocol

Some modifications to the task were included in an attempt to make it more user-friendly for a group of individuals with complaints of memory dysfunction. Abbreviated instructions were included on the screen during each trial to remind participants of task demands. During the study phase, “Study List: Read each word aloud and try to remember it” appeared along the top of the screen during each trial. During the test phase, “Test: Press ‘yes’ if you said this word aloud on the study list, ‘no’ if not; if a word appears twice on the test, then it was not on the study list” appeared in the same manner. These reminders were intended to reinforce task demands.

In addition, a lag interval pair of 0 and 1 was included and administered at the first session. The inclusion of a 0 lag was intended to assess participants’ understanding of the fundamentals of the task. In previous research, older adults showed reduced performance relative to young adults at lags as short as 3 or 4 intervening items (Jennings & Jacoby, 1997). One could question whether this poor performance was a reflection of a misunderstanding of task instructions and not a sign of impairment. By including a lag of 0 (hence, the repeated word immediately followed itself, such as 'light-light'), we intended to determine if poor performance at short lags was due to a lack of understanding of task demands. Since past research has shown that older adults perform at a level comparable to young adults at a lag of 0 (Jennings & Jacoby, 1997), it seemed unlikely that impaired performance was due to a lack of understanding, but for purposes of clarity it was useful to assess this directly in our sample.

It is unclear from previous work how the task was introduced to participants. For our purposes, instructions for the task appeared on the computer screen (as in Jennings & Jacoby, 2003) but were also read aloud to participants and summarized by the examiner. Any questions were clarified, as needed, before each session. In addition, the instructions were read aloud and briefly reviewed at the start of each training day. We found this to be particularly important in our sample of individuals with memory complaints.

Cognitive Test Battery

Generalizability – Proximal effects

False Fame Task – The false fame task is a well-recognized paradigm based on Jacoby's (1991) process dissociation framework. It has been used extensively in a variety of research studies (e.g., Jacoby et al., 1989; Mayes, 1995). It was chosen for inclusion in this study as a means of assessing proximal effects of training. Since the paradigm is similar to the one used to develop the training program, we expected any generalizability of training effects to be apparent on this measure. The methods used for our purposes closely followed those of Jacoby, Woloshyn, and Kelley (1989) and are described below.

Much like the training paradigm used in this study, the false fame task places the processes of familiarity and recollection in opposition. During the first phase of the task, participants read aloud a list of 40 non-famous names presented on a computer screen. The names appeared in white, 20-point font in the center of a black screen for approximately two seconds with a one second blank interval (i.e., black screen) between items. The first letters of both the first and last name were capitalized. A test phase followed during which participants read a mixed list of 60 new famous, 30 old nonfamous, and 30 new nonfamous names from the computer screen and made a fame judgement on each item. Participants were correctly told that the names on the first list were non-famous and, hence, that if they recognized a name from the first list, then that name was definitely *not* famous. Participants were also informed that the famous names were not extremely famous, like Harrison Ford or Pierre Trudeau, and that they would not be asked to describe what a named person had done to become famous. As in all false fame paradigms, these instructions were intended to encourage subjects to use familiarity as a basis for their fame judgements.

Feedback on false fame is generally not included in experimental investigations. However, as the purpose of including a false fame task in this study was to assess for proximal effects of training, we decided to provide oral feedback in an attempt to more closely mirror the training procedure. During the training program, the computer beeped and displayed the word ‘CORRECT’ when a correct response was given and was silent when an error is made. Feedback during the false fame task was provided orally by the examiner. Paralleling the feedback provided during training, the examiner said "correct" when an accurate response was given and was silent when an error was made on a trial-by-trial basis. The inclusion of feedback was intended to provide a context similar to that supplied during training.

After the fame-judgment task, a 20-item recognition test was administered consisting of the first five and the last five names read during the first phase of the experiment along with 10 new nonfamous names. Participants were given the list of pseudorandomized names on a sheet of paper and asked to circle the ones they had read in the first phase of the task.

The majority of names used in our project were developed for use in a previous study (Gruppuso & Lindsay, unpublished stimuli) and were matched according to length of the first and last name, sex indicated by the first name, nationality of origin of the last name, and oddity or commonness of the name. For example, the famous name Wyatt Earp was matched with the nonfamous names Lemar Bint and Cliff Talbot; the famous name Yasir Arafat was matched with the nonfamous names Omar Obadiah and Mohamed Nazar. An additional list of 60 names was created (20 famous names based on Jacoby, unpublished data; 40 nonfamous names generated to match famous names based on the Gruppuso & Lindsay criteria). Following the procedure of Jacoby and colleagues (1989), the order of presentation of the names was randomly generated with the restriction that in the test and recognition phases no more than three names of one type ("old" nonfamous, "new" nonfamous, famous for test phase; first 5, last 5, new nonfamous for recognition) could appear before one name of each of the other types. The names were presented in a fixed order across participants with different names used for each visit. The order of presentation of the sets (version 1, 2, 3) was counterbalanced across participants. This task took approximately 15 minutes to administer.

One version of the task was randomly selected for pilot testing on a group of six undergraduate volunteers. Pilot testing occurred in a group setting with the stimuli appearing via video projector on a large screen. Participants were asked to read the names silently as they were presented during study and to circle their responses during the fame judgment task.

As the false fame effect (i.e., falsely identifying as famous those items previously presented) is not generally seen in healthy controls except under conditions of divided attention, a distractor task was administered at encoding following the procedure of Jacoby et al. (1989). Specifically, a long string of numbers was presented via audiotape and participants were instructed to identify the total number of occurrences of strings of three consecutive odd numbers. The listening task used followed the procedure of Craik (1982). Namely, forty-three sequences of three odd numbers occurred within a list of 224 randomly generated numbers. Restrictions used to construct the list were that a minimum of 1 and a maximum of 8 numbers must occur between the end of one and the beginning of the next target sequence. Also, not more than three even numbers could occur in a sequence. Fourteen target sequences were presented as determined by the length of presentation of the study names.

Results under conditions of divided attention were consistent with the false fame effect as participants falsely identified as famous previously presented nonfamous names at a higher rate (39%) as compared to nonfamous names presented for the first time (29%). Identification of famous names was within an acceptable range (58%).

Generalizability – Distal effects

Buschke Selective Reminding Test (SRT) – The SRT is a commonly used clinical measure (Capruso & Levin, 1992) that assesses verbal learning and memory during a multiple-trial word list-learning task (Spreen & Strauss, 1998). It was selected for use in this study as a means of assessing distal effects of training on an instrument frequently used in the population of interest. Though we did not expect to detect training effects, we included this measure as a means of testing the limits of training. If training effects were to generalize to clinical measures of memory, it would be reasonable to detect improvement on a measure of list-learning.

Following standard administration procedures, a list of 12 unrelated words was presented over 12 selective reminding trials or until the participant was able to recall the entire list on three consecutive trials. Each subsequent learning trial involved the selective presentation of only those items that were not correctly recalled on the immediately preceding trial. Following completion of the learning trials, a cued-recall trial was presented in which the first two or three letters of each word were shown on individual index cards and the participant was asked to generate the list word. A multiple-choice recognition trial was then administered where 12 index cards were presented each containing a list word, a synonym, a homonym, and an unrelated distractor word. Participants were asked to identify the list word on each card. Following a 30-minute delay, an additional free-recall trial was administered. A variety of scores can be calculated on this task. For our purposes, the total recall over learning trials (the most stable index per Sass et al., in press; as cited in Spreen & Strauss, 1998) and the number of words correctly identified following a delay were analyzed.

Alternate lists were used at each session for a total of three distinct versions. The alternate forms, developed by Hannay and Levin (1985; as cited in Spreen & Strauss, 1998), have been shown to be of equivalent difficulty for elderly subjects (Masur et al., 1989) and patients with medically refractory epilepsy (Sass et al., in press, Westerveld et al., 1994; all as cited in Spreen & Strauss, 1998). When patients with seizures underwent multiple administrations of alternate forms on four consecutive days, there were no significant practice effects (Sass et al., in press; Westerveld et al., 1994; all as cited in Spreen & Strauss, 1998), though with normal individuals there does appear to be a non-specific practice effect with repeated administration of alternate forms (Clodfelter et al., 1987; Hannay & Levin, 1985; Loring & Panaicolaou, 1987; all as cited in Spreen & Strauss, 1998). Alternate form reliability coefficients are significant and tend to be moderate in magnitude (Clodfelter et al., 1987; Hannay & Levin, 1985; Morgan, 1982; Ruff et al., 1988; Sass et al., in press; Westerveld et al., 1994; all as cited in Spreen & Strauss, 1998). Administration of this task took approximately 30 minutes at each session.

Other measures of memory

Source Monitoring Task – Source monitoring paradigms force participants to indicate the source of learned information (Johnson et al., 1993). One means of conceptualizing the training effects seen on the Jennings and Jacoby paradigm is in terms of source monitoring – that participants are improving their ability to identify the source of learned information. During training, participants are not only told that repeated words in the test phase were not on the original list but are also given feedback on an item-by-item basis. This feedback could help them solidify over time a distinction between the two sources of learned information (study phase vs. test phase). Although Jennings and Jacoby have argued that the changes seen in training are better explained by a dual process model of memory as opposed to a source memory model (Jennings & Jacoby, 2003), this distinction has not been tested empirically.

To assess the role of source monitoring in the training paradigm, a source monitoring task was administered pre- and post-training to establish effects of training on this alternate measure. A word list paradigm was used such that participants listened to two taped word lists being read aloud by two different female voices. This basic procedure has been used previously to assess the effect of perceptual cues on performance (Johnson, De Leonardis, Hashtroudi, & Ferguson, 1995).

Two distinct word lists consisting of 15 items each were read aloud by two female voices at an approximate rate of one word every two seconds. The lists were presented via audiotapes prepared in a recording studio. Each voice was arbitrarily assigned a name. Prior to presentation of the lists, participants were told that they would be asked to identify not only the words, but also which of the two voices read the words. Presentation of both lists was followed by a recognition test requiring participants to indicate whether or not a word was heard earlier and if so, who spoke it.

To control for effects of practice across study visits, three versions of the task were created with the order of lists and voice administration predetermined. Word lists from alternate forms of the Rey Auditory Verbal Learning Test (Majdan et al., in press; Geffen et al., 1994; as cited in Spreen & Strauss, 1998) were used as well as the accompanying recognition tests that included an equal number of items from each originally presented list along with distractor items. Administration of the three versions was counterbalanced across participants. Outcome measures included the number of items correctly identified as well as the source of those items. This task took approximately five minutes to administer.

Materials were pilot tested on undergraduate volunteers for each of the three versions (7, 11, and 6 participants for versions 1, 2, and 3, respectively). On average, 72.5% of presented items (76.2%, 70.6%, 71.7%) and 64.0% of sources (65.7%, 60.0%, 66.2%) were correctly identified in our healthy volunteers. The order of presentation of the three versions was counterbalanced across participants.

Attention

Letter-Number Sequencing – To perform well on the training program, participants must identify study words while recognizing words that are repeated during the test phase. This task could include holding a repeated item in mind for a sufficient amount of time while completing a different task (namely, correctly identifying other items presented during the lag interval). One could argue that this type of ability is related to working memory. Improvement over time, then, could be conceptualized as an increased facility with retaining and identifying repeated items. To test this hypothesis, a Letter-Number Sequencing task, a subtest of the Wechsler Memory Scale-III (Wechsler, 1997) was administered pre- and post-training. Letter-Number Sequencing assesses auditory working memory by requiring participants to sequentially order a series of numbers and letters orally presented in a specified random order (Wechsler, 1997). Participants were asked to first remember the numbers and letters read aloud and then reorganize the numbers into ascending order and the letters into alphabetical order. Strings progressively increased in size from two to eight items each with three trials at each string length. The task was discontinued if a participant failed all three trials of a given string length. Trials completed correctly were scored one point for a total possible score of 21. This task took approximately five minutes to administer.

For purposes of repeated testing, two alternate versions of the task were created. Strings of numbers and letters (from 2 to 8 items each) were selected in groups of three to match the original version. As our goal was to maintain the relative difficulty of the task while diminishing the effect of practice, replacement numbers and letters were drawn from similar positions in the alphabet or along the number line to match the original items in terms of the degree of mental manipulation required to correctly reorganize the stimuli. For example, the number 4 was replaced by 3 or 5 or the letter L by J or K. Since numbers or letters at the beginning or end of the alphabet/number line provide clear anchors, 1, 9, A, and Z were generally not replaced. The letters I, O, and U were not used as these items do not appear in the original task. Once the stimuli were selected at each string length, the order of each set was shuffled using randomly generated numbers (www.randomizer.org) to be different from the original version. Order of presentation of each version was counterbalanced across participants.

Conners' Continuous Performance Task (CPT) – The Conners' CPT (1995) provides a measure of sustained attention or vigilance in addition to monitoring impulsivity and inattention (Spreen & Strauss, 1998). This task was included to assess whether improvements on measures of attention could explain training effects. The Conners' CPT was chosen as it combines a measure of vigilance, necessary for concentrating on the training task for 15-minute segments, with an inhibition component. Since the training paradigm requires the suppression of feelings of familiarity, poor performance could relate to difficulties in inhibition. By measuring inhibition pre- and post-training, we can assess whether training leads to improved inhibitory control. Participants sat in front of a computer screen and were required to press a designated key for any letter except 'X.' Six blocks were presented sequentially with 60 trials per block. Of these 60 trials, 54 were hits (e.g., letters other than 'X') and six were rejection trials (e.g., 'X') requiring inhibition of a response. Across the task, participants were presented with 324 hit trials and 36 rejection trials. The interstimulus interval (1, 2, or 4 seconds) varied in blocks of 20 trials. The number correct, omission errors, commission errors, and reaction times were recorded. Outcome measures of interest included errors of commission, which provide a measure of impulsivity, and variability (change in SE across continuously administered blocks), which measures consistency over time and provides an estimate of attentiveness. This task took approximately 14 minutes to administer and was administered at each testing session.

Order of administration

To reduce bias due to sequence of administration, the cognitive test battery was presented in one of two predetermined orders: 1) False Fame, Buschke SRT Immediate, Letter-Number Sequencing, Conner's CPT, Buschke SRT Delayed, Source Monitoring Task, or 2) Buschke SRT Immediate, Letter-Number Sequencing, Conner's CPT, Buschke SRT Delayed, Source Monitoring Task, False Fame. Each participant completed tasks in the same order across sessions and task order was counterbalanced across participants.

Metacognitive measures

Following the post-training session, participants were asked general questions about the way they completed the training task and any strategies they developed or used over the course of the study. Similar to the previous study (Jennings & Jacoby, 2003), participants were asked whether or not they noticed improvement over time and, if so, to explain any such improvement (i.e., What do you think caused your performance to improve over time?).

Analyses

Analyses will consist of comparisons at both the group and individual level to test the hypotheses previously stated. To reiterate: 1) Do the training effects found in a previous study (Jennings & Jacoby, 2003) generalize to a new population? 2) Does training impact other measures of memory? and, 3) Are there patterns of change across measures of memory and attention that lend insight into possible mechanisms underlying training or suggest directions for future research?

Replication of Jennings and Jacoby (2003)

We will assess whether training effects found in a previous study generalize to the current sample through a series of analyses. Firstly, to determine whether training led to improved performance, we will compare the highest lag achieved at day 1 to the highest lag on day 7 using a paired *t*-test. As the program only progresses when criteria are met, this analysis will determine whether participants, on average, were able to meet criteria at higher lags post-training as compared to before training. Paralleling the procedure of Jennings and Jacoby (2003), this comparison will be made between session 3 (day 1) and session 28 (day 7) to allow for acclimation to task demands on day 1.

Secondly, to determine whether the frequency of improvement in our sample is comparable to that seen in the Jennings and Jacoby study we will dichotomize our sample into two groups, improved and not improved, and run a Chi-square analysis. As Jennings and Jacoby did not provide information regarding how they classified improved versus unimproved, we will set our cut-off by excluding the lowest 25% of performers. This value will be adjusted to accommodate the Jennings and Jacoby classification (i.e., is our cut-off consistent with the way Jennings and Jacoby divided their own sample?) as well as clinical judgement of meaningful improvement. Since performance on this paradigm has not been tested directly against measures of everyday memory, there is no guidance in the literature as to what degree of change would be considered "clinically meaningful." For our purposes, we determined that since differences between normal and aging performance have been consistently seen across small lags (e.g., Jennings & Jacoby, 1997), we would expect clinically meaningful improvement to incorporate improvement above a lag of 4.

To further test our group-based criterion, we will conduct regression analyses at the level of individual subjects to determine whether the rate of change in lag interval (e.g., the highest lag achieved) across sessions is statistically significant. If the rate of change per session is not significant for any given individual, that individual will be classified as 'not improved.' The redundancies in the classification system are intended to preclude arbitrarily labelling someone as "improved."

Lastly, to assess whether the magnitude of change in the current sample is consistent with the Jennings and Jacoby study, we will perform a one-sample *t*-test comparing change in participants in this sample to the mean level of change in the Jennings and Jacoby sample.

Power analyses for primary hypothesis – Power analyses were conducted using g-power, a general power analysis program available via the web (Erdfelder, Faul, & Buchner, 1996). As improvement following training was our primary outcome measure, power analyses were conducted based on the paired t-test of performance on day 7 as compared to day 1. Using means and standard deviations provided by Jennings and Jacoby, we estimated the standard deviation of the mean difference, $\sigma = 16$. Since change on the paradigm would likely be large to lead to proximal or distal effects on other measures, we assumed meaningful change to be equal to 15. Based on these values, a large effect size was assumed, $f = .938$. Given this effect size, 10 participants would lead to power of .86 for this analysis at $p < .05$.

Performance on measures of memory and attention

To determine whether training impacts other measures of memory and whether there are patterns of change across measures of memory and attention that could lend insight into possible mechanisms of training, we will conduct separate repeated measures analyses for each of the five tasks administered comparing performance over time for both the experimental and control groups. As the focus is on determining tasks that were impacted by training, we will include in analyses only those experimental subjects that are classified as "improved" following training by the criteria described in the previous section (see Replication of Jennings and Jacoby (2003)). For those instances where a main effect of task is detected, *post-hoc* comparisons will be made comparing performance posttraining to the mean level of performance at sessions 1 and 2 (baseline and pretraining).

Recognizing that multiple analyses will greatly increase the likelihood of making a type I error, a multivariate approach was not adopted for the following reasons: 1) a multivariate approach could obscure interpretability and 2) each of the tasks noted above was conceptualized as measuring separate, independent constructs that could reflect distinct cognitive processes affected by training. Due to the exploratory nature of these analyses and to preserve power, a p-value of .05 was maintained for all analyses.

Results

Participants

Data from one male experimental participant with self-reported severe TBI were excluded from analyses due to computer problems during training. Data from one female control participant were excluded due to reports of concussion in childhood in combination with poor performance on a memory measure (CVLT; Learning total T-score = 37, Trial 1 z-score = -1.5, Delayed recall z-score = -.5) at screening.

A summary of demographic information is provided in Table 2.

Memory abilities

All participants completed the CVLT, a well-validated and widely used clinical measure of memory, at an initial screening visit. Scales of interest included the total number of words correctly identified over five learning trials, learning on Trial 1, and recall following a delay. Average performance of each group is presented in Table 2. Based on the criteria of consistently poor performance (at least 1 SD below the mean compared to a large normative sample) across various scales, six of the ten ABI participants were classified as showing impaired memory performance.

Replication of Jennings and Jacoby (2003)

Training effects in the study sample were significant, t (9) = -3.94, $p < .01$, Cohen's $d = 1.25$, estimated $\delta = 1.13$. Participants improved, on average, from only meeting criteria at lags of 1.7 ($SD = .67$) to meeting criteria at lags of 20 ($SD = 14.97$). There was no evidence to suggest that the magnitude of change in the sample was different from previous findings, as the mean level of change (18.3 lags) was not statistically different from Jennings and Jacoby's mean level of change (26 lags), t (9) = -1.66, $p > .10$. Power for this analysis was quite low due to our small number of observations ($N = 10$).

To determine frequency of improvement, we set a cut-off score at a raw difference of 6 lags, the 25th percentile of the frequency distribution of change in the sample. This interval was considered to fall in the clinically meaningful range as it incorporated change above a lag interval of 4 (see Analyses section above for further description of "clinically meaningful" change in this context). Using this cut-off, we classified eight individuals in the study sample as improved and two as unimproved. As applied to the Jennings and Jacoby sample, this criterion identified nine individuals as improved and three as unimproved; those identified as improved and unimproved according to this criterion were consistent with those identified by Jennings and Jacoby.

To further assess the significance of change at the individual level, the highest lag achieved was regressed onto session number for each subject. These analyses were intended to exclude any individuals classified as 'improved' based on our group level comparison that would not be considered to show meaningful improvement at the individual level. For example, our criterion of the lowest 25% of the frequency distribution of change in our sample could have classified as 'improved' individuals whose performance changed by only 2 or 3 lags. Though arguably conforming to our operational definition of clinically meaningful change (i.e., improvement above a lag of 4), we would not consider these individuals to be improved at the individual level. Therefore, individual regression analyses were conducted. Results of these analyses are presented in Table 3. No subjects were identified as unimproved on the basis of this analysis, as the slope of each regression line was found to be significantly different from zero, $p < .001$, for all participants. We, thus, maintained our division of improved versus unimproved noted above and classified eight individuals in our sample as improved and two as unimproved.

Once we had identified the ratio of improved to unimproved in our sample, we compared this frequency to that found in the Jennings and Jacoby study. Results were not significant (Fisher's exact test, $p = .594$). There is no evidence to support that the frequency of improvement in our sample is different from the frequency of improvement in the Jennings and Jacoby study. A summary of training data is provided in Table 4. Performance on the training paradigm over time is provided for the improved and unimproved samples in figures 1 and 2, respectively.

Overall accuracy on the training paradigm

Post hoc analyses of accuracy on the training paradigm were conducted to further explore training effects. As the training program includes a recognition test at each session, we calculated an overall accuracy score by averaging the proportion of correct responses to studied items (old) and the proportion of correct responses to nonstudied items (new). Results were analysed separately for those individuals classified as "improved" as compared to those individuals classified as "unimproved" based on the criteria noted above.

Overall accuracy was regressed onto session number to assess change in accuracy over time. The unimproved group showed a decrease in accuracy across sessions that was statistically reliable, $F(1, 54) = 4.58, p < .05$. Each increase in session number resulted in a decrease of a fraction of a percent in accuracy, $B = -.00205, p < .05$. Overall, session number accounted for 8% of the variance in accuracy ($r = -.28$) in the unimproved group. Results for the improved group showed a similar pattern with a trend towards significance, $F(1, 222) = 2.94, p = .088$. Each increase in session number resulted in a decrease of a fraction of a percent in accuracy, $B = -.00132, p = .088$, with session number accounting for 1% of the variance in accuracy ($r = -.11$) for the improved group. Further analyses suggest that this effect is driven by a decrease in correct recognition of previously studied items over time. When the proportion correct on studied items (old) and the proportion correct on nonstudied items (new) are simultaneously entered as predictors of session number, the proportion correct on studied items (old) is the only significant predictor both for the unimproved group, $t = -2.16, p < .05$, and the improved group, $t = -2.34, p < .05$. Controlling for the proportion of studied items (old) correctly identified, the proportion of nonstudied items (new) correctly identified is not a significantly related to session number.

Generalizability – Proximal effects

As noted above, since the focus of these and subsequent analyses was on determining tasks that were affected by training, we included in analyses only those experimental subjects who were classified as "improved" following training using the criteria previously discussed (see Replication of Jennings and Jacoby (2003)). Eight experimental and nine control subjects were therefore included in analyses.

Proximal effects of training were assessed by examining the occurrence of false fame for previously presented items across the three study visits (Baseline, Pretraining, Posttraining). A repeated measures ANOVA was performed with the number of errors on old nonfamous names as the dependent variable and study visit as the independent variable. Means for each group are provided in Table 5. Assumptions of sphericity were met (Greenhouse-Geisser E = .94). A moderate effect of session was found, $F\,(2,17) = 4.68, p < .05$, partial $\eta^2 = .24$, with a trend towards an interaction between session and group, $F\,(2,17) = 2.98, p = .07$, partial $\eta^2 = .17$. See figure 3. Contrasts of performance across the three sessions suggest that a significant decrease in false fame errors occurred at session 3 as compared to the average for sessions 1 and 2, $F\,(1,17) = 5.92, p < .05$, partial $\eta^2 = .27$.

To determine whether changes in false fame were representative of an overall pattern affecting the identification of nonfamous names as famous, we ran a separate repeated measures ANOVA on the number of errors for new nonfamous names. Means for each group are provided in Table 5. A Greenhouse-Geisser correction was used as assumptions of sphericity were not met (Greenhouse-Geisser E = .80). There was no significant effect across sessions, $F(1.61, 17) = 1.57, p > .05$, and no significant interaction, $F(1.61, 17) = .785, p > .05$.

Historically, these types of data have most often been analysed with methodology similar to that used above. However, some have argued that analyses of this kind are inappropriate for describing data collected in a pre-test/post-test design (e.g., Dugard & Todman, 1995). To this end, an alternate analysis of the same data is presented using an analysis of covariance (ANCOVA). Based on the graph presented in figure 3, we determined that there were no differential group effects between the baseline and pretraining sessions. Performance at the pretraining session was then selected as the covariate to control for effects of practice. Analyses were conducted using a hierarchical linear regression procedure with post-training performance as the dependent variable. The test for a common regression coefficient showed a non-significant interaction between group (experimental, control) and the covariate, R^2 *change* = .01, *F change* (1, 13) = .22, $p = .65$, when controlling for each of these variables independently. Using a common regression coefficient, then, there was no effect of group on post-training performance when controlling for pre-training performance, R^2 *change* < .00, *F change* (1, 14) < .00, $p = .95$.

To further address proximal generalizability, we looked at the correlation between the amount of improvement on the training paradigm to the occurrence of false fame before and after training. Difference scores were computed for each measure. Correlation between the two measures was weak, *Pearson's* $r = -.24, p > .05$.

Generalizability – Distal effects

More distal effects of training were assessed by examining performance on the Buschke SRT, a commonly used clinical measure of memory performance. We conducted a repeated measures ANOVA with the number of words recalled across 12 learning trials as the dependent variable and study visit as the independent variable. Means for each group are provided in Table 5. Assumptions of sphericity were met (Greenhouse-Geisser E = .94). A moderate and statistically reliable interaction between task and group was found, $F\ (2, 17) = 5.179, p < .01$, partial $\eta^2 = .26$, so results were analysed separately. See Figure 4. In the experimental group, no reliable change in performance was noted over time, $F\ (1.67, 8) = .846, p > .05$. In the control group, a large change over time was detected, $F\ (1.34, 9) = 10.83, p < .01$, partial $\eta^2 = .58$, and contrasts showed that performance improved significantly at the last session as compared to the mean level of performance at previous sessions, $F\ (1,9) = 14.37, p < .01$. Similar analyses conducted for the number of words recalled following a delay showed no significant effect over time, $F\ (1.79, 17) = .174, p > .05$, and no significant interaction, $F\ (1.79, 17) = 1.031, p > .05$.

Other measures of memory and attention

Results of all other ANOVA's run on key variables of interest (total score on letter-number sequencing; number of items recalled on a source monitoring task; number of sources correctly identified on a source monitoring task; errors of commission on a continuous performance measures; and an index of variability over time on a continuous performance measure) were not significant. Summary statistics are provided in Table 6.

Data for the experimental group were also analysed graphically for individual subjects in an attempt to identify potential patterns of change across training. No clear pattern emerged in this exploratory analysis.

Discussion

Replication of Jennings and Jacoby (2003)

The primary goal of the current project was to determine whether training effects found in a group of aging adults (Jennings & Jacoby, 2003) generalize to a population of brain-injured individuals. Brain-injured individuals generally show a similar pattern of memory impairment to the population previously studied in that impairment is found on consciously controlled processing with automatic uses of memory intact (Ste-Marie, Jennings, & Finlayson, 1996). Our current sample showed a pattern of performance consistent with such a prediction, as all but one of our brain-injured individuals falsely identified an unusually large number of repeated items as 'previously studied,' even when only one or two items intervened between repetitions. Based on process dissociation research (e.g., Jacoby, 1991), this finding suggests that our participants had an intact sense of familiarity (repetition led to a greater likelihood of identifying a word as 'studied') that was unregulated by conscious recollection (they were not able to consciously recollect that the item was repeated).

To determine the efficacy of training in our sample, we compared performance before training (day 1) to performance after training (day 7). As in the Jennings and Jacoby sample, we found improved performance over time that was statistically reliable. In a further analysis, we found no evidence to suggest that the magnitude of change in our sample was different from that in the Jennings and Jacoby sample, although the power for this analysis was quite low. Looking at the frequency distribution of change, the Jennings and Jacoby sample showed a more dichotomous distribution with some individuals showing large improvements and others showing none. In contrast, several individuals in our sample showed more moderate degrees of improvement. We can only speculate as to the reason for differences between the two groups. Perhaps the reduced level of performance is related to more wide-sweeping cognitive problems in the ABI sample; perhaps performance in the two groups would be more consistent with additional sessions; perhaps our small sample sizes limited the potential for change of varying magnitudes. Regardless of the reason for individual differences, the average effectiveness of the training program appears to be consistent in both groups.

Equally important as the magnitude of change in our sample was the frequency of improvement. If only one or two individuals showed improvement, the utility of the training program would be brought into question and support for the idea of general change through training would be reduced. By creating dichotomous groups of 'improved' versus 'unimproved' participants, we were able to compare the frequency of improvement across the two samples. In both groups, the frequency of individuals considered 'improved' was high (75-80% improved) and we found no evidence to suggest that the groups differed in terms of the frequency of improvement.

To further explore training effects, we assessed change in accuracy on the recognition test over time for both the improved and unimproved groups. Overall, there was a general trend towards a decrease in accuracy over time for both groups that was statistically reliable in the unimproved group. This decrease seemed to be accounted for by diminished ability to recognize studied items (old). Although overall accuracy was not reported for the Jennings and Jacoby sample, these findings are consistent with the authors' observation that recognition of previously studied items decreased with training. They postulated that the decrement in performance over time was due to the high volume of words encountered during training.

Findings of decreased accuracy over time on the training paradigm itself are concerning. If what is being trained is a memory process involved in the recognition of previously-learned items, we would expect overall accuracy on the recognition test to improve. However, one could also argue that overall recognition performance is not the focus of intervention. Instead, the intervention is focused specifically on recognition of repeated items. In this case, we may not expect general improvement on recognition performance as overall performance may obscure results. This argument, however, does not dismiss concerns regarding the decrement in accuracy over time and caution should be used in assumptions of generalizability to general memory processes. These issues are further highlighted in our discussion below of results from other measures of memory performance.

Despite concerns surrounding issues of generalizability, results support a replication of the Jennings and Jacoby findings in a new sample of memory-impaired individuals. Although limited by low power due to small sample sizes, these preliminary analyses suggest that training is equally effective at improving performance in the aging and ABI samples in terms of both frequency and magnitude.

Implications for a dual-process model of memory

The finding of replication in a new sample of memory-impaired individuals provides converging evidence to support a dual-process model of memory. As predicted, our ABI group showed an increased likelihood of identifying repeated items as study words prior to training. From the perspective of Jacoby's dual-process model of memory, these errors occur due to automatic influences of memory that are unregulated by conscious recollection of seeing the word previously in the test list. In normal memory processing, conscious recollection allows for regulation of errors since feelings of familiarity are balanced by recollection that the word is repeated. Another possibility that could lead to the pattern of results is that training led to suppression of feelings of familiarity. Based on current results, either interpretation is plausible. Despite the reason for change over time, by showing improvement to a level seen in healthy, young adults a certain degree of flexibility in processing required for completion of this task is suggested.

While it is unclear what the mechanism is underlying change (e.g., restoration of damaged function, substitution of intact function), the fact that change occurs over time provides an end result of improved performance on this type of opposition procedure, which relies directly on the ability to consciously recollect repeated words (or, to suppress intact feelings of familiarity). From a theoretical standpoint, the reduction in errors seen in our training sample over time suggests that improvements are being made to the process that allows for conscious recollection. Another possibility is that training leads to suppression of feelings of familiarity over time. Whatever the underlying mechanism for change, an overall shift in how participants complete the task is detected across session. So, unlike other attempts to improve memory by creating habitual, overlearned responses (e.g., Wilson et al., 1994), it is the training of a memory *process* that is proposed to be improving over time with this procedure. General conclusions, however, regarding the degree in improvement in memory processes can not be made based on these limited findings. Additionally, findings of decreased accuracy in recognition on the training program over time bring into question the extent of improved memory performance, even on the paradigm itself, cautioning against broad-based conclusions regarding improved general memory processing.

Proximal effects of training

While the theoretical framework of Jacoby's dual-process model of memory provides a context for interpreting training results, we were also interested in the degree of generalizability that was detectable on other measures of memory. Proximal effects of training were assessed by administering a similarly-designed false fame paradigm. As the false fame paradigm is based on the same theory of memory processing used to design the training program, we hypothesized that any effects of training on a general memory process would impact this similarly-structured task.

Analyses focused on the occurrence of false fame errors, a parallel outcome measure to that being 'trained.' Overall, a decrease in the number of false fame errors was detected across sessions with a significant change occurring at session 3 as compared to the average performance at sessions 1 and 2. Although this effect appears to be driven by performance of the experimental group (see Figure 3), the group by task interaction showed only a trend towards significance, suggesting that a similar pattern of performance was seen in the control group. As the power of this analysis was at chance (.54), the likelihood of detecting a significant interaction was quite low. An alternate analysis of these same results using ANCOVA suggest that substantial caution should be used in interpreting generalizability as no group difference was detected. Additionally, floor effects may have impacted the ability to detect differences between groups. Generally speaking, the occurrence of false fame errors decreased in the experimental group to such a degree that performance was nearly equivalent to that of controls.

One could argue that change on this one outcome measure was part of an overall pattern of decreased responses to nonfamous names. To test this hypothesis, we looked at the occurrence of errors on new nonfamous items across sessions. Unlike false fame errors, no change in errors on new nonfamous items were detected.

The pattern of results on false fame suggests that training impacted fame errors such that individuals were less likely to falsely identify previously seen names as famous following training. This improved level of performance does not seem to be part of an overall shift in responses to nonfamous names. Although the impact of training is only weakly supported (e.g., nonsignificant interaction between task and group, nonsignificant results when assessed using ANCOVA), these preliminary findings support the possibility that training results in a general change that affects other similarly structured tasks. However, change on false fame shows only a small correlation with improvement on training, bringing into further question the idea that training impacts performance on false fame.

Distal effects of training

For training to be practical in treating memory dysfunction, we would have to see generalization to other measures of memory performance. More distal effects of training would further support the idea that we are improving a *process* of memory and not just improving performance on one specific type of task. To assess change on more distal measures of memory, we included a list-learning task that is commonly used as a clinical tool to describe memory functioning (Buschke SRT). No change was detected on this task following training. In fact, while control subjects benefited from repeated practice and showed improved performance over time, the performance of experimental participants remained static, suggesting that they were unable to benefit from practice to the same degree as their counterparts.

In retrospect, this measure may not have been the best choice for determining distal effects of training. While the training task employs a list-learning format, the focus in training is not at improving recognition of previously studied words but instead on consciously recollecting the repetition of items in the test phase. In fact, consistent with previous findings, recognition performance of studied items actually decreased across training, likely due to the large volume of words encountered. A better approximation of generalization of findings would have been a measure of memory slips in daily life. Since the type of memory trained has been linked to everyday phenomena such as repeatedly telling the same story or forgetting appointments, monitoring performance on these day-to-day tasks may have been more sensitive to training effects.

Generalizability to other measures of memory and attention

Generalizability of training was also examined through other measures of memory and attention. Memory performance was evaluated with a source monitoring task. No effects following training were noted either in terms of the number of items recalled or the source of learned information. Although training could be conceptualized as requiring memory for the source of learned information (i.e., study phase or test phase), no generalization was detected on this task. Measures of attention and impulsivity (CPT; variability in reaction time, errors of commission) were variable over time and showed no clear change following training, suggesting that changes in attention cannot account for improvement on training. Similarly, no change was noted on a measure of working memory (Letter-number sequencing) following training suggesting training does not impact working memory performance. In these preliminary analyses, there is no evidence to support that training impacts other measures of memory and attention.

Limitations of the current project and suggestions for future research

Training

Over the course of the current project, we detected limitations to the training program that made interpretation of results difficult. Firstly, while participants can improve performance over time by meeting criteria at increasing lags, there is no possibility of worsening performance. When a subject meets criteria at a particular lag pair, the program progresses forward but at no time does the lag pair decrease if criteria are not met. Although the program was designed so that participants were always working on a new lag interval as well as one where the criterion has already been met, we are still faced with possible outcomes at the end of training that are limited to improved or consistent. By modifying the program to allow for flexibility in performance, perhaps by having the program drop back to a previous lag interval if criteria were not met at either lag, we would be in a better position to track performance over time. Additionally, in terms of clinical utility, such a modification would ensure that training occurred at the point of difficulty for that individual subject.

To further enhance clinical utility, another suggested modification to the program would be to strengthen the criteria used to define ‘normal’ performance. As the program currently stands, participants have to meet criteria for a given lag pair only once to progress to the next difficulty level. By setting a more stringent criteria (e.g., must meet criteria three times in a row), we could ensure that participants are showing improved performance and not just a chance occurrence of 'normal' performance. While this may decrease the overall training effect seen across the 28 sessions, we would be more secure in concluding that performance improved over time. An alternate way of providing more stringent criteria would be to require a certain level of performance on studied words before the program progresses. This would eliminate those instances where guessing is used to meet criteria and progress to the next lag pair.

Generalizability

Our ability to detect changes across various measures of memory and attention was severely limited by our sample size. This decreased the power of our analyses to such a degree that it was often at a level less than chance. In an attempt to move away from the assumptions of statistical hypothesis testing, we also visually inspected graphical representations of performance for each of our 'improved' subjects. No specific patterns emerged across participants following training. This finding could be caused by a variety of factors including: 1) training may be specific to the task administered and may not show generalizability to other measures of memory and attention, 2) improvement following training may occur through different pathways for various individuals so that finding one pattern of performance is unlikely, 3) we may have selected measures that were insensitive to the types of changes training affects. Reasons for a lack of generalizability are unclear. However, results of our preliminary analyses suggest that training did not generalize to measures that did not follow an opposition procedure format and only weakly generalized to similarly-structured tasks.

Future research

Due to the limitations of the current project, particularly in terms of issues affecting interpretation of generalization results, further research is required before firm conclusions can be made regarding the utility of this paradigm as a training procedure for memory dysfunction. Modifying the training program to more flexibly respond to performance, selecting generalization tasks that more closely mirror the training process, and increasing sample size are all recommendations for future research in this area.

Implications for clinical practice

At this point in time, only limited conclusions can be made regarding the clinical utility of this training procedure. Further research is warranted, however, as the training program has been found to be effective at improving performance on repeated lags in both aging and ABI groups and has shown suggestions of generalizability to a related task following a similar procedure. As memory dysfunction is a common occurrence in a variety of populations seen in a clinical setting (ABI, dementia, schizophrenia, substance abuse), the development of interventions that improve memory abilities are greatly needed. If this procedure shows utility over time at improving memory processing, it would be a valuable addition to the field of cognitive rehabilitation.

Perhaps the most relevant aspect of this project to the clinical researcher is that it provides an important first step in translating emerging cognitive theory into clinical practice. As with other intervention sciences, cognitive rehabilitation is faced with ever-increasing pressure to provide empirically-validated treatment. By using the cognitive literature as a basis for treatment planning, clinical researchers are provided with guidelines for how an intervention should be implemented (e.g., procedures), guidance on the types of clients who would likely benefit from treatment, recommendations on the types of assessment that would be helpful in screening potential patients, and are put in a more secure position to interpret treatment outcomes.

For example, in our study, we had a wealth of information at our disposal regarding a dual-process model of memory and proposed mechanisms for the type of memory that was captured through this paradigm. Although results were not seen on other measures of memory in this preliminary analysis, we are provided with clear guidance on where to expect improvements and how to direct future research attempts. This is not to say that cognitive theory translates directly into practice, but instead that it provides a framework for further hypothesis generation and empirical testing.

As all explorations of this type, we faced a variety of challenges in implementing our program. Firstly, as this training paradigm had never before been applied outside the aging population, our initial application was the first of its type in a new treatment population. We were, thus, unable to rely on previous research to predict those individuals who would likely benefit from training. For example, due to the high functioning level of the aging group previously studied (community dwelling, better than average years of education), we had originally sought out only mildly- to moderately-injured individuals as we predicted that their pattern of impairments would most closely mirror those seen in the aging population previously studied. Much to our surprise, after pilot testing the project on a several mild-TBI cases, we determined that these individuals did not show the pattern of impairment necessary for training; namely, they did not show difficulty at small lags. We, therefore, extended our search to include a moderately- to severely-impaired sample.

Similarly, despite a literature base for the type of memory captured by this paradigm, generalizability of training to other measures of memory had never been attempted. Before a treatment program can be implemented in clinical practice, we must validate that it leads to clinically-meaningful improvements in the areas it proposes. As generalizability had not been previously tested, we were left to make predictions as to the types of tasks that may be impacted by training without any previous work to provide guidance. Though we were able to hypothesize as to the types of tasks potentially impacted by training, much work is still needed to more clearly define the types of tasks that would most likely detect training effects. Before such tasks are identified, we cannot rule in or out the clinical effectiveness of treatment.

Lastly, another critical step in translating theory into intervention is providing evidence to support that training impacts the types of memory problems clinicians will likely encounter in their practice and will treat problems affecting everyday life. As this training paradigm provides focused treatment of a type of memory commonly disrupted in clinical populations, it is possible that it will benefit a clinical population. At the current stage of research, however, making predictions about the potential impact on everyday memory problems is not possible. One of the strengths of the paradigm, though, is that the focus is on the procedure used in training and not in imparting specific bits of information. One could imagine, then, modifications that would follow the same general procedure but would allow for flexibility in presentation format and lead to tailored intervention for particular patients. For example, perhaps a version could be created that focused on visually-presented stimuli? Or detecting repetition of a sentence or story? Perhaps words could be presented auditorally instead of on a computer screen?

Each of these modifications would preserve the context of training while allowing for provision of services to a wide audience of memory-impaired individuals. At this point in time, however, these are only far-reaching speculations

Conclusions

In conclusion, this study provided a replication of training results seen in a previous study (Jennings & Jacoby, 2003) and generalization to a new population of memory-impaired individuals. These findings are of theoretical significance as they support a dual-process model of memory and the idea that memory processing is flexible and can be modified over time.

We found some evidence to support that training generalized to a similarly designed paradigm, though these results were only weakly supported. Overall, a decrease in the mean number of fame errors was detected following training. We did not find generalization to other measures of memory or attention in these preliminary analyses, though we were limited by small sample sizes and reduced power. These findings may suggest that training is specific to this one type of memory procedure or, that we were not able to detect effects due to a lack of sensitivity in the measures used. Further research is necessary to determine generalizability of training effects. Suggested modifications to procedures for future research include changing the training program to respond to improved and reduced performance; modifying the tasks used to detect change following training; and increasing sample sizes.

These preliminary analyses of the efficacy and generalizability of the Jennings and Jacoby paradigm offer a model for translating cognitive science into clinical practice and provide a significant step towards developing a theoretically-driven rehabilitation program. However, this first step has been a small one and much work is required before conclusions can be drawn regarding the utility of this paradigm as a clinical tool.

References

Arciniegas, D., Adler, L., Topkoff, J., Cawthra, E., Filley, C.M., & Reite, M. (1999). Attention and memory dysfunction after traumatic brain injury: Cholinergic mechanisms, sensory gating, and a hypothesis for further investigation. *Brain Injury, 13(1),* 1-13.

Baddeley, A.D. (1992). Memory theory and memory therapy. In B.A. Wilson & N. Moffat (Eds.), *Clinical Management of Memory Problems* (2nd Edition, pp. 1-31). San Diego, CA: Singular Publishing Group, Inc.

Baine, D. (1989). Memory and mnemonics. *Mental Retardation & Learning Disability Bulletin, 17(2),* 23-33.

Bellezza, F.S. & Reddy, B. (1978). Mnemonic devices and natural memory. *Bulletin of the Psychonomic Society, 11(5),* 277-280.

Brooks, N., McKinlay, W., Symington, C., Beattie, A., & Campsie, L. (1987). Return to work within the first seven years of severe head injury. *Brain Injury, 1(1),* 5-19.

Buchner, A., Erdfelder, E., & Vaterrodt-Plunnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General, 124(2),* 137-160.

Butters, M.A., Soety, E.M., & Glisky, E.L. Memory Rehabilitation. In P.J. Snyder & P.D. Nussbaum (Eds.), *Clinical Neuropsychology: A Pocket Handbook for Assessment* (pp. 450-466). Washington, D.C.: American Psychological Association.

Capruso, D.X., & Levin, H.S. (1992). Cognitive impairment following closed head injury. *Neurologic Clinics, 10(4),* 879-893.

Craik, F.I.M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffmann, & E. van der Meer (Eds.) *Cognitive research in psychology* (pp. 152-161). Berlin: Verlag der Wissenschaften.

Curran, T., & Hintzman, D.L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(3),* 531-547.

Cycowicz, Y.M. (2000). Memory development and event-related brain potentials in children. *Biological Psychology, 54(1-3),* 145-174.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test: Adult Version Manual.* San Antonio, TX: The Psychological Corporation.

Dodson, C.S., & Johnson, M.K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General, 125(2),* 181-194.

Dugard, P., & Todman, J. (1995). Analysis of pre-test–post-test control group designs in educational research. *Educational Psychology, 15(2)*, 181-198.

Erdfelder, E., Faul, F., & Buchner, A. (1996). G*Power: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1-11.

Evans, J.J., Wilson, B.A., Needham, P., & Brentnall, S. (2003). Who makes good use of memory aids? Results of a survey of people with acquired brain injury. *Journal of the International Neuropsychological Society, 9,* 925-935.

Evans, R.W. (1992). The postconcussion syndrome and the sequelae of mild head injury. *Neurologic Clinics, 10(4),* 815-847.

Glisky, E.L. (2002, September). Is there a case to be made for treating memory impairment? In P.W. Halligan (Chair) *Effectiveness of rehabilitation for cognitive deficits.* Conference conducted in Cardiff, Wales.

Glisky, E.L., & Glisky, M.L. (2002). Learning and memory impairments. In P.J. Eslinger (Ed.) *Neuropsychological interventions: Clinical research and practice* (pp.137-162). New York, NY: Guilford Press.

Gruneberg, M.M. (1992). The practical application of memory aids: Knowing how, knowing when, and knowing when not. In M.M. Gruneberg, & P.E. Morris (Eds.), *Aspects of memory, Vol 1: The practical aspects* (2nd Edition, pp.168-195). Florence, KY: Taylor and Francis/Routledge.

Gruppuso, V., & Lindsay, D.S. [Recollection becomes familiarity over time]. Unpublished stimuli.

Gruppuso, V., Lindsay, D.S., & Kelley, C.M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23(2),* 259-278.

Harris, J.E. (1992). Ways to help memory. In B.A. Wilson & N. Moffat (Eds.), *Clinical Management of Memory Problems* (2nd Edition, pp. 59-85). San Diego, CA: Singular Publishing Group, Inc.

Hay, J.F., & Jacoby, L.L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology and Aging, 14(1),* 122-134.

Hay, J.F., Moscovitch, M., & Levine, B. (2002). Dissociating habit and recollection: Evidence from Parkinson's disease, amnesia and focal lesion patients. *Neuropsychologia, 40(8),* 1324-1334.

Hellawell, D.J., Taylor, R., & Pentland, B. (1999). Cognitive and psychosocial outcome following moderate or severe traumatic brain injury. *Brain Injury, 13(7),* 489-504.

Hertel, P.T., & Milan, S. Depressive deficits in recognition: Dissociation of recollection and familiarity. *Journal of Abnormal Psychology, 103(4),* 736-742.

Hintzman, D.L. (1990). Human learning and memory: Connections and dissociations. *Annual Review of Psychology, 41,* 109-139.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

Jacoby, L.L., Jennings, J.M., & Hay, J.F. (1996). Dissociating automatic and consciously controlled processes: Implications for diagnosis and rehabilitation of memory deficits. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M.K. Johnson (Eds.), *Basic and Applied Memory Research: Theory in Context,* (Vol. 1, pp. 161-193). Mahwah, NJ: Erlbaum.

Jacoby, L.L., Toth, J.P., & Yonelinas, A.P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122(2),* 139-154.

Jacoby, L.L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118(2),* 115-125.

Jennings, J.M., & Jacoby, L.L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation, 13(4)*, 417-440.

Jennings, J.M., & Jacoby, L.L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology and Aging, 12(2),* 352-361.

Johnson, M.K., De Leonardis, D.M., Hashtroudi, S., & Ferguson, S.A. (1995). Aging and single versus multiple cues in source monitoring. *Psychology and Aging, 10(4),* 507-517.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28.

Kee, D.W. (1994). Development differences in associative memory: Strategy use, mental effort, and knowledge access interactions. *Advances in Child Development and Behavior, 25,* 7-32.

Kinsella, G., Murtagh, D., Landry, A., Homfray, K., Hammond, M., O'Beirne, L., Dwyer, L., Lamont, M., & Ponsford, J. (1996). Everyday memory following traumatic brain injury. *Brain Injury, 10(7),* 499-507.

Klein, M., Houx, P.J., & Jolles, J. (1996). Long-term persisting cognitive sequelae of traumatic brain injury and the effect of age. *The Journal of Nervous and Mental Disease, 184(8),* 459-467.

Laatsch, L.K., Thulborn, K.R., Krisky, C.M., Shobat, D.M., & Sweeney, J.A. (2004). Investigating the neurobiological basis of cognitive rehabilitation therapy with fMRI. *Brain Injury, 18(10),* 957-974.

Lincoln, N.B., & Brooks, N. (1992). Assessment for rehabilitation. In B.A. Wilson & N. Moffat (Eds.), *Clinical Management of Memory Problems* (2nd Edition, pp.32-58). San Diego, CA: Singular Publishing Group, Inc.

Lindgren, M., Hagstadius, S., Abjornsson, G., & Orbaek, P. (1997). Neuropsychological rehabilitation of patients with organic solvent-induced chronic toxic encephalopathy: A pilot study. *Neuropsychological Rehabilitation, 7(1),* 1-22.

Mandler, G. Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252-271.

Mateer, C.A., Sohlberg, M.M., & Crinean, J. (1987). Focus on clinical research: Perceptions of memory function in individuals with closed-head injury. *Journal of Head Trauma and Rehabilitation, 2(3),* 74-84.

Mayes, A.R. (1995). Assessment of familiarity and recollection in the false fame paradigm using a modified process dissociation procedure. *Journal of the International Neuropsychological Society, 1(5),* 469-482.

Moffat, N. (1992). Strategies of memory therapy. In B.A. Wilson & N. Moffat (Eds.), *Clinical Management of Memory Problems* (2nd Edition, pp.86-119). San Diego, CA: Singular Publishing Group, Inc.

Penades, R., Boget, T., Lomena, F., Bernardo, M., Mateos, J.J., Laterza, C., Pavia, J., & Salamero, M. (2000). Brain perfusion and neuropsychological changes in schizophrenic patients after cognitive rehabilitation. *Psychiatry Research, 98(2),* 127-132.

Ricker, J.H., Rosenthal, M., Garay, E., DeLuca, J., Germain, A., Abraham-Fuchs, K., & Schmidt, K.U. (2002). Telerehabilitation needs: a survey of persons with acquired brain injury. *Journal of Head Trauma Rehabilitation, 17(3),* 242-250.

Roediger, H.L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning & Memory, 6(5),* 558-567.

Rosenthal, R., & Rosnow, R.L. (1984). *Essentials of Behavioral Research: Methods and Data Analysis.* New York: McGraw-Hill.

Russo, R., Cullis, A.M., & Parkin, A.J. (1998). Consequences of violating the assumption of independence in the process dissociation procedure: A word fragment completion study. *Memory & Cognition, 26(4),* 617-632.

Sattler, J.M. (2001). *Assessment of Children: Cognitive Applications* (4th Edition). San Diego, CA: Jerome M. Sattler, Publisher.

Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501-518.

Schacter, D.L., Chiu, C.Y.P., & Ochsner, K.N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16,* 159-182.

Schacter, D.L., Glisky, E.L., & McGlynn, S.M. (1990). Impact of memory disorder on everyday life: Awareness of deficits and return to work. In D.E. Tupper & K.D. Cicerone (Eds.), *The Neuropsychology of Everyday Life: Assessment and Basic Competencies,* (pp.231-257). Boston: Kluwer Academic Publishers.

Schacter, D.L., Rich, S.A., & Stampp, M.S. (1985). Remediation of memory disorders: Experimental evaluation of the spaced-retrieval technique. *Journal of Clinical and Experimental Neuropsychology, 7,* 79-96.

Schaechter, J.D., Kraft, E., Hilliard, T.S., Dijkhuizen, R.M., Benner, T., Finklestein, S.P., Rosen, B.R., & Cramer, S.C. (2002). Motor recovery and cortical reorganization after constraint-induced movement therapy in stroke patients: A preliminary study. *Neurorehabilitation and Neural Repair, 16(4),* 326-338.

Scruggs, T.E., & Mastropieri, M.A. (1984). Improving memory for facts: The "keyword" method. *Academic Therapy, 20(2),*159-166.

Small, S.A. (2001). Age-related memory decline. *Archives of Neurology, 58(3),* 360-364.

Sohlberg, M.M., & Mateer, C.A. (1989). Training use of compensatory memory books: A three stage behavioural approach. *Journal of Clinical and Experimental Neuropsychology, 11,* 871-891.

Sohlberg, M.M., & Raskin, S.A. (1996). Principles of generalization applied to attention and memory interventions. *Journal of Head Trauma Rehabilitation, 11(2),* 65-78.

Spreen, O., & Strauss, E. (1998). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (2nd Edition). New York: Oxford University Press.

Ste-Marie, D.M., Jennings, J.M., & Finlayson, A.J. (1996). Process dissociation procedure: Memory testing in populations with brain damage. *The Clinical Neuropsychologist, 10(1),* 25-36.

Tate, R.L. (1997). Beyond one-bun, two-shoe: recent advances in the psychological rehabilitation of memory disorders after acquired brain injury. *Brain Injury, 11(12),* 907-918.

Taub, E., Crago, J.E., Burgio, L.D., Groomes, T.E., Cook, E.W., 3rd, DeLuca, S.C., & Miller, N.E. (1994). An operant approach to rehabilitation medicine: Overcoming learned nonuse by shaping. *Journal of the experimental analysis of behavior, 61(2),* 281-293.

Taub, E., & Uswatte, G. (2003). Constraint-induced movement therapy: Bridging from the primate laboratory to the stroke rehabilitation laboratory. *Journal of Rehabilitation Medicine, 41 (Supplement),* 34-40

Warrington, E.K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia, 12,* 419-428.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (3rd Edition)/ Wechsler Memory Scale (3rd Edition) Technical Manual.* San Antonio, TX: The Psychological Corporation.

Wilkinson, G.S. (1993). *WRAT3 Administration Manual.* Delaware: Wide Range.

Wilson, B.A. (1987). *Rehabilitation of memory.* New York, NY: Guilford Press.

Wilson, B.A., Baddeley, A.D., Evans, J., & Shiel, A. (1994). Errorless learning in the rehabilitation of memory impaired people. *Neuropsychological Rehabilitation, 4,* 307-326.

Wittenberg, G.F., Chen, R., Ishii, K., Bushara, K.O., Taub, E., Gerber, L.H., Hallett, M., & Cohen, L.G. (2003). Constraint-induced therapy in stroke: Magnetic-stimulation motor maps and cerebral activation. *Neurorehabilitation and neural repair, 17(1),* 48-57.

Wood, L.E., & Pratt, J.D. (1987). Pegword mnemonic as an aid to memory in the elderly: A comparison of four age groups. *Educational Gerontology, 13(4),* 325-339.

Table 1. Outline of study visits.

	Cognitive Testing	**Training**	**Additional Procedures**
Screening	• CVLT • WRAT3 Reading • Block Design • Vocabulary		• Medical History
Baseline	• Full Cognitive Battery		
		1-Week Inactivity	
Pretraining	• Full Cognitive Battery		
Days 1-7		• 4 sessions per day	
Posttraining	• Full Cognitive Battery		• Metacognitive measures • Debriefing

Table 2. Demographic information on study participants.

	Experimental	Control
Age	46.2 (11.8)	19.3 (2.7)
Minimum	28	17
Maximum	59	26
Gender	5 male, 5 female	2 male, 7 female
Education	12.43 (2.57)[a]	13.22 (.83)
CVLT performance		
Learning total (T-score)	37.7 (10.65)	57.0 (5.98)
Trial 1 (z-score)	-1.1 (.96)	-.33 (.90)
Delayed recall (z-score)	-1.4 (.91)	.39 (.60)
Age at Injury	35 years (15.4)	.
Minimum	15	
Maximum	55	
Time post Injury	11.5 years (11.4)	.
Minimum	2.1	
Maximum	38.9	
Pending Litigation	2 participants	.
FSIQ$_{EST}$	95 (12)	119 (14)
WRAT3 Reading	93 (9)	110 (8)

Note. Numbers provided are mean (*SD*); Age, Education, Age at Injury, and Time post Injury all provided in years; Learning total = total words recalled across learning trials; Delayed recall = delayed free recall; FSIQ$_{Est.}$ = Estimated Full Scale IQ; WRAT3 Reading = Standard Score on WRAT3 Reading subtest

a. Missing education on three experimental participants.

Table 3. Regression of highest lag achieved onto session number for each subject.

Subject #	Total Difference	B	*P*	R	R^2
12	2	.101	<.001	.908	.825
11	3	.096	<.001	.884	.781
16	7	.184	<.001	.900	.810
7	11	.394	<.001	.955	.913
8	11	.336	<.001	.891	.795
17	13	.436	<.001	.899	.809
14	22	.746	<.001	.963	.927
13	38	1.84	<.001	.970	.941
15	38	1.68	<.001	.971	.943
20	38	1.89	<.001	.961	.924

Note: difference scores provided represent highest lag on day 7 (session 28) minus highest lag on day 1 (session 3)

Table 4. Summary of training data as compared to Jennings and Jacoby (2003).

	Aging[a]	ABI
Day 1 (session 3)	1.92(0.66)	1.70 (.67)
Day 7 (session 28)	27.92(16.77)	20.00 (14.97)
Difference	26	18.3
# Improved	9[b]	8[b]
# Unimproved	3[b]	2[b]

Note: Numbers provided for day 1, day 7, and difference are mean (*SD*)

a. Aging data from Jennings and Jacoby, 2003

b. Fisher's exact test, $p = .594$

Table 5. Mean level of performance for each group.

Task	Session #	Group	Mean (SD)
<u>False Fame</u>			
- Old nonfamous	1	Experimental	10.00 (6.19)
		Control	3.89 (3.92)
	2	Experimental	9.25 (6.52)
		Control	2.44 (2.19)
	3	Experimental	5.00 (5.10)
		Control	2.78 (1.72)
- New nonfamous	1	Experimental	7.50 (5.90)
		Control	15.33 (6.26)
	2	Experimental	8.75 (5.52)
		Control	13.44 (6.77)
	3	Experimental	5.50 (4.60)
		Control	12.89 (6.57)
<u>Buschke SRT</u>			
- Learning	1	Experimental	88.37 (23.86)
		Control	121.00 (8.69)
	2	Experimental	92.13 (18.39)
		Control	124.89 (12.57)
	3	Experimental	88.88 (21.91)
		Control	132.78 (5.61)
- Delay	1	Experimental	7.00 (3.02)
		Control	11.56 (.73)
	2	Experimental	7.13 (3.40)
		Control	11.78 (.44)
	3	Experimental	7.50 (2.73)
		Control	11.33 (.71)

Note: Experimental improved only (*N*=8) versus control (*N*=9)

Table 6. Results of analyses of other measures of memory and attention.

Letter-Number Sequencing: Total Score

Descriptives			Statistics				
Session	Group	Mean (SD)		*F*	df	*p*	η^{2a}
1	Experimental	9.38 (3.11)	Main Effect:	1.48^b	$(1.78, 17)^b$	$.24^b$	$.09^b$
	Control	13.22 (2.64)	Interaction:	1.36^b	$(1.78, 17)^b$	$.27^b$	$.08^b$
2	Experimental	9.00 (2.33)					
	Control	13.33 (3.00)					
3	Experimental	10.75 (2.60)					
	Control	13.33 (2.55)					

Source Monitoring: Item Recall

Descriptives			Statistics				
Session	Group	Mean (SD)		*F*	df	*p*	η^{2a}
1	Experimental	17.50 (6.61)	Main Effect:	.003	(2,17)	.99	.00
	Control	23.11 (3.18)	Interaction:	1.76	(2, 17)	.19	.11
2	Experimental	18.00 (5.78)					
	Control	22.56 (4.07)					
3	Experimental	16.00 (3.30)					
	Control	24.44 (3.40)					

Source Monitoring: Source Recall

Descriptives			Statistics				
Session	Group	Mean (SD)		*F*	df	*p*	η^{2a}
1	Experimental	9.50 (4.93)	Main Effect:	.814	(2,17)	.45	.05
	Control	18.67 (3.32)	Interaction:	1.99	(2,17)	.15	.12
2	Experimental	10.25 (4.33)					
	Control	20.33 (5.79)					
3	Experimental	8.63 (1.85)					
	Control	21.89 (4.91)					

CPT: Errors of Commission

Descriptives			Statistics				
Session	Group	Mean (SD)		*F*	df	*p*	η^{2a}
1	Experimental	9.00 (6.11)	Main Effect:	2.49	(2,16)	.10	.15
	Control	5.00 (3.57)	Interaction:	.836	(2,16)	.44	.06
2	Experimental	6.00 (3.65)					
	Control	4.22 (3.23)					
3	Experimental	5.71 (3.82)					
	Control	4.11 (4.59)					

CPT: Variability Index

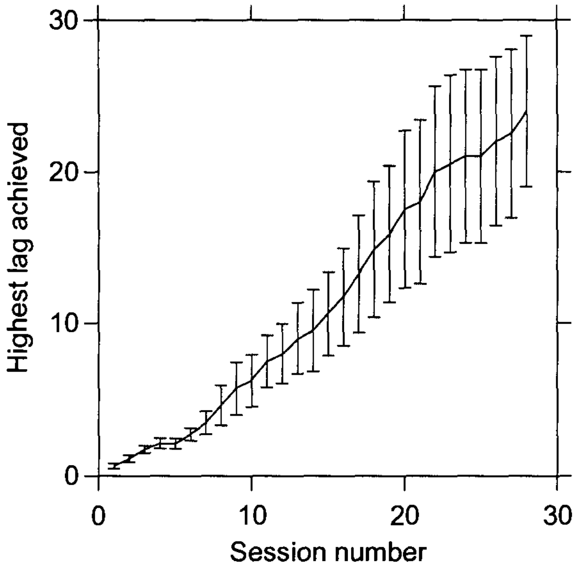
Descriptives			Statistics				
Session	Group	Mean (SD)		*F*	df	*p*	η^{2a}
1	Experimental	7.23 (2.49)	Main Effect:	$.844^b$	$(1.54, 16)^b$	$.42^b$	$.06^b$
	Control	5.93 (2.08)	Interaction:	1.49^b	$(1.54, 16)^b$	$.25^b$	$.10^b$
2	Experimental	5.91 (2.23)					
	Control	5.65 (2.51)					
3	Experimental	7.31 (4.82)					
	Control	4.92 (2.35)					

Note: Results reported for improved experimental (*N*=8) and controls (*N*=9); Source Monitoring: Item Recall = number of items correctly recalled/30; Source Monitoring: Source Recall = number of sources correctly recalled/30; Data missing for one experimental on CPT (migraine); CPT = continuous performance test; Errors of commission reported in raw values; Variability = variability in reaction times

a. Partial η^2

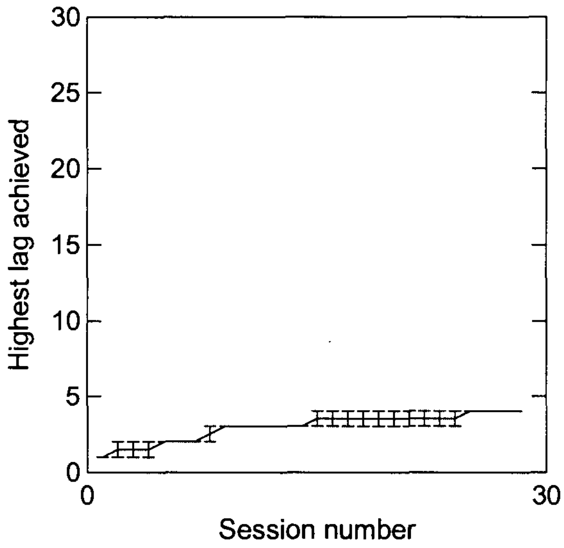
b. Greenhouse-Geisser corrected values

Figure 1. Performance over time of improved sample.



Note: $N = 8$

Figure 2. Performance over time of unimproved sample.



Note: $N = 2$

Figure 3. Number of old nonfamous names falsely identified as famous.

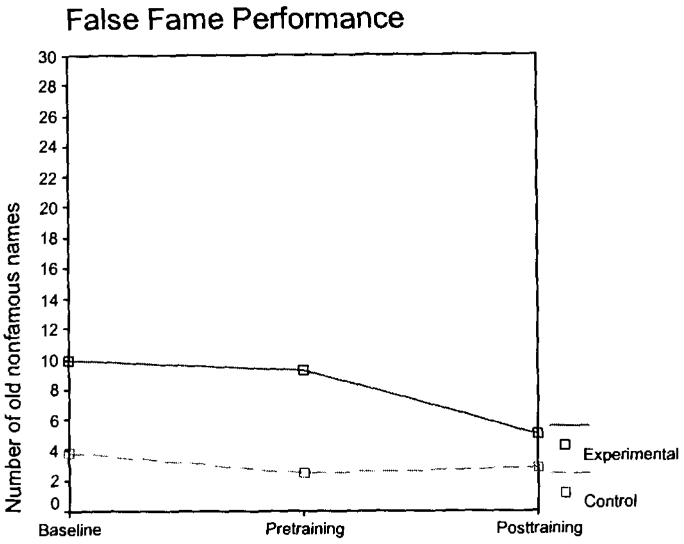


Figure 4. Learning performance on the Buschke SRT.