Errorless Learning and Spaced Retrieval in Cognitive Rehabilitation:  
A Contrast of Principles  

Erica L. Middleton, PhD  
Myrna F. Schwartz, PhD  

Introduction  
In neuropsychology research, there is increased exploration of errorless learning (EL) treatments, where the guiding principle is that errors committed by participants during treatment can be learned. Thus, the priority of these approaches is to eliminate errors during treatment. In the most common form of EL training in cognitive rehabilitation, errors are avoided by preventing a participant from attempting to retrieve target responses from long-term memory. For example, in EL training of naming impairments in aphasia (Conroy, Sage, & Lambon Ralph, 2009b; Fillingham, Sage, & Lambon Ralph, 2005a, 2005b, 2006; McKissock & Ward, 2007), the target name for a depicted object is provided on each training trial (the word is seen/heard for repetition) in a manner that preempts attempts at retrieving the name from long-term lexical memory. Investigations of the errorless learning approach in neuropsychology originally focused on patients with severe anterograde amnesia, but more recently investigations have extended to other memory-impaired populations (e.g., Alzheimer’s disease) and well as acquired aphasia.  

Errorless learning techniques typically achieve errorless performance by eschewing retrieval practice, practice with retrieving target information from long-term memory. However, this runs counter to core principles derived from psychological research on testing and spacing effects in learning and memory: (1) powerful learning arises when individuals are provided opportunities to retrieve target information from long-term memory (see Roediger & Karpicke, 2006a for review). Specifically, retrieval practice changes memory in ways that makes retrieval of the same information in the future more likely; (2) the benefits from retrieval practice to long-term performance are maximized when repeated retrieval attempts of items are distributed or spaced over time rather than massed (i.e., spacing effect); (3) the synergy of these phenomena is understood to reflect the importance of effortful retrieval for ensuring long-term success. This paper provides a brief discussion of these core learning principles and how they bear on errorless learning research in amnesia and aphasia. Discussion focuses on
the scope and relevance of error learning in these domains, particularly when weighed against the potentially great therapeutic value of retrieval practice.

**Testing and Spacing Effects in Learning and Memory**

A host of studies demonstrate that retrieving information from long-term memory changes memory, bolstering successful retrieval in the future (e.g., Allen et al., 1969; Carpenter & DeLosh, 2005; Carrier & Pashler, 1992; Hogan & Kinstch, 1971; Izawa, 1970; Jacoby, 1978; Karpicke & Roediger, 2007a; 2007b; 2008; Roediger & Karpicke, 2006b; Slamecka & Graf, 1978; Wheeler and Roediger, 1992). An early demonstration of the learning that arises from retrieval practice showed that retrieving a target from memory (e.g., retrieve “COLD” given “HOT-_______”) rather than reading a target (e.g., read “COLD” in “HOT-COLD”) increased the likelihood of successfully retrieving that target later (a “generation” effect; Slamecka & Graf, 1978). In the testing literature, empirical demonstrations of retrieval practice effects begin with an initial study period, giving participants an opportunity to commit target information to memory. This is followed by further study opportunities or tests in lieu of more study, where participants practice retrieving target information. Retrieval practice effects are demonstrated when training involving tests outperforms training involving more study on final measures of performance. To illustrate, in Karpicke and Roediger (2007b) the task was learning a list of 40 words. During training, the list was studied (S) three times followed by a recall test (T) of the list (SSST), or a study opportunity was replaced by a test opportunity (STST). Each sequence (e.g., SSST) was repeated five times during training. On a final recall test one week later, recall in STST (68%) outperformed recall in SSST (57%) even though the word list was studied five more times in the SSST condition. In a variant of this design (Karpicke and Roediger, 2008) final performance after 1-week for training involving tests (80%) greatly outperformed training involving more study (36%), demonstrating the powerful effects of retrieval practice.

Retrieval practice effects are maximized according to the schedule of the retrieval opportunities, with distributed (i.e., spaced) schedules consistently outperforming massed schedules (e.g., Balota et al., 2006; Carpenter & DeLosh, 2005; Cull, 2000). To illustrate, in Karpicke and Roediger (2007b) the task was learning to associate vocabulary words and their definitions (e.g., sobriquet—nickname). Training started with an initial study trial followed by three retrieval practice attempts (cued-recall) administered according to different schedules: (1) massed (0-0-0), the study trial was
followed immediately by three successive tests; (2) fixed distributed, five items intervened between the study trial and subsequent tests (5-5-5). Though during training the massed condition outperformed fixed, final cued recall performance assessed after 2-days was 60% (fixed) versus 19% (massed). Thus, the impact of retrieval practice on delayed measures is strongly influenced by the schedule with which retrieval for an item is repeatedly practiced. Literally hundreds of studies—almost without exception—have demonstrated that spacing retrieval practice over time outperforms non-spaced (massed) schedules, particularly when critical performance measures are administered after a delay (i.e., spacing effect; see Cepeda et al., 2006 for review).

At odds with the priority of errorless treatments, the testing and spacing literature would suggest that eliminating conditions that usually lead to non-trivial error rates (i.e., effortful retrieval states) omits powerful learning experiences. In fact, the general understanding in the spaced retrieval literature is that conditions that increase difficulty and lower performance during treatment generally promote the best long-term performance (see Schmidt & Bjork, 1992, for discussion). A compelling illustration of this was provided by Pashler et al. (2003). In that study, the task was learning to associate novel vocabulary words (Eskimo) and their English translation. On Day 1, participants encountered each item three times, first for study, followed by two retrieval attempts (Test1 and Test2) where the Eskimo word was presented and participants attempted to generate the English word (i.e., cued-recall). When retrieval failed they were provided the English word as feedback. Test1 always followed initial study after 2 intervening items. The key experimental manipulation involved the lag between Test1 and Test2 (Test1-Test2 lag), which was 2, 4, 8, 16, 24, or 32 items. One day later, performance on all items was assessed in a final cued-recall test.

A predictable outcome was that as Test1-Test2 lag increased, performance on Test2 decreased. The remarkable finding was that on the final cued-recall test, performance increased as a function of Test1-Test2 lag the previous day. This demonstrates the operation of effortful retrieval—though performance on Test2 at Day 1 steadily decreased as a function of Test1-Test2 lag (increased lag=increased difficulty), the effect of Test1-Test2 lag reversed on the final cued-recall test (increased difficulty=increased retention). In a second experiment, Pashler et al. increased the Test1-Test2 lag to 96 items. Long lags still outperformed short lags on the final-cued recall test, but final-cued recall began to drop slightly around a lag of 64 items. Pashler et al.’s methods illustrate a means to experimentally diagnose an optimal lag or “sweet
spot” (a notion revisited later). This is where retrieval is maximally difficult to ensure the best long-term retention, but not so difficult that retrieval failure rate during training begins to offset return on final performance measures. Overall, this work is a compelling demonstration of the importance of effortful retrieval for the long-term persistence of training effects. Hints of the operation of these powerful learning principles are apparent in errorless learning investigations of amnesia, the next topic of discussion.

**Errorless Learning in Amnesia**

Most neuropsychological studies of errorless learning have focused on individuals with amnesia resulting from acquired brain injury or neurodegenerative conditions, including dementia of the Alzheimer type (DAT; see Clare and Jones, 2008, for review on errorless learning in amnesia). Amnesia, here, refers to an impairment within long-term, explicit, episodic memory that impedes the conscious recall of new information.

Early work on EL treatments in amnesia was conducted by Baddeley and colleagues (Baddeley & Wilson, 1994; Wilson, Baddeley, Evans, & Shiel, 1994). In a seminal study by Baddeley and Wilson (1994) people with severe amnesia studied words under EL conditions (e.g., I'm thinking of a word AR_ _ _ _ and it's ARTIST) or errorful (EF) conditions. Under EF conditions, the participant was encouraged to generate guesses about the target (e.g., ARCHES) before being given the answer. In this and many future studies incorporating these methods, EF training was designed to ensure errors, for example, by switching the target if it was guessed in initial trials (e.g., participant generates the correct response “ARTIST” but they are told the target was “ARCHES”). Baddeley and Wilson found that people with amnesia demonstrated uniformly worse performance on stem completion tests after EF training compared to EL training. The authors attributed the result to the relatively intact implicit memory and dysfunctional explicit memory systems in amnesia. Whereas implicit memory enabled the learning of errors and correct training responses, impaired explicit memory prevented their participants from differentiating errors from correct responses at final test.

Since this early work, a number of additional studies—also focusing on participants with severe amnesia—have reported an EL advantage (Hunkin et al., 1998; Page et al., 2006; Riley & Heaton, 2000; Squires et al., 1997; Wilson et al., 1994). However, a number of caveats regarding the superiority of EL techniques in memory-impaired populations has emerged, which have to do with how memory is probed, the severity of memory impairment, and the longevity of treatment effects. Firstly, EF
methods can produce comparable (or better) learning, relative to EL methods in patients with mild to moderate memory impairment, such as those in the early stages of DAT (Bier et al., 2008; Dunn & Clare, 2007; Haslam et al., 2010; Metzler-Baddeley & Snowden, 2005; Riley et al., 2004). Secondly, an EF advantage is often demonstrated in tasks that go beyond implicit learning (i.e., tasks requiring explicit recall of information such as sequential information required for navigating a route; Evans et al., 2000; Kessel, van Loon, & Wester, 2007; Riley et al., 2004). Thirdly, EF conditions may not have fared well in the early studies on errorless learning in amnesia because standard EF conditions were artificially errorful. That is, they often incorporated elements that specifically increased the rate of errors and would not realistically be adopted by clinicians: (1) a standard method was to switch the correct response for an alternative if the participant managed to guess correctly (at least on initial trials) to ensure EF training conditions induced errors; (2) participants generally were encouraged to guess and produce a minimum number of errors (e.g., 3 per trial) when otherwise they might have refrained from responding. Lastly, there is very little data on the persistence of treatment effects after EL and EF training methods, with some indication that EF methods may be more robust against forgetting (Hunkin et al., 1998; Squiers et al., 1997). For example, Hunkin et al., (1998) reported significant forgetting on treatment targets across a 48-hour delay trained with EL methods but not with EF methods.

In the literature on errorless learning, there is a growing number of studies showing that memory-impaired individuals benefit from regular opportunities to retrieve target information from long-term memory (Dunn & Clare, 2007; Laffan et al., 2010; Tailby & Haslam, 2003). For example, compared to standard EL treatment, Tailby and Haslam (2003) found a greater treatment benefit when they supplemented the traditional stem-completion training task with a cue to assist retrieval (e.g., “I’m thinking of a five letter word that begins with BR and it’s a food made from flour, yeast, liquid and is baked...”). Subsequent studies have confirmed the functional relevance of retrieval practice in memory impaired populations, particularly when its effects are enhanced by efficacious schedules of learning. In Sumowksi, Chiaravalloti, and DeLuca (2010), individuals with memory impairment from multiple sclerosis enjoyed superior treatment benefits after spaced testing versus spaced study, demonstrating the importance of retrieval practice in this population. In two studies with Alzheimer’s patients, Balota et al., (2006) found an advantage for fixed-distributed and expanding test schedules over massed test schedules, demonstrating the importance of distributing retrieval practice
over time. Finally, in a direct comparison between training involving spaced retrieval practice and standard errorless learning, Haslam, Hodder, and Yates, (2011) found that an expanding test schedule reliably outperformed errorless learning on final performance measures in memory-impaired individuals (from acquired brain injury and dementia). Overall, the value of prioritizing errorless learning experiences during treatment is not clear, particularly when weighed against the therapeutic potential of (effortful or spaced) retrieval practice. We now turn to a brief discussion of errorless learning interventions in aphasia. As with amnesia, here we also find important limitations to the scope of error learning effects.

Errorless Learning in Aphasia

Aphasia is a primary disorder of spoken and written language. In aphasia, a ubiquitous complaint is difficulty in producing individual words, particularly content words such as nouns and verbs. Such naming problems can manifest as sound or word errors (e.g., “sild” is produced for the target “shield,” or “cat” for “dog”), exceptionally long production latencies, or a failure to produce any response (i.e., omission). The growing interest in errorless methods in aphasia is motivated by the possibility that the act of attempting and failing to produce accurate speech is a learning event in its own right, such that similar errorful outcomes in the future are made more likely.

The possibility of error learning in speech has profound implications for aphasia and its treatment. Not only can aphasic speech be rife with difficulty and errors, the canonical form of treatment for naming problems in aphasia has the patient endeavoring to name single or multiple items from pictures, a scenario which can repeatedly induce stymied or errorful productions. Indeed, recent work with healthy speakers suggests that errorful retrieval states leading to tip-of-the-tongue (Warriner & Humphreys, 2008) or sound errors (Humphreys, Menzies, & Lake, 2010) may be learned (i.e., errors are self-reinforcing). However, experimental comparisons of errorless and errorful treatments in aphasia have revealed no systematic detriment from making errors during treatment, with both methods proving efficacious for the rehabilitation of naming impairments (Conroy et al., 2009b; Fillingham et al., 2005a, 2005b, 2006; McKissock & Ward, 2007).

Briefly, the canonical errorless (EL) intervention for picture naming involves presenting an object name for repetition in the presence of the depicted object. This contrasts with errorful (EF) approaches, which permit errors and typically involve confrontation naming or naming assisted by onset cueing. Though both methods are
efficacious, there is a variety of evidence suggestive of a role for spaced retrieval practice in impacting treatment outcomes. Firstly, in a meta-analysis where the authors classified existing word retrieval treatments as EL or EF (Fillingham et al., 2003), EF treatments were generally associated with greater generalization and longer retention of treatment effects. Secondly, EF methods fared at least as well as EL methods in a number of studies even though EF methods were set at a disadvantage from lack of feedback (Fillingham et al., 2005a; 2005b; cf. McKissock & Ward, 2007). In Fillingham et al., (2005a), for example, each item was trained 90 times—in EL training the object name was provided with the picture for repetition, whereas the object name was never provided in EF treatment (only the onset of the name was provided). Thirdly, in an experimental contest where EL methods were compared to an EF method that would be expected to create high rates of successful yet effortful retrieval practice (Fillingham et al., 2006), the EF method was associated with greater retention of treatment benefits. In that study, the EF method involved accumulating cueing, where on each training trial failure to name was followed by cues of increasing specificity until naming was achieved (e.g., first phoneme/letter provided → first syllable provided → full word provided for repetition). On immediate post-therapy measures both methods generally fared equivalently. However, on long-term measures, accumulating cueing fared better, with nine participants demonstrating reliable long-term retention of treatment gains, versus only five for errorless learning. In sum, there are patterns of results in research on errorless learning in aphasia that are suggestive of greater therapeutic potential of EF methods versus EL methods, particularly when designed to induce successful yet effortful retrieval practice.

Synthesis and Conclusion

Retrieval practice effects and spacing effects are among the most powerful and robust phenomena in cognitive psychology, proving almost universal in their relevance to performance in countless domains and across a broad range of populations. In light of this knowledge, an important strategy for future research on errorless learning is to design well-controlled experimental contests between EL and EF methods. However, it is of critical importance to compare EL methods against methods incorporating retrieval practice when its effects are maximized. To illustrate, in the context of naming treatments in aphasia, for an individual or group of individuals with comparable naming impairment severity, an optimal lag to guide the construction of spaced schedules of learning could be derived by adapting the Pashler et al., (2003) paradigm described.
above. For example, pilot work could establish the optimal lag between an initial repetition trial and a subsequent test (naming, without any cues) that was associated with the best long-term performance on final naming measures. Once identified, in a second phase of research, the optimal lag would be incorporated into a distributed schedule of learning, with errorless training events (i.e., repetition) and retrieval practice events (e.g., naming test) programmed according to the same schedule. The key contrast would involve a comparison of average naming success on items assigned to the EL or retrieval practice training conditions. It would be important to assess naming success at different retention intervals, as the merits of retrieval practice methods are most transparent in terms of the long-term retention of treatment effects.

To conclude, a great deal more research is needed to determine if, for whom, and to what extent errorless learning approaches in cognitive rehabilitation should be prioritized by clinicians. The imperative of future investigations of errorless treatments should be to evaluate their long-term efficacy against “errorful” methods designed to maximize the robust learning effects of spaced retrieval practice.

References


