Aging, Neuropsychology, and Cognition, 2015
Vol. 22, No. 2, 244–258, http://dx.doi.org/10.1080/13825585.2014.915918

Repetition-lag training to improve recollection memory in older people with amnestic mild cognitive impairment. A randomized controlled trial

Maurice Finn[a,b]* and Skye McDonald[b]

[a]*Aged Care Services, Royal North Shore Hospital, University of New South Wales, St Leonards, Australia;* [b]*School of Psychology, University of New South Wales, St Leonards, Australia*

(*Received 24 September 2013; accepted 12 April 2014*)

The results of a randomized controlled trial of repetition-lag training in older adults with amnestic mild cognitive impairment (aMCI) are reported. A modified repetition-lag training procedure with extended encoding time and strategy choice was used. The training required discriminating studied words from non-studied lures that were repeated at varying intervals during the test phase. Participants were assessed pre/post using untrained measures of cognition and self-report questionnaires. Primary outcome measures were recall of unrelated word pairs both immediately following presentation and following a delay. Secondary outcomes were a measure of attention, cognitive flexibility, and visual working memory. Participants were also asked to report on the frequency of cognitive failures and mood before and after training. Participants ($N = 31$) were randomized into either the treatment or a no-contact control group and attended the clinic twice per week over a four week period. Twenty-four participants completed the study (twelve in each group). Results indicated that the training group improved at recalling unrelated word pairs after a delay. There were no significant effects of training on other outcomes, self-reported cognitive failures or mood. The results are discussed along with suggestions for future research.

Keywords: elderly; cognitive training; repetition-lag training; amnestic mild cognitive impairment; randomized

Amnestic mild cognitive impairment (aMCI) involves significant impairments on objective measures of memory compared with age-matched peers but with intact activities of daily living (Petersen, 2004). Impairments can be restricted to memory alone or include deficits in other areas of cognition such as visuospatial ability, executive functions and language. The risk of conversion to dementia is high in this group (Ravaglia et al., 2006), which underscores the need to find ways to improve or maintain memory functioning. Attempts to improve memory in aMCI often involve teaching memory strategies (Troyer, Murphy, Anderson, Moscovitch, & Craik, 2008), an approach which may be well suited to a particular recall task, i.e., remembering the name of an acquaintance (Hampstead, Sathian, Moore, Nalisnick, & Stringer, 2008). While these interventions can be effective in providing a strategy for the intended situation, they may be limited in scope and rely on an unreliable memory for implementation. An alternative approach seeks to remediate underlying memory processes that are impaired in aMCI and which may contribute to

*Corresponding author. Email: maurice.finn@health.nsw.gov.au
This article was originally published with errors. This version has been corrected. Please see Erratum (http://dx.doi.org/10.1080/13825585.2014.929366).

memory failures in a range of situations. By identifying these processes, it may be possible to provide training interventions that are both more efficient and more likely to transfer to other recall situations that rely on the trained process.

There is an extensive literature on memory processes in both normal elderly and in memory-impaired groups. One area of investigation, dual-process theory (Jacoby, 1991; Mandler, 1980) has drawn a distinction between controlled (conscious, effortful) and automatic (unconscious, implicit) processes. The recognition memory literature has further distinguished between the automatic process "familiarity" (for example, knowing you have seen someone before without conscious recollection as to where) and a controlled process known as "recollection" (retaining memory of the contextual information, e.g., they were the person who sold you tickets to a concert). There is evidence that normal elderly, aMCI and Alzheimer's disease groups (Adam, Van der Linden, Collette, Lemauvais, & Salmon, 2005; Jennings & Jacoby, 1997; Smith & Knight, 2002) are less able to engage processes underlying recollection, instead tending to rely more on familiarity. This suggests that recollection might be a suitable target for remediation via systematic training that encourages both deeper encoding and retention of this information over a period of time.

Repetition-lag training

One method of training recollection is repetition-lag training (RLT), a form of recognition memory training that encourages recollection over mere familiarity. The RLT procedure requires the learning of a series of words and subsequent discrimination of those words from unstudied lures. The task differs from a straightforward recognition task in that the lures are repeated at varying intervals known as lags, and the length of the lag increases as performance improves.

Initial studies (Jacoby, Jennings, & Hay, 1996; Jennings & Jacoby, 2003) using RLT involved normal elderly and suggested that problems in recollection on the RLT were (1) related to self-reported memory complaints, (2) that recollection could be improved with training, and (3) the training effect was driven by incremented difficulty. Incremented difficulty occurs when the lags are systematically increased once a criterion level of performance is reached. This progressive increase in task difficulty is thought to force participants to develop and/or practice strategies to recollect at longer intervals. Jennings and Jacoby (2003) found that when intervals were systematically increased in this way, recollection improved more than when the intervals were varied randomly.

Transfer of training

There is encouraging evidence of transfer following RLT. For example, Jennings, Webster, Kleykamp, and Dagenbach (2005) reported gains on n-back task (a measure of working memory), the Source discrimination task (visual long-term memory), Digit-symbol substitution (processing speed), and the Self-ordered pointing test (non-spatial executive working memory). In another study using a modified form of RLT (Lustig & Flegal, 2008) gains were found on Trails B (cognitive flexibility) along with a reduction in self-reported memory errors as measured by the Everyday Memory Questionnaire. In a group with Alzheimer's disease (Boller, Jennings, Dieudonné, Verny, & Ergis, 2012) transfer gains were seen on an n-back task (working memory), RL/R1 (episodic recognition memory), DMS 48 task (visual recognition memory), and the source recognition task (recognition memory). One conference report found gains in a group with aMCI on the

California Verbal Learning Test – Second Edition (CVLT-II) measures of immediate and delayed, free and cued recall (Jennings et al., 2006). Encouragingly, the transfer effects observed in these studies do not appear to be due to expectancy effects or social contact, as active control groups completing a recognition practice task did not show similar gains. Potential transfer gains in aMCI following RLT are worthy of further investigation, given that recollection is impaired in this population and that successful training may have beneficial effects across a range of recall situations encountered in everyday life.

Optimizing RLT for older adults

Under the original RLT procedure, encoding time is limited to 2 seconds per word. Bissig and Lustig (2007) found that under a modified RLT procedure, with open-ended encoding time, the time spent on encoding was the major determinant of performance rather than factors such as age or crystallized intelligence. This appeared to be due to how the time at encoding was used. Participants who spontaneously utilized a self-generated strategy to aid learning performed better than those who did not or those who simply relied on repetition. The observation that a self-generated memory strategy may improve performance, echoes work by Derwinger, Stigsdotter-Neely, MacDonald, and Backman (2005), where elderly participants who came up with their own strategy retained numerical information for longer periods compared with peers who were given a specific mnemonic strategy to aid recall. This finding was subsequently tested by Lustig and Flegal (2008) who increased encoding time to 14 second, and instructed participants to either make a sentence out of each word, or to think about the meaning of the word, without offering a specific strategy. They found that allowing participants to choose their own strategy during extended encoding aided performance compared with being instructed on which strategy to use. Given this, it seems reasonable to suggest that optimizing performance of an aMCI sample on the RLT procedure could involve encouraging participants to think of some way to remember the words (while leaving the exact method for doing so open) and extending the encoding period for each word to allow sufficient time for this to occur.

Taken together, the existing research literature suggests that RLT targets a process known to be impaired in aMCI, and that there is some evidence of transfer following training. The adapted procedure with extended encoding time allows time for participants to generate their own strategy. However, these effects have yet to be tested in an aMCI group despite calls to do so (Anderson et al., 2008; Ranganath, Flegal, & Kelly, 2011; Shumaker, Legault, & Coker, 2006).

Aims

(1) To assess the ability of aMCI participants to complete a RLT program modified to allow extended encoding time and strategy choice. It was hypothesized that aMCI participants would be able to increase their lag level across the course of training. To ensure that improvements on the training task did, in fact, reflect improvements in recollection it was also hypothesized that there would be no change in either accuracy or response bias across training. If recollection alone is improving then accuracy (proportion of "hits," i.e., saying yes to a studied word; minus proportion of false alarms, responding "yes" to a new word) should not change across training. Such a finding would suggest that improvements during training are due to a change in the ability to correctly discriminate studied words from

repeated lures (recollection). Changes in response bias could indicate that participants simply learned to respond "no" more often without truly recollecting if the word was from the studied list or a new word.

(2) To measure transfer gains in recollection following training. It was hypothesized that aMCI participants would show improved recollection following RLT. Recollection was operationalized as recall of unrelated word pairs, both immediately following presentation and again after a delay. The Verbal Paired Associates Test (from the Wechsler Memory Scale – Fourth Edition, Wechsler, 2009) contains both related ("sock-shoe") and unrelated word pairs ("laugh-stand"). This task is known to be sensitive to the kinds of memory impairment seen in aMCI (Perri, Carlesimo, Serra, & Caltagirone, 2005), and recall of unrelated word pairs may be a more sensitive measure of recollection than related word pairs (Amieva, Rouch-Leroyer, Fabrigoule, & Dartigues, 2000).

(3) To extend findings by other researchers that RLT may improve working memory (measured here by Symbol Span) and cognitive flexibility (Trails B). Finally, we explored the effect of training on self-reported everyday cognitive failures (Cognitive Failures Questionnaire) and on mood (DASS-21). It was thought that RLT may result in reduced frequency of self-reported cognitive failures. A measure of mood was included to gauge the impact (either positive or negative) of RLT on mood.

Method

Participants

Community dwelling clients of Aged Care Services at Royal North Shore Hospital who had been assessed for memory problems in the Memory and Geriatric Speciality Clinics in the previous 12 months were nominally eligible to participate. These clinical assessments typically involved a physical exam, clinical history, detailed neuropsychological testing and cerebral imaging. Participants were excluded due to reasons such as having a diagnosis other than aMCI, co-morbid psychiatric problems, having moved outside the boundaries of the area health service, progression to dementia, or recent commencement on a cholinesterase inhibitor. Participants gave written informed consent in accordance with HREA requirements. As aMCI is an unstable diagnosis, to ensure that all participants continued to meet the criteria for aMCI the remaining potential participants were further screened as per the following criteria:

(1) Diagnosis of amnestic MCI (single or multiple domain). Participants with non-amnestic MCI were not eligible to participate. MCI was determined using standardized diagnostic criteria (Winblad et al., 2004). All participants had a recent (<12 months old) neuropsychological evaluation that indicated objective impairments in memory (at least 1.5 *SD* below age and education related normative data).

(2) Intact global cognitive functioning (score >23) on the Mini Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975); absence of untreated psychiatric illness or substance abuse problems; and absence of visual, auditory or motor impairment that would hinder use of a computer. Participants who were on stable doses of cholinesterase inhibitors for the duration of the study

were eligible to take part, but participants who commenced taking these medications during the study were excluded.

Participants with both aMCI Single domain (n = 8) and aMCI – Multiple domain (n = 23) were included in the study. Demographic and neuropsychological data for the 31 participants who were originally recruited into the study are shown in Table 1, as well as data for those who completed the study. Baseline comparisons (t-tests) revealed no significant differences between the treatment and control group on these variables.

Twenty-four participants completed the study, twelve in each arm, with four treatment and three control group participants either withdrawing or being excluded. The four withdrawals all occurred after randomization but prior to commencing the study. The primary reasons for withdrawal were unrelated medical or personal issues. Those who withdrew did not differ from those who completed the study in terms of age, sex, education, or depressive symptoms. In the treatment group, two participants were excluded, one because his memory functioning was not sufficiently impaired to meet the study criteria, despite a clinical presentation in keeping with aMCI. Another participant was excluded as she was unable to complete the training within the five week time-frame permitted. In the control group, one participant was excluded due to a diagnosis of non-amnestic MCI. The flow of participants through the study to the posttest assessment is shown in Figure 1. Data collection took place over the period July 2012 to May 2013.

Table 1. Demographic and neuropsychological data for entire sample and for each group.

	Groups				
	Whole sample (N = 31)	Treatment (n = 12)	Control (n = 12)	T-test	
Gender	9 female, 22 male M (SD)	4 female, 8 male M (SD)	3 female, 9 male M (SD)	T (22)	P
Age (years)	75.19 (6.3)	72.83 (5.7)	75.08 (7.5)	−0.824	0.419
Education (years)	13.71 (3.0)	13.75 (2.8)	13.67 (3.8)	0.061	0.952
MMSE (/30)	28.10 (1.5)	27.75 (1.3)	27.83 (1.9)	−0.122	0.904
Information	12.71 (1.1)	13.17 (0.6)	12.33 (1.5)	1.738	0.096
Logical memory I	21.39 (6.9)	21.50 (6.6)	21.33 (5.1)	−0.069	0.946
Logical memory II	6.81 (6.1)	7.42 (7.1)	5.58 (4.4)	0.761	0.455
LM recognition	14.81 (4.4)	14.44 (4.0)	15.20 (4.9)	−0.473	0.640
Visual reproduction I	27.03 (6.7)	28.50 (5.9)	27.58 (7.41)	0.334	0.741
Visual reproduction II	10.32 (6.7)	11.08 (6.0)	9.42 (9.4)	0.517	0.610
VR recognition	4.29 (1.6)	4.13 (1.5)	4.47 (1.8)	−0.567	0.575
Digits forward	9.77 (1.8)	9.53 (1.9)	10.00 (1.8)	−0.677	0.504
Digits backward	7.57 (1.5)	7.40 (1.6)	7.73 (1.4)	−0.583	0.564
Vocabulary	41.55 (9.6)	40.07 (9.1)	42.93 (10.2)	−0.792	0.436
Block design	32.00 (10.6)	33.44 (10.7)	30.47 (10.7)	0.768	0.448
Symbol digit	34.43 (10.4)	36.20 (11.7)	32.67 (9.1)	0.922	0.365
Naming	50.97 (9.6)	50.81 (10.4)	51.13 (9.0)	−0.091	0.928
Visual scanning	30.43 (8.7)	30.38 (8.9)	30.47 (8.9)	−0.024	0.981
Color naming	36.81 (5.9)	37.20 (6.2)	36.33 (5.7)	0.371	0.714
Word reading	23.96 (3.1)	24.07 (3.2)	23.83 (3.0)	0.190	0.851
Inhibition	81.56 (37.0)	75.27 (34.3)	89.42 (40.1)	−0.986	0.334
Inhibition-switching	104.26 (44.2)	102.40 (42.9)	106.58 (47.6)	−0.240	0.812

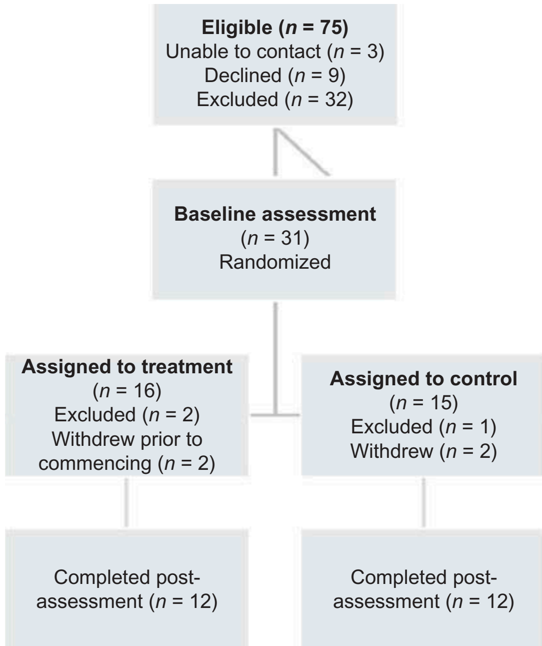

Figure 1. Participant flow through the study.

Design and randomization

The study used a randomized controlled design. Participants were randomized in blocks consisting of eight participants, (four allocated to treatment, four allocated to controls). Randomization for each block was achieved by having an independent person place eight cards with either "Treatment" or "Control" (four of each one) written on them into opaque envelopes which were then sealed. At the completion of baseline testing, an envelope was selected and assigned to the next participant on a list of screened participants. A total of 16 participants were randomly assigned to the treatment group and 15 to the control group.

Materials and procedure

The repetition-lag training procedure was broadly similar to that outlined by Lustig and Flegal (2008). Computer files that provided the repetition-lag training algorithm were kindly provided by Kristin Flegal. Words were sourced from the MRC Psycholinguistic database (located on UWA website http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). From this, word lists for the experiment were randomly generated via an Excel macro. Each list had 30 words and two lists were used for each round of training (30 words to be studied, 30 words as lures). Word lists were balanced in terms of word length $F(1,46) = .992$, $P = .489$ and frequency $F(1,46) = 1.232$, $P = .139$. The experiment was generated using E-Prime software (version 2.0; Psychological Software Tools, Inc.). Within the constraints set by the lag-length, the rate of repeated lures was random, and was randomized uniquely for each participant.

Each training session comprised four rounds of training with the exception of the final session which had three. Within each session there was an initial study phase where 30 words were presented one at a time for 14 seconds each, followed by a test phase where

Table 2. Lags and criterion level to reach for each RLT level.

Level	Lag 1	Lag 2	Criterion to move to next level (%)
1	1	2	85
2	1	3	85
3	2	4	85
4	2	8	83
5	4	12	83
6	4	16	83
7	8	20	83
8	12	28	83
9	12	32	83
10	16	36	83
11	16	40	83

the studied words were re-presented, mixed in with non-studied lures. The lures were repeated at systematically varying intervals. All participants began with a lag of 1 + 2 (i.e., 15 of the 30 lures were repeated after a lag of 1 intervening word, the other 15 lures were repeated following an interval of 2 words) and were able to work their way up to a maximum lag of 16 + 40. Instructions were given on-screen to think about the meaning of each word to help them remember it but participants were free to choose their own method of doing so. Participants were also warned that some words would be repeated in the test phase. To motivate participants, it was explained that the target to reach was a minimum of 85% studied words correctly recognized and 85% of unstudied words correctly rejected. Participants were instructed to write down their percentage score for both studied words and repeated lures on a scoresheet provided. However, the lag interval was programmed to increase each time participants were able to correctly reject *repeated* lures at the criterion level. Lags increased as follows: 1 + 2, 1 + 3, etc. At the end of each round percentage scores for both studied words and for lures were displayed on the screen. The lags for each level of training are displayed in Table 2. As initial piloting with a small sample (N = 4) of people with aMCI had indicated that the threshold levels of lures correctly rejected (96% and 93%) used in Lustig and Flegal (2008) to move up to the next lag were too high for this impaired group, modified criterion levels of 85% and 83% of unstudied words correctly rejected were used instead.

After a practice round on the training task, training was then completed over the following six sessions. A final session was devoted to post-treatment measures. The control group attended twice to complete pre and post-treatment measures.

Primary outcome measures

The primary outcome measures comprised both immediate and delay scores from the Verbal Paired Associates Test (Wechsler Memory Scale – Fourth Edition, Wechsler, 2009). While all participants had prior neuropsychological testing, none had previous exposure to this task.

VPA I – This subtest assesses verbal memory for associated word pairs. After 10 word pairs are read to the participant, the first word of each pair is read and the participant is asked to provide the word that it was paired with. The same list of word pairs is presented over four trials, in a different order each time. If a participant responds incorrectly, the correct word pair is re-stated. Correct responses are affirmed ("*That's right*") by the examiner. Six of the word pairs are unrelated words, (e.g., "*laugh-stand*") four of the word-pairs are related words ("*sock-shoe*"). Recall of unrelated word pairs was used as the measure of transfer of recollection. In our study, the score is the sum of correctly recalled unrelated word-pairs across all four trials.

VPA II – The delayed condition assesses long-term recall of the word pairs. After a delay of 20–30 minutes, the first word of each pair is again read and the participant is asked to state the word that goes with it. Only a single trial is administered and no feedback is given during this test.

Secondary outcome measures

Symbol Span: taken from the Wechsler Memory Scale – Fourth Edition (Wechsler, 2009). This subtest assesses visual working memory using novel visual stimuli. Beginning with two symbols, abstract visual symbols are exposed for 5 seconds. In the test phase, the participant has to correctly recall not only the correct symbols from distractor items, but also the order in which they were presented from left to right. The number of symbols presented increases by one at intervals as the test progresses. Higher scores indicate better visual working memory. Responses to each item are scored either zero, one or two points in accordance with the manual and the total score is the sum of the item scores.

Number Sequencing and Number-Letter switching: These tests are taken from the Delis–Kaplan Executive Function System Test Battery (D-KEFS; Delis, Kaplan, & Kramer, 2001) and mirror the well-known Trail Making Part A and Part B. In Number Sequencing, the participant is asked to draw a line connecting numbers in order from low to high as quickly as possible without making mistakes, and is a measure of attention. In Number-Letter switching, the task is to switch between connecting numbers and letters, in order, from lowest to highest, e.g., 1-A, 2-B, 3-C etc., and is a measure of cognitive flexibility. Data collected for each test is the time taken (seconds) to complete the task. Higher scores indicate poorer performance on these measures.

Everyday cognitive functioning was measured by the Cognitive Failures Questionnaire (Broadbent, Fitzgerald, Cooper, & Parkes, 1982) which is a reliable and validated (Wallace, Kass, & Stanny, 2002) 25 item self-report measure of the frequency of various cognitive failures. In this study participants were asked to report on cognitive failures in the past week. Each item is scored on a five point scale indicating how often a specific problem has occurred ranging from 4 = "very often" to 0 = "never". The scores are then added to arrive at a total score (Possible range 0–100). Higher scores indicate more frequent cognitive failures.

Mood was measured using the Depression Anxiety and Stress Scale (Lovibond & Lovibond, 1995) 21 item version. This shortened version has been shown to have excellent reliability (internal consistency for each subscale of at least .87) and concurrent validity, being highly correlated with well-established measures of depression and anxiety such as the Beck Depression Inventory and the Beck Anxiety Inventory (Antony, Bieling, Cox, Enns, & Swinson, 1998). High scores indicate increased symptoms of depression, anxiety and stress.

Data analysis

All analyses were carried out using SPSS Statistics – version 21. Planned analyses included the following:

(1) A paired samples *t*-test to check that participants progressed through the training levels by completion of training.

(2) A one-way repeated-measures ANOVA to check that levels of accuracy on the recognition task did not change across training.
(3) A one-way repeated-measures ANOVA of response bias across the six sessions of training.
(4) Treatment effects were assessed using 2 (group) × 2 (pre and post) repeated measures ANCOVAs, for each of the primary and secondary outcome variables with diagnosis (single or multiple domain aMCI) included as a covariate. A treatment effect was defined as a significant group × time interaction.

Results

Training task data

Data for hits, false alarms, accuracy and response bias are presented in Table 3.

Training outcome

To assess the effectiveness of recollection training, the RLT level reached (possible range 1–11) after three rounds of the first training session ($M = 1.41$, $SD = 0.66$) was compared with the level reached by the end of the final session ($M = 6.75$, $SD = 3.01$) using a paired samples *t*-test. This showed a significant improvement in RLT level between the first and last day of training, $t(11) = -6.65$, $P < .001$. This indicates that participants improved on the task itself over the course of training. All participants were able to progress at least two further lag levels beyond the level reached at the end of the first session, and eight of the twelve participants progressed at least four levels.

Accuracy

To check that the task improvements were actually due to improved recollection rather than simply due to practice effects, i.e., increased accuracy, a one-way repeated measures ANOVA was conducted on accuracy across all six sessions of training. The results indicated no significant change in accuracy ratings over the course of training $F(5,270) = .409$, $P = .842$.

Response bias

As the probability of correct responses is not equally distributed between "yes" and "no" responses in our study, an analysis of response bias was conducted. There were 30 studied

Table 3. Proportion of hits, false alarms, accuracy and estimates of bias (β) for each training session.

Training session	Hits	False alarms	Accuracy	Bias
Day 1	.74 (.13)	.19 (.15)	.54 (.24)	−.01 (.29)
Day 2	.73 (.16)	.22 (.15)	.51 (.19)	−.02 (.21)
Day 3	.72 (.16)	.17 (.12)	.55 (.23)	−.06 (.13)
Day 4	.70 (.16)	.14 (.11)	.55 (.23)	−.02 (.11)
Day 5	.69 (.18)	.14 (.11)	.54 (.23)	−.04 (.09)
Day 6	.70 (.15)	.11 (.11)	.58 (.21)	−.06 (.15)

Note: Numbers in brackets are standard deviations.

words and 60 unstudied words (30 new and 30 repeated) implying that responding "yes" will be correct 1/3 of the time, while "no" will be correct 2/3 of the time. It could therefore be argued that changes on the task are not due to true improvements in recollection but merely reflect participants learning to press "no" more often. The statistic β was calculated as a measure of bias using the method outlined in Stanislaw and Todorov (1999). To assess for changes in response bias over the course of all six sessions of training a one-way repeated measures ANOVA was conducted on response bias as quantified by β. This indicated no systematic change in response bias by session $F(5,270) = .615$, $P = .689$.

Outcome data

All outcome variables were checked for outliers, normality and sphericity. There were no outliers detected using boxplots, and Mauchly's test of Sphericity was non-significant in all cases. Where some variables were positively skewed, tests of skewness and kurtosis were within reasonable bounds (Skewness and kurtosis values were both within ± bounds set by multiplying std. error of the relevant statistic by two).

Correlations between the baseline variables and the outcome measures were obtained. Baseline MMSE score was not significantly correlated with both primary outcomes and nor were other baseline variables such as age, years of education. While a chi-square test confirmed that diagnosis (single vs multiple domain aMCI) did not differ between groups $\chi^2(2, N = 24) = 0.75$, $P = .38$, point biserial correlations revealed that participants with single domain aMCI performed better on both primary outcomes as well as on a measure of cognitive flexibility, Number-Letter switching (correlations ranging from .413 to .423). It was then included in subsequent analyses as a covariate.

Primary outcomes

Mean scores both at baseline and posttest, for the two groups are detailed in Table 4.

There was a main effect of Time (practice effect) on immediate recall of unrelated word pairs but no main effects of Group on any of the measures. Details of main, practice

Table 4. Performance on primary and secondary outcomes.

	Training *M* (*SD*)		Control *M* (*SD*)	
	Pre	Post	Pre	Post
Primary outcomes				
Verbal Paired Associates I				
Unrelated	5.42 (3.7)	7.75 (4.2)	6.50 (5.8)	8.67 (7.2)
Verbal Paired Associates II				
Unrelated	1.50 (1.3)	2.42 (1.7)	2.25 (1.9)	2.08 (1.9)
Secondary outcomes				
Number sequencing	49.67 (22.6)	42.67 (15.8)	45.42 (12.8)	46.92 (22.7)
Number-Letter switching	128.92 (47.5)	120.42 (48.2)	141.33 (54.4)	115.42 (49.8)
Symbol Span	14.42 (4.4)	16.83 (3.3)	14.17 (6.0)	14.75 (6.2)
DASS-21				
Depression subscale	7.50 (8.6)	10.00 (8.0)	4.83 (6.7)	5.50 (6.1)
Anxiety subscale	4.16 (5.8)	6.83 (5.7)	1.83 (2.16)	3.83 (4.5)
Stress subscale	8.33 (6.9)	10.83 (8.4)	6.16 (6.8)	6.50 (6.2)
Cognitive Failures Questionnaire	35.83 (15.0)	34.75 (12.0)	28.42 (13.5)	31.42 (9.0)

Table 5. Summary results from ANCOVA of outcome measures.

Domain	Outcome measure	Main effect of group F value (P)	Main effect of time (practice effect) F value (P)	Group by time (treatment effect) F value (P)
Immediate auditory recall	VPA I unrelated	0.01 (.904)	9.19 (.006*)	0.08 (.781)
Delayed auditory recall	VPA II unrelated	0.02 (.965)	1.97 (.182)	4.52 (.046*)
Attention	Number sequencing	0.03 (.852)	0.18 (.677)	1.59 (.220)
Cognitive flexibility	Number-Letter switching	0.36 (.554)	1.79 (.195)	1.15 (.294)
Visual working memory	Symbol Span	0.57 (.457)	0.68 (.418)	0.89 (.355)
Everyday cognition	CFQ	1.81 (.192)	0.62 (.437)	0.47 (.500)
Mood – DASS 21	Depression	1.51 (.232)	3.48 (.075)	1.16 (.291)
	Anxiety	2.27 (.146)	7.59 (.012*)	0.15 (.698)
	Stress	1.41 (.247)	1.94 (.177)	1.14 (.297)

Note: *Significant $P \leq .05$.

and interaction effects for all measures are provided in Table 5. On the measures of recollection, recall of *unrelated* word pairs immediately following presentation was not significantly affected by training. However, on the delayed measure, results of a repeated-measures ANCOVA revealed a significant group × time interaction with the training group improving their recall of unrelated word pairs, while the control group declined. The eta squared statistic (.17) indicated a modest effect size. Recall of *related* word-pairs (which were not expected to change) was not affected by training either immediately following presentation $F(1,21) = 1.75$, $P = .200$ or after a delay $F(1,21) = .35$, $P = .361$.

Secondary outcomes

There were no significant group by time interactions on measures of either attention or cognitive flexibility as measured by Number sequencing and Number-Letter switching. Nor was there a significant interaction on the measure of visual working memory (Symbol Span). Inspection of means suggested that the lack of significant group by time interactions on both Number sequencing and Symbol Span could have been due to low power, so further exploratory analyses were carried out. Paired samples *t*-tests revealed that for Symbol Span the training group improved their performance following training $t(11) = -2.68$, $P = .021$, while the control group did not $t(11) = -.367$, $P = .720$. Results for Number sequencing were non-significant for both groups. Both self-reported everyday cognitive failures and mood were unaffected by training, with no significant group by time interactions.

Discussion

To the author's knowledge this is the first study using the modified RLT program described by Lustig and Flegal (2008) in a group of people with aMCI. The study builds on an existing body of research which suggests that RLT improves recollection on the task itself, and may also transfer to untrained measures.

The modified RLT program proved effective as participants were able to increase the lag at which they could correctly discriminate unstudied from studied words to criterion levels of accuracy. These improvements cannot be attributed to practice effects nor to changes in response bias suggesting that training was effective at improving recollection on the task itself. Taken together, these findings suggest that the modified RLT procedure may be a robust method of training recollection in aMCI.

This study aimed to test the efficacy of RLT in transferring to an untrained measure of recollection and found an improvement in delayed recall of unrelated word-pairs following training. This finding is preliminary and requires further research. However, it is notable as delayed recall is known to be significantly impaired in aMCI, more so than immediate recall (Perri et al., 2005) and problems with delayed recall are predictive of conversion to Alzheimer's disease in memory clinic patients (Morris, Heyman, Mohs, & Hughes, 1989; Smith & Rush, 2006; Takayama, 2010). Thus, a group of aMCI participants who are likely to be at increased risk of progression to dementia improved in an area of memory known to be impaired following just six sessions of RLT. However, the other primary outcome, immediate recall of unrelated word pairs, did not improve with training. While there are many differences in how the immediate and delayed trials of the VPA are administered and presumably in the cognitive processes engaged at these two phases, it is plausible to suggest that training effects occurred only on the delayed test because RLT emphasized delayed recollection over more short-term retention. The immediate recall task also appears to have been too sensitive to practice effects with both treatment and control groups improving.

None of the secondary outcome measures showed strong evidence of transfer of training. Lustig and Flegal (2008) previously reported gains following RLT training on a measure of cognitive flexibility (Trails B). This was an unexpected finding and they speculated that it could be because RLT actually trains executive control processes mediated by the prefrontal cortex. However, it was not replicated in this study using a similar measure (Number-Letter switching) suggesting that it may have simply been an isolated finding. While previous research has reported gains in measures relating to working memory such as the Self-Ordered Pointing Test, n-back, and Digit Symbol Substitution Test (Jennings et al., 2005), we did not find strong evidence of transfer to a measure of visual working memory (Symbol Span), however there is evidence that this may have been due to low power.

The self-report data showed no effect of training on either frequency of cognitive failures or on mood. As improvements in recollection are known to be related to everyday memory it was disappointing to observe no change in self-reported cognitive failures, despite training and transfer data indicating improved recollection. One possible reason is that the measure was not sensitive enough to memory lapses due to recollection failure, instead being designed to look at cognitive failures more generally (sample items include, "*Do you have trouble making up your mind?" "Do you daydream when you ought to be listening?" "Do you drop things?*" that may be unrelated to lapses in recollection). In our study mood was not affected by training. This suggests that while RLT did not boost mood, exposure to cognitive challenges via training does not appear to have an adverse effect on self-reported mood.

This study had a number of strengths. It included independently diagnosed and objectively assessed aMCI participants who were re-screened on entry to the study. The study used a robust randomized controlled trial design to control for potential bias in observing treatment effects. Further, it made use of an established training procedure, optimized for older people and further adapted on the basis of piloting with an aMCI sample. The study was designed to test the hypothesis that RLT would improve recollection not just on the task itself, but also transfer to an untrained measure of recollection. Primary outcome measures were carefully selected and a conservative number of analyses planned. Secondary analyses were conducted to assess the transfer to other cognitive domains reported in previous research and to provide some insight into the impact on frequency of cognitive failures in daily life.

Despite these strong design features, several important qualifications remain. Firstly, although there were only a few outcome measures, a Type I error cannot be ruled out, and the results need to be replicated in a larger trial. Secondly, the size of the improvement is modest, and may not be clinically meaningful. On average following training participants were able to recall one more word pair than prior to training. Also, the improvement was not evenly distributed across all participants. A sub-group (those who scored 0 at baseline) did not improve, with the improvement occurring in those who scored between 1 and 5 at baseline. The other qualification is that a study of this size is necessarily limited in its ability to see effects (raised Type II error), due to low statistical power. Another limitation is that an active control group was not employed, meaning that the results could be due to factors such as experimenter contact. This latter concern is ameliorated by the finding that previous research with active control groups using RLT have found that recognition practice with an equivalent amount of contact time does not produce the same training or transfer gains as RLT (Boller et al., 2012).

The effect of changing the criterion to move from one lag to another is uncertain. In piloting, four aMCI participants were unable to reach the criterion of 96% correct rejections after four rounds of training at entry level (lag interval of 1 + 2) so the threshold was reduced to 85%. In the training data this meant that all participants were able to progress to at least lag level 3 (2 + 4). An alternative to changing the criterion would have been to reduce the number of words in the studied list. However this was avoided as the extent of improvement on the RLT task itself in an aMCI group was unknown. Most studies reporting on this have used normal elderly participants, so reducing the number of words (which necessarily reduces the longest lag possible) may have created an artificial ceiling on training gains. This was borne out in our data where several participants reached the highest RLT level possible by completion of training.

The findings here suggest that RLT may be useful for improving recollection in a group who are impaired in this ability. With the modified procedure, participants are encouraged to use deep meaning-based processing to encode novel stimuli in a way that should aid recollection. Ideally, future studies should assess transfer of recollection on a range of measures, including ones sensitive to everyday recollection failures experienced in aMCI. This may help to provide a clearer view as to which aspects of recollection memory are improved by RLT. Finally, the variable response to training highlights a need to further explore the role of impaired familiarity in response to training in aMCI. Studies with aMCI and Alzheimer's disease groups have reported mixed findings with regard to whether or not familiarity is preserved. Some have reported problems with recollection alone (Guerdoux, Dressaire, Martin, Adam, & Brouillet, 2012; Serra et al., 2010), while others find problems with both recollection and familiarity (Ally, Gold, & Budson, 2009; Wolk, Signoff, & DeKosky, 2008). At the least, those who have impairments in familiarity (on top of problems with recollection) are likely to have more difficulty completing the RLT procedure. If so, the RLT procedure may need to be further adapted to cater for the level of impairment in familiarity at baseline, either by pre-training to improve familiarity, or by shortening the length of the word lists. Overall, there are several avenues to explore so that RLT can be more precisely targeted to the heterogeneous clinical presentations seen in aMCI.

Acknowledgments

The authors would like to acknowledge Kristy Delmas, Kristin Flegal, Oren Griffiths, Paul Jackson, Jillian Patterson, and Lauren Roche for assistance provided in carrying out the study.

References

Adam, S., Van der Linden, M., Collette, F., Lemauvais, L., & Salmon, E. (2005). Further exploration of controlled and automatic memory processes in early Alzheimer's disease. *Neuropsychology*, *19*, 420–427. doi:10.1037/0894-4105.19.4.420

Ally, B. A., Gold, C. A., & Budson, A. E. (2009). An evaluation of recollection and familiarity in Alzheimer's disease and mild cognitive impairment using receiver operating characteristics. *Brain and Cognition*, *69*, 504–513. doi:10.1016/j.bandc.2008.11.003

Amieva, H., Rouch-Leroyer, I., Fabrigoule, C., & Dartigues, J. F. (2000). Deterioration of controlled processes in the preclinical phase of dementia: A confirmatory analysis. *Dementia and Geriatric Cognitive Disorders*, *11*, 46–52. doi:10.1159/000017213

Anderson, N. D., Ebert, P. L., Jennings, J. M., Grady, C. L., Cabeza, R., & Graham, S. J. (2008). Recollection and familiarity-based memory in healthy aging and amnestic mild cognitive impairment. *Neuropsychology*, *22*, 177–187. doi:10.1037/0894-4105.22.2.177

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. *Psychological Assessment*, *10*, 176–181. doi:10.1037/1040-3590.10.2.176

Bissig, D., & Lustig, C. (2007). Who benefits from memory training? *Psychological Science*, *18*, 720–726. doi:10.1111/j.1467-9280.2007.01966.x

Boller, B., Jennings, J. M., Dieudonné, B., Verny, M., & Ergis, A. M. (2012). Recollection training and transfer effects in Alzheimer's disease: Effectiveness of the repetition-lag procedure. *Brain and Cognition*, *78*, 169–177. doi:10.1016/j.bandc.2011.10.011

Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Psychology*, *21*, 1–16. doi:10.1111/j.2044-8260.1982.tb01421.x

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis–Kaplan executive function system*. San Antonio, TX: Psychological Corporation.

Derwinger, A., Stigsdotter-Neely, A., MacDonald, S., & Backman, L. (2005). Forgetting numbers in old age: Strategy and learning speed matter. *Gerontology*, *51*, 277–284. doi:10.1159/000085124

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. doi:10.1016/0022-3956(75)90026-6

Guerdoux, E., Dressaire, D., Martin, S., Adam, S., & Brouillet, D. (2012). Habit and recollection in healthy aging, mild cognitive impairment and Alzheimer's disease. *Neuropsychology*, *26*, 517–533. doi:10.1037/A0028718

Hampstead, B. M., Sathian, K., Moore, A. B., Nalisnick, C., & Stringer, A. Y. (2008). Explicit memory training leads to improved memory for face-name pairs in patients with mild cognitive impairment: Results of a pilot investigation. *Journal of the International Neuropsychological Society*, *14*, 883. doi:10.1017/S1355617708081009

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541. doi:10.1016/0749-596X(91)90025-F

Jacoby, L. L., Jennings, J. M., & Hay, J. F. (1996). Dissociating automatic and consciously controlled processes: Implications for diagnosis and rehabilitation of memory deficits. *Basic and Applied Memory Research: Theory in Context*, *1*, 161–193.

Jennings, J. M., Carello, E. A., Dagenbach, D., Rapp, S. R., Brenes, G. A., & Atkinson, H. H. (2006, April). *Improving recollection in individuals with mild cognitive impairment: Training and transfer effects*. Paper presented at the Biannual Cognitive Aging Conference, Atlanta, GA.

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*, 352–361. doi:10.1037/0882-7974.12.2.352
Jennings, J. M., & Jacoby, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation*, *13*, 417–440. doi:10.1080/09602010244000390
Jennings, J. M., Webster, L. M., Kleykamp, B. A., & Dagenbach, D. (2005). Recollection training and transfer effects in older adults: Successful use of a repetition lag procedure. *Aging, Neuropsychology and Cognition*, *12*, 278–298. doi:10.1080/138255890968312
Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney: Psychology Foundation.
Lustig, C., & Flegal, K. E. (2008). Targeting latent function: Encouraging effective encoding for successful memory training and transfer. *Psychology and Aging*, *23*, 754–764. doi:10.1037/a0014295
Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271. doi:10.1037/0033-295X.87.3.252
Morris, J. C., Heyman, A., Mohs, R. C., & Hughes, J. P. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Archives of Neurology*, *48*, 278–281.
Perri, R., Carlesimo, G. A., Serra, L., & Caltagirone, C. (2005). Characterization of memory profile in subjects with amnestic mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, *27*, 1033–1055. doi:10.1080/13803390490919317
Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. doi:10.1111/j.1365-2796.2004.01388.x
Ranganath, C., Flegal, K. E., & Kelly, L. L. (2011). Can cognitive training improve episodic memory? *Neuron*, *72*, 688–691. doi:10.1016/j.neuron.2011.10.022
Ravaglia, G., Forti, P., Maioli, F., Martelli, M., Servadei, L., Brunetti, N., … Mariani, E. (2006). Conversion of mild cognitive impairment to dementia: Predictive role of mild cognitive impairment subtypes and vascular risk factors. *Dementia and Geriatric Cognitive Disorders*, *21*, 51–58. doi:10.1159/000089515
Serra, L., Marco Bozzali, M., Cercignani, M., Perri, R., Fadda, L., Caltagirone, C., & Carlesimo, G. A. (2010). Recollection and familiarity in amnesic mild cognitive impairment. *Neuropsychology*, *24*, 316–326. doi:10.1037/a0017654
Shumaker, S. A., Legault, C., & Coker, L. H. (2006). Behavior-based interventions to enhance cognitive functioning and independence in older adults. *JAMA*, *296*, 2852–2854. doi:10.1001/jama.296.23.2852
Smith, G., & Rush, B. K. (2006). Normal aging and mild cognitive impairment. In D. K. Attix & K. A. Welsh-Bohmer (Eds.), *Geriatric neuropsychology: Assessment and intervention* (pp. 27–55). New York, NY: Guilford Press.
Smith, J. A., & Knight, R. G. (2002). Memory processing in Alzheimer's disease. *Neuropsychologia*, *40*, 666–682. doi:10.1016/S0028-3932(01)00137-3
Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*, 137–149. doi:10.3758/BF03207704
Takayama, Y. (2010). A delayed recall battery as a sensitive screening for mild cognitive impairment: Follow-up study of memory clinic patients after 10 years. *Journal of Medical & Dental Science*, *57*, 177–184.
Troyer, A. K., Murphy, K. J., Anderson, N. D., Moscovitch, M., & Craik, F. I. (2008). Changing everyday memory behaviour in amnestic mild cognitive impairment: A randomised controlled trial. *Neuropsychological Rehabilitation*, *18*, 65–88. doi:10.1080/09602010701409684
Wallace, J. C., Kass, S. J., & Stanny, C. J. (2002). The Cognitive Failures Questionnaire revisited: Dimensions and correlates. *Journal of General Psychology*, *129*, 238–256. doi:10.1080/00221300209602098
Wechsler, D. (2009). *Wechsler memory scale – fourth*. San Antonio, TX: Psychological Corporation.
Winblad, B. I., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., … Petersen, R. C. (2004). Mild cognitive impairment – Beyond controversies, towards a consensus: Report of the international working group on mild cognitive impairment. *Journal of Internal Medicine*, *256*, 240–246. doi:10.1111/j.1365-2796.2004.01380.x
Wolk, D. A., Signoff, E. D., & DeKosky, S. T. (2008). Recollection and familiarity in amnestic mild cognitive impairment: A global decline in recognition memory. *Neuropsychologia*, *46*, 1965–1978. doi:10.1016/j.neuropsychologia.2008.01.017