The Effects of Mixed Working Memory Training on Subsequent Training Gains among Older Adults
NCT05672771
May 26, 2023 (Rev. February 2, 2024)

Study Protocol

Objectives. While an intellectually active and socially integrated lifestyle shows promise for promoting cognitive resilience with aging, the mechanisms underlying any such effects are not well understood (Stine-Morrow et al., 2014, 2021; Stine-Morrow & Manavbasi, 2022). The aim of the current project was to test the implications of the "Mutualism" hypothesis (Savi et al., 2019; van der Maas et al., 2006; van der Maas et al., 2017), which suggests that intellectual function emerges out of the reciprocal influence of growth in abilities as they are exercised in the ecology of everyday life. Such a view stands in contrast to the Principle of Direct Transfer (Thorndike, 1906) in suggesting that improvement in one component will enhance the *modifiability* of a related component. In other words, while conventional conceptualizations of transfer imply immediate generalization, mutualism assumes that practice in the novel skill is required for benefits to be manifest. An additional aim was to test the idea that mutualistic effects will be enhanced by more diverse training in related skills, relative to single-component training (Brown et al., 2014; Schmidt & Bjork, 1992).

To this end, participants engaged in two phases of training. The target task in Phase 2 was the complex reading span, which has been found to be predictive of standardized tests of verbal ability and measures of language comprehension (Baddeley, 2012; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Payne & Stine-Morrow, 2017; Peng et al., 2022; Stine & Wingfield, 1987). As a WM task, the reading span requires the concurrent management of sentence processing and a memory load. Participants were randomly assigned to one of four groups that varied in the type of training in Phase 1. In the Different Mixed (DM) condition, participants practiced lexical decision span and category span tasks, two different tasks similar to the reading span in requiring the individual to manage a memory load while processing verbal material but with different sorts of processing components. In the Different Single (DS) condition, participants practiced only the lexical decision span, which provided less diverse experience with tasks that depend on WM. These were contrasted with two controls, a Same Task (ST) condition in which participants practiced the reading span, exactly the same as the target task, and a Placebo Control (PC) condition, in which participants made speeded judgments to the same verbal stimuli as in the lexical decision task but with no memory component, and hence, no demand to simultaneously manage processing and memory (i.e., WM). We measured both how quickly the target task in Phase 2 was acquired, as well as pre- to posttest change in practiced and unpracticed WM tasks and episodic memory tasks, as a function of prior experience in Phase 1.

According to the Principle of Direct Transfer, training in the reading span in Phase 2 should be accelerated by conditions in Phase 1 that share the most elements with Phase 2 training. Thus, Principle of Direct Transfer predicts an early advantage for the ST group relative to the other three groups that would be sustained through acquisition, as well as some advantage for the two other WM training groups, the DM and DS conditions, relative to the PC condition. The Principle of Mutualism, on the other hand, predicts accelerated acquisition for the DS, and especially the DM, conditions, relative to the PC condition, on par with the ST condition. This is because even though the DS and DM tasks share fewer elements with the target task, the Mutualism principle predicts that the acquisition of the elements that are shared enhances plasticity of the unpracticed elements of the target task. Critically, Mutualism would predict that any advantage of the DS and DM conditions would not be immediately apparent but only emerge with practice.

The two theories also make contrasting predictions for pre- to posttest change. Principle of Direct Transfer would predict greater change in the practiced WM tasks relative to the unpracticed WM tasks for the ST, DS, and DM groups relative to the PC control. Specifically, the ST group — who exclusively practiced the reading span (RS) task in both phases of training would be expected to show the greatest

pre- to posttest improvement in RS, but the least improvement in all other span tasks. The DS group (who practiced the lexical decision span (LDS) in Phase 1 and RS in Phase 2) and the DM group (who practiced the category span (CatS) and LDS) in Phase 1 and RS in Phase 2), however, would be expected to show some improvement in the trained tasks according to their collective exposure across the two phases (i.e., DS=DM on RS; DS>DM on LDS; DM>DS on CatS), and less improvement in the untrained tasks, and even less in the episodic memory task (which would represent far transfer). According to the principle of Mutualism, performance is most likely to be enhanced when elements of the task are acquired and then the target task containing those elements is practiced. This situation is embodied in the DM condition in which the individual acquires multiple elements of the WM task that are then practiced in the context of a new task. Thus, the Principle of Mutualism predicts the largest gains from pre- to posttest for the DM group, and to a lesser extent for the DS group, who had less diverse exposure to elements of the WM task relative to the DM group. Because there was no opportunity to practice the untrained WM tasks or episodic memory task, the Principle of Mutualism does not predict improvement in these areas.

Collectively then, Principles of Direct Transfer and of Mutualism make quite different predictions for performance in the successive enrichment paradigm and in pre- to posttest change. Most importantly, the former predicts the reading span in Phase 2 will be acquired the most quickly and show the largest pre- to posttest gains by the ST group, while the latter predicts faster acquisition and largest pre- to posttest gains by the DM and (to a lesser extent) the DS group.

Design. The experimental design involved two phases of working memory training, with 10 days of activities in each phase. Each day, participants engaged in 4 8-min blocks of training activities (so a total of 32 minutes per day). Participants were encouraged to complete the 10 days of activities across a 2-week period. During Phase 2 all participants engaged in the same training with a reading span task. The training in Phase 1 was manipulated so that participants in the Placebo Control (PC) group practiced a speeded lexical decision task with no decision component, the Same Task (ST) group practiced with the same reading span task as in Phase 2, the Different Single (DS) group practiced a lexical decision span task, and the Different Mixed (DM) group practiced both the lexical decision span task and a category span task. Thus, the critical tests are whether the four groups differ in (a) pre- to posttest gains on the span tasks, and in (b) training gains in the reading span task in Phase 2 training.

Methods. All tasks in both phases were adaptive, starting at level one (one decision with one item to be remembered, or in the PC condition, simply one decision), and then increased or decreased depending on decision and memory accuracy combined. Set size was increased by one when performance was over 90%, decreased by one when performance was less than 70%, and otherwise, stayed the same.

In the reading span, participants read a series of either semantically congruent sentences or syntactic prose sentences (e.g., As the ship gets better, your child needs to develop this oven) for which they make sentence acceptability judgments on each sentence. Participants make an acceptability judgment by pressing a Yes button or No button at the bottom of the screen. Once participants make a decision, the sentence disappears, and participants are presented with accuracy feedback. At the end of the set, participants are cued to recall the last word of each sentence in the order in which they were presented. The cued recall screen consisted of a set of empty text boxes that participants can press and then type their responses via an on-screen keyboard, with no time limit. Acceptable sentences are adapted from two sources. The Nelson and Narens (1980) general information question norms provided 244 sentences. The other source was the Manually Annotated Sub-Corpus (MASC) of the Open American National Corpus (Ide et al., 2013), which provided 301 sentences that ranged widely in topic, length, and syntactic structure. In addition, 346 unacceptable sentences were adapted from the syntactic prose conditions in earlier studies by Lee and Federmeier (2011) and Payne et al. (2015). Unacceptable sentences have syntactically well- formed sentence frames, but contain no coherent message-level semantics. All sentences ranged between 60 and 90 characters, and all sentence final words were between 4 and 9 characters.

In the lexical decision span task, participants are presented with a set of letter strings constituting words (e.g., seek) or non-words (e.g., ceek) and are cued to decide whether or not each string formed a word or not. The letter strings are presented for 4 s. At any point within this interval, participants decide if

the letter string was a word or non-word by pressing a Yes button or No button at the bottom of the screen. Once participants make a decision, the letter string disappears, and participants are presented with accuracy feedback. Following each lexical decision, an unrelated single letter is presented for 1500 ms for participants to recall at the end of the set. At the end of each set, participants are cued to recall each of the letters in the order in which they were presented. The cued recall screen consists of a set of empty text boxes that participants can press and then type their responses via an on-screen keyboard. Participants have no time limit to enter their recall responses at the prompt. A total of 9,000 common and proper nouns and 10,000 phonologically regular and pronounceable non-words were generated from the English Lexicon Project database (Balota et al., 2007). Word/non- word strings ranged in length between 4 and 9 characters (for word stimuli: log word frequency range: 5 to 13.67).

In the category span, participants are presented with a semantic category at the top of the screen (e.g., weather) along with a set of single words for which they made validity judgments (e.g., humidity - Yes; chocolate - No). Each trial consisted of the category and target word presented for 4 s. At any point within this duration, participants can decide if the target matches or does not match the category by pressing a Yes button or No button at the bottom of the screen. Once a participant makes a decision, the target word disappears and s/he presented with accuracy feedback. The program would then progress to the next trial within the set. At the end of each set, participants are cued to recall each of the words in the order in which they were presented. The cued recall screen consists of a set of empty text boxes that participants can press and then type their responses via an on-screen keyboard. Participants have no time limit to enter their recall responses at the prompt. Categories and exemplars were drawn from the Van Overschelde et al. (2004) category norms. The final stimulus set contains a total of 69 unique categories and over 1500 unique words. Items were drawn randomly such that, within a set, each word had an equal probability of belonging to the presented category or not. Across training sessions, items were rotated through such that all categories had to be selected at least once before a particular category could be repeated again.

In the placebo control, participants engaged with a speeded lexical decision task. This was an active control matched to the other conditions in materials and engagement (speed increased and decreased adaptively to performance) but without the memory component.

Electronic tablets were delivered to the participants' homes along with detailed instructions and a log book, using contact-free delivery during the COVID-19 pandemic. Participants kept an iPad for each phase of training for about two weeks and could complete their 10 sessions of training on any days during the two weeks.

The pre- (administered Prior to Phase 1) and posttest (administered after Phase 2) batteries were programmed in Qualtrics and administered online. The testing session began with the experimenter on the phone to consent the participant and provide guidance in initiating the protocol. The experimenter was available by phone throughout to answer any questions throughout the session. The testing session generally required 1.5-2 hours. To the extent possible, the pretest-posttest interval was held constant (M = 34.5 days, SD = 7.0). There were no differences among the groups, F<1.

Outcome measures included rate of learning during Phase 2 and change in cognitive measures from pre- to posttest.

Statistical Analysis Plan

The primary outcomes were (a) standard unit change in cognitive measures from pre- to posttest (i.e., posttest score - pretest score) / standard deviation at pretest) for reading span and a composite of work and (b) daily mean reading span score achieved across the 10 days of training in Phase 2. For the latter variable, calculations were based on scores from Trial 2 to the end of the training block (we discarded Trial 1 because this trial only included one element to remember and thus, elicited uniform performance across participants). Pre- to posttest change was analyzed using one-way ANOVA to test differences among the four groups with post hoc Bonferroni comparisons to correct for the number of

comparisons. Acquisition of the reading span in Phase 2 (change across the 10 days of practice) was analyzed with mixed-effects linear regression using the lme4 and lmerTest packages for R version 4.2.2 (R Core Team, 2022). The PC Group was set as reference in comparisons with other groups in an analysis with Training Day and Phase 1 Training Group as the fixed factors, and by-subject random intercept, using the restricted maximum likelihood (REML) method of estimation. The design was powered at .9 to detect an effect size of .43 for differences between the groups with an alpha of .05, and to detect an effect size of .14 for group differences in change across 10 days (Faul et al., 2007), though of course, there estimates of power for linear mixed-effects models are controversial. There was no correction for the number of tests in this latter analysis.

Informed Consent Form
(approved April 12, 2022)

Welcome to The Adult Learning Lab!

This letter describes the research project in which you have been asked to participate. It is being conducted by Professors Liz Stine-Morrow (Educational Psychology), Aron Barbey (Psychology), Brad Sutton (Electrical Engineering), and Dan Morrow, all faculty at the Beckman Institute. The purpose of this project is to better understand how to maintain cognitive resilience through participation in mentally stimulating activities. Specifically, we are testing the benefits of engaging attention and memory by playing specially designed games on an electronic tablet.

For this study, we will be mailing an iPad to your home, along with instructions about how to use it to play the games. At any point if you require further information, you can email the lab member you have been in contact with or our lab at tall@education.illinois.edu, and we can arrange a phone call or video meeting with you. You will be randomly assigned to engage in a particular schedule of mental games. We ask that you do your best to play these games in a focused way for about 32 minutes a day, 5 days a week for 4 weeks. At the halfway point of your training (after two weeks), you will mail the iPad to us in a prepaid addressed envelope that we will send you with the iPad. We will update your training program and send it back to you, along with a new set of training instructions, so that you can complete the next two weeks of training. Again, we will include a prepaid addressed envelope in which you can return the iPad.

Prior to starting the training, you will be asked to provide some information about your background and to participate in a variety of tasks that deal with memory, attention, language understanding, and problem solving. For example, you will be asked to recall short lists of words, to describe the meanings of words, and make quick decisions about whether two patterns are the same or different. You will be able to participate in these activities on your home computer by clicking on a link that we send you. After you are finished with the iPad activities, you will do a similar set of activities on your home computer. Each time, these activities will take about 2 hours. The link for the first set of activities will appear once you consent to participation by signing this form, and after the 4 weeks of training, we will send you another link.

None of the cognitive tasks you will be asked to perform involve painful or stressful events. During the online sessions, you may take short breaks when you are at instructions pages before each task if you start to feel tired. If you do so, make sure not to exit the page that you are taking the survey on. The instructions will tell you when a task is timed. Once the task starts, the timer will automatically start and the page will move ahead once your time is up.

Participation is completely voluntary, and you are, of course, free to discontinue at any time. If you are from the Educational Psychology subject pool, you will receive one credit per hour of participation in the

tasks before and after training. If you start a session but do not complete it, you will receive partial payment (1 credit per hour) for the time you participated for.

If you are a paid subject, you will receive an electronic gift card to make purchases on Amazon. You will receive $20 for each testing session before and after the iPad training. If you start a session but do not complete it, you will still receive partial payment for that session ($5 per half-hour). There will be no payments for the time you spend playing the mental games on the tablet. However, if you complete all the iPad training (a total of 20 days of training), you will receive an additional $20 bonus for participating. We will email you your Amazon e-code, which is the electronic gift card code, after completion of all the sessions (pre-test, 4 weeks of training and post-test session). If you choose to withdraw sooner, you will be sent the code at the time you let us know of your withdrawal. Please allow for a week after completing post-test and mailing us the iPad, for the e-code to be emailed to you. We believe that this research will be beneficial in clarifying the nature of adult learning and the conditions that help us to maintain our cognitive vitality as we grow older, and we deeply appreciate your participation!

The results of your participation are kept strictly confidential. Neither your name nor any identifying information will appear on any of your answer sheets from tasks you complete. Your signed online consent form will be completely separate from the link to the cognitive tasks you complete. All your materials will be labeled with a unique identifier code; a key that connects your identity with this code will be secured in a separate location, accessible only to lab personnel. In addition, you should know that we never report individual results; research reports are always based on statistics characterizing the group. Test results (with a unique identifier code and no personally identifying information) will be secured in the lab for seven years after publication of findings. The key that connects your identity with the identifier code will be retained until the completion of this funded project.

We will use all reasonable efforts to keep your personal information confidential, but we cannot guarantee absolute confidentiality. When this research is discussed or published, no one will know that you were in the study. But, when required by law or university policy, identifying information may be seen or copied by:
• The Institutional Review Board that approves research studies;
• The Office for Protection of Research Subjects and other university departments
that oversee human subjects research;
• University and state auditors responsible for oversight of research; and
• Federal regulatory agencies such as the Office of Human Research Protections
in the Department of Health and Human Services; and
• The National Institutes of Health

Please feel free to ask questions about our research as you continue your participation in the program. If you have any questions please feel free to contact Dr. Stine-Morrow in the Department of Educational Psychology (eals@illinois.edu or 217-244-2167). (You will be emailed a copy of this form so that you can refer to this information later.) If you have any questions about the rights of research participants, feel free to contact the Institutional Review Board (IRB) at the University of Illinois either by phone (217-333-2670) or by e-mail (irb@illinois.edu). As a participant in this project, you will receive an electronic tablet that is the property of the University of Illinois. You are welcome to use it as much as you like while you have it (even if not directly for your mental exercises), but you must return the tablet when you return for posttest at the end of the study. By typing your full name in the field below, you are indicating that you have read and understand the above information and provide your consent to participate in this study.

[Note that we were approved to conduct this study with university students, but ultimately did not do so.]