Training attentional control in older adults

Anna MacKay-Brandt[1]
Washington University in St. Louis, St. Louis, MO, USA

Abstract

Recent research has demonstrated benefits for older adults from training attentional control using a variable priority strategy, but the construct validity of the training task and the degree to which benefits of training transfer to other contexts are unclear. The goal of this study was to characterize baseline performance on the training task in a sample of 105 healthy older adults and to test for transfer of training in a subset ($n = 21$). Training gains after 5 days and extent of transfer was compared to another subset ($n = 20$) that served as a control group. Baseline performance on the training task was characterized by a two-factor model of working memory and processing speed. Processing speed correlated with the training task. Training gains in speed and accuracy were reliable and robust ($ps < .001$, $\eta^2 = .57$ to $.90$). Transfer to an analogous task was observed ($ps < .05$, $\eta^2 = .10$ to $.17$). The beneficial effect of training did not translate to improved performance on related measures of processing speed. This study highlights the robust effect of training and transfer to a similar context using a variable priority training task. Although processing speed is an important aspect of the training task, training *benefit* is either related to an untested aspect of the training task or transfer of training is limited to the training context.

Keywords

Training; Older adults; Aging; Attentional control; Working memory; Processing speed; Multitask training; Cognitive intervention

Early investigations of ways to enhance cognitive performance in later life focused on intelligence (e.g., Willis & Schaie, 1986) and memory (see Verhaeghen, Marcoen, & Gossens, 1992, for a meta-analysis). More recently other aspects of cognition have also been the object of training studies. Kramer, Larish, and Strayer (1995) began a program of research designed to improve executive function in older adults. They used a variable priority strategy developed by Gopher, Weill, and Siegel (1989) that required participants to frequently shift their processing priorities between two or more concurrently performed tasks. The benefit of the variable priority training approach over a fixed division between two tasks was demonstrated by faster mastery of the task and better performance on a transfer task of a similar dual task nature for both young and older adults under the variable priority condition.

Kramer et al. (1995) argued against the position that age-related differences in dual task performance (as it applied to observed training benefits) supported either the complexity hypothesis (McDowd & Craik, 1988) or a general slowing account of age-related changes in performance (Cerella, 1985, 1990; Salthouse, 1985). Both theories suggest that age effects

Address correspondence to: Anna MacKay-Brandt, Division of Cognitive Neuroscience, Taub Institute, Columbia University, 630 W 168th St., PHG 18 Box 16, New York, NY 10032, USA. am3609@columbia.edu.
[1]Anna MacKay-Brandt is now at Columbia University.

on dual task can be explained in terms of simpler aspects of the task, either as a function of single task performance in the case of the complexity hypothesis or in terms of a general slowing factor. Using an elegant experimental design that compared training on single task components as well as dual task under the two training conditions described in the previous paragraph, Kramer demonstrated that training benefit was specifically associated with task coordination or attentional control. Subsequent studies from this group have continued to demonstrate the specific benefits of such training (Bherer et al., 2005, 2006; Bherer, Kramer, & Peterson, 2008; Kramer, Larish, Weber, & Bardell, 1999) and have identified regions of the prefrontal cortex thought to underlie aspects of attentional control associated with task performance (Erickson et al., 2007).

In addition to the original experimental evidence and the found associations with brain regions thought to subserve attentional control, a strong test of construct validity to support Kramer's theoretical position would come from a correlational analysis using independent tests of cognitive constructs posited to have a relationship with the training task and to underlie the training benefit. Previous work has established working memory and response-distractor inhibition as two empirically derived constructs that represent aspects of executive function related to attentional control (Kane & Engle, 2003; Miyake et al., 2000). Further, individual differences in processing speed are known to account for large portions of variance on other, more complex cognitive tasks (Salthouse, 1996) and should be included when considering relationships among cognitive measures. These three constructs – working memory, response-distractor inhibition, and processing speed – have been described using a latent variable approach that decreases the influence of task specific aspects of any single measure by considering only the common variance among several individual tests. This approach is ideally suited to identify specific aspects of cognition that may play a key role in complex task performance.

An important aspect of training is the implication that transfer of a trained skill to other contexts is possible. In their review of several decades of work on transfer of training, Barnett and Ceci (2002) highlighted stark contrasts of scholarly opinion ranging from ubiquitous transfer to no evidence of any form of transfer. They pointed out that researchers do not share the same definitions of transfer. Hence, they made a distinction between *far transfer* as transfer to a dissimilar context and *near transfer,* which involves a more similar context. Variable priority training benefits studied by Kramer et al. (1995, 2004) focused on near transfer to other dual task contexts. In earlier work using the similar *emphasis change* paradigm, Gopher, Weil, and Bereket (1994) demonstrated far transfer of training benefit to flight simulator and flight performance in groups of pilots. Far transfer to tasks that do not share a similar context with the training but do share an assumption of a relationship to a similar cognitive construct (such as tasks that indicate processing speed or working memory ability), serve to empirically validate theoretical assumptions about what is trained and the degree to which training is related to an underlying aspect of more complex cognition. Far transfer in this form has not been tested for Kramer's variable priority training paradigm, although other training studies have used this approach to empirically validate hypotheses about specific cognitive constructs that influence observed training effects. For example, Basak, Boot, Voss, and Kramer (2008) recently examined the training effect of a real-time strategy videogame on tests of executive control and visuospatial skills.

Typically training tasks are complex and involve numerous cognitive processes. Indeed, because of the multiple activities involved I will use the term *multitask* instead of dual task. The probability of achieving successful far transfer is greatly increased if processes common to the training task and the transfer task are identified. Little research, however, has examined the particular processes that may be affected by multitask training. One goal of the present study was to establish the construct validity of the training task through a

correlational analysis with previously established indicators of candidate cognitive constructs. Specifically, the training task was hypothesized to be related to two aspects of executive function-working memory and inhibition. A relationship between performance on the training task and working memory was hypothesized because Kramer et al. (1995) cited memory load as an important aspect that led to training benefit. Further, working memory capacity is often described as a measure of executive control (Kane & Engle, 2003). There is both theoretical and empirical motivation for a hypothesized relationship between inhibition and the training task.

Kramer et al. (1995) also suggested that variable priority training affects task coordination or task switching ability. Taken together with findings from a factor analytic study (Friedman & Miyake, 2004) that identified measures of inhibitory control as strongly related to task switching, I proposed an underlying response-distractor inhibition construct to be related to the training task. Response-distractor inhibition, as defined empirically in Friedman and Miyake's (2004) factor analysis, was strongly associated with task switching, the cognitive process that Kramer et al. (1995) considered to improve with training.

Further, many cognitive tasks have nontrivial relationships with processing speed and associations among tasks are often explained by shared variance with measures of speed (Salthouse, 1996). Therefore, processing speed was also included in the measurement model. In addition to the factor analytic and correlational analyses, both near and far transfer tasks were examined in a research design that included a control condition that allowed consideration of practice effects. In summary, the hypotheses tested in this study were the following: (a) To what degree are working memory, response-distractor inhibition, and processing speed related to the training task? (b) Is multitask variable priority training replicable in a new sample of healthy older adults? (c) Is there transfer of training to a novel multitask? (d) In those cognitive constructs found to be related to the training task in Hypothesis 1, do the individual cognitive tests show training-related improvement at post-test compared with a control group?

METHOD

Participants

Participants were 105 older adults aged 59 to 70 years ($M = 65.49$) recruited from the Washington University Psychology Department's older adult volunteer pool. Exclusion criteria included history (past or current) of neurological illness (e.g., stroke, transient ischemic attack, traumatic brain injury, dementia, Parkinson's disease, loss of consciousness for more than 5 minutes, seizures), psychiatric illness (e.g., major depression, bipolar disorder, schizophrenia), alcohol∕or substance abuse, alcohol and∕or substance abuse, chemotherapy or radiation therapy, cardiac bypass surgery, myocardial infarction, colorblindness, uncorrectable severe vision deficits, uncontrolled diabetes, uncontrolled thyroid disorders, or major surgery within the past month. In addition, participants were excluded if they were taking any neuroleptic or psychoactive medication at the time of testing. Three individuals were not eligible due to cardiac bypass surgery, and one individual was not eligible due to Parkinson's disease.

In a telephone screening call, participants were administered the Short Blessed Orientation and Memory Scale (Katzman et al., 1983) as a screening for possible dementia. Participants with a score of six or greater were not included; no one was excluded. Participants were also screened for depression using the Geriatric Depression Scale (Yesavage & Brink, 1983). Participants with a score of 11 or greater were not included; 4 participants were excluded on this basis and provided contact details for an outpatient psychological counseling center. Of the eligible participants, approximately 40% were not able to participate due to scheduling

conflicts with the training protocol (i.e., 1 hour per day for 5 consecutive days, 2 hours the week prior, and 2 hours the week following training).

All enrolled participants reported good health and had at least 20⁄30 corrected vision. The recruitment procedure was to first enroll all participants into the training study, with random assignment into training or control group. Upon completion of the training phase of the study an additional 64 participants were recruited into a normative group for a single testing session. Age and education for the training, control, and normative groups were not significantly different ($ps > .05$). Mean (*SD*) age in years for the training group was 64.75 (2.27), the control group 66.13 (3.88), and normative group 64.71 (2.88). Mean (*SD*) education in years was 15.85 (2.80) for the training group, 16.05 (2.35) for controls, and 16.38 (2.82) for the normative group.

Study design

In Session 1 all 105 participants received one of two versions of the training task and a battery of cognitive tests. The 41 training and control group participants then attended 1-hour training sessions for 5 consecutive days (Sessions 2 to 6). The following week in Session 7 these 41 participants repeated the cognitive battery from Session 1 along with a variant of the training task they performed in Session 1.

The 20 people in the training condition practiced the training task (see description in section on Measures) at each training session. The 21 people in the training study control condition received semi-structured internet training at each of the 5 daily sessions to parallel the experience of the training group. To allow a test of near transfer two versions of the training task were administered such that one half of the participants in the training condition received Version A at the first testing session and for the 5 training sessions and then the variant training task (Version B) at the final session. The other half of the training group received Version B at first session and for subsequent training and then Version A at the final session. Similarly the 21 people in the training study control condition and the participants ($n = 64$) who did not participate in the training portion of the study received either Version A or B at the initial session and then (for the control condition) the variant version at the final session.

Measures

The initial battery included measures of processing speed, working memory, inhibition, task switching, and the multitask used in training. The measures of inhibition were omitted from data analysis and this report because of failure to observe the expected effect on the Stroop task[2] and very poor reliability estimates that ranged from 0.00 to 0.06 in this older adult sample in contrast with those reported previously for college samples (Friedman & Miyake, 2004; Miyake et al., 2000). Reasons for these difficulties could not be identified. Thus, the measurement model examined in the data analysis included only two factors (Processing Speed and Working Memory).

The three measures of processing speed included a measure of simple reaction time, Pattern Comparison (Salthouse & Babcock, 1991) and Digit Symbol (Wechsler, 1997). These tasks were selected as hypothesized indicators of a Processing Speed factor such that the shared variance among the tests would represent processing speed and thus statistically control for the task impurity arising from the multidimensionality of each individual test considered

---

[2]The Stroop task requires naming the color of a presented word. The word was presented in colors red, blue, green, or yellow. The word was either a color name (e.g., red) or a neutral word (e.g., bank). The Stroop effect is the increase in response latency to name the color of the word when the word is a color name compared with a neutral word.

alone (see Salthouse, 1996, 2000, for a discussion of processing speed indicators in factor analysis and see also Miyake et al., 2000, for a similar discussion of factor analytic analysis to address the task-impurity problem in studies of executive functions). In the computerized task of simple reaction time, participants made a manual response when an arrow appeared on the computer screen; 10 practice trials were administered before the 40 test trials. The score was the median RT for correct responses. Pattern Comparison required participants to judge the similarity of pairs of patterns of symbols. Three practice items were administered before the 30 test items. The score was the number of items correctly completed within 60s. Digit Symbol required the speeded transcription of symbols based on a number key. The score was the number of items correctly completed in 120 s.

Three measures of working memory included Operation Span (Ospan; Unsworth, Heitz, Schrock, & Engle, 2005), Reading Span (Turner & Engle, 1989), and the two-back task (Kane, Conway, Miura, & Colflesh, 2007). Ospan required participants to solve an equation to determine if the provided answer was correct or false. After each computation a letter was presented for later recall. Each component of the task was given separately during a practice session. Speed to solve math problems was determined during the practice, and 2.5 *SD* was then added to create a time limit for the experimental session. If a participant did not respond before the time limit that item was scored as an error. There were three trials for each of 3 to 7 trial lengths yielding a total of 75 math problems to perform and 75 letters to recall. The order of set size was random for each participant. Accuracy on the math portion of the task was held to the standard of 85% correct to qualify a trial as valid for the letter component to be analyzed. The Ospan score (Unsworth et al., 2005) was the number of perfectly recalled (including order of presentation) letter trials. Administration of the Reading Span task (Rspan) was the same as for OSpan, but participants read a sentence and determined if it made sense or not. After each sentence a letter was presented for later recall. The Rspan score was obtained in the same way as the Ospan score. The two-back test, required the participant to press a key to indicate if a single viewed letter was the same (3 key on keypad) or different (4 key on keypad) from the letter presented two prior. The participant used the pointer finger from each hand to execute the task. Single lowercase letters on the computer screen were presented one at a time until a response was made or 800 ms elapsed (maximum intertrial interval was 800 ms); 18 practice items were presented before 72 test items. The score was percentage correct.

Given the hypothesized relationship between dual task performance and task switching (Kramer et al., 1999), a switching task adapted from the work of Minear and Shah (2008) was administered. A number letter pair (e.g., 31A) was presented centrally on the computer screen. In the first block of 10 practice and 60 test trials participants attended only to the number. The cue *odd* appeared in the upper left corner of the screen to indicate the participant was to press a left key (D) if the number was odd; the cue *even* appeared in the upper right corner of the screen to indicate the participant should press a right key (K) if the number was even. Responses were made with the pointer fingers of each hand. In the second block of 10 practice and 60 test trials the participant indicated if the letter was a consonant or a vowel. The cues *consonant* and *vowel* were presented in the left and right corners of the computer screen, and the same D and K keys were used for responses. These two tasks were combined to represent the no-switch condition. In the third block of 10 practice and 60 test trials both odd versus even and consonant versus vowel judgments were required. The participant monitored the cue labels and responded to the number or letter based on the cues (odd and even or consonant and vowel); however, the pattern of cue switch was predictably every two trials. RTs of correct responses for each participant across all blocks were converted to standardized scores. The participants' score on this task was the average standardized RT for the switch (third) condition.

Training task

Overview—Game Maker 7.0 software (YoYo Games Ltd, 2007) was used for task development; administration was on a Sony PCG-72GL laptop computer. The display was 20×27 cm. Responses were recorded via three number keys (1, 2, and 3) on the keyboard located in the upper left position of the keyboard, the space key (located centrally), and an external infrared mouse positioned on a large mouse pad to the right of the computer. This arrangement was the same regardless of hand dominance, and both hands were used for responding (left hand for the numbers and the right hand for the mouse). All participants were comfortable with the mouse in the right hand and did not express a preference for a left-sided mouse.

As shown in the upper portion of Figure 1, the training task consisted of two distinct tasks performed concurrently, one task presented in the center of the screen and one task presented peripherally. Participants were free to give a response for either task. Task administration lasted approximately 45 minutes. The number of trials depended on participant response. These tasks were developed based on tasks used by Kramer et al. (1995). The tasks were designed to have both memory and visuospatial demands in addition to the key aspects of variable priority and individualized feedback included in the original dual task training studies conducted by Kramer et al. (1999).

Version A—In the central number task a 4×4-cm white box with 2.5×1.5-cm blue numbers was presented centrally on the display. Participants pressed one of three number keys (1, 2, or 3) to indicate the correct value when the displayed numbers (presented serially in the center of the white box) were summed. Summing was continuous so that the displayed number advanced with each key press by the participant. The space bar could be used to advance to the next number without entering a value. If 1500 ms elapsed without a key press, the number advanced, and the response was logged as a miss. Numbers were presented from a pseudorandom list generated with constraints that the summed value equal 1, 2, or 3 in all instances, that no numbers repeat consecutively, and that all presented numbers were single digits. Negative values were included in the presentation list but not among possible solutions. Response categories included correct, wrong, miss, and spacebar. For each response the latency was recorded, and after each response the timer was reset. Two scores from this task were used in the data analyses: number of correct key presses per minute and median RT for correct key presses.

In the peripheral flower task a 6×4-cm flower image was presented at a random location on the screen (constrained not to appear in the center of the screen or reappear in the same location consecutively) and remained in that location until the participant 'cut' it with the scissor cursor or 1500 ms elapsed, after which time the flower disappeared from the screen and reappeared at another location. The scissor cursor was a 6×5-cm pair of cartoon scissors with a central *X* placed between the blades of the scissors for proper positioning over the flower. The mouse controlled the position of the scissors, and the left button of the mouse changed the image of the scissors to close. The category of response (*hit* if over the flower, *clicked not on flower* if over any other location on the screen) was recorded, along with the time interval from the presentation of the flower until the response. If the flower was not accurately cut after 1500 ms, the flower disappeared, and a *miss* was logged. Only hits and misses reset the timer. If the participant clicked not on flower, the error was logged, but the trial did not advance, and the timer continued until either a hit or miss (>1500 ms) occurred. Two scores from this task were used in the data analyses: number of correct mouse presses per minute and median RT for correct mouse presses.

Version B—An alternate version of the training task was developed to test for transfer to a similar task with different stimuli. Version B had the same properties as Version A, but the specific materials differed such that in the central position an alphabet task was substituted for a number-summing task. Participants entered the distance a letter was away from the previously presented letter in the alphabet (e.g., H is 3 units away from K). The same variables were collected for data analysis: number of correct key presses per minute and median RT for correct key presses. In the peripheral task of Version B, the scissors were replaced with a leg, and the flower was replaced with a soccer ball. The object was to kick the soccer ball through the goal (aiming was not necessary). An additional difference between the versions was that the soccer ball was a slowly moving target requiring the coordination of the leg location with the soccer ball's trajectory, whereas the flower was stationary. The two scores from this task were the same as in Version A: number of correct mouse presses per minute and median RT for correct mouse presses.

Variable priority—In addition to the central and peripheral tasks, a third element of task prioritization with feedback monitoring was included in the multitask training. Two rectangular bars of equal height (4.5 cm) were stacked on top of each other in the upper left corner of the display. The top bar represented feedback and priority for the central task, and the bottom bar represented the same for the peripheral task. The bars were labeled and color coded such that the central task was colored green and red and the peripheral task was colored blue and yellow. The stacked bars varied in length depending on the priority condition. They were generated to create yoked priorities indicating the relative emphasis of the two embedded tasks: 20:80, 40:60, 50:50, 60:40, and 80:20. Each priority condition was set to a fixed time interval of 3000 ms and repeated 12 times in a pseudorandom order constrained so that no priority was repeated consecutively.

Feedback was displayed using the color contrast of green or red and blue or yellow. Feedback was calculated based on the RT of the previous five correct responses for each task separately. This value was then modified based on the priority level and the participant's baseline performance from the last five responses. If participants performed at the same level as baseline (proportional to the priority level at the time) then the bar was either completely green (for the number-summing task) or completely blue (for the flower-cutting task). If they were slower, then the amount of green (or blue) in the bar was less and the red (or yellow) was revealed. Participants were instructed to attempt to keep the bar completely filled with green or blue. Fixed priority changes influenced the relative amount of effort required to keep each bar filled (with green or blue). See the display of feedback in the upper left hand corners of Figure 1.

Statistical analyses

In order to evaluate Hypothesis 1 confirmatory factor analysis (CFA; LISREL; Jöreskog & Sörbom, 2005) was used to aggregate the baseline normative ($N = 105$) data in terms of constructs with uniquely shared variance rather than based on task specific similarities (to avoid the task-impurity problem). The predicted factor structure was based on previous exploratory and CFA studies with the same or similar indicator variables (Friedman & Miyake, 2004; Salthouse, 1996; Salthouse & Meinz, 1995) that established relationships among these constructs of processing speed, working memory, and aspects of inhibition. The model fit for the CFA was tested, estimated reliability was evaluated, and the nature of each indicator variable was reviewed for adherence to established expectations. For example, means were reviewed to determine if basic predictions of longer RTs for incongruent versus congruent conditions of response-distractor tasks were met. To the extent that the necessary *a priori* expectations were not met, adjustments were made to subsequent analyses. Multitask performance measures were then tested within the CFA for association

with the obtained factors. In addition, given Kramer et al.'s suggestion that multitasking is related to task switch/task coordination, a measure of task switching was also characterized as it related to the factors obtained in the CFA.

Mixed analyses of variance (SPSS, 2001) including task version as a between subjects independent variable and time (six sessions) as a within subjects independent variable were conducted for each multitask performance measure to test for improvement across sessions and any effect of version on training gains (Hypothesis 2).

Near transfer of training improvements (Hypothesis 3) was tested using an alternate version of the training task in session 7. The effect of version was again tested within these analyses of variance by including both task version and group (training vs. control) as between subjects independent variables for each performance measure.

Hypothesis 4 was a test of far transfer of training-related improvement on the multitask to improvement on other cognitive measures. This hypothesis was exploratory because the limited size of the training and control groups ($n = 41$ combined) made it difficult to detect a small effect and effects of far transfer are often small, if observed (Barnett & Ceci, 2002). Training-related performance improvement was examined by comparing group (training vs. control) performance differences at post test for each of the tasks only for tests in which the cognitive construct was found to be uniquely related to the training task, thus limiting the probability of a Type I error and adopting a conservative standard. It is important, however, to keep in mind that at the task-level the indicator variables (i.e., individual tests) of a construct also reflect variance associated with other cognitive constructs and task-specific processes. For example, variance in Digit Symbol reflects not only processing speed but also individual differences in memory (for number/symbol pairing) and motor control. The multidimensionality can be seen in the simple correlations among the individual tests reported in the Appendix.

RESULTS

Construct validation

Means, standard deviations, and the intercorrelations for the measures from the cognitive battery for the 105 participants at the initial session are shown in the Appendix. Confirmatory factor analysis for a two-factor model of Processing Speed and Working Memory as described in the methods section provided a reasonably good fit to the data, $\chi^2(8) = 11.34$, $p = .18$; RMSEA = 0.07, 90% confidence interval was 0.00 to 0.14); CFI = 0.98; the two factors were correlated ($r = .64$). The standardized factor loadings for the individual measures are shown in a diagram of the measurement model (Figure 2). When the switching task was entered into the two-factor measurement model it was significantly related to Processing Speed (0.46) but not to Working Memory (0.03); see Table 1. The reliability estimate for the switch task from this analysis was 0.77.

The correlations between performance on the training multitask and the individual cognitive measures at the initial session are shown in Table 2. Measures from the central task were modestly correlated with the cognitive measures, although those from the peripheral task were mainly not significant. To determine the amount of initial training performance variance unique to each of the cognitive factors, correct key press per minute, mouse press per minute, key RT, and mouse RT were entered into the two-factor CFA measurement model. Only correct key RT from the central task was significantly associated with a factor, its standardized factor loading on Processing Speed was 0.58, see Table 1.

Effect of training and near transfer

A preliminary set of analyses examined the effect of version of the training task. Mixed analyses of variance including task version (A or B) as a between subjects independent variable and time (the initial testing session plus the five training sessions) as a within subjects independent variable were conducted for each of the four measures (number of correct key and mouse presses per minute and median RT for correct key and mouse presses). For each of the four measures each participant's scores from the pre-test and the five training sessions (within participant performance) were first converted to z scores to eliminate individual differences in speed; therefore the between subjects effect of version always produced an F of 0. In all four analyses the sphericity assumption was violated, and the Greenhouse–Geisser correction was applied. Degrees of freedom were rounded to whole numbers for the F ratios reported here.

There was no interaction between version and time (all $Fs < 1.0$). That is, the pattern of change over the sessions did not vary as a function of version of the task. Therefore the means and standard deviations for the standardized scores from the four measures collapsed over version are shown in Table 3. There were very large effects of training (time) on each measure in the predicted direction of improved performance over time. In the central task (key press) participants increased the number of correct keys pressed per minute over the six sessions, $F(2, 33) = 159.16$, $p < .0001$, $\eta^2 = .90$. Median RT per key press response decreased over time, $F(3, 45) = 24.09$, $p < .0001$, $\eta^2 = .57$. For the peripheral task (mouse hits), participants increased the number of correct mouse hits per minute over the six sessions, $F(2, 37) = 128.10$, $p < .0001$, $\eta^2 = .88$. Median RT per mouse hit decreased over time, $F(2, 30) = 61.11$, $p < .0001$, $\eta^2 = .77$.

To test near transfer, analyses of variance including task version and group (training vs. control) as between subjects independent variables were conducted for each performance measure (correct key press per minute, mouse hit per minute, key RT, and mouse RT) obtained on the *novel* multitask administered at the final testing session (post-test). Scores were standardized within version collapsing across training and control groups.

There was a main effect of group for key press per minute, $F(1, 37) = 5.20$, $p < .05$, $\eta^2 = .12$; mouse press per minute, $F(1, 37) = 7.81$, $p < .01$, $\eta^2 = .17$; and mouse RT, $F(1,37) = 4.13$, $p < .05$, $\eta^2 = .10$, but no interactions between version and group ($ps > .30$). That is, the difference between groups did not vary as a function of the version of the task. Hence, the means shown in Table 4 are collapsed across version. The training group demonstrated more key presses per minute, more mouse presses per minute, and faster mouse RT on the novel multitask than the control group. Although the training group had faster key RTs than did the control group, the difference was not significant, $F(1, 37) = 0.95$, $p = .33$.

Far transfer

Given that some of the measures from the training task were correlated at pre-test with the Processing Speed factor and the switching task, mixed analyses of variance comparing the training and control groups at pre-test and post-test were conducted for the three measures from the Processing Speed factor (Simple RT, Pattern Comparison, and Digit Symbol) and the switching task. See Table 5 for means and standard deviations.

For simple RT there were significant effects of group, $F(1, 39) = 3.95$, $p = .05$, $\eta^2 = .098$, and session (pre-test to post-test) for simple RT, $F(1, 39) = 16.50$, $p = .001$, $\eta^2 = .29$, that were qualified by a significant × Session interaction, $F(1, 39) = 4.09$, $p = .05$, $\eta^2 = .10$. Unexpectedly, both groups were slower (rather than faster) at post-test, but the slowing was less pronounced in the training group (298 vs. 310 ms) than in the control group (308 vs. 344 ms).

For Pattern Comparison the group effect was not significant, $F(1, 39) = 1.87$, $p = .18$, but session was, $F(1, 39) = 4.36$, $p = .04$, $\eta^2 = .10$. The Group × Session interaction was not significant, $F(1, 39) = 0.48$, $p = .49$. Both groups improved slightly.

For Digit Symbol the group effect approached significance, $F(1, 39) = 3.27$, $p = .08$, $\eta^2 = .08$, and the effect of session was significant, $F(1, 39) = 5.31$, $p = .03$, $\eta^2 = .12$. The Group × Session interaction also approached significance, $F(1, 39) = 3.47$, $p = .07$, $\eta^2 = .08$. As predicted, scores improved at post-test and more so in the training group (52.50 vs. 63.70) than the control group (64.81 vs. 66.00). This effect, however, may be accounted for by the unusually low pre-test score in the training group.

Finally, there were no significant effects for the switching task: group, $F(1, 39) = 2.32$, $p = .14$; session, $F(1, 39) = 1.42$, $p = .24$; Group × Session, $F(1, 39) = 0.32$, $p = .58$. In general, there was little evidence of transfer of the training effect to related measures.

DISCUSSION

Substantially improved performance was obtained after training on multitasks proposed to enhance attentional control. Compared with a control group that underwent a comparable amount of training on use of the internet, the effect of training using a variable priority strategy accounted for 10 to 17% of the variance when near transfer was tested with a different version of the training task. These results are congruent with those reported previously (Bherer et al., 2005, 2006, 2008; Kramer et al., 1995, 1999) in other settings using different tasks and research designs.

Although it is clear that older adults can improve their performance on multitasks, what are the processes that are affected by the training? Despite modest associations of the training task with the individual tasks representing working memory, in terms of latent variables the training task was associated only with processing speed, not with working memory. This indicates that working memory is involved in the training task only to the extent that both working memory and the training multitask are related to processing speed. It appears that the training multitask does not assess any other aspects of cognitive function that are required in the shared variance among the tests of working memory used in this study. In other words, the statistical control of a latent variable approach to address the task-impurity problem demonstrated that the observed simple correlations between working memory tasks and the training tasks resulted from shared individual differences in processing speed rather than variance that was specifically attributable to the working memory construct.

The switch task, which was expected to assess processes similar to those involved in switching back and forth between the two components of the variable priority training task, also was associated with processing speed but not working memory, again despite modest associations of the switch task with the individual working memory tests. The same reasoning can be applied to it. Its modest association with tests of working memory arises because both the switch task and the tests of working memory are related to speed of processing.

This finding of unique variance associated with processing speed but not with working memory supports Salthouse's (1996) explanation for the relationship between measures of processing speed and other aspects of cognition which may not have an overt relationship to speed. According to his theory, the shared variance among the measures of processing speed and working memory may be explained by a time limitation for certain mental processes such that only early stages of processing are completed due to slowed execution; thus the computation is not fully carried out. Alternatively, there is a necessary simultaneity of

processing for certain tasks, and slowing could disrupt this interdependence among the processing streams necessary to carry out the computation successfully.

Although performance on the training multitask improved with training, by and large far transfer was not obtained. Even though the training task shared variance with measures of processing speed, whatever beneficial effect training had on the multitask, it did not translate to improved performance on related measures of speed. Similarly, there was no directly observable training benefit on the switch task. Such a result is not surprising, as far transfer of the benefits of cognitive training has been difficult to achieve (e.g., Ball et al., 2002; Edwards et al., 2002; Willis et al., 2006). This study highlights an important distinction between the characterization of the training task to understand the cognitive processes associated with baseline performance and the additional identification of specific processes that improve with training and thus may be observed in improved performance on other tasks. A complete characterization of a training task into these general performance and training-specific aspects should increase the likelihood of observing far transfer.

Perhaps performance on the training multitask and the switch task are related to other, additional aspects of executive function distinct from working memory and speed. The attempt to relate the training multitask or the switch task to inhibition failed in the current study because of difficulties with the measures chosen to represent the response-distractor inhibition construct reported by Freidman et al. (2004). Previous studies of undergraduates reported higher reliability estimates but also used more trials for each task (e.g., Freidman & Miyake, 2004 and Miyake, 2000). Feasibility in terms of the length of the cognitive battery limited the number of trials per task in this study, and this change may have reduced reliability for some of the measures. It would certainly be worth another attempt to examine the inhibition construct as a potential component of the multitask and to determine if it is influenced by variable priority training. The current approach was to systematically examine a small set of potential underlying cognitive constructs selected based on previous empirical work. In general, a better understanding of the various aspects of cognition that are involved in performing both components of the multitask and in alternating between them would provide a major step forward in understanding what training does and does not affect. The next step then would be to examine transfer to functional outcomes rather than other laboratory measures of cognition. Of course, functional measures will likely have some of the same methodological obstacles that were encountered in the laboratory cognitive measures, including the need for ground work to address test and construct validity.

This study replicated a robust effect of variable priority training and transfer to a similar multitask context in a group of healthy older adults. Further, it demonstrated the importance of accounting for practice effects as evidenced by improved performance in the control group with differential benefit in the training group on only select measures of the training task. The specification of the training task is an important step to understand training benefit and transfer. It is also key for theoretically and empirically motivated hypothesis generation of targets of far transfer. This study used an empirically based approach to examine the relationships between the training task and cognitive constructs thought to be related to it. It demonstrated the importance of distinguishing between task- and construct-level relationships, with the former subject to misinterpretation due to task-level impurity.

Further, this study highlights the important distinction between the identification of training task processes related to performance in general versus the processes linked to training gains and targeted for testing transfer effects. Given the difficulties in achieving observable far transfer, it is important to not only cast a wide net to identify the cognitive processes responsible for training gains but to do so with the specificity so as to distinguish among interrelated processes and effects that are not specific to training gains.

Acknowledgments

APPENDIX

Means, standard deviations, and correlations of measures in cognitive battery

Measure	2	3	4	5	6	7	*M*	*SD*
1. Simple RT	−0.21	−0.31	−0.20	−0.15	−0.28	−0.23	293.55	46.21
2. Pattern Comparison		0.37	0.34	0.44	0.12	0.26	27.67	4.78
3. Digit Symbol			0.26	0.29	0.13	0.32	64.46	15.76
4. Operation Span				0.73	0.30	0.24	28.02	19.90
5. Reading Span					0.30	0.28	29.83	19.31
6. Two Back						0.26	60.54	7.22
7. Switch							0.91	0.16

Note: Correlations 0.20 are significant at .05. The switch task was not included in the confirmatory factor analysis.

REFERENCES

Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, et al. Effects of cognitive training interventions with older adults: A randomized controlled trial. Journal of the American Medical Association. 2002; 288:2271–2281. [PubMed: 12425704]

Barnett SM, Ceci SJ. When and where do we apply what we learn? A taxonomy for far transfer. Psychological Bulletin. 2002; 128:612–637. [PubMed: 12081085]

Bherer L, Kramer AF, Peterson MS, Colcombe S, Erickson K, Becic E. Training effects on dual-task performance: Are there age-related differences in plasticity of attentional control? Psychology and Aging. 2005; 20:695–709. [PubMed: 16420143]

Bherer L, Kramer AF, Peterson MS, Colcombe S, Erickson K, Becic E. Testing the limits of cognitive plasticity in older adults: Application to attentional control. Acta Psychologica. 2006; 123:261–278. [PubMed: 16574049]

Bherer L, Kramer AF, Peterson MS. Transfer effects in task-set costs and dual-task cost after dual-task training in older and younger adults: Further evidence for cognitive plasticity in attentional control in late adulthood. Experimental Aging Research. 2008; 34:188–219. [PubMed: 18568979]

Edwards JD, Wadley VG, Myers RS, Roenker RL, Cissel GM, Ball KK. Transfer of a speed of processing intervention to near and far cognitive functions. Gerontology. 2002; 48:329–340. [PubMed: 12169801]

Erickson KI, Colcombe SJ, Wadhwa R, Bherer L, Peterson MS, Scalf PE, et al. Training-induced functional activation changes in dual-task processing: An FMRI study. Cerebral Cortex. 2007; 17:192–204. [PubMed: 16467562]

Friedman NP, Miyake A. The relations among inhibition and interference control functions: A latent-variable analysis. Journal of Experimental Psychology: General. 2004; 133:101–135. [PubMed: 14979754]

Gopher D, Weil M, Bereket Y. Transfer of skill from a computer game trainer to flight. Human Factors. 1994; 36:387–405.

Gopher D, Weill M, Siegel D. Practice under changing priorities: An approach to training of complex skills. Acta Psychologica. 1989; 71:147–178.

Jöreskog, KG.; Sörbom, D. Lisrel 8.72. Scientific Software International; Chicago, IL: 2005.

Kane MJ, Engle RW. Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. Journal of Experimental Psychology: General. 2003; 132:47–70. [PubMed: 12656297]

Kane MJ, Conway ARA, Miura TK, Colflesh GJH. Working memory, attentional control, and the n-back task: A question of construct validity. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2007; 33:615–622.

Kramer AF, Larish JF, Strayer DL. Training for attentional control in dual task settings: A comparison of young and old adults. Journal of Experimental Psychology: Applied. 1995; 1:50–76.

Kramer, AF.; Larish, J.; Weber, T.; Bardell, L. Training for executive control: Task coordination strategies and aging. In: Gopher, D.; Koriat, A., editors. Attention and performance XVII. MIT Press; Cambridge, MA: 1999. p. 617-652.

Kramer AF, Bherer L, Colcombe SJ, Dong W, Greenough WT. Environmental influences on cognitive and brain plasticity during aging. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2004; 59:M940–M957.

Minear M, Shah P. Training and transfer effects in task switching. Memory and Cognition. 2008; 36:1470–1483.

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. Cognitive Psychology. 2000; 41:49–100. [PubMed: 10945922]

Salthouse TA. Processing-speed theory of adult age differences in cognition. Psychological Review. 1996; 103:403–428. [PubMed: 8759042]

Salthouse TA. Aging and measures of processing speed. Biological Psychology. 2000; 54:35–54. [PubMed: 11035219]

Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. Developmental Psychology. 1991; 27:763–776.

Salthouse TA, Meinz EJ. Aging, inhibition, working memory, and speed. Journals of Gerontology: Psychological Sciences. 1995; 50B:297–306.

SPSS. For Windows, Rel. 11.0.1. SPSS Inc; Chicago, IL: 2001.

Turner ML, Engle RW. Is working memory capacity task dependent? Journal of Memory and Language. 1989; 28:127–154.

Unsworth N, Heitz RP, Schrock JC, Engle RW. An automated version of the operation span task. Behavior Research Methods. 2005; 37:498–505. [PubMed: 16405146]

Verhaeghen P, Marcoen A, Goossens L. Improving memory performance in the aged through mnemonic training: A meta-analytic study. Psychology and Aging. 1992; 7:242–251. [PubMed: 1535198]

Wechsler, D. Wechsler Memory Scale. 3rd ed.. Psychological Corporation; San Antonio, TX: 1997.

Willis SL, Schaie KW. Training the elderly on the ability factors of spatial orientation and inductive reasoning. Psychology and Aging. 1986; 1:239–247. [PubMed: 3267404]

Willis SL, Tennstedt SL, Marsiske M, Ball K, Elias J, Koepke KM, et al. Long-term effects of cognitive training on everyday functional outcomes in older adults. Journal of the American Medical Association. 2006; 296:2805–2814. [PubMed: 17179457]

YoYo Games Ltd. *Game Maker* (Version 7.0) [Software]. Dundee, Scotland, UK: 2007. Available from http://www.yoyogames.com/gamemaker
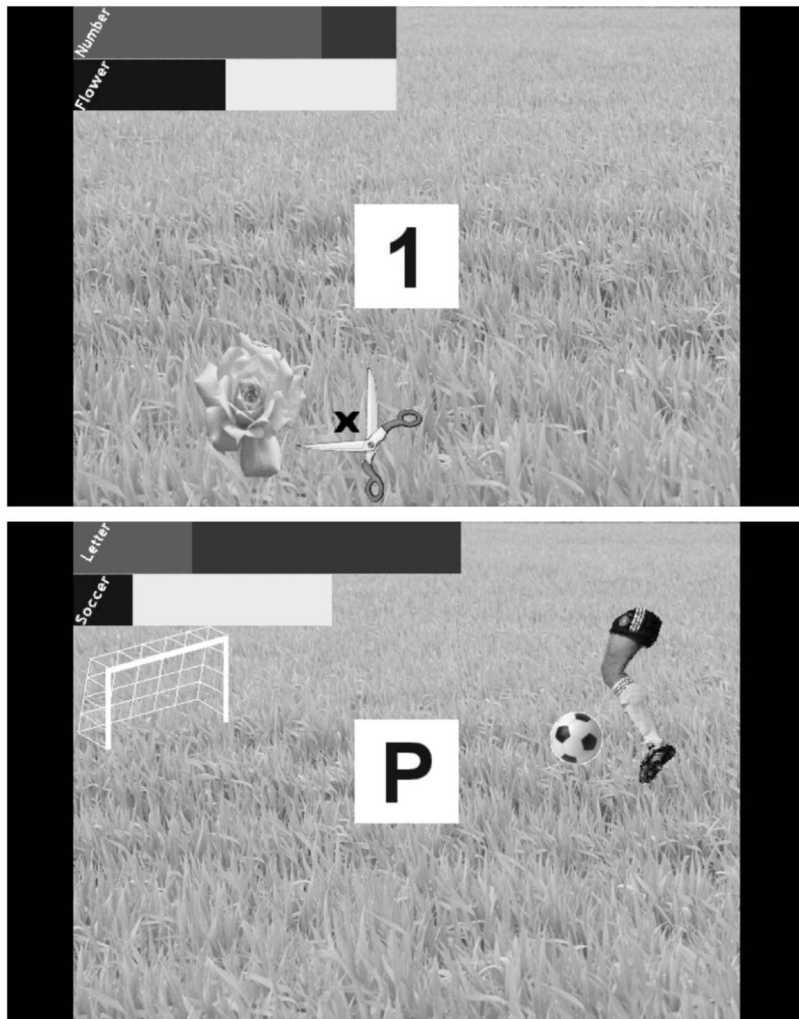
Figure 1.
The upper figure shows version A of the multitask, and the lower figure shows version B.
The central task is the number in version A or letter in version B in the center of the screen.
The peripheral task is the flower cutting in version A and the soccer ball kicking in version
B.
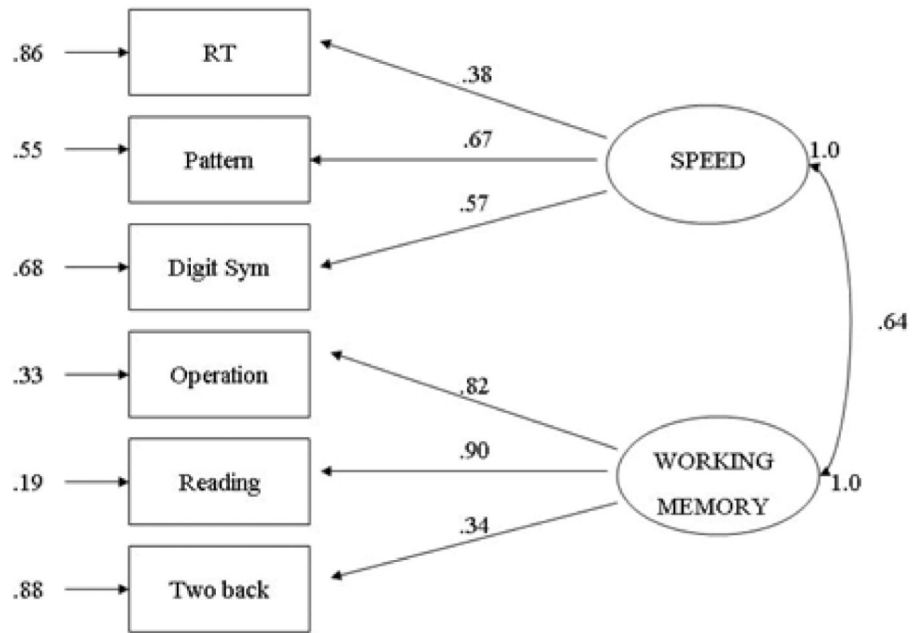
Figure 2.
Estimated two-factor model. Single-headed arrows have standardized factor loadings next to them (all significant at the .05 level). The numbers on the far left at the ends of the smaller arrows are error terms. Squaring these terms gives an estimate of the variance for each task that is not accounted for by the latent construct (i.e., task-specific variance). The curved, double-headed arrow indicates a significant correlation between the latent variables 'working memory' and 'speed'.

TABLE 1

Standardized factor loadings for training task performance and switching on the latent variables working memory and speed

Measure	Working memory	Speed
Switch	0.03	−0.46[*]
Training task		
Key press	0.27	0.27
Mouse press	0.01	0.28
Key RT	0.11	0.58[*]
Mouse RT	0.13	−0.08
relax		

[*] $p < .05$.

TABLE 2

Correlation of measures in cognitive battery with multitask performance at initial assessment (N = 105)

Measure	Key press	Mouse press	Key RT	Mouse RT
Simple RT	−0.26	−0.15	0.28	0.06
Pattern *Comparison*	0.27	0.13	−0.32	−0.12
Digit Symbol	0.23	0.01	−0.27	−0.10
Operation Span	0.39	0.25	0.29	−0.13
Reading Span	0.34	0.20	−0.18	−0.14
Two Back	0.45	0.22	−0.24	−0.08
Switch	0.33	0.16	−0.10	−0.14

Note: Correlations 0.20 are significant at .05.

TABLE 3

Means (and standard deviations) of standardized (within each participant) scores of the training group ($n = 20$) on the four multitask measures from the initial pre-test and training sessions 1 to 5

Session	Key press	Mouse press	Key RT	Mouse RT
Pre-test	−1.64 (0.17)	−1.71 (0.35)	1.29 (0.64)	1.64 (0.65)
1	−0.52 (0.26)	−0.36 (0.47)	0.38 (0.71)	0.31 (0.67)
2	0.11 (0.33)	0.19 (0.33)	−0.04 (0.69)	−0.29 (0.37)
3	0.41 (0.33)	0.37 (0.29)	−0.30 (0.46)	−0.38 (0.25)
4	0.71 (0.26)	0.70 (0.17)	−0.51 (0.55)	−0.58 (0.23)
5	0.92 (0.40)	0.81 (0.28)	−0.81 (0.65)	−0.71 (0.29)

TABLE 4

Means (and standard deviations) of the training and control groups' standardized performance measures on the novel multitask at post-test

Measure	Training (*n* = 20)	Control (*n* = 21)
Key press	0.34 (1.00)	−0.33 (0.87)
Mouse hit	0.41 (.98)	−0.39 (0.84)
Key RT	−0.15 (0.83)	0.15 (1.11)
Mouse RT	−0.31 (0.74)	0.30 (1.11)

TABLE 5

Means (and standard deviations) for the training and control group on four cognitive measures at pre-test and post-test

Time	Training	Control
Simple RT		
Pre-test	298.33 (32.99)	308.02 (50.47)
Post-test	310.12 (30.18)	343.70 (41.01)
Pattern Comparison		
Pre-test	29.20 (4.30)	27.33 (3.62)
Post-test	30.20 (4.46)	28.33 (3.31)
Digit Symbol		
Pre-test	52.50 (12.02)	64.81 (16.67)
Post-test	63.70 (17.52)	66.00 (15.27)
Switch		
Pre-test	0.92 (0.20)	0.98 (0.10)
Post-test	0.90 (0.10)	0.94 (0.12)