In-Home Cognitive Training with Older Married Couples: Individual versus Collaborative Learning

Jennifer A. Margrett and Sherry L. Willis
Jennifer A. Margrett, West Virginia University, Department of Psychology, 53 Campus Drive, P.O. Box 6040, Morgantown, West Virginia 26506;

Abstract
Research has demonstrated that older adults' cognitive performance can be enhanced via formal intervention (e.g., Gratzinger, Sheikh, & Friedman, 1990; Kliegl, Smith, & Baltes, 1990; Willis & Schaie,1986) as well as more informal intervention including collaboration, or working with a partner (e.g., Dixon & Gould, 1996; Margrett & Marsiske, 2002; Staudinger & Baltes, 1996).The current study investigated the effects of an inductive reasoning training program adapted for in-home use (Willis & Schaie, 1986, 1994) among older adults assigned to individual training (n = 30), collaborative training (n = 34), or a no-treatment control group (n = 34). The training consisted of ten sessions, and all participants completed a pretest followed by a posttest six weeks later. Findings suggest that older adults could effectively "train themselves" without the guidance of a formal instructor. The results, however, did not indicate immediate added benefit in reasoning performance for collaborative versus individual training using the current reasoning program.

In-Home Cognitive Training with Older Married Couples: Individual versus Collaborative Learning

Cognitive plasticity in later adulthood has been demonstrated using a variety of experimental approaches. These approaches include formal cognitive training, extensive practice without instruction, and most recently collaborative learning. While significant improvement in cognitive performance has been shown employing each of these procedures, the approaches have differed in conception and in experimental protocol. Moreover, research focusing on combining or comparing the approaches is very limited. In this study the facilitative effects of collaborative cognition are examined in the context of a formal cognitive training intervention.

Recently, collaboration as a mechanism for enhancing cognitive functioning in older adulthood has received increasing attention (for discussion see Meegan & Berg, 2002 and Strough & Margrett, 2002). Collaborative cognition research paradigms which have demonstrated positive effects of dyadic interaction include studies of collaborative storytelling and prose recall (Dixon & Gould, 1996, 1998; Gould & Dixon, 1993; Gould, Osborn, Krein, & Mortenson, 2002; Gould, Trevithick, & Dixon, 1991), wisdom-related performance (Staudinger & Baltes, 1996), social-advice giving (Margrett & Marsiske, 2002, 2004), comprehension of everyday printed materials such as medication schedules (Margrett & Marsiske, 2002, 2004), and errand planning/everyday life management (Berg, Meegan, & Johnson, 1995; Margrett & Marsiske, 2002, 2004).

Studies of collaborative cognition have varied not only in the nature of the task, but also in the extensiveness of collaborative interaction required and the familiarity of the partners. In their 1996 study, Staudinger and Baltes examined adults' advice giving on wisdom-related tasks and the effect of varying degrees of collaborative interactive effort. In two collaborative conditions (discussing a life dilemma with a familiar social partner, or imagining a discussion with a familiar social partner), participants provided higher quality solutions compared to other conditions where task demands were less explicitly collaborative.

Collaborative studies suggest that partner familiarity (i.e., working with a significant other compared to a stranger) may be an important factor in collaborative process and outcome, although studies have yielded equivocal results (e.g., Gould, Osborn, Krein, & Mortenson, 2002). In earlier work Gould, Dixon, and their colleagues found that older unfamiliar dyads (i.e., strangers) spent more time in task-unfocused social interactions, and this actually impeded prose recall performance (Gould, Kurzman, Dixon, 1994). Thus, in the prose memory research collaboration may come with costs, particularly reflected in an increase in social "getting to know you" behaviors among unfamiliar dyads. Additionally, Margrett and Marsiske (2002, 2004) found that although older unfamiliar dyads benefited from collaborating with a partner on an everyday cognitive task which involved comprehension of everyday printed materials, familiar dyads (i.e., spouses) profited significantly more from their collaboration on this particular task. Using this same data set, Kimbler (2004) found differential patterns of collaborative behavior for familiar spouses and unfamiliar strangers.

Despite the overall positive results found for cognitive collaboration in later life, research investigating recall suggests that interactive or "transactive" memory may come at a cost. Andersson and Rönnberg (1995, 1996) observed diminished memory recall in the collaborative condition compared to the individual condition and suggested that time and efforts devoted to the collaborative and social aspects of the task may have impeded recall strategies.

Summarizing the recent literature, it appears that the effects of cognitive collaboration may be positive and facilitative of elders' everyday cognitive performance on a variety of tasks (i.e., prose recall, wisdom-related advice giving, comprehension of printed materials, route planning/everyday life management). At the same time, positive effects appear to be facilitated by having familiar social partners, explicit collaborative instructions, and tasks which do not rely on immediate memory recall. This growing body of research suggests several avenues for future study. First, the efficacy of collaborative cognition has typically been examined over only one or two sessions. Further research is needed to investigate the long-term effects of cognitive collaboration. Secondly, to date most studies have focused on relatively complex cognitive tasks (e.g., social advice giving, route planning, everyday life management) rather than basic mental abilities (e.g., reasoning). Thirdly, cognitive collaboration has been investigated as an informal intervention (i.e., instructions regarding optimization of collaboration have not been provided and non-training materials have been used). Research has not investigated the combined utility of collaboration within a formal cognitive training protocol.

Cognitive Training

Formal cognitive training as a mechanism for enhancing performance on mental abilities exhibiting early age-related decline has a longer research history than collaborative cognition. Significant group differences in mental ability performance in comparison to a control group have been shown for abilities such as verbal memory (Gratzinger, Sheikh, Friedman, & Yesavage, 1990; Kliegl, Smith, & Baltes, 1990), inductive reasoning (Willis & Schaie, 1986, 1994), speed of processing, and text processing (Meyer, Young, Bartlett, 1989). Inductive reasoning training effects have been on the order of 0.25 to 0.50 standard deviation units, and over half of participants have shown reliable improvement as a result of training (Willis & Schaie, 1986, 1994).

Although experimental protocols have varied, many intervention studies have involved multiple training sessions that focused on training of strategies known to enhance performance on the target ability and have been directed by a trainer. The training sessions have typically been highly structured with segments of the session devoted to teaching of strategies, practice on use

of strategies with feedback from the trainer, and assessment of performance on the target ability or skill. Few formal cognitive intervention studies have investigated the feasibility of "trainer-less" programs; however, some prior work examining memory training is promising. For instance, in a study of a video-administered mnemonic training program West and Crook (1992) found evidence of trainee benefit. In an initial study of a short-term self-administered memory training program which was an analog to a longer, trainer-led program, Woolverton, Scogin, and Shackelford (2001) also found benefits of the training, however, effects were less robust compared to the original training protocol. Additionally, Blackburn, Papalia-Finlay, Foye, and Serlin (1988) found that older adults benefited more from a cognitive training program which utilized self-instructed interactive group sessions with peers rather than sessions led by a trained instructor. As these prior studies suggest, additional research is needed to investigate the social context of the training experience and its effect on training outcome.

Collaborative Cognitive Training

As discussed earlier, collaborative research in later life has typically focused on joint completion of relatively complex everyday problem-solving tasks (e.g., prose recall, social dilemmas). Indeed, collaboration on social or everyday tasks may depend on domain-specific knowledge with one spouse having knowledge of this type. In addition, long-term married couples are likely to have had prior experience in solving these types of problems and draw upon this prior experience. In the current context, older adults are likely to find that solving inductive reasoning problems are just as cognitively complex but less familiar, thus collaboration on these tasks may facilitate their completion.

Prior research on peer tutoring or peer learning in academic settings may be a useful framework for study of collaboration on structured cognitive tasks. A peer learner/tutor performs tasks such as modeling problem solutions, giving feedback, monitoring the problem solving process and giving procedural support (e.g., how to mark answer choices, pacing of task, page orientation). A key component of the current reasoning training program is acquiring and using strategies specific to the reasoning problems that have been identified in prior research (Kotovsky & Simon, 1973). However, prior research with older learners indicates that they can acquire cognitive strategies, but do not tend to spontaneously use them in problem solving (e.g., Craik & Jacoby, 1996; Kausler, 1994). Of particular interest then is whether collaboration facilitates initial learning, subsequent use (reinforcement), and transfer of these strategies more effectively than traditional training protocols which focus on the individual learner. Since strategy use is associated with higher cognitive functioning, enhanced strategy use should be manifest in cognitive performance at posttest.

Thus, collaboration could serve to enhance the effects of training programs via the tutorial and reinforcing nature of the interaction during training sessions. In addition to the possible positive influence on strategy acquisition, collaboration with a familiar partner could serve to improve participant adherence and enjoyment throughout the duration of a training protocol thereby improving performance.

Rationale for the Current Study and Specific Aims

While it is evident that cognitive training and collaborative cognition have potential benefits for older adults, limitations of each approach remain and the utility of combining these two approaches has not been examined. Thus, this study sought to investigate the efficacy of adapting a successful formal cognitive training protocol for in-home use by older adults and to compare the effects of individual training versus collaborative training with a familiar social partner (i.e., spouse).

The current study addressed several limitations of previous research on cognitive intervention and collaboration in older adulthood: (1) The training protocol involved increased ecological validity as (a) training occurred in participants' homes without a formal trainer and (b) half of the training participants collaborated on the sessions with their spouse; (2) For participants in the collaborative training condition, collaboration was explicitly promoted and specific opportunities for collaboration were provided; (3) Participants in the collaborative training condition collaborated over a sustained period of time (i.e., ten sessions); and (4) Collaboration on a highly structured, systematic, and purposeful task (i.e., basic mental ability strategy training and practice) was examined.

The current study addressed two primary research questions, namely the investigation of the degree of group and individual-level training gain. First, group-level differences between the three treatment groups in reasoning performance assessed immediately after training were examined. Secondly, the degree of reliable individual-level change in performance from pretest to posttest was investigated. It was expected that the two training groups would significantly outperform the Questionnaire Only Group. Additionally, it was hypothesized that the collaborative group would benefit more from the training protocol compared to the individual group. Hypothesized greater retention and use of learned strategies by the collaborative trainees was based on several potentially positive aspects of collaboration including deeper encoding of strategies during initial collaborative learning experiences and greater spousal reinforcement of strategy use after training. The same pattern of results was expected for the individual-level analyses.

Method

Participants

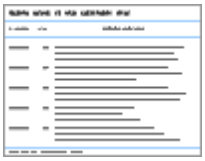
Participants were part of a convenience sample recruited through local media and organizations to participate in a study of cognitive training, however, the study was not intended nor advertised as for those with cognitive or memory problems. The sample included fifty three older couples (N = 106) who were screened prior to study enrollment to ensure that dyads were community-dwelling, reported no limitations in self-care activities (i.e., bathing, dressing, personal hygiene), and had been married for fifteen or more years to their current spouse. Independent, well elders were recruited for two reasons: (1) The training protocol provided less instructional support and required more self-direction compared to prior trainer-guided programs, and 2) a caregiver-care recipient relationship between spouses was avoided as it was expected to bias the dynamics of collaborative interaction.

Fifty three couples were enrolled in the study; however, four of those couples were not included in the present analyses. Two couples that were randomly assigned to a treatment condition did not participate after initial screening, and one couple did not participate after the pretest assessment due to one spouse's poor health. Additionally, one couple in the Individual Training condition was dropped from the current analyses as one partner in this couple completed the pretest and training, but did not participate in the immediate posttest assessment. Thus the final sample totaled 98 individuals (49 dyads). This attrition rate (8%) is comparable to other cognitive training studies using community-based, volunteer samples (e.g., Kliegl, Smith, & Baltes, 1990). T-tests comparing the two couples with demographic data with the sample of 49 couples completing all assessments did not reveal any significant differences in age, education, or marital length between the two groups.

The mean age of the final sample (*N* = 98 individuals) was 71.43 years (*SD* = 5.85; range = 61 – 89) and participants had been married an average of 46.53 years (*SD* = 8.27; range = 15 – 61). The average educational level was 15.95 years (*SD* = 3.06; range = 12 –22) with 76% of participants reporting an educational level of more than twelve years. The median yearly income was $47,644 (range = $18,000 – $50,000+). Using a five-point Likert scale (1 = "Very Good/Happy" to 5 = "Very Poor/Unhappy"), participants rated their physical health as "Good" to "Moderately Good" (*M*= 2.38), their mental health as "Good" (*M*= 1.84), and their life satisfaction as "Very Happy" (*M* = 2.06).

Table 1 depicts the demographic characteristics of participants in each treatment condition. Analysis of variance tests were conducted to examine group differences. The treatment groups (i.e., Questionnaire Only, Individual Training, Collaborative Training) did not significantly differ on any demographic characteristic. Possible gender and within-couple differences were also investigated. Results indicated that overall men tended to be older (Male *M* = 72.55; Female *M* = 70.31; [$F(1, 96) = 3.72$, $p = .06$]) and report more years of education (Male *M* = 17.10; Female *M* = 14.80; [$F(1, 96) = 16.28$, $p < .0001$]) than women in the sample. Next, the within-couple differences (i.e., husband's value minus wife's value) on demographic characteristics were investigated to identify any group differences in spousal variation. These analyses did not reveal any significant differences in spousal variation between groups.

Table 1



Participant Demographic Characteristics as a Function of Group Assignment

Measures

Inductive Reasoning Measures

Inductive reasoning reflects the ability to infer general rules or principles from specific instances of serial tasks and to produce the next element in a series. Three reasoning measures were administered at the pretest and posttest sessions. It is important to note that (a) while the items administered during the pre and posttests used the same strategies emphasized during training, none of the items used in training were included in the pre or posttest assessments, and (b) prior studies of practice only without training of strategies or without feedback on performance have demonstrated reduced effects compared to studies employing an intervention (e.g., Hofland, Willis, & Baltes, 1981).

Two scores were created from each measure: (1) the number of correct responses and (2) the accuracy of responses. The accuracy of responses was calculated by dividing the number of correct responses by the total number of correct, incorrect, and omitted responses. Thus, this accuracy estimation is quite stringent as the denominator represents all items that the respondent considered during the test administration.

Letter Series Test

The Letter Series Test (Blieszner, Willis, & Baltes, 1981) is a 20-item measure which was used to assess respondents' ability to identify the pattern in a series of letters and to generate the next letter in each series ($\alpha = .91$). Series ranged from seven to fifteen letters and the test time was 4.5 minutes. This test has been used in previous training studies to assess individuals' reasoning ability (e.g., Ball, et al., 2002).

Word Series Test

The Word Series Test (Schaie, 1985) is a 30-item test, which required participants to identify the pattern in a vertical series of related words (i.e., days of the week, months of the year) and to produce the next word in the series. The test time was 6 minutes. This test has also been used in previous training studies to assess individuals' reasoning ability (e.g., Ball, et al., 2002).

Letter Sets Test

The Letter Sets Test (Ekstrom, French, Harman, & Derman, 1976) is a 15-item test, which required respondents to identify the pattern in five sets of letters and decide which letter set did not adhere to the pattern. Each letter set consisted of four letters, and the four letter sets that were alike adhered to the same rule or pattern. The test time was 7 minutes. The original test was designed for individual administration; however, for purposes of the current study participants completed this measure jointly with their spouse. This test was chosen to compliment the other two reasoning measures (which have been used in prior work to assess inductive reasoning performance in the context of a training intervention) by assessing performance on a measure utilizing a different type of stimulus. Couples completed this measure in a separate room by themselves and were instructed to collaborate to determine a common response per problem. Scores on this test represent the dyad's score (i.e., one measure was completed and turned in by each spousal pair).

Demographics

As part of a presession homework package, participants completed a paper-and-pencil measure assessing demographic information as well as self-reported ratings on several dimensions including physical and mental health and life satisfaction.

Reasoning Training Materials

The inductive reasoning program in the current study is well established and has been used in the Seattle Longitudinal Study (e.g., Willis & Schaie, 1986, 1994) and more recently in the ACTIVE randomized clinical control trial (Ball, et al., 2002). The training involves ten sessions and in the prior training studies was conducted under trainer direction with either individuals or groups of participants. Training sessions involved two major sections: (1) Practice on basic series problems with an emphasis on the use of strategies shown in prior research to facilitate solution of reasoning problems. Strategies included visual scanning of the item, saying the item aloud in order to hear the pattern, underlining repeated elements in the series, inserting slashes between patterns, and placing tick marks above each skip in a pattern, and (2) Practice on everyday tasks that involved serial problem solving (e.g., understanding a bus schedule, completing a mail-order form, comprehension of a medication schedule). A timed assessment of individual inductive reasoning performance using letter series concluded each session. Total session time was 60 to 75 minutes.

Two modifications were made to the existing training program in order to accommodate completion of materials without a trainer and to promote collaboration in the Collaborative Training condition. First, in order to facilitate self-directed learning in participants' homes in the absence of a trainer, participants received a training booklet and (unlike the original training protocol) an accompanying answer booklet for each session. Each participant in the Individual Training condition received their own answer booklet, whereas Collaborative Training couples shared one answer booklet for each session. The second modification to the training materials involved the identification of sections of material to be completed by the couple together in the collaborative condition. Materials in the Collaborative Training condition included explicit written directions and accompanying pictorial icons indicating sections of the material to be completed collaboratively (i.e., instructions, example items, and checking of answers) and

sections to be completed individually (i.e., practice items and the final timed assessment at the end of each session).

Procedure

Couples were recruited and screened for eligibility during a brief phone interview. The screening criteria were: 1) Sixty years of age or older; 2) No self-reported limitations in activities of daily living or mental impairment, and 3) A minimal marital duration of at least 15 years (Carstensen, Levenson, & Gottman, 1995). Couples were randomly assigned to one of three conditions: (1) Individual Training group (n = 30 individuals), (2) Collaborative Training group (n = 34 individuals), and (3) Questionnaire Only Group (n = 34 individuals). Consistent with prior training studies (i.e., ACTIVE clinical trial: Ball, et al., 2002; Seattle Longitudinal Study: Willis & Schaie, 1986, 1994), couples received compensation for parking and a small honorarium based on their hours of effort (i.e., $40 for the two training groups and $20 for the non-training group).

Pretest and Posttest Assessments

All participants completed a three-hour group pretest assessment session and a two-hour immediate posttest group session which was conducted within one week of the final training session. During both assessments participants completed the inductive reasoning measures. Participants in all three groups were assessed under the same conditions – in the laboratory, in small groups with a trained proctor. Each group testing session contained a *combination* of participants from each of the three treatment conditions and the length of time between pre- and posttests was consistent across the three groups (i.e., posttest occurred 6–7 weeks after the baseline assessment). The same assessment procedure used in the Seattle Longitudinal and ACTIVE training studies was followed. Hence the assessment protocol followed what had been done in trainer-guided training research in order to maximize comparability of findings.

Reasoning Training Protocol and Use of a Control Group

The training protocol consisted of ten sessions. A researcher introduced participants in both training conditions to the training protocol at Session 1 to ensure participants understood the conceptual and procedural elements of the training program. Participants in both training conditions were encouraged to contact the researcher at any time during training with any questions that might arise during the course of training. In addition, after Session 5 materials from the first five sessions were collected and the training protocol was discussed with the researcher in person. At this time participants were also given materials for Sessions 6–10.

Individual Training

Participants in the Individual condition completed all materials individually for sessions two through ten in their home. As a measure of adherence participants signed a statement at the beginning of each session indicating that they would complete the materials alone and would not discuss training with their spouse. Spouses in the Individual condition were encouraged to complete sessions at differing times in the home or at a minimum to work in separate rooms during training sessions.

Collaborative Training

Trainees in the Collaborative condition completed all sessions with their spouse. At the beginning of each session participants signed a statement indicating that they would complete the materials in conjunction with their spouse. Couples in the Collaborative training condition were videotaped during Sessions 1, 5, and 10. Videotaping of Sessions 1 and 10 were conducted at a central testing location (e.g., laboratory) and Session 5 was videotaped either in participants' homes or at a central testing location depending on couple preference.

Questionnaire Only Control Group

Several factors weighed in on the decision to use a no treatment control group. First, when exploring a new intervention procedure the standard comparison group is a no treatment control. It is only after numerous replications of an intervention are conducted and the magnitude of treatment effects are established across multiple studies that alternative comparison groups are most useful. Second, given the possibility of practice effects on novel cognitive tasks for elders, a no treatment control was particularly important to demonstrate that training effects were larger than practice effects.

Results

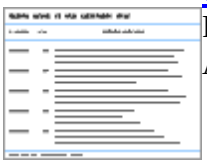
Prior to the analyses, the treatment of each dependent variable and mean performance scores are discussed. Next, the first set of analyses examined mean-level group differences in performance on the inductive reasoning measures. The second set of analyses investigated individual-level change, specifically the proportion of individuals in each group showing reliable pretest-posttest change on the reasoning measures.

Reasoning Test Performance Scores

The total raw number of correctly answered items varied across reasoning tests at Pretest [Letter Series (M = 8.11; SD = 3.39), Word Series (M = 14.11; SD = 4.63), and Letter Sets (M = 8.76; SD = 2.90)] and Posttest [Letter Series (M = 10.71; SD = 3.77), Word Series (M = 16.48; SD = 4.91), and Letter Sets (M = 10.22; SD = 2.56)] assessments. There was substantial variation in the number of correctly answered items given the maximum item limit of each test: Letters Series = 20; Word Series = 30; Letter Sets = 15. To facilitate the interpretability of training effects across reasoning tests, the total number of correct items is expressed in T-score metric (M = 50; SD = 10) and was used in the first set of analyses. For each measure, scores were standardized for the entire sample and the pretest mean and standard deviation were used as the base. This procedure was used to preserve differences between the pretest and immediate follow-up assessments.

Accuracy scores for the reasoning measures were computed, resulting in a proportional score: (Total Number of Correct Items) / (Total Number of Correct Items + Total Number of Incorrect Items + Total Number of Omitted Items). Accuracy scores were not standardized, and thus represent the raw proportional scores. The accuracy proportion varied across reasoning tests at Pretest [Letter Series (M = .62; SD = .22), Word Series (M = .82; SD = .15), and Letter Sets (M = .77; SD = .19)] and Posttest [Letter Series (M = .74; SD = .19), Word Series (M = .86; SD = .14), and Letter Sets (M = .68; SD = .17)] assessments. As evident by these means, participants tended to be very accurate in their responses, resulting in less variation in accuracy scores compared to the number of correctly answered items. Both t-score and accuracy means and standard deviations depicted by treatment group and assessment occasion are presented in Table 2.

Table 2



Reasoning Performance (t-scores) as a Function of Group Assignment and Assessment Occasion

Training Effects: Group Differences

A series of repeated measures analyses of variance (ANOVA) tests was conducted to examine training group differences in reasoning performance. Specifically, for each reasoning measure two 3 (Condition: Questionnaire Only, Individual Training, Collaborative Training) x 2 (Occasion: Pretest, Posttest) repeated measures ANOVAs were conducted with the variables

number of items correctly answered and proportion of items accurately answered serving as the dependent variables. Condition (i.e., treatment group) represented the between-subjects factor and Occasion (i.e., time of assessment) was the within-subjects factor.

Letter Series Test

Total Number of Correct Items

A significant main effect of Occasion, [$F(1, 95) = 93.77$, $p < .0001$, $\eta^2 = .50$], and Condition, [$F(2, 95) = 7.40$, $p < .01$, $\eta^2 = .14$], were observed. These main effects were qualified by a significant Occasion x Condition interaction, [$F(2, 95) = 23.91$, $p < .0001$, $\eta^2 = .36$]. Examination of the univariate ANOVAs indicated that three treatment groups' performance did not differ at pretest, however, the average group performances at the posttest significantly differed, [$F(2, 95) = 21.33$, $p < .0001$, $\eta^2 = .31$]. The post hoc comparisons revealed that the two training groups' posttest performance (Individual $M = 62.06$; Collaborative $M = 62.26$) was significantly greater than the Questionnaire Only group ($M = 49.24$), but did not differ significantly from each other.

Accuracy

There were significant main effects of Condition, [$F(2, 95) = 4.43$, $p < .05$, $\eta^2 = .04$], and Occasion, [$F(1, 95) = 43.38$, $p < .0001$, $\eta^2 = .32$]. Consistent with the previous analyses examining the number of correct items, these main effects were qualified by a significant Occasion x Condition interaction, [$F(2, 95) = 13.59$, $p < .0001$, $\eta^2 = .07$]. Examination of the univariate ANOVAs indicated that the three groups did not differ at pretest, however, there was a significant group difference in posttest accuracy, $F(2, 96) = 16.80$, $p < .0001$, $\eta^2 = .26$]. The accuracy of both the Individual (.81) and Collaborative (.81) training groups' responses at the posttest was significantly greater than that of the Questionnaire Only (.60) group.

Word Series Test

Total Number of Correct Items

Analyses of group performance on the Word Series Test revealed a significant main effect of Occasion, [$F(1, 95) = 45.33$, $p < .0001$, $\eta^2 = .32$]. The main effect was qualified by a significant Occasion x Condition interaction, [$F(2, 95) = 3.66$, $p < .05$, $\eta^2 = .07$]. Examination of the univariate ANOVAs revealed that the three groups did not differ at pretest, however, there was a significant difference in the posttest performance of the treatment groups, [$F(2, 95) = 3.73$, $p < .05$, $\eta^2 = .07$]. The performance of the Questionnaire Only group ($M = 51.22$) was significantly less than the performances of the Individual ($M = 57.47$) and Collaborative ($M = 56.94$) training groups at the posttest.

Accuracy

A significant main effect of Occasion, [$F(1, 95) = 11.12$, $p < .01$, $\eta^2 = .11$], and a significant Occasion x Condition interaction, [$F(2, 95) = 3.22$, $p < .05$, $\eta^2 = .06$], were observed. Examination of the univariate ANOVAs revealed a significant difference in the treatment groups' posttest performance, [$F(2, 95) = 3.66$, $p < .05$, $\eta^2 = .07$]. The post hoc comparisons revealed that the accuracy of both the Individual (.90) and Collaborative (.88) training participants' responses at the posttest was significantly greater than that of the Questionnaire Only (.81) participants.

Letter Sets Test

Next, analysis of group performance on the Letter Sets Test was conducted. As dyads completed this measure jointly and one performance score was produced, the number of observations equals the number of dyads with complete data (i.e., $N = 49$).

Total Number of Correct Items

The repeated measures ANOVA revealed a significant main effect of Occasion, [$F(1, 46) = 21.32$, $p < .0001$, $\eta^2 = .32$]. The main effect of Condition was indicated a trend, but was not significant, [$F(2, 46) = 2.47$, $p < .10$, $\eta^2 = .10$] and the Occasion x Condition interaction was not significant, [$F(2, 46) = 2.06$, $p = .14$, $\eta^2 = .08$]. Examination of the post hoc pairwise comparisons revealed that for the *total sample*, the posttest test scores ($M = 55.16$) were significantly greater than the pretest scores ($M = 50.07$; $p < .001$).

Accuracy

The repeated measures ANOVA revealed a significant main effect of Occasion, [$F(1, 46) = 12.59$, $p < .001$, $\eta^2 = .22$]. The main effect of Condition indicated a trend, but was not significant, [$F(2, 46) = 2.94$, $p < .10$, $\eta^2 = .11$]. The Occasion x Condition interaction was not significant, [$F(2, 46) = 1.30$, $p = .28$, $\eta^2 = .05$]. Examination of the post hoc comparisons revealed that for the *total sample*, the posttest test accuracy scores ($M = .69$) were significantly less than the pretest accuracy scores ($M = .77$; $p = .001$).

Effects of Gender, Age, and Education as Covariates on Group Performance

In order to determine if relevant demographic variables impacted training gain, a series of repeated measures ANCOVAs analogous to the analyses reported above was conducted including age status (Young-Old: 61–71 years, Older: 72–89 years), education status (Less-Educated: 12–16 years, More-Educated: 17 years or greater), and gender (Male, Female) as covariates. Dichotomous age and educational categories were based on the respective median cutoffs. For the Letter Sets Test, dyadic average age and education were used to determine the respective status categories and sex was not included in analyses as the dyad produced one score. The same pattern of results was found in analyses with and without covariates.

Summary of Group-Level Results

In general, the findings of the group-level analysis demonstrate the beneficial effects of training. Both the individual and collaborative training groups significantly outperformed the control group on the Letter and Word Series measures at the posttest in terms of number of correctly answered items and the accuracy of response. The two training groups however did not significantly differ in their performances on these measures at the posttest assessment, indicating that there was not differential benefit for the collaborative training group at the group level.

For the Letter Sets test, which was completed collaboratively, the results indicated an effect of assessment occasion. For the total sample, the number of correctly answered items was significantly greater at the posttest. In contrast, the accuracy of response for the total sample was significantly less at the posttest. As previously mentioned, initial accuracy scores tended to be quite high and may indicate either true collaboration between partners or a problem in terms of possible ceiling effects.

Training Effects: Reliable Change at the Individual level

In addition to mean-level group differences, change in reasoning performance was also examined at the individual level. The proportion of participants showing significant change in reasoning performance was investigated. Raw scores were used in these analyses so that the degree of change (i.e., single item differences) relative to each of the reasoning measures could be assessed. Significant, ‘reliable’ change in performance from pretest to posttest was defined in terms of one standard error of measurement (SEM; Dudek, 1979; Schaie, & Willis, 1986) with participants classified as having “Improved” if their posttest performance was equal to or greater than one SEM above their pretest performance. Participants were classified as “Stable” if their posttest performance was less than one SEM above their baseline pretest performance. One standard error of measurement was chosen as it provides a stringent confidence interval around

each participant's pretest scores from which to assess change in performance. Using this method allowed us to assess 'reliable' change which exceeded (and was not likely due to) a random fluctuation in score at retest. The average standard error of measurement for the Letter Series, Word Series, and Letter Sets Tests were 3.04, 3.21, and 2.19, respectively. The proportion of individuals classified as "Stable" or "Improved" on the inductive reasoning measures is depicted in Table 3. Chi-square analyses were used to examine the independent effects of treatment group, gender, educational status, and age status on change status classification. As the Letter Sets Test was completed by dyads, sex effects were not examined for this measure.

Table 3

Average Reasoning Performance Change over the Pretest-Posttest Interval and Percentage of Participants Demonstrating Reliable Change

Letter Series Test

Total Number of Correct Items

For the entire sample 37.76% of participants demonstrated reliable improvement on this measure. Chi-square analyses revealed a significant effect of treatment group on change status classification. There was a significant difference in proportion of participants showing reliable gain by treatment group [^{2} (2, N = 98) = 23.03, p < .0001]. The majority of participants in the Questionnaire Only group (94.12%) remained stable while a large proportion of persons in the Individual (50.00%) and Collaborative (58.82%) training groups exhibited reliable improvement.

Accuracy

The majority of the sample did not improve in the accuracy of performance over the pre-posttest interval (Stable: 67.35%). However, there was a significant difference in proportion of participants showing reliable gain by treatment group [^{2} (2, N = 98) = 11.62, p < .01]. Within the Questionnaire Only group 11.76% of participants improved. In comparison, 36.67% of persons in the Individual group and 50.00% of persons in the Collaborative group exhibited reliable improvement. No significant differences in change status were found based on sex or education. There was a trend suggesting that a greater proportion of old-old adults demonstrated reliable improvement [^{2} (1, N = 98) = 3.48, p < .10].

Word Series Test

Total Number of Correct Items

For the Word Series Test, 40.82% of the participants in the total sample were classified as improved. Chi-square analyses did not reveal any group differences in change status classification, however, there was a noticeable, but nonsignificant difference between treatment groups. In the Questionnaire Only group, 29.41% of individuals demonstrated reliable improvement, compared to 46.67% in the Individual and 47.06% in the Collaborative training groups.

Accuracy

For the Word Series Test, most participants (73.47%) remained stable in their accuracy of performance over the pre-posttest interval. Chi-square analyses did not reveal any group differences in change classification. The difference in classification for each treatment group was not significant, however, the two training groups demonstrated a greater proportion of persons who improved: Questionnaire Only = 17.65%; Individual Training = 26.67%; Collaborative Training = 35.29%).

Letter Sets Test

Total Number of Correct Items

The majority of persons in the total sample (65.31%) remained stable on this measure. Within the Questionnaire Only group, 29.41% of participants were classified as improving. In comparison, 33.33% of the Individual Training participants and 41.18% of the Collaborative Training participants showed reliable improvement. Chi-square analyses did not reveal any differences in change status classification based on treatment group, education status, and age status.

Accuracy

There was little change in dyad's accuracy on the Letter Sets Test (i.e., 95.92% were stable). Across the three treatment groups, 5.88% of dyads demonstrated improvement in the Collaborative Training and Questionnaire Only groups and none of the Individual Training dyads demonstrated improvement on this measure. Chi-square analyses did not reveal any differences in change status classification based on treatment group, education status, and age status.

Summary of Individual-Level Results

Using a rather rigorous criterion to assess individual-level change (i.e., standard error of measurement), the results from the Letter Series Test confirm the group-level results. A greater proportion of participants in the two training groups evinced significant improvement in the number of correctly answered items and accuracy on this measure of inductive reasoning. The findings for the Word Series Test showed the same pattern of findings, although the results were not significant. Results for the Letter Sets Test, which couples completed collaboratively, suggested some benefit of collaborative training in terms of the number of correctly answered items, however, these results were not significant.

Discussion

This study represents an initial effort to examine the intersection of formal cognitive training and collaborative cognition in older adulthood. Using a pretest-posttest/no-treatment control group design, the utility of adapting a proven formal inductive reasoning training protocol to a more ecologically valid and more economic form (i.e., in-home use by married couples) was examined. Additionally, the current study investigated the effects of individual versus collaborative learning within a training context.

Training gain at the group level was assessed by two methods (i.e., number of correct items, accuracy of responses). Performance on three inductive reasoning measures (i.e., Letter Series, Word Series, Letter Sets tests) was assessed. For the Letter and Word Series tests, analysis of variance tests revealed group differences in posttest performance for the three groups: The two training groups demonstrated significantly greater pretest-posttest gain on the measures of inductive reasoning compared to the control group. These effects were maintained even when demographics (i.e., age, education, sex) were included in the analyses as covariates. The same pattern of results held for both performance measures (i.e., number of correct items, accuracy of response).

Results of analyses examining group performance differences on the Letter Sets test revealed a non-significant trend for the effect of training group in terms of number of correct items and accuracy of response. It should be noted that the initial accuracy on the Letter Sets and Word Series tests was substantially higher compared to initial accuracy on the Letter Series test and accuracy on Letter Sets actually tended to decline over the pre-posttest interval. The decline in accuracy on the collaborative measure may have reflected that there was indeed greater collaboration at posttest than at pretest. Collaboration would take time and thus could reduce the number of items attempted by the couple. Since accuracy is computed as the number of correct

items divided by the number of items attempted, a reduction in the total number of items attempted due to collaboration might reduce the accuracy score.

Using a more stringent criterion to assess the degree of individual change, reliable improvement was defined as performance change greater than or equal to one standard error of measurement (SEM). Using this criterion of change, the most dramatic group differences were observed on the Letter Series Test for both number correct and accuracy. A significantly greater proportion of participants in both the individual and collaborative training groups demonstrated reliable improvement in performance compared to the Questionnaire Only group. There was a similar trend for Word Series Test performance; however, these findings were not statistically significant. Using the SEM criterion for change, findings for the Letter Sets test were not significant. Examples and items in the training sessions were most similar to Letter Series Test items, thus it is not surprising that the most substantial improvement was observed on performance on this test. In addition, respondents' initial accuracy was much higher on the Word Series and Letter Sets tests compared to the Letter Series test thus leaving less room for improvement in accuracy on this measure.

A related question warranting additional research is the transfer of training to everyday cognitive tasks. Some prior work (e.g., Allaire & Marsiske, 1999; Willis, Jay, Diehl & Marsiske, 1992) has demonstrated a relationship between reasoning ability and cognitively demanding everyday tasks, however, recent work linking cognitive intervention to everyday cognitive tasks has been equivocal (Ball, et al., 2002). The issue of training transfer is of the utmost importance as the ultimate goal of most cognitive training protocols is to affect positive change in the maintenance/enhancement of cognitive abilities which are believed to impact elders' functioning in daily life. Transfer of training to a measure of comprehension of everyday printed materials (e.g., medication labels; Willis & Marsiske, 1997) is being explored in the present study.

Comparison of the Two Training Groups

Contrary to expectations, results from the current study examining immediate post-intervention training gains did not suggest that collaborative training with a spouse yielded greater training gains immediately after training compared to individual training. Collaborative trainers did not benefit significantly more in terms of the number of correct items and degree of accuracy from pretest to posttest compared to their counterparts in the individual training condition; however, it is possible that these two groups differed in other ways. For instance, it is unknown whether collaborative and individual learning in later life may differ in the type and timing of cognitive benefit.

First, it is possible that the context of learning and learning application may be most effective when congruent (i.e., collaborative learning and collaborative cognitive performance). Saczynski, Margrett, and Willis (2004) examined objective strategy use (i.e., strategy markings on the test page) within this sample at immediate and delayed posttests. While strategy use did not differ for individual and collaborative groups at immediate posttest, strategy use was better maintained at delayed posttest for the collaborative group. This lends support to the possibility that the benefits of collaborating with a familiar partner may show delayed effects, particularly when completing measures in a similar context. Perhaps an advantage of collaboration is more evident in the long-term maintenance of abilities—a point in time which was not captured in the current analyses or assessment occasions. For example, spouses who completed the reasoning training together may practice and reinforce training principles and strategies long after training has been completed, thus collaboratively trained persons may show longer term benefits. Both individual and collaborative groups reached a high level of performance immediately after

training. The effects of training with a partner may be most beneficial after some dissipation of the initial training gain.

In terms of learning pace, it is also possible that differences between the two training groups may be evident in strategy acquisition and reasoning performance during the training protocol. Perhaps persons in the collaborative training condition may have exhibited training benefit earlier in the training session sequence. In contrast, persons in the individual training condition may have reached their peak level of performance later in the session sequence. In other words, the benefit of collaborating with one's spouse may be most advantageous to initial learning and practice attempts and less beneficial to sustained practice efforts once individual asymptotes had been reached. Both the possibility of differential initial learning and more long-term maintenance warrant further investigation and future analyses within this data set will examine the training trajectories of each treatment condition across the ten training sessions and over long assessment periods.

Additionally, individual and collaboratively trained participants may differ in other ways such as their subjective experiences of training as well as the degree of training transfer to a measure of everyday functioning. It will be possible to also test these hypotheses in the current study. For example, as previously mentioned couples trained collaboratively may be more likely to utilize training principles and strategies in their everyday life. This additional practice and extension of the training may be evident in measures assessing transfer of training. It is possible that while cognitive collaboration may be beneficial in general, perhaps collaborating on this very structured reasoning training program was not conducive to reaping collaborative benefits in this sample of relatively high-functioning community-dwelling older adults.

An additional caveat to consider when interpreting the current findings is the effect of married partners' relationship quality on the learning context and subsequent training gains. Further research is needed to investigate relationship quality as a potential mediating factor affecting the relation between training condition and cognitive outcome. Prior collaborative research suggests complex relationships between cognitive outcomes and both the nature of the partners' relationship (e.g., partner familiarity: Dixon, et al., 1997, Gould, et al., 1991, Gould, et al., 1994, Margrett & Marsiske, 2004; partner gender: Margrett & Marsiske, 2002) as well as the collaborative interaction (e.g., Gould, et al., 1991, Gould, et al., 1994; Kimbler, 2004). As noted by Carstensen, Levenson, and Gottman (1995) and Dixon and Gould (1996), the "expertise" developed in long-term relationships may facilitate the process and product of cognitive collaboration.

A practical limitation of the current study was the fact that the individual training condition was comprised of individuals within the same spousal dyads. Precautions were taken to reduce the possibility that spouses in the individual condition completed training sessions together; however, it is possible that some couples may have discussed the training program. This study also intentionally targeted well, community-dwelling older adults. Thus, the current findings may not be generalizable to at-risk populations of older adults. Additional studies examining the effectiveness of in-home, "self-paced" cognitive interventions are needed. With more impaired populations, it is possible that a higher degree of support (e.g., telephone contact) by the interventionist may be needed, or alternatively, collaborative support by a more-able partner may be required. Despite the prevalence of well elders in the current study the results are promising in that relatively highly educated persons still benefited from the training program.

Conclusions

The current study demonstrated success in adapting a ten-session inductive reasoning protocol used previously with a formal trainer for in-home use by older adults without a formal trainer. This study confirmed that older adults could successfully "train themselves" on an in-home cognitive training program. Many older adults in the two training conditions were able to substantially enhance their inductive reasoning performance. These findings are promising in that greater adaptation of formal cognitive training intervention protocols may be possible. Increased use of more self-directed training could reach a greater number of older adults. Findings from the current study did not demonstrate differential benefit for persons training alone compared to persons training with their spouses. However, it is possible that the benefits of cognitive collaboration observed in other gerontological studies may be evident at different points of the training protocol or in other aspects of cognitive training (e.g., subjective experience, transfer of training effects). Perhaps, the benefits of cognitive collaboration may be borne out over time and may be evident after an extended period. These possibilities provide rich avenues for future research.

Acknowledgments

The research reported in this study was supported by a post-doctoral research award from Division 20 of the American Psychological Association and the Retirement Research Foundation awarded to Jennifer Margrett. The writing of an earlier draft of this article was supported by a postdoctoral research fellowship (#T32-MH-18904) from the National Institute of Mental Health awarded to The Pennsylvania State University. The training program used in the current study (Willis & Schaie, 1986, 1994) was mostly recently updated in the context of the ACTIVE clinical trial (Ball, et al., 2002; Jobe, et al., 2001). We would like to thank Joan Irwin, Dawna Kasper, Mimi Lutz, Mary Markowski, Mike Nealon, Amy Roth, and Danielle Schmidheiser for their technical assistance with this project. We would also like to gratefully acknowledge the couples that participated in this study. Portions of this manuscript were presented at the 54th annual meeting of the Gerontological Society of America, Chicago, IL.

Contributor Information

Jennifer A. Margrett, West Virginia University, Department of Psychology, 53 Campus Drive, P.O. Box 6040, Morgantown, West Virginia 26506;

Sherry L. Willis, The Pennsylvania State University, Gerontology Center, 0405 Marion Place, University Park, Pennsylvania 16802-6504;

References

1. Allaire JC, Marsiske M. Everyday cognition: Age and intellectual ability correlates. Psychology & Aging. 1999;14(4):627–644. [PMC free article] [PubMed]
2. Andersson J, Rönnberg J. Recall suffers from collaboration: Joint recall effects of friendship and task complexity. Applied Cognitive Psychology. 1995;9:199–211.
3. Andersson J, Rönnberg J. Collaboration and memory: Effects of dyadic retrieval on different memory tasks. Applied Cognitive Psychology. 1996;10:171–181.
4. Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL. Journal of the American Medical Association. 2002;288(18):2271–2281. [PMC free article] [PubMed]
5. Berg, C. A., Meegan, S. P., & Johnson, M. M. S. (November, 1995). *Collaborative interaction styles in young and old married couples* Paper presented at the annual meeting of the Gerontological Society of America: Los Angeles, CA.

6. Blackburn JA, Papalia-Finlay D, Foye BF, Serlin RC. Modifiability of figural relations performance among elderly adults. Journal of Gerontology: Psychological Sciences. 1988;43(3):P87–P89.
7. Blieszner R, Willis SL, Baltes PB. Training research in aging on the fluid ability of inductive reasoning. Journal of Applied Developmental Psychology. 1981;2:247–265.
8. Carstensen LL, Levenson RW, Gottman JM. Emotional behavior in long-term marriage. Psychology and Aging. 1995;10(1):140–149. [PubMed]
9. Craik, F. I. M., & Jacoby, L.L. (1996). Aging and memory: Implications for skilled performance (pp. 113–137). In W. Rogers & A. Fisk, (Eds.), *Aging and skilled performance: Advances in theory and application*s. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
10. Dixon RA, Fox DP, Trevithick L, Brundin R. Exploring collaborative problem solving in adulthood. Journal of Adult Development. 1997;4(4):195–208.
11. Dixon, R. A., & Gould, O. N. (1996). Adults telling and retelling stories collaboratively. In P.B. Baltes & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspective on the social foundation of cognition* New York: University Press.
12. Dixon RA, Gould ON. Younger and older adults collaborating on retelling everyday stories. Applied Developmental Science. 1998;2(3):160–171.
13. Dudek FJ. The continuing misinterpretation of the stand error of measurement. Psychological Bulletin. 1979;86(2):335–337.
14. Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests* (Rev. Ed.). Princeton, NJ: Educational Testing Service.
15. Gould ON, Dixon RA. How we spent our vacation: Collaborative storytelling by young and old adults. Psychology and Aging. 1993;8(1):10–17. [PubMed]
16. Gould O, Kurzman D, Dixon RA. Communication during prose recall conversations by young and old dyads. Discourse Processes. 1994;17:149–165.
17. Gould ON, Osborn C, Krein H, Mortenson M. Collaborative recall in married and unacquainted dyads. International Journal of Behavioral Development. 2002;26(1):36–44.
18. Gould ON, Trevithick L, Dixon RA. Adult age differences in elaborations produced during prose recall. Psychology & Aging. 1991;6(1):93–99. [PubMed]
19. Gratzinger P, Sheikh JI, Friedman L, Yesavage JA. Cognitive interventions to improve face-name recall: The role of personality trait differences. Developmental Psychology. 1990;26(6):889–893.
20. Hofland BF, Willis SL, Baltes PB. Fluid intelligence performance in the elderly: Intraindividual variability and conditions of assessment. Journal of Educational Psychology. 1981;73:573–586.
21. Jobe JB, Smith DM, Ball K, Tennstedt SL, Marsiske M, Rebok GW, Morris JN, Willis SL, Helmers K, Leveck MD, Kleinman K. ACTIVE: A cognitive intervention trial to promote independence in older adults. Controlled Clinical Trials. 2001;22(4):453–79. [PMC free article] [PubMed]
22. Kausler, D.H. (1994). *Learning and memory in normal aging* San Diego, CA: Academic Press.
23. Kimbler, K. (2004). *Effective verbal behaviors during cognitive collaboration by older adult partners* Unpublished master's thesis, West Virginia University, Morgantown, WV.

24. Kliegl R, Smith J, Baltes PB. On the locus and process of magnification of age differences during mnemonic training. Developmental Psychology. 1990;26(6):894–904.
25. Kotovsky K, Simon HA. Empirical tests of a theory of human acquisition of concepts for sequential patterns. Cognitive Psychology. 1973;4(3):399–424.
26. Margrett JA, Marsiske M. Gender differences in older adults' everyday cognitive collaboration. Collaboration in later life [Special Issue] International Journal of Behavior Development. 2002;26(1):45–59.
27. Margrett, J. A. & Marsiske, M. (2004). *Collaboration on everyday problem solving tasks by older adults* Manuscript submitted for publication.
28. Meegan S, Berg CA. Contexts, functions, forms, and processes of collaborative everyday problem solving in older adulthood. International Journal of Behavioral Development. 2002;26(1):6–15.
29. Meyer, B. J. F., Young, C. J., & Bartlett, B. J. (1989). *Memory improved: Reading and memory enhancement across the life span through strategic text structures* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
30. Saczynski, J. S., Margrett, J. A., & Willis. S. L. (in press). *Older adults' strategic behavior: Effects of individual versus collaborative cognitive training*.
31. Schaie, K. W. (1985). *Manual for the Schaie-Thurstone Adult Mental Abilities Test (STAMAT)* Palo Alto, CA: Consulting Psychologists Press.
32. Staudinger UM, Baltes PB. Interactive minds: A facilitative setting for wisdom-related performance? Journal of Personality and Social Psychology. 1996;71(4):746–762.
33. Strough J, Margrett JA. Introduction. Collaboration in later life. [Special Issue] International Journal of Behavior Development. 2002;26(1):2–5.
34. West RL, Crook TH. Video training of imagery for mature adults. Applied Cognitive Psychology. 1992;6(4):307–320.
35. Willis SL, Jay GM, Diehl M, Marsiske M. Longitudinal change and prediction of everyday task competence in the elderly. Research on Aging. 1992;14(1):68–91. [PMC free article] [PubMed]
36. Willis, S. L., & Marsiske, M. (1997). *Manual for the Everyday Problems Test* University Park, PA: Department of Human Development and Family Studies, The Pennsylvania State University.
37. Willis SL, Schaie KW. Can decline in adult intellectual functioning be reversed? Developmental Psychology. 1986;22(2):223–232.
38. Willis, S. L., & Schaie, K. W. (1994). Cognitive training in the normal elderly. In F. Forette, Y. Christen, & F. Boller (Eds.), *Plasticité Cérébrale et Stimulation Cognitive* Paris: Foundation Nationale de Gérontologie.
39. Woolverton M, Scogin F, Shackelford J, Black S, Duke L. Problem-targeted memory training for older adults. Aging, Neuropsychology, & Cognition. 2001;8(4):241–255.