Long-term Maintenance of Retest Learning in Young Old and Oldest Old Adults

Lixia Yang1,2 and Ralf T. Krampe2,3

1Department of Psychology, Ryerson University, Toronto, Ontario, Canada.
2Department of Psychology, University of Leuven, Belgium.
3Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.

This study examined the maintenance of retest learning benefits in young old and oldest old adults over an 8-month period in 3 cognitive abilities: reasoning, perceptual-motor speed, and visual attention. Twenty-four young old (aged 70–79 years, M = 74.2) and 23 oldest old adults (aged 80–90 years, M = 83.6) who participated in a previously published study (Yang, L., Krampe, R. T., & Baltes, P. B. [2006]. Basic forms of cognitive plasticity extended into the oldest-old: Retest learning, age, and cognitive functioning. Psychology and Aging, 21, 372–378) returned after an 8-month delay to complete 2 follow-up retest sessions. The results demonstrated that both young old and oldest old groups maintained about 50% of the original retest learning benefits. This extends the earlier findings of substantial long-term cognitive training maintenance in young old adults to a context of retest learning with oldest old adults, and thus portrays a positive message for cognitive plasticity of the oldest old.

Key Words: Retest learning—Long-term maintenance—Oldest old—Young old—Reasoning—Perceptual-motor speed—Visual attention.

SUBSTANTIAL research evidence suggests that older adults’ performance on a variety of cognitive tasks can be improved through cognitive training (e.g., Ball et al., 2002; Salthouse, 2006; Schaie & Willis, 1986; Thompson & Foth, 2005), and the improvement can be maintained for a long period of time, ranging from 5 months (Günther, Schafer, Holzner, & Kemmler, 2003) to 2–5 years (Willis et al., 2006) and up to 7 years (Willis & Nesselroade, 1990). Previous studies on the long-term maintenance effect used primarily a tutor-guided training program, where a tutor explains rules and teaches effective strategies, and focused mainly on young old adults in their 60s or 70s. Inspiringly, one recent study showed that even the oldest old in their 80s were able to improve their cognitive performance through a self-guided retest paradigm in which they were repeatedly tested without receiving any guidance or feedback (Yang, Krampe, & Baltes, 2006). However, it is unclear whether the oldest old will be able to maintain this retest learning for an extended period of time. The current study addressed this issue by investigating the long-term maintenance of retest learning in oldest old adults as compared with young old adults.

The literature suggests different predictions concerning this question. Considering the accelerated cognitive declines (Baltes & Mayer, 1999) and the severely limited potential to learn new memory skills among very old adults (Singer, Lindenberger, & Baltes, 2003), the oldest old may have a limited capacity to maintain training or retest benefits. However, the substantial retest learning among the oldest old promisingly suggests that they may still reserve basic forms of cognitive plasticity by reactivating and practicing available skills (Yang et al., 2006), and thus the oldest old may still be able to maintain retest learning for a longer period of time.

Previous work revealed that when the two training phases were closely spaced temporally (e.g., 1 year), training effects (for the training group) or retest effects (for the noncontact control group) at the later phase appeared to be cumulative to the learning achieved at the initial phase (Willis & Nesselroade, 1990). Therefore, we expect a larger retest learning effect at our 8-month follow-up phase than at our original retest phase. Additionally, this effect may be disproportionally more pronounced for the young old considering the evidenced age-related decline in retest learning effects (Yang et al., 2006).

Inspired by these questions, this follow-up study of Yang and colleagues (2006) aims to examine the long-term maintenance of retest learning in oldest old adults compared with young old adults. To be consistent with our previous study (Yang et al.), we pragmatically consider participants aged 80 years and onward as the oldest old and those aged 70–79 as the young old. Specifically, we will address three questions: (a) Can the oldest old maintain retest learning over an 8-month interval? (b) Do the two age groups differ in the magnitude of maintenance? and (c) Do they differ in cumulative retest learning effects?

METHOD

Participants

Forty-seven of the 68 participants in the study by Yang and colleagues (2006) returned for this follow-up study, consisting of 24 young old (aged 70–79 years, M = 74.2, SD = 2.7) and 23 oldest old adults (aged 80–90 years, M = 83.6, SD = 3.2). All participants reported to be of good health,
without serious vision or hearing problems. The two age groups did not differ in education, vocabulary, scores on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), and cognitive functioning status in their own age group as indexed by the percentile ranks (see Yang et al.) with respect to the norms in the Berlin Aging Study (Baltes & Mayer, 1999), $p > .39$. There were no attrition effects because the returned ($n = 47$) and nonreturned participants ($n = 21$) did not differ in their performance on any pretest cognitive measures or demographic variables ($p > .14$), either collapsed across age groups or separately within each age group.

**Materials**
At each of the two follow-up retest sessions, participants completed the same retest measures as those used in the original study (Yang et al., 2006), consisting of four inductive reasoning tests: (a) Cultural Fair Test (Form A, part 1; Weiß, 1987); (b) Raven’s Progressive Matrices (Raven, 1962); (c) Letter Series (Blieszner, Willis, & Baltes, 1981) and Number Series (Blieszner et al.); and (d) a perceptual-motor speed test, the Digital Symbol Substitution Test (Wechsler, 1981), as well as a visual attention test, the D2 Test (Brickenkamp, 1994).

**Procedure**
The original study involved one pretest session, six retest sessions (Sessions 1 to 6), and one posttest session. In this follow-up study, two 1-hr retest sessions (Follow-up Sessions 1 and 2) were conducted after an 8-month delay. These two follow-up sessions were spread over a 1-week period, with a minimum 2-day interval in between. The six retest measures were given to groups of 2–5 participants at each follow-up session. The order of the tasks was counterbalanced across sessions and participants. The result patterns reported below remain the same when the task counterbalance order was included as an independent variable. Task order does not interact with any other factor; thus, we omit it from the final analyses to simplify our main results.

**Results**
With a central focus on long-term maintenance effects, we reported two main sets of results: (a) long-term maintenance effects, by comparing the original retest learning at Session 6 with maintenance at the two follow-up sessions, over baseline performance at Session 1 and (b) cumulative learning effects, by comparing the learning effects at the original (Yang et al., 2006) with the follow-up phases. To obtain comparable scales, all the raw scores (i.e., number of correct solutions) on each retest measure were transformed into T scores, standardized to Session 1 ($M = 50; SD = 10$). A composite reasoning score was then calculated from the four reasoning measures at each session.

**Long-term Maintenance of the Retest Learning Benefits**
The retest learning effect size was estimated by subtracting the mean T scores at baseline Session 1 from that at Session 6, at Follow-up Session 1, and at Follow-up Session 2, respectively, and then dividing the resulting retest gain scores by the baseline standard deviation at Session 1, for each of the three ability domains (Salthouse & Tucker-Drob, 2008; see Figure 1). The analyses on effect size scores allow for direct comparison of different ability domains (Willis et al., 2006). The 2 (age) $\times$ 3 (ability domain) $\times$ 3 (session) analysis of variance (ANOVA) on the effect size scores revealed significant main effects of all three
variables: age, $F(1, 45) = 9.76, MSE = 0.92, p < .01, p^2 = 0.18$; ability domain, $F(2, 90) = 3.29, MSE = 0.69, p < .05, p^2 = 0.07$; and session, $F(2, 90) = 99.60, MSE = 0.10, p < .001, g^2 = 0.69$. The young old ($M = 0.94, SD = 0.33$) benefitted and maintained more than the oldest old ($M = 0.65, SD = 0.31$); the retest gain was larger in perceptual-motor speed ($M = 0.94, SD = 0.55$) than in reasoning ($M = 0.71, SD = 0.49$), $p < .05$, and marginally larger than in visual attention ($M = 0.74, SD = 0.52$), $p = .05$, whereas the latter two did not differ ($p = .81$). Finally, the original retest gain at Session 6 ($M = 1.07, SD = 0.39$) was larger than the maintenance at Follow-up Session 1 ($M = 0.53, SD = 0.33$) and Follow-up Session 2 ($M = 0.80, SD = 0.42$), $ps < .001$, and the maintenance at Follow-up Session 2 was larger than that at Follow-up Session 1, $p < .001$. Overall, the maintenance at Follow-up Session 1 (mean effect size = 0.53 baseline SD units) and at Follow-up Session 2 (mean effect size = 0.80 baseline SD units) is about 50% and 75%, respectively, of the original retest learning (mean effect size = 1.07 baseline SD units). No significant interaction was revealed, $Fs < 1.64, ps > .20$.

Within each ability domain, the retest learning gain at Session 6 was significantly correlated with maintenance at the two follow-up sessions, $rs = .52–.63, ps < .001$. This pattern was confirmed with the regression analyses, suggesting that the performance at the follow-up sessions reflected largely the maintenance of the original retest learning.

Despite the high intercorrelations of T scores among the three ability domains at baseline ($rs > .66$) and across sessions ($rs = .66–.79$), the correlations of the retest effect sizes among domains are largely nonsignificant ($rs < .31, median r = .16$) at Session 6 and the two follow-up sessions.

**Cumulative Learning Effects**

Although the two age groups maintained an equal amount of retest learning, the young old may have the advantage of showing pronounced cumulative learning effect. To test this hypothesis, we conducted a 2 (age) × 3 (ability domain) × 2 (phase: original vs. follow-up) mixed-model ANOVA on the learning effects indexed by the gain T scores across two sessions (i.e., Session 2–Session 1 for the original phase and Follow-up Session 2–Follow-up Session 1 for the follow-up phase). Overall, the learning effect was larger for young old ($M = 3.45, SD = 1.66$) than for oldest old adults ($M = 2.32, SD = 1.39$), $F(1, 45) = 6.37, MSE = 14.09, p < .05, g^2 = 0.09$. The main effect of ability domain, $F(2, 90) = 4.21, MSE = 22.10, p < .05, g^2 = 0.09$, indicated that the learning effect did not differ between perceptual-motor speed ($M = 3.76, SD = 3.52$) and visual attention ($M = 3.13, SD = 3.32$), but both domains demonstrated larger learning effects than did reasoning ($M = 1.81, SD = 2.47$), $ps < .05$. The ability domain by phase interaction was significant, $F(2, 90) = 3.60, MSE = 17.26, p < .05, g^2 = 0.07$. The learning effect in reasoning and visual attention did not change across phases, but the learning in perceptual-motor speed dropped from the original ($M = 4.86, SD = 5.49$) to the follow-up phase ($M = 2.65, SD = 4.82$), $p = .05$. In summary, the results showed little support for cumulative retest effects in both age groups.

**DISCUSSION**

Together with our previous work (Yang et al., 2006), this follow-up study suggested that retest learning effects are substantial and can endure for at least 8 months, even among the oldest old. Despite the overall advantage of the young old in retest benefits, the two age groups maintained an approximately equal amount of original retest learning effects (i.e., about 50%). This is consistent with the estimate that retest effects can be detected for up to 7 or more years (Salthouse, Schroeder, & Ferrer, 2004) and extends the pattern to an oldest old population. Strikingly, the maintenance in reasoning ($0.45 SD$) is comparable to the immediate training gain ($0.48 SD$) in a tutor-guided training paradigm across 10 sessions, or the amount of expected age-related decline over 14 years ($0.42 SD$) in older adults without dementia (Ball et al., 2002). The maintenance in speed ($0.66 SD$) is more than four times that of the expected age-related decline over 2 years ($0.16 SD$) in the elderly adults without dementia (Ball et al.).

The finding of weak correlations of retest effect sizes among ability domains extends recent findings of low correlations of retest effect sizes even within a specific ability domain (Salthouse & Tucker-Drob, 2008) and the well-established result of highly ability-specific benefits in cognitive training literature (e.g., Ball et al., 2002). The conservative interpretation is that the retest effect is largely ability- and/or task-specific. However, given the problems in reliabilities for difference scores that we used to calculate effect size, and the small sample size, our data may not evaluate specific versus general gains or maintenance. Nevertheless, the high correlations of T scores among domains suggest that the different domains tax largely similar cognitive abilities. In any event, the finding of substantial maintenance effect in the oldest old is striking.

Little cumulative retest effect is detected, presumably because the number of follow-up sessions (i.e., two) may not be sufficient to show cumulative effect. In contrast, the learning effect in speed tends to drop over time. We speculate that retest learning in the speed measure is largely driven by item-general familiarity or procedural effects that are supposed to be more important at earlier sessions (Yang, Reed, Russo, & Wilkinson, 2009).

One limitation of this study is that the sample includes individuals from the upper two thirds of the respective age cohorts and thus constrains the generalizability of the results (Yang et al., 2006). Nevertheless, this study leads to a novel finding that even the oldest old are able to maintain a significant portion of retest learning for up to 8 months.
suggested that the basic form of plasticity is reserved, robust, and durable, even in the oldest old.

FUNDING
This study is funded primarily by the Max Planck Institute for Human Development in Berlin. It also receives partial support through the Ryerson New Faculty SRC Development Fund awarded to Lixia Yang and BOF grant OT25/05 from the Research Council of KU Leuven and FWO-grant G.0379.06 awarded to Ralf T. Krampe.

ACKNOWLEDGMENTS
We thank Anita Günther for her assistance with data collection. Many ideas and initiatives in shaping this project came from Prof. Paul B. Baltes. We acknowledge and greatly appreciate his invaluable input.

CORRESPONDENCE
Address correspondence to Lixia Yang, PhD, Department of Psychology, Ryerson University, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada. Email: lixiy@ryerson.ca

REFERENCES


Weiß, R. H. (1987). Grundintelligenztest Skala 2 CFT 20 [Basic Intelligence Test, Scale 2, CFT 20], Champaign, IL: The Institute for Personality & Ability Testing.


Received November 22, 2008
Accepted July 12, 2009
Decision Editor: Rosemary Blieszner, PhD