

## On the Locus and Process of Magnification of Age Differences During Mnemonic Training

Reinhold Kliegl, Jacqui Smith, and Paul B. Baltes  
Max Planck Institute for Human Development and Education  
Berlin, Federal Republic of Germany

The focus of this study was on developmental reserve capacity in old age as revealed by testing-the-limits. We examined (a) the time course of training-related magnification of age differences in serial word recall and (b) predictability of training gains by pretest individual differences in cognitive abilities. In 20 sessions, young ( $n = 18$ ) and old ( $n = 19$ ) adults were taught to recall lists of 30 words using the Method of Loci. Age differences were magnified early in practice at long presentation times (20 s and 15 s per word) and later at 5 s per word. Regression of posttraining scores on various pretraining abilities revealed significant effects of digit symbol substitution. Also, consistent with the assumption of age-related decline in developmental reserve capacity, the unique variance in serial word recall associated with age group became more salient as the training unfolded.

In a previous study, we used a testing-the-limits approach to examine age differences in developmental reserve capacity for memory performance during adulthood (Kliegl, Smith, & Baltes, 1989). As a consequence of learning how to use a specific mnemonic technique (i.e., the Method of Loci; Bower, 1970), young and older adults were able to substantially improve their memory performance in serial word recall. Alongside this demonstration of substantial performance plasticity in both young and older adults, however, age differences in mean level of recall were magnified. In addition, whereas before training the performance ranges of young and older adults were nested, after training very few older and young adults overlapped in their performance. We obtained this result in two experiments using different training programs. Figure 1 displays the Age  $\times$  Training interaction and the ranges of performance obtained in Experiment 2 of that earlier study.

The data set of this experiment also contains information relevant to issues that, because of theoretical emphasis and limitations of space, we could not elaborate on in the previous article and are therefore addressing here. Specifically, we had two major objectives. The first objective was to chart the acquisition

functions of older and young adults at several points during the training program to understand better the conditions of magnification of age differences. Specifically, we wanted to determine whether the magnification was due to instruction in the mnemonic device or, instead, was a function of subsequent practice. Immediately after instruction, the impact of performing the memory task with a different set of cognitive operations (associated with the Method of Loci) should be most visible. Thus, it is possible that the Age  $\times$  Training interaction obtained after the entire training schedule (i.e., at posttest, see Figure 1) was already present at an earlier stage in the acquisition process.

Our second objective was to examine the question of predictability of posttraining differences. The central question was whether the posttraining differences in memory performance represented "novel" variance associated with developmental reserve capacity or whether these differences could be easily detected and predicted prior to training with indicators of what we have labeled *baseline performance* (Kliegl & Baltes, 1987). We hypothesized that in the context of the Method of Loci training study, three ability or skill clusters were particularly promising as predictor variables: (a) pretraining memory performance, (b) fluid-type measures of intelligence, and (c) a subjective rating of vividness of mental images. Serial word recall was the criterion task throughout the training. Pretraining memory performance was assessed with a measure of serial word recall and a measure of cued recall. Initial ability in cued recall was hypothesized to be predictive of serial word recall because after instruction in the Method of Loci, participants were assumed to use implicit encoding and retrieval cues (i.e., mental landmarks) to perform the memory task (Bellezza, 1981). Thus, in terms of cognitive processing, serial word recall should be more similar to cued recall after instruction in the mnemonic device. Indicators of fluid intelligence (digit symbol substitution and mental rotation) were included because training-related magnification of age differences could conceivably be due to general age-related decline associated with this cluster of abilities (Cat-

---

This research was conducted in the context of the project Expertise and Cognitive Aging, codirected by Paul B. Baltes and Reinhold Kliegl. We express our gratitude to Annette Rentz and Werner Scholtysik for research assistance and coordination involving the entire project. We also thank the following research assistants who helped us in the present study: Gregor Caregnato, Jackie Durfee, Sabine Edler, Anita Günther, and Dirik Sahm. We are especially grateful to our study participants for their patience, interest, and persistence. We also appreciate helpful comments by Ulman Lindenberger, Ulrich Mayr, Timothy A. Salthouse, Laura A. Thompson, and two anonymous reviewers on an earlier version.

Correspondence concerning this article should be addressed to Reinhold Kliegl, Max Planck Institute for Human Development and Education, Lentzeallee 94, D-1000 Berlin 33, Federal Republic of Germany.

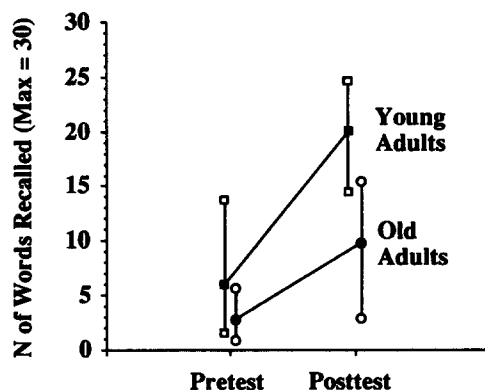


Figure 1. Age  $\times$  Time of Assessment interaction for serial recall of words. (Data represent average recall assessed for presentation times of 20 s, 15 s, 10 s, 5 s, 3 s, and 1 s per word. Bar lines indicate range of scores. Reprinted from "Testing-the-Limits and the Study of Adult Age Differences in Cognitive Masticity of a Mnemonic Skill," by R. Kliegl, J. Smith, and P. B. Baltes, 1989, *Developmental Psychology*, 25, pp. 247–256. Copyright 1989 by the American Psychological Association. Reprinted by permission.)

tell, 1971; Horn, 1982). The last predictor variable, vividness of mental images, was chosen because the Method of Loci mnemonic technique is assumed to require the generation of mental images linking to-be-remembered material with landmarks. Subjective ratings of vividness of mental images might be predictive of the ability to work with the mnemonic device.

The following predictions were made. First, we predicted that a major share of the variance following training would be detectable early on and could be predicted from pretraining information. This prediction was based primarily on the view that training programs, even if they involve a new set or coordination of cognitive skills, as is true for the Method of Loci, need to be founded on extant abilities or skills. The second prediction, however, was that some new variance would be generated by training of the testing-the-limits kind (P. B. Baltes, 1987; Kliegl & Baltes, 1987). Specifically, we maintained that performance near limits of developmental reserve capacity revealed by the testing-the-limits approach carries information about latent cognitive potential not available prior to training.

The basic data set used to examine these questions was collected in the context of the second experiment reported in Kliegl et al. (1989). The present analyses extend information about the training effects reported in the earlier article by focusing on the process of acquisition during the training program. In addition, we report new information on psychometric measures collected prior to training, which were administered to predict individual differences in acquisition and in the final level of performance achieved.

## Method

In the following sections, we describe the design and the training program. The study was cross-sectional and involved a total of 20 sessions. Subjects were paid DM 20 (approximately \$13) for each of the 20 sessions.

## Participants

The final study sample consisted of 18 young adults (11 men, 7 women; mean age = 23.9 years, range = 19–29 years) and 19 older adults (8 men, 11 women; mean age = 71.7 years, range = 65–80 years) from Berlin, Federal Republic of Germany. Two young adults dropped out for lack of interest, and 1 young adult and 2 older adults dropped out for health-related reasons. All of the participants were volunteers and physically able to come to the laboratory. Reported subjective health was judged as above average (on a 5-point self-report scale: young adults,  $M = 4.2$ ,  $SD = 0.7$ ; older adults,  $M = 3.9$ ,  $SD = 0.7$ ; this difference was not statistically significant).

The young adults were university students specializing in various fields of study who had completed 13 years of school and from 1 to 6 years of university studies. Their average Hamburg-Wechsler Intelligence Test (HAWIE) IQ, assessed in the first session, was 117.5 ( $SD = 6.8$ ).

Because we aimed for a conservative design favoring the older adults, we were interested in comparing young adults with a positively selected sample of older adults to minimize issues of negative cohort differences in education, brain-related pathology, and lack of practice in mental testing. Four months prior to this experiment, the older adults had participated in a cognitive training study with items defining the fluid ability factor of figural relations (P. B. Baltes, Sowarka, & Kliegl, 1989). On the pretest measure of intelligence given in that study (i.e., the German version of the Adult Development and Enrichment Program battery; P. B. Baltes, Dittmann-Kohli, & Kliegl, 1986; P. B. Baltes & Willis, 1982), they had scored above the 60th percentile of the sample. Their average HAWIE IQ prior to the present experiment was 125.4 ( $SD = 8.0$ ); their average level of education was 12.5 years ( $SD = 3.0$ ).

Although the older adults had a significantly higher age-corrected IQ than the young adults,  $t(35) = 3.2$ ,  $p < .01$ , the age groups were equal in verbal raw scores ( $M = 63.3$ ,  $SD = 3.8$ , for young adults, and  $M = 63.1$ ,  $SD = 6.1$ , for older adults). Moreover, the older adults had significantly lower raw performance scores than the young adults, with means of 54.0 ( $SD = 7.1$ ) versus 62.2 ( $SD = 6.4$ ), respectively,  $t(35) = 3.7$ ,  $p < .01$ . Thus, the older sample's intelligence profile was consistent with the pattern of normal aging.

## Design of Test and Training Program

The test and training program comprised a total of 20 sessions per participant (for an overview of the program, see Table 1). In the first four sessions, baseline data were collected, including a standardized test of intelligence and an assessment of memory performance on serial word recall, cued recall, and recognition tasks. The next 12 sessions (Sessions 5 to 16) comprised training and intermittent assessment of the target mnemonic skill (i.e., of serial word recall performance). In Sessions 5 and 6, participants were instructed in the Method of Loci. The effect of this initial instruction was tested in Session 7. The subsequent nine sessions consisted of six individualized adaptive practice sessions designed to provide a learning environment supportive of further optimization of the Method of Loci and three intermittent assessments under standardized conditions. The final four sessions were set aside for posttraining assessment (a repeat of baseline memory measures with different materials, Sessions 17 to 19) and debriefing (Session 20).

## Apparatus and Material

Memory tests and training were administered on Apple IIe computers. Two sets of 18 lists were assembled for assessment of serial word recall, cued recall, and recognition at pretest and posttest. Each list

Table 1  
*Experimental Schedule*

Phase	Session	Type of meeting	Type of session
Pretest	1	I	Psychometric assessment
	2	G	Serial word recall (6)
	3	G	Cued recall (6)
	4	G	Recognition (6)
Training	5	G	Instruction in method of loci
	6	I	Instruction in method of loci
	7	G	Assessment of serial recall (6)
	8	I	Adaptive practice (4)
	9	I	Adaptive practice (4)
	10	G	Assessment of serial recall (6)
	11	I	Adaptive practice (4)
	12	I	Adaptive practice (4)
	13	G	Assessment of serial recall (6)
	14	I	Adaptive practice (4)
	15	I	Adaptive practice (4)
	16	G	Assessment of serial recall (6)
Posttest	17	G	Serial word recall (6)
	18	G	Cued recall (6)
	19	G	Recognition (6)
	20	G	Debriefing

*Note.* The sessions were devised as either an individual (I) or a group (G) meeting, both with a tutor. Psychometric assessment lasted 2 hr, the other sessions 1–1.5 hr. Number of lists (30 words each) administered in a given session is given in parentheses.

consisted of 30 concrete nouns. Presentation order of sets was counter-balanced across test occasions. Lists within sets and words within lists were administered in an invariant order.

Two random orders of a single set of 360 concrete nouns were constructed for individualized, adaptive practice. Participants were exposed to both forms in a counterbalanced order. Thus, 24 lists with 30 concrete nouns in each list were used.

Thirty Berlin landmarks constituted the mental map for the Method of Loci mnemonic. The landmarks (well-known buildings and places) were arranged in a geographically meaningful order. Lists of these landmarks were used as retrieval cues during some test sessions.

### *Procedure for Instruction in the Method of Loci (Sessions 5 and 6)*

Instruction in the Method of Loci occurred in two sessions. In the first session, the Method of Loci was described according to Bower (1970). Participants were shown color photographs of the 30 landmarks and asked to build their own associations for each location. They then practiced using the technique to recall three 10-word lists. Successive sets of 10 landmarks were used for this practice. Verbal labels of landmark cues were visible during encoding and recall. The tutor evaluated the memory images and emphasized the need to create interactive, dynamic images or thoughts. Participants were told that they had to overlearn the list of 30 landmarks for the next session; no specific instruction was given in how to accomplish this task.

In the following session, participants were asked to recite the map of 30 landmarks. Training proceeded only if participants could recite the map in correct order within 90 s. All participants met the criterion. (This check of the mental map also occurred at the beginning of adaptive training sessions.) Practice was identical to the previous session except that no location/landmark cues were used during encoding and retrieval; participants had to use their mental map to perform the task.

Words were read by the tutor at a rate that was comfortable for the participant but that emphasized the need to generate images as fast as possible.

### *Procedure for Assessment of Cognitive Abilities*

*Psychometric measures (Session 1).* Participants were administered the HAWIE, the German version of the Wechsler Adult Intelligence Scale. The HAWIE was used as a screening device to select only mentally fit older adults. In addition, participants were administered the Card Rotation Test of Ekstrom, French, and Harman (1976), the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1972, 1973), and a short biographical questionnaire.

*Memory ability at pretest and posttest (Sessions 2 to 4 and 17 to 19).* Participants' memory ability was assessed using serial word recall, cued recall, and recognition tasks<sup>1</sup> in separate experimental sessions. In each of these sessions, six lists of 30 concrete nouns were presented at times of 20 s, 15 s, 10 s, 5 s, 3 s, and 1 s per word, in that order. In sessions of serial word recall (i.e., Sessions 2 and 17), the to-be-remembered words were presented on the computer screen at encoding. In the recall phase, participants were asked to write remembered words on a sheet with 30 lines, preserving as closely as possible the absolute position in which words were presented. In assessment of cued recall (i.e., Sessions 3 and 18), labels of Berlin landmarks appeared above the to-be-remembered words on the computer screen during encoding. In the recall phase, these cues were provided again on the answer sheet. Maximum recall time per list was always 10 min. These test sessions were conducted with groups of 3 to 5 persons.

*Intermittent assessment of serial recall (Sessions 7, 10, 13, and 16).* The procedure described in the previous section was also used for intermittent assessment of serial word recall. One difference was that landmark cues were provided during recall only; no cues were presented during encoding. This variation ascertained that positional recall errors could be attributed to the incorrect application of the mental map during the time-constrained encoding phase.

### *Procedure for Adaptive Practice of Mnemonic Skill (Sessions 8, 9, 11, 12, 14, and 15)*

In each of six adaptive practice sessions, four lists of 30 words were presented. We call these practice sessions *adaptive* because the lists of words were administered at time intervals dependent on the participant's prior recall performance. Starting with 20 s per word, presentation time was decreased whenever a 50% success level was achieved on two consecutive word lists, that is, when 15 or more of the 30 words were recalled in correct serial position. To provide for similarity with pretest and posttest assessment procedures, the following presentation times were used in that sequence: 20 s, 15 s, 10 s, 5 s, 3 s, and 1 s per word.

As reported in Kliegl et al. (1989), all young and older participants reached criterion with 20 s and 15 s per word; all but 1 older adult also passed the 10-s threshold. All but 1 young adult passed the 5-s threshold, but only 11 older adults did. Whereas 72% of the young adults succeeded at 3 s per word, none of the older adults did so. Finally, 1 young person also met the 1-s criterion.

For all practice lists, encoding occurred without explicitly presented cues (landmarks); during recall, the computer prompted with a landmark and the participant provided the appropriate word, if possible.

<sup>1</sup> Young adults' recognition scores (Sessions 4 and 19) were at ceiling for the four slow presentation times. Therefore, these measures were not considered in the analyses.

Table 2  
Performance on Tests of Serial Word Recall

Time per word	Session of assessment											
	Pretest (2)		7		10		13		16		Posttest (17)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Young adults ( <i>n</i> = 18)												
20 s	7.6	6.2	25.2	5.1	27.6	3.5	27.7	2.8	28.2	2.1	28.7	1.9
15 s	8.6	5.6	22.6	7.5	25.9	4.5	23.7	7.1	24.5	7.4	27.2	3.3
10 s	7.3	7.8	19.3	8.5	22.8	6.0	23.4	6.7	23.3	5.0	23.4	5.6
5 s	4.3	2.4	9.3	5.1	14.5	7.4	18.7	7.0	19.4	6.3	19.3	6.2
3 s	4.8	3.7	6.8	5.5	8.7	4.5	9.3	6.6	10.3	5.4	16.4	8.1
1 s	2.0	1.6	2.2	1.4	2.4	1.6	3.2	2.1	3.9	4.2	4.4	2.4
<i>M</i>	5.8	3.1	14.2	3.9	17.0	3.7	17.7	4.2	18.4	3.1	19.9	3.3
Older adults ( <i>n</i> = 19)												
20 s	4.0	3.9	14.0	8.6	18.4	7.7	18.2	7.4	19.9	8.6	17.8	8.7
15 s	3.5	3.2	12.6	6.6	14.1	7.7	16.9	7.1	16.4	7.8	15.4	8.2
10 s	3.0	2.0	10.1	5.3	10.9	6.0	11.8	6.3	13.1	5.7	12.1	6.1
5 s	2.6	2.1	4.3	3.3	5.7	4.7	6.3	4.1	7.1	4.0	7.1	3.1
3 s	1.9	1.2	2.7	1.9	2.4	2.0	3.1	2.0	3.7	2.1	4.1	3.2
1 s	1.3	1.6	1.9	1.5	1.5	1.4	2.3	1.5	2.5	1.4	2.3	1.6
<i>M</i>	2.7	1.6	7.6	3.3	8.8	4.1	9.8	3.8	10.4	4.0	9.8	3.9

Note. Maximum score is 30. Pretest and posttest scores are graphed in Figure 1.

The tutor entered this response on the computer, and latencies to the tutor's first keypress were collected. A response had to be given within 20 s of the prompt. Responses were entered by the tutor to reduce effects of familiarity with computer keyboards and spelling skills. After each list, participants received feedback about the number of words recalled correctly. The computer displayed the landmarks, the to-be-remembered nouns, and the participant's responses.

Between lists, the tutor probed participants about their images or thoughts, offered evaluative comments, and reminded participants of the mnemonic principles. These interactions also served the purpose of keeping the participants engaged in the memory program. An effort was made to provide a relaxed atmosphere during training sessions. Participants were not rushed and could take breaks for coffee and cookies, as well as for conversation with the tutor.

## Results

Presentation of results is organized into the following four sections: (a) immediate effects of instruction in the Method of Loci, (b) the effect of continued practice in using the mnemonic technique, (c) the stability of individual differences in serial word recall within age groups across the training program, and (d) the predictability of memory performance by pretest serial word recall, cued recall, digit symbol substitution, mental rotation, and VVIQ. Means (and standard deviations) of serial word recall at pretest, at intermittent assessments during training, and at posttest, broken down by age group and presentation time, are summarized in Table 2. Means (and standard deviations) for markers of cognitive and intellectual ability are given in Table 3. With the exception of the VVIQ, age groups differed significantly on these marker tests.

### Immediate Effect of Instruction in the Method of Loci

Participants' memory ability in serial word recall was tested prior to (Session 2) and immediately after (Session 7) instruction in the Method of Loci. An Age Group (2)  $\times$  Time of Assessment (2: Sessions 2 and 7)  $\times$  Presentation Time (6) mixed-model analysis of variance (ANOVA) with time of assessment and presentation time as within-subject factors was carried out on serial word recall.<sup>2</sup>

Young adults recalled significantly more words than did older adults,  $F(1, 35) = 35.9$ ,  $p < .01$ . Performance immediately after instruction was significantly higher than performance at pretest,  $F(1, 35) = 118.4$ ,  $p < .01$ . Moreover, young adults benefited significantly more from the instruction than did older adults:  $F(1, 35) = 8.7$ ,  $p < .01$ , for the Age Group  $\times$  Time of Assessment interaction. As shown in Table 2, young adults increased their recall, averaged across all presentation times, from 5.8 to 14.2 words; the analogous values for older adults were 2.7 and 7.6 words.

There were also significant effects associated with presentation time,  $F(5, 175) = 86.4$ ,  $p < .01$ , the interaction between age group and presentation time,  $F(5, 175) = 8.6$ ,  $p < .01$ , the interaction between time of assessment and presentation time,  $F(5, 175) = 37.1$ ,  $p < .01$ , and the interaction involving the three

<sup>2</sup> Presentation time was confounded with order of lists within the experimental session (i.e., the 20-s rate was administered first; the 1-s condition was last). Thus, proactive interference or fatigue may have exacerbated differences between presentation times. Although interpretations must be qualified in this respect, the dominant source of variance is in all likelihood differences in encoding time.

Table 3  
*Psychometric Measures Assessed Prior to Training*

Group	Cued recall (number correct)		Digit symbol substitutions (scaled score)		Mental rotation (number correct)		VVIQ (rating)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Young adults ( <i>n</i> = 18)	13.2	5.2	13.5	1.6	55.8	15.8	2.1	0.5
Older adults ( <i>n</i> = 19)	6.4	2.0	10.7	1.8	31.7	14.7	1.8	0.5

*Note.* Except for VVIQ (Vividness of Visual Imagery Questionnaire), differences between age groups were significant at  $p < .01$ .

factors of age group, time of assessment, and presentation time,  $F(5, 175) = 2.4$ ,  $p < .05$ . As can be seen in Table 2 (Columns 1 and 2), the effect of instruction was larger for long than for short presentation times. Young adults improved more than older adults, especially for the long presentation times. In six post hoc ANOVAS, one for each presentation time, the Age Group  $\times$  Time of Assessment interaction was significant only for scores based on the 20 s and 15 s per word presentation times ( $p < .008$ ; adjusted  $\alpha = .05/6$ ). As can be seen in Table 2, there was restriction in the range of older adults' pretest scores, particularly for short presentation times. The results of the post hoc ANOVAS indicated that the age magnification was largest for the long times. Thus, the potential floor effect apparently reduced rather than generated the interaction. The three-way interaction is also shown as part of Figure 2. Improvement in the number of words recalled from pretest (Session 2) to the first test after instruction in the Method of Loci (Session 7) is displayed as a function of age group and three levels of presentation time. (The particular aggregation of times was chosen because of age-differential changes across subsequent assessments reported below)

In summary, a training-related magnification of age differences was observed immediately after instruction primarily for slow (20-s and 15-s) presentation times. For presentation times at which participants could deploy their newly acquired mnemonic strategy, the young adults benefited more than the older adults.<sup>3</sup>

#### *Cumulative Effect of Practicing the Method of Loci*

The next question was whether continued practice of the Method of Loci would result in a further increase (or reduction) in the age differences. For this analysis, the focus was on the four intermittent tests of serial word recall during training (Sessions 7, 10, 13, and 16). Data were analyzed with a mixed-model ANOVA with age group (2) as a between-subjects factor and session (4) and presentation time (6) as within-subjects factors. Young adults recalled more words than did older adults,  $F(1, 35) = 48.1$ ,  $p < .01$ , there was significant improvement across sessions,  $F(3, 105) = 17.6$ ,  $p < .01$ , and recall was significantly related to presentation time,  $F(5, 175) = 254.5$ ,  $p < .01$ . The two-way interactions between age group and presentation time,  $F(5, 175) = 13.9$ ,  $p < .01$ , and between session and presentation time,  $F(15, 525) = 2.6$ ,  $p < .01$ , were also significant. The inter-

pretation of these interactions was further qualified by the significance of the three-way interaction,  $F(15, 525) = 2.9$ ,  $p < .01$ .

To determine the source of the three-way interaction, six post hoc ANOVAS were conducted separately for the various presentation times, with alpha adjusted to  $.05/6 = .0083$ . In only one of the six analyses, that for the 5-s presentation time, did we obtain a significant interaction between age group and session,  $F(3, 105) = 8.5$ ,  $p < .001$ ; for the other five analyses,  $F$  values for this interaction were below 1.8. The interaction between age group and session for the 5-s presentation time is shown as the middle section of Figure 2. Also shown are the average acquisition functions for scores based on presentation times longer than 5 s (i.e., 20 s, 15 s, and 10 s) and for scores based on presentation times shorter than 5 s (i.e., 3 s and 1 s).

For the 5-s presentation time, the difference between age groups increased significantly from Session 7 to Session 10 and from Session 10 to Session 13. For the final segment, from Session 13 to Session 16, both groups exhibited equivalent increments (0.7 and 0.8 for young and older adults, respectively). No such age-differential change was observed for the other presentation times. Additional analyses showed that of the 12 growth functions (i.e., 6 presentation times  $\times$  2 age groups), only the slope for the young adults' 5-s time was different.

#### *Stability of Individual Differences Across the Training*

An important issue was whether mnemonic training would affect the stability of interindividual differences in serial word recall across the training program. Table 4 contains the correlations among the various tests of serial word recall for young and older adults prior to and during the training program.

The pattern of correlations indicated three important results. First, reliability coefficients were high (i.e.,  $> .80$ ) for all measures. Second, serial word recall at pretest did not correlate significantly with postinstruction assessments of serial word recall; this was true for young and for older adults. Third, the new rank ordering of participants after instruction in the mne-

<sup>3</sup> One may wonder whether older adults would have shown an equal growth rate if presentation times longer than 20 s had been chosen. Although definite evidence is unavailable, it is likely that this would not have been the case because under self-paced conditions (no fixed presentation times), older adults and young adults were shown to use similar times of 22–24 s (Kliegl et al., 1989).

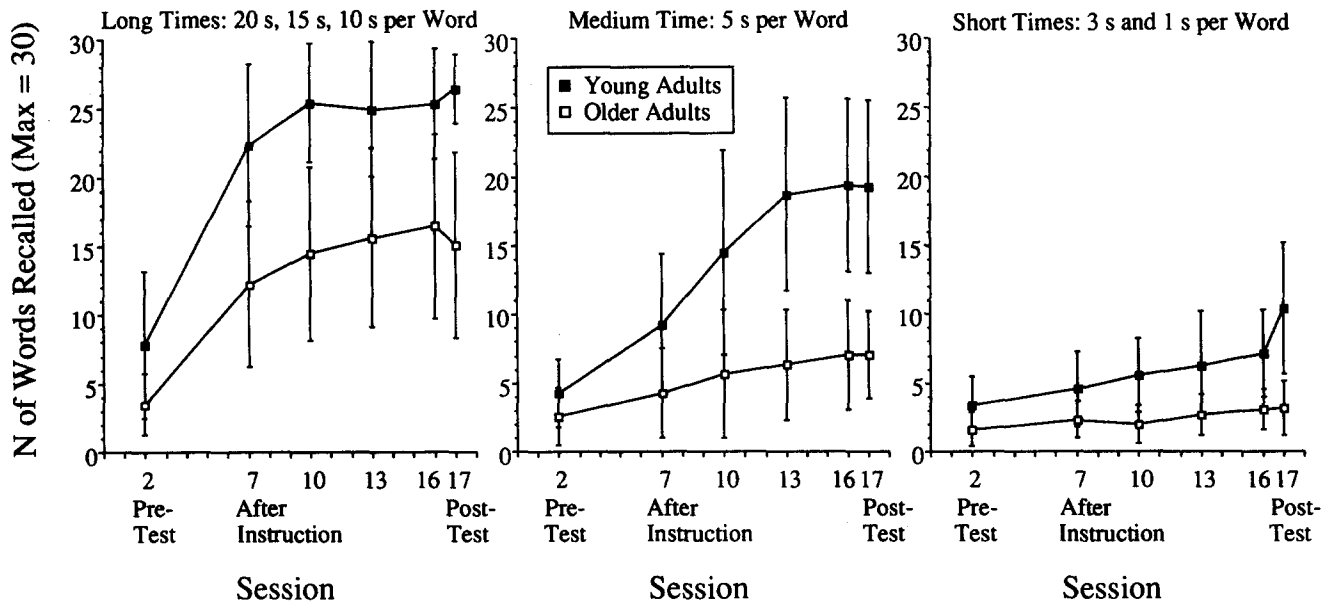


Figure 2. Acquisition functions of the mnemonic skill for young and older adults at different levels of presentation time. (Curves in the left panel are averaged across scores based on 20-s, 15-s, and 10-s presentation times; in the right panel, scores from the 3-s and 1-s times were averaged. Bar lines indicate  $\pm 1$  SD. For content of the sessions, see Table 1.)

monic device was stable across subsequent practice sessions: Correlations between successive assessments were always larger than .60, and the final (Sessions 13 to 17) intersession stability coefficients were all larger than .70.

In summary, within-group individual differences in serial word recall at pretest were not predictive of later memory ability in serial word recall. Given the high reliability of pretest scores, we can rule out measurement problems as the source of this result. Thus, as expected in our cognitive engineering approach to memory skill (Kliegl & Baltes, 1987), instruction in the mnemonic device led to a reshuffling of participants within age groups with respect to performance in serial recall of words. This new ordering was stable soon after basic instruction and remained stable for the rest of the training program.

#### *Predictive Validity of Cognitive Abilities for Serial Word Recall*

The next questions to be addressed were Which of the other measures of cognitive ability assessed prior to training would predict serial word recall at the various stages of practice? and Would these variables account for the added age difference induced by training? Five predictor variables were used: pretest serial word recall, cued recall, digit symbol substitution, card rotation, and vividness of mental imagery. The zero-order correlations of these variables are also part of Table 4.

Separate hierarchical multiple regression analyses were conducted for each assessment of serial word recall. The five predictors were entered in a first step. (For the regression of pretest serial word recall on cognitive abilities, of course, only four predictors were used.) In a second step, a dummy variable coding age group (young, older) was entered. If the addition of this

variable to the regression equations significantly increased  $R^2$ , then the age-group difference in posttest memory performance (i.e., serial word recall using the Method of Loci) would not be accounted for by initial interindividual differences in cognitive abilities.

In Table 5, the results of the six regression analyses are summarized. For each analysis, standardized regression coefficients (i.e., semipartial correlation coefficients),  $R^2$ , and the incremental  $R^2$  after adding the age-group variable are displayed.

*Serial word recall at pretest.* As already mentioned in the previous section, serial word recall prior to training was not predictive of serial word recall after instruction in the mnemonic device. Thus, as expected in our cognitive engineering approach (i.e., introducing the Method of Loci to perform the task), the training substantially affected the way participants performed the memory task.

*Cued recall.* In contrast to pretest measures of serial word recall, individual differences in cued recall prior to training were predictive of serial word recall after instruction. Although this predictor was significant in only one of the analyses (i.e., in Session 7), the coefficients across the assessments (with the exception of Session 10) were of comparable magnitude. This result suggested that mnemonic instruction, as anticipated, induced the use of implicit cues (i.e., the landmarks of the mental map). Because in cued recall, landmark cues were explicitly provided simultaneously with to-be-remembered nouns on the computer screen, the conceptual similarity between cued recall and serial word recall following instruction was obviously of importance. The relevance of cued recall as a predictor of mnemonic skill immediately after instruction was also evident in two post hoc hierarchical regression analyses. With pretest serial word recall and age group partialled out, cued recall ac-

Table 4  
Correlations Between Variables Used in the Regression Analyses

Variable	Prior to instruction					After instruction (SWR test sessions)				
	SWR	CR	DSS	MR	VVIQ	7	10	13	16	17
Young adults ( $n = 18$ )										
Prior to instruction										
Serial word recall (SWR)	(.98)									
Cued recall (CR)	.37	(.97)								
Digit symbol substitution (DSS)	.39	.53*	(.94) <sup>a</sup>							
Mental rotation (MR)	.54*	.35	.21	(.85) <sup>b</sup>						
Vividness of Visual Imagery Questionnaire (VVIQ)	.28	.02	-.09	-.06	(.74) <sup>c</sup>					
After instruction (SWR test)										
Session 7	.38	.52*	.35	.33	-.27	(.98)				
Session 10	.39	.37	.55*	.23	-.06	.63*	(.97)			
Session 13	.24	.47*	.38	-.01	.05	.59*	.83*	(.98)		
Session 16	-.01	.27	.19	-.23	.16	.38	.60*	.86*	(.99)	
Session 17	.25	.57*	.21	.10	.06	.60*	.66*	.84*	.75*	(.93)
Older adults ( $n = 19$ )										
Prior to instruction										
SWR	(.92)									
CR	.43*	(.81)								
DSS	-.05	.14	(.94) <sup>a</sup>							
MR	.26	.24	.38	(.85) <sup>b</sup>						
VVIQ	.13	.21	.17	.16	(.74) <sup>c</sup>					
After instruction (SWR test)										
Session 7	.11	.29	.18	.52*	.06	(.96)				
Session 10	.24	.26	.52*	.48*	.31	.61*	(.97)			
Session 13	.16	.13	.38	.51*	.17	.75*	.86*	(.97)		
Session 16	.15	.30	.60*	.37	.03	.46*	.83*	.77*	(.96)	
Session 17	.09	.07	.68*	.48*	.10	.60*	.77*	.87*	.74*	(.97)

Note. Values in parentheses on the diagonal reflect Spearman-Brown split-half (odd-even) reliability coefficients.

<sup>a</sup> Oswald and Fleischman (1986). <sup>b</sup> Ekstrom et al. (1976). <sup>c</sup> Marks (1973).

\*  $p < .05$ .

counted for a significant 7% unique variance of mnemonic skill in Session 7. In contrast, after controlling for mnemonic skill in Session 7 and age group, no unique contribution of cued recall to later assessments of mnemonic skill was observed.

**Digit symbol substitution.** Digit symbol substitution, the first measure selected to index fluid intelligence, was unrelated to memory performance at pretest and immediately after instruction. However, after participants had practiced their mnemonic skill for two more sessions, a significant relation emerged. This interpretation of the effect of digit symbol substitutions was corroborated by post hoc hierarchical regression analyses, which showed a pattern of results opposite to that obtained for cued recall: no significant change in the  $R^2$  of mnemonic skill after instruction and a significant change of 4% in the  $R^2$  of posttest mnemonic skill. In both cases, age group and prior level of mnemonic skill were partialled out. This result is in accordance with our general theoretical expectations: Mnemonic training should bring about a much clearer relation between an experimental task of memory functioning and a psychometric indicator of fluid intelligence than would be apparent at pretest (P. B. Baltes & Kliegl, 1986; Kliegl & Baltes, 1987).

**Card rotation.** Individual differences in card rotation, an-

other indicator of fluid intelligence, showed a pattern of relations opposite to that for digit symbol substitution and contrary to our expectations. Card rotation was a significant predictor of pretest performance in serial word recall, was still significant after two instructional sessions, but dropped to a nonsignificant level as participants moved through the training program. Post hoc regressions were in agreement with this interpretation. A significant change in  $R^2$  (6%) because of card rotation (with age group and pretest serial word recall partialled out) was observed only for mnemonic skill after instruction. We do not have a good explanation for this result. It is interesting, however, that card rotation correlated significantly with four of five measures of serial word recall in the older adults but with none in the younger adults (see Table 4). Thus, the predictive validity of this variable may be different for different ages.<sup>4</sup>

**Vividness of mental images.** The subjective rating of partici-

<sup>4</sup> For all regression analyses, we also tested the interaction between age group and psychometric markers for significance by adding multiplicative terms in a third step to the equation. A significant change in  $R^2$  would have indicated different regression slopes for the two age groups. None of these statistical tests was significant.

Table 5  
*Results of Six Hierarchical Multiple Regression Analyses of Serial Word Recall on Cognitive Abilities Assessed Prior to Training*

Measure	Session					
	Pretest	7	10	13	16	17 (posttest)
Predictors						
Serial word recall	—	.08	.11	.04	.04	.01
Cued recall	.32	.39*	.13	.32	.28	.30
Digit symbol substitution	.02	.13	.47*	.34*	.47*	.40*
Card rotation	.39*	.33*	.21	.15	.07	.23
VVIQ	.17	-.11	.08	.08	.10	.08
$R^2$	.49*	.60*	.66*	.56*	.54*	.67*
Change in $R^2$ due to age group	.00	.03	.04	.05	.09*	.10*

Note. Values are standardized regression coefficients (i.e., semipartial correlations). Predictors were entered in the first step; age group was entered in the second step. VVIQ = Vividness of Visual Imagery Questionnaire.

\*  $p < .05$ .

pants' vividness of mental images, assessed with the VVIQ (Marks, 1972), did not predict serial word recall at any of the assessments. This was also true for zero-order correlations of this variable with all other variables used in the study (see Table 4).

*Effects of age independent of cognitive abilities.* If individual differences in the cognitive abilities assessed prior to training and investigated above were sufficient to account for the difference in serial word recall between young and older adults, then the addition of a variable coding age group in a second step should not significantly improve  $R^2$ . As indicated in the last row of Table 5, the amount of variance uniquely associated with the difference between age groups increased across the training program. Significant changes (5%) in  $R^2$  due to age group were also obtained in the last two assessments of mnemonic skill when serial word recall after instruction was partialled out instead of pretest serial word recall. This result suggested that processes of aging, not indexed by the cognitive measures used in the present study, may gain in relevance in later stages of practice. The results also suggested, as expected in our testing-the-limits approach aimed at unravelling age-specific facets of developmental reserve capacity (Kliegl & Baltes, 1987), that factors associated with chronological age became more salient as the training unfolded.

### Discussion

This study addressed two major questions. The first concerned the acquisition functions of older and young adults. The focus was on a better understanding of the magnification of age differences in the course of a mnemonic training program involving the use of the Method of Loci. The second question concerned the predictability of posttraining differences and whether the training program revealed novel information about developmental reserve capacity not available at pretest.

#### *Course of Magnification and Its Prediction*

In the time frame of a 20-session study (8 sessions of which were devoted to mnemonic training), most of the training-re-

lated magnification of age differences in serial word recall was due to the initial instruction in the use of the Method of Loci. Subsequent practice saw a further magnification of the age difference for the 5-s presentation time. This added magnification associated with continued practice was primarily a consequence of the fact that only young adults were able to use the Method of Loci under this condition.<sup>5</sup>

What about the predictability of posttraining differences? There was evidence for effects specific to the stage of learning and for a change in the compositional nature of the task. Instruction in the Method of Loci and subsequent practice led to an increase in the variance ( $R^2$ ) accounted for by pretest assessment of cognitive abilities. In addition, the magnitudes of predictive relations changed because of the process of cognitive engineering of serial word recall. Whereas pretraining scores were related to card rotation and cued recall, the posttraining outcome identified digit symbol substitution as the most powerful indicator (Table 5). Such an outcome could be expected because digit symbol substitution indexes multiple intellectual abilities such as visuomotor coordination, selective attention, psychomotor speed, and memory.

Training likely induced the use of more general intellectual resources and, in addition, reduced the impact of task-specific performance factors such as familiarity with test situations. After instruction, participants approached the memory task in a more standardized manner than before (Kliegl & Baltes, 1987; Kliegl, Smith, Heckhausen, & Baltes, 1987). This training-related increase in predictive power and the stage-depen-

<sup>5</sup> Age-differential development spurts in the acquisition of mnemonic skill could covary with progress in the training schedule. Presentation time in training sessions was increased according to recall levels. Young adults moved somewhat faster through the initial practice sessions with long presentation times than did older adults. Nevertheless, at the intermittent test in Session 10, 53% of older adults ( $n = 10$ ) had practiced the memory task under the 5-s condition at least once; the comparative value for young adults was 72% ( $n = 13$ ). Thus, very early in the training, the majority of young and older adults were practicing the skill with 5-s presentation times per word.



dent changes in predictive relations can be interpreted as adding to a growing body of literature suggesting that testing-the-limits or learning-test procedures lead to a more precise assessment of cognitive ability constructs than do standard one-time assessments (e.g., M. M. Baltes & Kindermann, 1985; Carlson & Wiedl, 1978; Embretson, 1987; Guthke, 1982).

What about the role of chronological age? Did the testing-the-limits procedure highlight the status of chronological age? The amount of variance uniquely associated with age group was found to increase in the course of the training. This demonstration of novel variance of learning potential associated with adult age differences at high levels of skill would seem to corroborate the usefulness of the testing-the-limits approach and the rationale of searching for age differences in latent potential at the level of developmental reserve capacity (P. B. Baltes, 1987; Kliegl & Baltes, 1987). At the same time, we also need to acknowledge that other measures of cognitive ability could have been administered prior to training (e.g., measures of learnability as suggested by Guthke, 1982), some of which might have accounted for the variance associated with the age differences in developmental reserve capacity that emerged during training. Thus, it is an open question whether developmental reserve capacity can be captured without extensive training by other means of assessment. In future research, alternative predictor constellations need to be used (including, for example, a paired-associate learning task).

#### *Would Further Practice Reduce or Enhance Age Differences?*

Assessment of developmental reserve capacity by means of a theory-guided training program should, in the long run, lead to an identification of asymptotic performance limits and to a better quantification of individual differences in latent potential, or what we call developmental reserve capacity (Kliegl & Baltes, 1987). Moreover, if performance limits are true indicators of maximum developmental potential, individual differences in general and age differences in particular should not be reversible at late stages of practice (i.e., near limits of performance).

The present data set, because of its limitations in extent of training, does not offer a final answer to the question of whether older adults have a lower level of developmental reserve capacity than do young adults. However, the data do suggest a trend. If, as we hypothesize, older adults have a definite loss in developmental reserve capacity, the following three expectations should be met. First, benefits of mnemonic skill should be visible sooner for long than for short presentation times because execution of the new set of cognitive operations will initially take extra time. With practice, the mnemonic skill should be transferable to shorter and shorter presentation times. This assumption is also consistent with extant theories about acquisition of cognitive skill (e.g., Anderson, 1982). Second, for a given presentation time, efficient application of the mnemonic skill should occur first for young adults and only at a later point in time—if at all—for older adults. Third, for any given presentation time and with increasing probability for shorter presentation times, the asymptotic level of performance should be lower for older than for young adults. The combina-

tion of these three expectations implies a continuous magnification of age differences and increasing evidence for age differences in developmental reserve capacity.

What is the evidence given by the present data? Consistent with the first expectation, benefits of instruction were initially restricted to long presentation times. In addition, consistent with the second expectation, young adults were the first to reach a high level of effective use of the Method of Loci (Figure 2), whereas older adults clearly performed below ceiling, even with long presentation times. Furthermore, a developmental spurt for a short time interval (5 s) could be observed only for young adults. Moreover, with further practice, developmental spurts such as the one observed for 5-s intervals may be likely for young adults for shorter times as well. As a matter of fact, there is some added evidence (tentative as it is): The young adults' change from Session 16 to Session 17 (i.e., between the final two assessments of serial word recall; see Figure 2 and Table 2) indicated that at least some of them were starting an acceleration of memory performance at the 3-s time condition.

The third prediction, lower asymptotic levels for older adults compared with young adults, can hardly be answered with a 20-session training program. Using linear models of growth, one can extrapolate somewhat from changes in response to practice subsequent to instruction. For long presentation times (20 s, 15 s, 10 s), young adults were functioning close to ceiling; older adults did not exhibit major increments after Session 10, although there was definite room for improvement. Based on the present data, the maximum level is limited roughly by the acquisition function for long times in Figure 2. This function is linear from Session 10 to Session 16. Assuming that it would continue to take an estimated four sessions to increase average recall by one word, another 40 sessions would be required for the older adults to reach the current level of young adults for long presentation times. Most likely, however, with more sessions, the session-to-session improvement would start to decrease if mnemonic skill acquisition follows the power law of practice reported for other cognitive and motoric skills (e.g., Newell & Rosenbloom, 1981).

The alternative expectation to decreasing gains—that is, growth spurts for older adults for long presentation times and for longer times than those included in the present study—seems unlikely. This is because it has been shown in other research in our laboratory that older adults prefer presentation times of 22–24 s even under conditions in which they can dictate time of presentation in a self-paced manner (Kliegl et al., 1989). With such long intervals, older participants reported that they were not further able to improve their mental images or thoughts. At the same time, they were concerned that the resultant increase in time between encoding and retrieval contributed to an increased likelihood of forgetting their mental products. Thus, it appears unlikely that the lack of presentation times longer than 20 s per word was the limiting factor in the level of mnemonic skill reached by older adults.

In summary, we argue that it is likely that the magnification of age differences obtained would continue if participants were asked to continue practice—especially if one were to compute the age difference as integral over all presentation times between, say, 1 s and 25 s. We interpret this as further evidence for robust age differences in developmental reserve capacity. The

stability of individual differences in memory ability after instruction and the increasing amount of unique variance associated with chronological age are also in line with this interpretation. Thus, it seems implausible, on the basis of the present data, to construct a scenario in which age differences would decrease with further practice. But this reasoning, of course, needs to be tested with more extensive training studies. In the long run, as argued earlier (Kliegl & Baltes, 1987), our expectation will have to be evaluated against strict criteria of asymptotic levels of performance.

### *On the Locus of Aging-Related Decline in Developmental Reserve Capacity*

In this concluding section, we would like to discuss the possible significance of the present findings for general theories of cognitive aging. Integrative accounts of cognitive aging that focus on aging loss assume that age-related differences are the reflection of limited resources in some critical capacity.<sup>6</sup> Three accounts prevail. A first one is that older adults have a reduced developmental reserve capacity in fluid intelligence (P. B. Baltes, 1987; Horn, 1982). A second account is the view that a reduction in the rate of processing is critical (Salthouse, 1985). Third, task complexity—irrespective of the specific components involved—has been proposed as the central issue (P. B. Baltes, Dittmann-Kohli, & Dixon, 1984; Craik, 1983; Hasher & Zacks, 1979).

Because these theoretical orientations aim at explaining the dominant findings of age-related decline in cognitive processing, the present study can be placed into each of these frameworks and, fortunately or unfortunately, offers some evidence for each of these perspectives. Because of the evolving correlation between digit symbol substitution and the training task (see Table 5), it could be argued that the present findings contribute to the body of data suggesting that fluid intelligence is a critical factor. This interpretation is only tenuous because of the finding that the other indicator of fluid intelligence, card rotation, showed a different pattern of relations to mnemonic skill.

Digit symbol substitution was also proposed as a prime indicator of mental processing rate (Salthouse, 1985). Therefore, the training-related increase in the predictive validity of this measure can be interpreted as support for age-related decline in the processing rate. More direct evidence was provided by two other results. First, the fact that older adults were not able to perform adequately with short presentation times (i.e.,  $\leq 5$  s) gives credence to the position that speed of generating mental images or thoughts is of importance. The second source of evidence is drawn from the Age Group  $\times$  Presentation Time interaction obtained immediately after instruction. Magnification of age differences was especially pronounced for long presentation times. This result is consistent with the processing-rate theory of cognitive aging if we assume that recall is tied to the number of processing cycles and that younger adults, because of relatively short-duration mental processing cycles, execute more of those cycles than do older adults in the same amount of time.

Task complexity arguments are also relevant. By introducing the Method of Loci, task complexity was most likely increased. At encoding, for example, participants spent the available time per word in generating and elaborating their mental images or

thoughts and in successively accessing mental landmarks in correct serial order. Also consistent with the task complexity account is the view that participants, as they responded to the demands of the Method of Loci, increased their degree of self-initiated processing, which in turn required a larger and larger share of reserve capacity (Craik, 1983; Kliegl et al., 1989).

The determination of limits of developmental reserve capacity is, of course, only a first step toward explaining age and other interindividual differences. We also need to attempt to identify the component processes that are critical in accounting for the observed group differences in cognitive performance. By engineering a cognitive skill in the laboratory, one can control the task-relevant knowledge and strategies that are critical for achieving and maintaining high levels of performance. Such control should be advantageous in isolating aging-related deficits in component processes.

<sup>6</sup> We need to emphasize that our general theory of cognitive aging does not state that all cognitive developmental reserve capacity decreases with age (P. B. Baltes, 1987; P. B. Baltes & Kliegl, 1986). Our prediction about aging loss is restricted to what we call the "mechanics" of cognitive functioning.

### References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Baltes, M. M., & Kindermann, T. (1985). Die Bedeutung der Plastizität für die klinische Beurteilung des Leistungsverhaltens im Alter [The significance of plasticity for clinical assessment of performance in old age]. In D. Bente, H. Coper, & F. Kanowski (Eds.), *Hirnganische Psychosyndrome im Alter: Vol. 2. Methoden zur Objektivierung pharmakotherapeutischer Wirkung* [Organic brain disease in old age: Vol. 2. Methods for objective assessment of pharmacotherapeutic effects] (pp. 171–184). Berlin: Springer-Verlag.
- Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, 23, 611–626.
- Baltes, P. B., Dittmann-Kohli, F., & Dixon, R. A. (1984). New perspectives on the development of intelligence in adulthood: Toward a dual-process conception and a model of selective optimization with compensation. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 6, pp. 34–76). San Diego, CA: Academic Press.
- Baltes, P. B., Dittmann-Kohli, F., & Kliegl, R. (1986). Reserve capacity of the elderly in aging-sensitive tests of fluid intelligence: Replication and extension. *Psychology and Aging*, 1, 172–177.
- Baltes, P. B., & Kliegl, R. (1986). On the dynamics between growth and decline in the aging of intelligence and memory. In K. Poeck, H.-J. Freund, & H. Gänshirt (Eds.), *Neurology* (pp. 1–17). Berlin: Springer-Verlag.
- Baltes, P. B., Sowarka, D., & Kliegl, R. (1989). Cognitive training research on fluid intelligence in old age: What can older adults achieve by themselves? *Psychology and Aging*, 4, 217–221.
- Baltes, P. B., & Willis, S. L. (1982). Plasticity and enhancement of intellectual functioning in old age: Penn State's Adult Development and Enrichment Program (ADEPT). In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 353–389). New York: Plenum Press.
- Bellezza, F. S. (1981). Mnemonic devices: Classification, characteristics, and criteria. *Review of Educational Research*, 51, 247–275.
- Bower, G. H. (1970). Analysis of a mnemonic device. *American Scientist*, 58, 496–510.

- Carlson, J. S., & Wiedl, K. H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency, 82*, 559-564.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London, B 302*, 341-359.
- Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Embretson, S. E. (1987). Improving the measurement of spatial aptitude by dynamic testing. *Intelligence, 11*, 333-358.
- Guthke, J. (1982). The learning test concept—An alternative to the traditional static intelligence test. *The German Journal of Psychology, 6*, 306-324.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356-388.
- Horn, J. L. (1982). The aging of human abilities. In B. B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 847-870). Englewood Cliffs, NJ: Prentice-Hall.
- Kliegl, R., & Baltes, P. B. (1987). Theory-guided analysis of mechanisms of development and aging through testing-the-limits and research on expertise. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 95-119). Norwood, NJ: Ablex.
- Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology, 25*, 247-256.
- Kliegl, R., Smith, J., Heckhausen, J., & Baltes, P. B. (1987). Mnemonic training for the acquisition of skilled digit memory. *Cognition and Instruction, 4*, 203-223.
- Marks, D. F. (1972). Individual differences in the vividness of visual imagery and their effect on functioning. In P. W. Sheehan (Ed.), *The function and nature of imagery* (pp. 83-108). San Diego, CA: Academic Press.
- Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64*, 17-24.
- Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.
- Oswald, W. D., & Fleischmann, U. (1986). *Nürnberger Altersinventar* [Nuremberg Aging Inventory]. Nuremberg, Federal Republic of Germany: Universität Erlangen-Nürnberg.
- Salthouse, T. A. (1985). *A theory of cognitive aging*. Amsterdam: North-Holland.

Received July 11, 1989

Revision received April 11, 1990

Accepted April 11, 1990 ■