Child Neuropsychology
1998, Vol. 4, No. 3, pp. 161-177

0929-7049/98/0403-161$12.00

Rey AVLT: Developmental Norms for Children and the Sensitivity of Different Memory Measures to Age*

Eli Vakil[1], Haya Blachstein[1], and Masha Sheinman[2]
[1]Bar-Ilan University, Ramat-Gan, Israel, and [2]Psychological and Counseling Services of the Israeli Ministry of Education (SHEFI), Ma'aleh Adumim, Israel

ABSTRACT

Nine hundred and forty-three children (487 boys and 456 girls, age range 8 to 17 years, divided into 10 age cohorts) were administered the Hebrew version of the Rey AVLT. Separate norms for boys and girls in each age group are reported. Sensitivity to age for 14 memory measures extracted from the Rey AVLT was analyzed. One of the most consistent findings across the different scores is that memory changes in the 8- to 10- year-old age range are more dynamic than changes in the 11- to 17- year-old age range. The results also show a significant and consistent advantage for girls over boys on most of the verbal memory measures. Detailed analyses of sensitivity to age of the different measures extracted from the Rey AVLT provide very useful information for the diagnostician testing auditory-verbal memory in children.

The Rey Auditory-Verbal Learning Test (AVLT) has been found to be affected by age, intelligence, and population type (Query & Berger, 1980; Query & Megran, 1983; Wiens, McMinn, & Crossen, 1988). The fact that a number of memory components are measured, such as immediate and delayed recall, learning rate, and recognition, clearly enhances the test's sensitivity as a diagnostic tool (Lezak, 1983; Query & Megran; Ryan, Rosenberg, & Mittenberg, 1984; Wiens et al., 1988).

Based on factor analysis studies with adult participants, scores extracted from the Rey AVLT reflect different aspects of verbal memory (Ryan et al., 1984). Moses (1989) grouped three of the Rey AVLT scores with Benton's Visual Retention Test (BVRT, Moses, 1989) scores in an 'immediate memory' factor. Smith, Ivnik, et al. (1992), and Smith, Ivnik, Malec, and Tangalos (1993) submitted scores from the Wechsler Adult Intelligence Scale-Revised (WAIS-R), Wechsler Memory Scale-Revised (WMS-R), and two scores from the Rey AVLT to factor analysis. One of the Rey AVLT scores was loaded on the 'learning' factor and the other on the 'retention' factor. In a more recent study by Vakil and Blachstein (1993), factor analysis produced one, two, or three factors, depending on the combination of scores included in the analysis. The basic factors that were identified were 'acquisition' and 'retention'. The latter could be further subdivided into 'storage' and 'retrieval', thus yielding a total of three factors.

Children with attention deficit hyperactivity disorder showed improvement on the Rey AVLT after drug therapy (Barrickman, Perry, Allen, & Kuperman, 1995). Several scores from the Rey AVLT differentiated between adolescents with confirmed neuropathologies and control participants (Powell, Cripe, & Dodrill, 1991). The Rey AVLT was found to make a unique contribution to the memory assessment of learning-disabled children (Talley, 1986).

---

* This research was funded by a grant from the Psychological and Counseling Services of the Israeli Ministry of Education (SHEFI).
Address correspondence to: Eli Vakil, Psychology Department, Bar-Ilan University, Ramat-Gan, 52900, Israel.
Accepted for publication: September 4, 1997.

Appropriate age norms are a crucial requirement for deriving full advantage from the test. In recent years several normative studies for the Rey AVLT were published. Most of these studies tested younger and older adults (Bleeker, Bolla-Wilson, Agnew, & Meyers, 1988; Geffen, Moar, O'Hanlon, Clark, & Geffen, 1990; Ivnik, Malec, Tangalos, & Petersen, 1990; Ivnik et al., 1992; Mitrushina, Satz, Chervinsky, & D'Elia, 1991; Savage & Drew Gouvier, 1992; Smith, Malec, & Ivnik, 1992; Vakil & Blachstein, 1997). To date, we are aware of only two published children's norms for the English version (Lezak, 1983) of the Rey AVLT (Bishop, Knights, & Stoddart, 1990; Forrester & Geffen, 1991). Bishop et al. tested 252 patients aged 5 to 16 years who were referred for neuropsychological assessment. Based on this sample of patients, norms for 10 scores extracted from the Rey AVLT are reported for each age group. The number of participants in each age group ranged from 9 to 33. Bishop et al. reported that three of the Rey AVLT scores (total number of words recalled in the first five trials, delay and recognition trials) correlated signifigantly with age. Overall, there was no gender effect. As pointed out by Forrester and Geffen, the major problem with this study is the fact that the norms are based on performance of patients with problems known to affect memory (e.g., head injury, epilepsy, learning disability, etc.).

Forrester and Geffen (1991) tested 80 boys and girls, aged 7 to 15, who were divided into four age groups, with 10 boys and 10 girls in each age group. Age effect was found only for some of the scores derived from the Rey AVLT (i.e., acquisition, retention, word span, and retrieval efficiency). As reported by Bishop et al. (1990), gender effect was not significant. The limitations of Forrester and Geffen's study are the overall small sample size and the related fact that each age group included children within a 2-year age range.

Most studies that have reported either adults' or children's norms only analyzed the overall effect of age on different memory scores, but have not reported systematically which age groups were specifically differentiated by which scores. In a previous study by the authors (Vakil & Blachstein, 1997), norms of the Rey AVLT Hebrew version were reported for 528 adults ranging in age from 21 to 91 years. This study reported on sensitivity to age, of 22 memory scores extracted from the Rey AVLT. Furthermore, for each score found to be significantly affected by age, follow-up analysis was conducted in order to detect the source of this effect, that is, which age groups specifically were differentiated by the particular score.

As in the previous study (Vakil & Blachstein, 1997), the primary goal of the present study was to compare age groups on the different memory scores extracted from the test, using raw as well as combined scores (i.e., difference between two raw scores or the sum of several raw scores). This will help to identify the measures most sensitive to change for each age group. Such analyses make a clear diagnostic and theoretical contribution. An additional goal was to publish norms for the Rey AVLT Hebrew version, based on a very large sample of children. Using the same Rey AVLT Hebrew version, Vakil and Blachstein (1994) suggested executing an additional trial of the Rey AVLT in order to test memory for temporal order. Norms for this score were analyzed and are reported as well.

METHOD

Participants

Nine hundred and forty-three children (487 boys and 456 girls) participated in this study. The ages of the sample population ranged from 8 to 17 years, divided into 10 age cohorts. Because the temporal order task requires writing the answer, we chose not to test children younger than 8 years old in order to ensure writing ability. Eight hundred and nineteen children (424 boys and 395 girls) were recruited for the study from a population of children in 14 public schools in central Israel (i.e., the greater Tel Aviv area). The Israeli Ministry of Education uses a scale by which all public schools in the country are ranked according to five criteria: parents' income, parents' education, family size, proportion of immigrants in the school, and distance from a major city. We were referred to public schools ranked in the middle range of this scale. Only four of the selected schools did not participate for technical reasons (e.g., another study was being conducted in one of the schools at the same time). After receiving approval from the school's principal, the class teacher was approached. The teacher or the principal was asked *not* to refer children with either very high or very low academic achievement to the study. Based on the teacher's judgement, children with learning disabilities, attention disorders, or those requiring special assistance in school, were also excluded. All the children were selected according to their birth dates, so that only those children whose birthdays fell within the 3 months prior to or after the testing date were tested. All participants were born in Israel and Hebrew was their mother tongue. Ten to 20 children participated from each school in each class year, who were sampled from different parallel classes of the same school year. Each school year was sampled from four to six schools. An additional 124 children (63 boys and 61 girls) in the same age groups from different parts of the country were tested on a voluntary basis, fulfilling the same selection criteria as the former group. In a preliminary analysis, this group did not differ from the rest of the sample on any parameter, so therefore the two samples were merged. The number of boys and girls in each age group is shown in Table 1.

Test and Procedure

Children were tested individually in a room allocated for this purpose, in their own schools, and during school hours. The children participated voluntarily in the study. Furthermore, they were told that they could stop at any time if they wished to do so. This happened with just a few children who claimed that they were tired. In addition to the Rey AVLT, the children were tested on a larger test battery that included: Vocabulary, Digit Symbol, Digit Span (subtests of the WISC-R), and the Trail Making Test. Some of the tests were administered during the 20-min delay in administration of the Rey AVLT, and the remainder following the Rey AVLT. The Vocabulary test was always administered following Rey AVLT in order not to expose the children to new words that might interfere with the word list of the Rey AVLT. This paper reports only the Rey AVLT results. The examiners in this project were 14 undergraduate psychology majors at Bar Ilan University, who were trained to administer and score the tests.

The Hebrew version of the Rey AVLT was used (Vakil & Blachstein, 1993). Administration was standard, as described by Lezak (1983). The test consists of 15 common nouns (e.g., drum, curtain, bell), which were read to the participants, at the rate of one word per second, in five consecutive trials (Trials 1 through 5); each reading was followed by a free recall task. In Trial 6, an interference list of 15 new common nouns was presented, followed by free recall of these new nouns. In Trial 7, without an additional reading, participants were again asked to recall the first list. Twenty minutes later, and again without an additional reading, participants were once more asked to recall the first list (Trial 8). Next, in Trial 9, they were given a list of 50 words (15 from the first list, 15 from the second, and 20 new common nouns) and were asked to identify the 15 first-list words. To measure the ability to remember temporal order, an extra trial (Trial 10) was added to the standard administration (Vakil & Blachstein, 1994; Vakil, Blachstein, & Hoofien, 1991). In Trial 10, which follows the recognition task, participants were presented with the 15 first-list words written in an order different from that originally presented. Participants were asked to write the words in their original order.

RESULTS

Fourteen scores were derived from the Rey AVLT (i.e., 8 raw and 6 combined scores). The combined scores presented are those frequently used in the literature and reported to reflect different memory processes (Geffen et al., 1990; Ivnik et al., 1992; Query & Megran,

Table 1. Number of Boys and Girls at Each Age Group.

	Age (years)									
	8	9	10	11	12	13	14	15	16	17
Boys	51	51	56	55	45	42	42	61	47	37
Girls	49	49	54	48	45	41	45	43	43	39
Total	100	100	110	103	90	83	87	104	90	76

1983; Ryan et al., 1984; Vakil & Blachstein, 1993, 1994; Wiens et al., 1988). In some reports, combined scores are expressed as ratio scores between two raw scores (see Forrester & Geffen, 1991), but most frequently these scores are reported in the literature as the difference between two raw scores. Thus in order to reduce redundancy, the combined scores reported in this study were only those calculated as difference scores. The analyses of these results are presented in four sections of memory domains, following the order of test administration: learning, interference, delayed recall and recognition, and temporal order judgment. A summary of the different scores is presented in Table 2. Some of these scores are partially redundant because the same raw scores might be computed in different combinations. However, all scores are reported and analyzed, because comparison of the different measures is of diagnostic value. Such a report could provide the diagnostician with useful information about whether a particular score is in the normal range and whether this score is sensitive to age. A comparison between scores of the same child could indicate whether the performance level is consistent across the different memory domains (i.e., at the normal range, below, or above it) or is inconsistent so that in one memory domain performance is at the normal range but it is above or below it in a different domain.

Norms are reported for the raw scores of the nine trials of the test and for the additional two temporal order measures (i.e., Trial 10). Means and standard deviations for these memory scores are presented in Tables 3a and 3b, for boys and girls, respectively. The performance of parametric statistical tests (e.g., ANOVA) assumes a normal distribution of the variable. Thus, Skewness and Kurtosis was tested for each variable. The underlying assumption of normal distribution was violated in only 2 of the 14 variables (i.e., $-1.96 < Z > +1.96$). For these two variables, nonparametric statistical analyses were applied (i.e., Kruskal-Wallis test followed by multiple Mann-Whitney tests). For the remaining variables, parametric statistical analyses were applied (ANOVA followed by Duncan procedure).

Table 2. Scores Presented in Order of the Test Administration.

Trial	Variable name	Code	Row	Combined	Description
Learning					
T1	Immediate memory	(T1)	*		
T5	Best learning	(T5)	*		
	Total learning	(TL)		*	(Trials 1+2+3+4+5)
	Learning rate	(LRd)		*	(Trials 5-1)
Interference					
T6	List B	(T6)	*		
	Proactive interference	(PId)		*	(Trials 1-6)
T7	List A	(T7)	*		
	Retroactive interference	(RId)		*	(Trials 5-7)
Delayed memory					
T8	Delayed recall	(T8)	*		
	Forgetting rate	(FRd)		*	(Trials 5-8)
T9	Recognition	(T9)	*		
	Retrieval efficiency	(REd)		*	(Trials 9-8)
Temporal order judgment					
T10	Absolute deviations	(AD)	*		(Sum of absolute deviations)
	Correlation score	(CO)	*		(Pearson-PM correlation)

Table 3a. Boys: Means and (SD) of the Raw Memory Scores for Each Age Group.

	Age group (years)									
Trial	8 (*n* = 51)	9 (*n* = 51)	10 (*n* = 56)	11 (*n* = 55)	12 (*n* = 45)	13 (*n* = 42)	14 (*n* = 42)	15 (*n* = 61)	16 (*n* = 47)	17 (*n* = 37)
T1(List A)	6.10	5.69	6.30	6.80	6.80	7.29	7.12	7.46	7.17	6.97
	(1.98)	(1.48)	(1.49)	(1.71)	(1.89)	(1.50)	(1.53)	(2.01)	(1.49)	(1.36)
T2	7.84	8.43	9.29	9.76	9.44	10.00	10.04	10.54	10.02	9.83
	(2.18)	(1.71)	(2.08)	(2.08)	(2.15)	(2.37)	(2.15)	(2.20)	(2.04)	(2.15)
T3	9.04	10.12	10.66	11.54	11.53	11.71	11.59	12.00	11.40	11.43
	(2.61)	(2.16)	(1.78)	(1.80)	(1.73)	(2.03)	(1.97)	(1.77)	(1.71)	(1.86)
T4	10.00	10.82	11.23	12.27	12.20	12.16	12.30	12.42	12.19	12.48
	(2.36)	(2.40)	(1.81)	(1.70)	(1.47)	(1.75)	(1.77)	(1.67)	(1.51)	(1.60)
T5	10.61	11.71	11.78	12.38	12.91	12.28	12.80	12.96	12.63	12.27
	(2.05)	(1.89)	(1.70)	(1.81)	(1.47)	(1.81)	(1.56)	(1.73)	(1.55)	(1.72)
T6(List B)	5.06	5.35	5.59	6.09	6.40	6.02	6.67	6.93	6.32	6.40
	(1.58)	(1.76)	(1.56)	(1.76)	(2.08)	(1.88)	(2.09)	(2.35)	(2.18)	(1.82)
T7(List A)	8.92	9.47	9.89	11.31	11.24	11.17	10.95	11.72	11.34	11.08
	(2.78)	(2.52)	(2.17)	(2.03)	(2.25)	(2.73)	(2.28)	(2.33)	(2.03)	(2.26)
T8(DR)	8.98	9.86	10.20	11.29	11.71	11.48	10.95	11.57	11.34	10.81
	(2.63)	(2.31)	(2.49)	(2.43)	(2.12)	(2.52)	(2.56)	(2.35)	(2.30)	(2.50)
T9(RC)[a]	13.75	13.27	13.84	14.07	14.40	14.24	14.26	13.93	13.98	14.14
	(1.65)	(3.03)	(1.04)	(1.03)	(0.81)	(1.10)	(1.36)	(1.48)	(1.36)	(1.05)
T10(AD)	37.17	32.92	29.71	28.40	26.18	28.86	27.56	28.51	31.37	28.69
	(18.59)	(15.27)	(16.50)	(17.48)	(16.22)	(17.44)	(15.44)	(14.97)	(14.98)	(14.25)
T10(CO)	0.63	0.70	0.77	0.75	0.78	0.75	0.77	0.76	0.73	0.75
	(0.27)	(0.21)	(0.16)	(0.23)	(0.20)	(0.23)	(0.19)	(0.17)	(0.21)	(0.19)

Note. DR = delayed recall; RC = recognition (hit rate); AD = absolute deviation; CO = correlation score.
[a] Because the distribution of this measure was found not to be normal, medians are presented for the age groups in ascending order (15, 14, 15, 14, 13, 13).

Learning

The learning curve of the different age groups, based on their performance in Trials 1 to 5, is analyzed first, because it takes into account all five learning trials. Four additional learning measures, extracted from the five learning trials, are then analyzed and reported in this section.

Learning curve (Trials 1 to 5)

A mixed design ANOVA was conducted to analyze the effect of age group (8 to 17 years), gender, and learning trial (1 to 5). The age group and gender are between-subjects factors and the learning trials are within-subjects factor. The three main effects reached significance, but because the Age group × Learning trial interaction reached significance, interpretation of these two main effects should be made cautiously: Age group, $F(9, 923) = 27.04$, $p < .001$ (as can be seen in Tables 3a and 3b; the older the group, the more words recalled); gender, $F(1, 923) = 17.89$, $p < .001$ (girls recalled more words overall than did boys); and learning trial, $F(4, 3692) = 2610.03$, $p < .001$, indicating that there is an overall increase in number of words recalled from trial to trial. The Age group × Gender interaction did not reach significance, $F(9, 923) = .42$, $p > .05$, indicating that the girls' advantage remained constant across all age groups. Therefore, the results for boys and girls were combined in the following analyses. The only interaction to reach significance was the Learning trial × Age group interaction, $F(36, 3692) = 2.31$, $p < .001$, indicating that age has a differential effect on the learning curve. The exact pattern of this difference will be analyzed later in this section. The source of the age-group effect was detected by the multiple comparison procedures conducted on the 'Total Learning' measure.

Multiple comparisons of age groups on each memory score were conducted using Mann-Whitney tests where a nonparametric statistic

Table 3b. Girls: Means and (SD) of the Raw Memory Scores for Each Age Group.

	Age group (years)									
Trial	8 (*n* = 49)	9 (*n* = 49)	10 (*n* = 54)	11 (*n* = 48)	12 (*n* = 45)	13 (*n* = 41)	14 (*n* = 45)	15 (*n* = 43)	16 (*n* = 43)	17 (*n* = 39)
T1 (ListA)	5.82	5.96	6.85	7.33	7.27	7.10	7.56	8.14	7.56	7.67
	(1.54)	(1.57)	(1.62)	(1.73)	(1.70)	(1.73)	(1.42)	(2.14)	(1.87)	(1.59)
T2	8.04	8.80	9.70	10.44	10.44	10.10	10.53	10.77	10.00	10.72
	(1.98)	(2.25)	(1.10)	(2.07)	(1.73)	(2.07)	(1.96)	(2.23)	(2.30)	(2.13)
T3	9.51	10.45	11.48	11.92	12.13	11.61	11.98	12.53	11.79	12.41
	(2.41)	(2.33)	(2.07)	(2.06)	(1.63)	(1.77)	(1.30)	(2.03)	(1.74)	(1.63)
T4	10.08	11.61	12.00	12.46	12.36	12.17	12.42	13.00	12.16	12.54
	(2.90)	(1.62)	(1.85)	(1.76)	(1.43)	(1.91)	(1.56)	(1.48)	(2.03)	(1.54)
T5	11.08	12.29	12.56	12.90	12.96	12.80	13.13	13.14	13.09	13.03
	(2.53)	(1.59)	(1.66)	(1.74)	(1.35)	(1.49)	(1.29)	(1.68)	(1.23)	(1.44)
T6(List B)	4.69	5.59	6.33	6.19	6.64	6.41	6.51	7.32	6.42	6.54
	(1.79)	(1.50)	(2.11)	(1.54)	(1.54)	(2.17)	(1.42)	(2.92)	(1.99)	(1.80)
T7(List A)	9.18	10.69	11.11	11.54	11.22	11.02	11.89	12.02	11.39	11.87
	(3.29)	(2.13)	(1.93)	(1.98)	(1.96)	(1.85)	(1.67)	(2.56)	(2.17)	(2.04)
T8(DR)	9.45	10.75	11.09	11.69	11.60	11.49	12.24	12.16	11.88	12.10
	(3.00)	(2.22)	(2.36)	(1.99)	(1.90)	(2.27)	(1.77)	(2.40)	(1.90)	(2.11)
T9(RC)[a]	13.59	14.10	14.05	14.67	14.15	14.17	14.00	14.16	14.16	14.20
	(2.45)	(2.30)	(2.18)	(0.59)	(2.36)	(1.11)	(2.32)	(1.53)	(1.02)	(2.46)
T10(AD)	31.23	26.96	27.04	24.65	23.12	30.52	24.66	28.79	30.37	26.21
	(16.10)	(12.67)	(14.06)	(15.10)	(15.76)	(16.27)	(10.20)	(18.00)	(16.76)	(11.64)
T10(CO)	0.72	0.79	0.79	0.80	0.79	0.73	0.81	0.76	0.72	0.79
	(0.20)	(0.17)	(0.18)	(0.21)	(0.21)	(0.22)	(0.12)	(0.22)	(0.23)	(0.15)

Note. DR = delayed recall; RC = recognition (hit rate); AD = absolute deviation; CO = correlation score.
[a] Because the distribution of this measure was found not to be normal, medians are presented for the age groups in ascending order (15, 14, 15, 14, 13, 13).

was required; otherwise Duncan procedure was used. In this section multiple comparisons conducted on four learning measures extracted from the five learning trials are reported. Age groups that were found significantly different from one another on these learning measures by applying multiple comparison procedures are shown in Table 4.

Total learning (TL)
This measure consists of the sum of words recalled in all five learning trials and is one of the most common scores used to reflect learning (Crossen & Wiens, 1988; Moses, 1989; Ryan, Geisser, Randall & Georgemiller, 1986; Ryan et al., 1984; Wolf, Ryan, & Mosnaim, 1983). Analysis of this score showed that the age groups were significantly different, $F(9, 933) = 26.56$, $p < .001$. Ivnik et al. (1992), and Smith et al. (1993) introduced an alternative measure to the total learning score which corrects for baseline differences in the initial trial. This measure is computed as [Total learning-(L1*5)], and it loaded on the 'learning' factor.

Immediate memory (Trial 1) (T1)
This is another score derived from the learning trials that has been used separately in previous factor analyses and found to load on the 'acquisition' factor (Vakil & Blachstein, 1993). This score was found to be sensitive to the effect of age, $F(9, 933) = 13.57$, $p < .001$.

Best learning (Trial 5) (T5)
Analysis of this score revealed a significant age group effect, $F(9, 933) = 14.86$, $p < .001$. The learning curve was analyzed as a continuum, using all five learning trials. An alternative combined score is commonly applied in the literature to reflect the learning rate, either as a difference score of Trials 1 and 5 or as a ratio between these two scores.

Table 4. Age Group Comparisons and Pearson Product-Moment Correlation.

	Learning measure				
Age group (years)[a]	TL	T1	T5	LRd	Total
8 vs.					
9	+	–	+	+	3
10	+	+	+	+	4
11	+	+	+	+	4
12	+	+	+	+	4
13	+	+	+	–	3
14	+	+	+	+	4
15	+	+	+	–	3
16	+	+	+	–	3
17	+	+	+	–	3
9 vs.					
10	+	+	–	–	2
11	+	+	+	–	3
12	+	+	+	–	3
13	+	+	+	+	4
14	+	+	+	–	3
15	+	+	+	+	4
16	+	+	+	+	4
17	+	+	+	+	4
10 vs.					
11	+	–	–	–	1
12	+	–	+	–	2
13	+	+	–	–	2
14	+	+	+	–	3
15	+	+	+	–	3
16	+	+	+	–	3
17	+	+	–	–	2
Total	24	21	20	9	74/96 = .77
Correlation with age for ages (8-10) (n = 310)	.34***	.16**	.27***	.13*	
Correlation with age for ages (11-17) (n = 633)	.06	.09*	.03	–.06	

Note. TL = Total learning (T1+T2+T3+T4+T5); LRd = Learning Rate (T5-T1).
[a] The comparisons between ages 11 to 17 are not presented since only few of the age groups differed significantly.
* = p < .05; ** = p < .01; *** = p < .001.

Learning rate (Trials 5–1) (LRd)
Most frequently, Trials 1 and 5 are presented as difference scores (Mitrushina et al., 1991). In the factor analysis study of the Rey AVLT scores (Vakil & Blachstein, 1993) this score maximally loaded on the 'acquisition' factor. This score was found to be sensitive to age effect, $F(9, 933) = 3.06$, $p < .001$. As can be seen in Table 4, the differences between the 8- to 10-year-old group and all the other age groups were specified, but not those for ages 11 to 17. The reason for not presenting these comparisons is that for the age range of 11 to 17 years, only a few of the age groups differed significantly on just two of the learning measures. The TL score significantly differentiated between the 15- year age group and the 11-, 12-, 13-, and 16- year age groups. The T1 score significantly differentiated between the 15-year age group and the 11-, 12-, and 13-year age groups.

The above conclusion that the 11- to 17-age groups are less differentiated than the 8- to 10-age groups, suggests that the effect of age on most of the Rey AVLT scores is nonlinear. Thus, this may be the source of the Age group × Learning trial interaction found in the analysis of the learning curve on Trials 1 to 5. In order to test this hypothesis, the learning curve of Trials 1 to 5 was analyzed separately for the younger (i.e., 8–10) and the older (i.e., 11–17) age groups. For the younger age groups, main effects of age group, $F(2, 307) = 20.68$, $p < .001$, and learning trials, $F(4, 1228) = 705.64$, $p < .001$, reached significance, as did the Learning trials × Age groups interaction, $F(8, 1228) = 4.75$, $p < .001$. In contrast, for the older age groups learning trials was the only significant effect, $F(4, 2504) = 1986.28$, $p < .001$. The main effect for age group and Age group × Learning trial did not reach significance, $F(6, 626) = 1.72$, $p > .05$, and $F(24, 2504) = 0.84$, $p > .05$, respectively.

Interference

The introduction of the interference list in Trial 6 enables the extraction of several interference measures. In this section the age effect on these different measures is reported.

List B (interference list) (Trial 6)

The analysis of this score revealed a significant age group effect, $F(9, 933) = 10.46$, $p < .001$.

List A (following the interference list) (Trial 7)

An age effect was found to be significant in the analysis of this score as well, $F(9, 933) = 13.35$, $p < .001$. As mentioned above, proactive and retroactive interference scores are derived from the comparison between interference list and first list learning trials scores.

Proactive interference (Trials 1–6) (PId)

In order to test the significance of the proactive interference effect, the effects of age group (8 to 17) and proactive interference (Trial 1 vs. Trial 6) were analyzed. Both main effects, but not the interaction between them, reached significance. The significant age group effect, $F(9, 933) = 17.03$, $p < .001$, indicates the advantage of some age groups over the others on the total number of words recalled, in Trials 1 and 6. The exact source of this effect was detected by follow-up analysis of age group effect for Trials 1 and 6. The significant proactive interference effect, $F(9, 933) = 127.55$, $p < .001$, suggests that significantly less words were recalled in Trial 6 than in Trial 1. The nonsignificant Age group × Proactive interference interaction, $F(9, 933) = 1.23$, $p > .05$, suggests that all age groups showed proactive interference to the same extent. Proactive interference can be expressed as a difference score. Consistent with the analysis above, it was insensitive to age, $F(9, 933) = 1.23$, $p > .05$.

Retroactive interference (Trials 5–7) (RId)

The retroactive interference effect was analyzed by testing the effect of age group (8 to 17) and retroactive interference (Trial 5 vs. Trial 7). Both main effects and the interaction between them reached significance. The significant age group effect, $F(9, 933) = 16.79$, $p < .001$, indicates the advantage of some age groups over the others, on the total number of words recalled, in Trials 5 and 7. The exact source of this effect was detected by follow-up analysis of age group effect for Trials 5 and 7. The significant retroactive interference effect, $F(9, 933) = 659.45$, $p < .001$, suggests that significantly less words were recalled in Trial 7 than in Trial 5. The significant Age group × Retroactive interference interaction, $F(9, 933) = 2.07$, $p < .05$, suggests that the different age groups were affected differently by retroactive interference. As above, retroactive interference can be expressed as a difference score (consistent with the analysis above) and it was found to be sensitive to age, $F(9, 933) = 2.07$, $p < .05$.

As for the learning measures, the older age groups (i.e., 11–17) were hardly distinguishable by the interference measures. The T6 score was the only measure sensitive to age in this age range. The 15-year age group was significantly different on this score from all the other age groups, except for the 14-year age group. Table 5 presents the groups in the 8- to 10-age range that were significantly different on the interference measures, based on multiple comparison procedures.

Delayed Memory

Delayed recall (Trial 8)

In the factor analysis study of the Rey AVLT scores (Vakil & Blachstein, 1993) this scores maximally loaded on the 'retention' factor. The analysis of this score revealed a significant age group effect, $F(9, 933) = 12.26$, $p < .001$. Table 6 presents the results of the follow-up analysis on this effect.

Forgetting rate or Delay effect (Trials 5 vs. 8) (FRd)

In order to test the significance of the delay effect, the effect of age group (8 to 17) and delay (Trial 5 vs. Trial 8) was analyzed. Both main effects, but not the interaction between them, reached significance. The significant age group effect, $F(9, 933) = 15.66$, $p < .001$, indicates the advantage of some age groups over the others on the total number of words recalled, in Trials 5

Table 5. Age Group Comparisons and Pearson Product-Moment Correlation.

	Interference measure				
Age group (years)[a]	T6	T7	PId	RId	Total
8 vs.					
9	+	+	–	–	2
10	+	+	–	–	2
11	+	+	–	+	3
12	+	+	–	–	2
13	+	+	–	–	2
14	+	+	–	–	2
15	+	+	–	+	3
16	+	+	–	–	2
17	+	+	–	–	2
9 vs.					
10	–	–	–	–	0
11	+	+	–	+	3
12	+	+	–	–	2
13	+	+	–	–	2
14	+	+	–	–	2
15	+	+	–	+	3
16	+	+	–	–	2
17	+	+	–	+	3
10 vs.					
11	–	+	–	–	1
12	–	+	–	–	1
13	–	–	–	–	0
14	+	+	–	–	2
15	+	+	–	–	2
16	–	+	–	–	1
17	–	+	–	–	1
Total	18	22	0	5	45/96 = .47
Correlation with age for ages (8-10) ($n = 310$)	.25***	.23***	–.09	–.03	
Correlation with age for ages (11-17) ($n = 633$)	.06	.04	–.01	–.03	

Note. PId = Proactive Interference (T1-T6); RId = Retroactive Interference (T5-T7).
[a]The comparisons between the ages 11 to 17 are not presented since only few of the age groups differed significantly.
* = $p < .05$; ** = $p < .01$; *** = $p < .001$.

and 8. The exact source of this effect was detected by follow-up analysis of age group effect in Trials 5 and 8. The significant delay effect, $F(9, 933) = 557.07$, $p < .001$, suggests that significantly less words are recalled on Trial 8 as compared to Trial 5. The nonsignificant Age group × Delay interaction, $F(9, 933) = 1.46$, $p > .05$, suggests that all age groups were affected similarly by the delay. This was confirmed by analysis of the difference score between Trials 5 and 8 which was not sensitive to age, $F(9, 933) = 1.46$, $p > .05$. When expressed as percent retention, this score loaded on the 'retention' factor (Smith, Ivnik et al., 1992; Smith et al., 1993).

Recognition (Trial 9)

In the factor analysis study of the Rey AVLT scores (Vakil & Blachstein, 1993) when the retention factor was further broken down to storage and retrieval, this score maximally loaded on the 'storage' factor. This score was analyzed using a nonparametric test (i.e., Kruskal-Wallis) because it did not fulfill the normal distribution assumption. Age effect was significant, χ^2 (9, $N = 942$) = 16.90, $p < .05$. A nonparametric signal detection measure is an alternative score used in the literature. This measure (0.5(1+HR-FP)) consists of the proportion of words correctly identified from list A, hit rate (HR), and the proportion of false positive responses (FP) (Geffen et al., 1990).

Retrieval efficiency (Trials 9–8) (REd)

As for the previous score, this score was analyzed using a nonparametric test (i.e., Kruskal-Wallis) because it did not fulfill the normal distribution assumption. The age effect was significant here as well, χ^2 (9, $N = 943$) = 81.11, $p < .001$.

Because the older age groups (i.e., 11–17) were barely distinguishable by the delay measures, Table 6 presents only the significant comparison for the younger age groups (i.e., 8–10). Among the older age groups the REd score was the only score that distinguished between the 15-year age group and the 11-, 12-, and 13-year-old groups.

Temporal Order

Two alternative scores for the supplementary temporal order trial were introduced by Vakil and Blachstein (1994). One score is *Absolute deviation (AD)* calculated as the sum of absolute distances between the rearranged order of the words and their original order. The second score is *Correlation score (CO)*, calculated for each subject, as Pearson product-moment correlation between the rearranged order and the original order (Tzeng, Lee, & Wetzel, 1979). In the factor analysis study of the Rey AVLT scores (Vakil & Blachstein, 1993) when the retention factor was further broken down to storage and retrieval, this score maximally loaded on the 'retrieval' factor. Both measures were sensitive to age: Absolute deviation, $F(9, 931) = 2.92$, $p < .01$, and correlation score, $F(9, 930) = 2.92$, $p < .01$. Consistent with the pattern above, the older age groups (i.e., 11–17) were not found to be significantly distinguishable from each other by using the temporal order measures. Thus Table 7 presents only comparisons for the 8- to 10-year age groups that reached significance by using multiple comparison procedures.

Proportions of Scores' Sensitivity – Summary

Table 8 presents more clearly the sensitivity contribution of the different scores of each memory domain to age. This table gives the summary of the proportional number of measures significantly differentiating between each age group for the above four domains of memory. As can be seen in this table, some age groups are more clearly differentiated than others by a number of measures. There is a consistent trend showing that the three younger groups (i.e., 8–10) are much more distinguishable from each other than the seven older groups (i.e., 11–17). As can be seen in Table 8, 188 out of 336 (55.95%) age comparisons were found to be significant for the three younger age groups, whereas only 17 out of 294 (5.78%) age comparisons were found to be significant for the seven older age groups. Furthermore, with increasing development from age 8 to age 10, the different measures' sensitivity to age gradually decreases. Eight-year-olds have the highest number of significant age comparisons across the different

Table 6. Age Group Comparisons and Pearson Product-Moment Correlation.

	Delayed measure				
Age group (years)[a]	T8	T9	FRd	REd	Total
8 vs.					
9	+	–	–	+	2
10	+	–	–	+	2
11	+	+	–	+	3
12	+	+	–	+	3
13	+	–	–	+	2
14	+	+	–	+	3
15	+	–	–	+	2
16	+	–	–	+	2
17	+	+	–	+	3
9 vs.					
10	–	–	–	–	0
11	+	–	–	+	2
12	+	+	–	+	3
13	+	–	–	+	2
14	+	–	–	+	2
15	+	–	–	+	2
16	+	–	–	+	2
17	+	–	–	+	2
10 vs.					
11	+	+	–	+	3
12	+	+	–	+	3
13	+	–	–	+	2
14	+	–	–	+	2
15	+	–	–	+	2
16	+	–	–	+	2
17	+	–	–	+	2
Total	23	7	0	23	53/96 = .55
Correlation with age for ages (8-10) (n = 310)	.22***	.05	.02	–.15**	
Correlation with age for ages (11-17) (n = 633)	.01	–.06	–.01	–.08*	

Note. FRd = Forgetting Rate (T5-T8); REd = Retrieval efficiency (T9-T8).
[a]The comparisons between the ages 11 to 17 are not presented since only few of the age groups differed significantly.
* = p < .05; ** = p < .01; *** = p < .001.

categories of memory (ranging from 8 of 14 possible comparisons, to 12 of 14). For nine-year-old children, there is a lower number of significant comparisons (ranging from 2 of 14 possible comparisons to 9 of 14). Ten-year-olds have the lowest number of significant comparisons (ranging from 6 of 14 possible comparisons to 7 of 14).

Pearson product-moment correlations were conducted between age and the different memory measures. This is an alternative method for assessing how these memory measures are related to age. The above conclusion that the older age groups (i.e., 11–17) are less distinguishable than the younger age groups (i.e., 8–10), suggests that the effect of age on most of the Rey AVLT scores is nonlinear. In order not to violate

Table 7. Age Group Comparisons and Pearson Product-Moment Correlation.

	Temporal order measure		
Age group (years)[a]	AD	CO	Total
8 vs.			
9	–	+	1
10	+	+	2
11	+	+	2
12	+	+	2
13	–	+	1
14	+	+	2
15	+	+	2
16	–	+	1
17	+	+	2
9 vs.			
10	–	–	0
11	–	–	0
12	+	–	1
13	–	–	0
14	–	–	0
15	–	–	0
16	–	–	0
17	–	–	0
10 vs.			
11	–	–	0
12	–	–	0
13	–	–	0
14	–	–	0
15	–	–	0
16	–	–	0
17	–	–	0
Total	7	9	16/48 = .33
Correlation with age for ages (8-10) (n = 310)	–.15**	.20***	
Correlation with age for ages (11-17) (n = 633)	.07	–.04	

Note. AD = absolute deviation; CO = correlation score.
[a]The comparisons between the ages 11 to 17 are not presented since only few of the age groups differed significantly.
* = $p < .05$; ** = $p < .01$; *** = $p < .001$.

the linearity assumption, a separate set of correlations was calculated for the younger and the older age groups. These results are reported in Tables 4 to 7, along with the respective results of the multiple comparison procedures. Generally, the results obtained by both methods of assessing age effect (i.e., comparison of age group and correlation with age) were consistent to each other. That is, when the correlation of a particular memory score with age was lower, fewer groups were distinguishable by that score and vice versa.

Finally, in order to detect the memory domain most affected by age, we compared the most sensitive scores of each memory domain to each other. The most sensitive score was preferred rather than the overall sensitivity of all the scores at each domain, because the latter can be

Table 8. Proportional Number of Significantly Different Measures.

	Memory domain				
Age group (years)[a]	Learning	Interference	Delayed	Temporal	Total
(8–10) vs. (9–17)					
8 vs.					
9	3/4	2/4	2/4	1/2	8/14
10	4/4	2/4	2/4	2/2	10/14
11	4/4	3/4	3/4	2/2	12/14
12	4/4	2/4	3/4	2/2	11/14
13	3/4	2/4	2/4	1/2	8/14
14	4/4	2/4	3/4	2/2	11/14
15	3/4	3/4	2/4	2/2	10/14
16	3/4	2/4	2/4	1/2	8/14
17	3/4	2/4	3/4	2/2	10/14
9 vs.					
10	2/4	0/4	0/4	0/2	2/14
11	3/4	3/4	2/4	0/2	8/14
12	3/4	2/4	3/4	1/2	9/14
13	4/4	2/4	2/4	0/2	8/14
14	3/4	2/4	2/4	0/2	7/14
15	4/4	3/4	2/4	0/2	9/14
16	4/4	2/4	2/4	0/2	8/14
17	4/4	3/4	2/4	0/2	9/14
10 vs.					
11	1/4	1/4	3/4	0/2	5/14
12	2/4	1/4	3/4	0/2	6/14
13	2/4	0/4	2/4	0/2	4/14
14	3/4	2/4	2/4	0/2	7/14
15	3/4	2/4	2/4	0/2	7/14
16	3/4	1/4	2/4	0/2	6/14
17	2/4	1/4	2/4	0/2	5/14
Total	74/96	45/96	53/96	16/48	188/336
(11–17) vs. (12–17)					
Total	8/168	5/168	3/196	1/84	17/294

Note. [a]The comparisons between the ages 11 to 17 are not presented since only few of the age groups differed significantly.

biased by the amount of redundancy among the scores chosen. The total possible comparisons between the ages 8 to 10 years is 24. As can be seen in Table 8, the most sensitive scores in the different memory domains are: Learning-Total learning (24/24 = 100%, correlation with age, $r = .34$, $p < .001$); Interference-Trial 7 (22/24 = 92%, correlation with age, $r = .23$, $p < .001$); Delay-Trial 8 (23/24 = 96%, correlation with age, $r = .22$, $p < .001$); Temporal order-Correlation score (9/24 = 38%, correlation with age, $r = .20$, $p < .001$).

DISCUSSION

In this study 943 children aged 8 to 17 years were evaluated on 14 different memory scores extracted from the Rey AVLT. The uniqueness of this study is not only in its large sample size but in the approach that we took in analyzing the data. In this study we systematically conducted follow-up analyses comparing every age group to all the other age groups, on all memory scores and a significant age group effect was demonstrated. These multiple comparisons enabled us to detect which score reliably distinguishes between which specific age groups. These analyses revealed a consistent finding across the different scores, that is, in the 8- to 17-year range the age effect on memory measures is not linear. More specifically, the younger age groups (i.e., 8–10 years) are much more distinguishable from each other than the older age groups (i.e., 11–17 years). This conclusion is based on multiple comparisons of the groups and correlations of the different scores with age. Forrester and Geffen (1991) also found that two scores (i.e., TL & REd) were sensitive to age when the younger age group (i.e., 7–8 years old) and the older groups (i.e., 9–15 years old) were compared. This finding suggests that whatever capabilities are required to cope optimally with the different demands of the Rey AVLT, such as storage capacity (Pascual-Leone, 1970) or strategies (Simon, 1974), these are stabilized at about the age of 11.

The discontinuity in the development of memory (i.e., ages 8–10 vs. ages 11–17) supports other studies in the literature. Several studies demonstrate that younger and older children differ in their ability to utilize an efficient strategy in relation to mnemonic task demands (for review see Lange, 1978). For example, Paris (1978) concludes that ‘Until the age of 7 or 8, children do not ordinarily elaborate and transform stimuli that are to be recalled later. Older children, 11, or 12 years of age, begin to rearrange items and construct additional relationships spontaneously, as adults commonly do’ (p. 153). The mental operations developed by the age of 11, such as utilization of strategy, planning, and categorization are attributed to frontal lobe functioning (Shimamura, 1995). Some researchers interpreted such results as an indication of the maturation of the frontal lobes by this age (Passler, Isaac, & Hynd, 1985; Welsh & Pennington, 1988). More direct evidence to the involvement of the frontal lobe in such a task as the Rey AVLT can be found in a recent PET study by Grasby et al. (1993). In this study when recall of a superspan list of 15 words (similar to the Rey AVLT) was compared to recall of a subspan list of five words, the maximal increases of rCBF were located in the prefrontal cortex bilaterally. In light of this argument it is logical to conclude that those scores that were not sensitive to age at all (i.e., proactive interference and forgetting rate) require minimum elaboration or utilization strategy as compared to the other scores. This point could be more conclusively investigated by testing patients with frontal lobe injury with the Rey AVLT.

In a similar study with older adults (Vakil & Blachstein, 1997) with the same analysis approach, two distinct segments in the older adults’ age range were detected as well. The group aged 20 to 59 years was less distinguishable than the 60- to 91-year age range. There is growing evidence suggesting that memory decline associated with age results from neural deterioration, not only of the temporal, but of the frontal lobes as well (Moscovitch & Winocur, 1992). Thus, our findings that the Rey AVLT is most sensitive to developmental changes up to 11 years old in children and from 60 years old in older adults, may reflect the maturation and the deterioration of frontal lobe functioning, respectively.

The comparison of the most sensitive score of each memory domain revealed that learning, interference, and delay are about equally affected by age. It is important to note that the most sensitive scores at each domain are not the scores found in factor analysis (Vakil & Blachstein, 1993) that are maximally loaded in the different factors. It should be remembered that the factor analysis was conducted on adults and quite possibly if conducted on children it might have produced different results. The consistency in the three memory domains (i.e., learning, interference, and delay) suggest that the normal developmental changes are parallel in these domains. This is very useful information for the diagnostician who finds inconsistency in performance among these domains. Such a case could be interpreted as an indication of a pathology.

Unlike the findings with adults (Vakil & Blachstein, 1997), in the present study temporal order, was found less sensitive to age. This finding can be interpreted as supporting Hasher and Zacks’ (1979) claim that temporal order is an ‘innate automatic’ process. Insensitivity to age is one of Hasher and Zacks’ criteria for an innate automatic task. However, the findings with adults are at odds with this claim. Thus, the age criterion is supported by the developmental component but not by the aging component. An alternative theoretical framework to interpret the dissociation between temporal order and the other memory scores is in terms of the distinction between 'item' (i.e., words) and 'source' (i.e., temporal order; Schacter, Harbluk, & McLachlan, 1984). This distinction is based on findings demonstrating a dissociation between amnesic and frontal lobe patients in which the latter suffer from amnesia to 'source' information but not to 'item' information. This would lead to a contradicting conclusion to the one presented above indicating that the components of the task which are more dependent on frontal lobe functioning will be the most sensitive to age. Further research is required with frontal lobe patients to resolve this issue.

Six raw scores and five combined scores (excluding the total learning score which is not a difference score and the two temporal order measures which do not have parallel combined scores) were analyzed in this study. A very telling finding is that the raw scores significantly distinguished in 77% of the 168 possible comparisons at the age range of 8–10 years, whereas the combined scores distinguished only 31% of the 120 possible comparisons. This suggests that at least for the memory changes in normal development despite the changes in the raw scores the relationships between scores (i.e., combined scores) remain relatively more stable. Thus, when changes in the relationships between scores are detected they could have a very important diagnostic value, indicating a possible abnormal memory changes.

In the present study, consistent with most studies of adults, girls' performance was superior to that of boys (Bleeker et al., 1988; Geffen et al., 1990; Vakil & Blachstein, 1997; but see Savage & Drew Gouvier, 1992 for different results). However, the two previous reports of children's Rey AVLT norms did not find such an effect (Bishop et al., 1990; Forrester & Geffen, 1991). A possible reason for the nonsignificant gender effect in the previous children's studies is the lack of statistical power in the analyses due to a small sample size. This possibility is further supported when the results of the present study are compared to those of Forrester and Geffen. In all cases where the two studies reached different results, the findings were such that they reached significance in the present study but not in that of Forrester and Geffen. This consistent pattern raises a strong possibility that the very large sample in the present study provides the sufficient statistical power which enables the detection of effects otherwise not identified in other studies. This study further stresses the importance of utilizing a large sample in order to exploit the full potential of the Rey AVLT for detection of the developmental effect.

Finally, the detailed analyses of sensitivity to age of the different measures extracted from the Rey AVLT provide very useful diagnostic information. For example, it could direct the diagnostician to the most sensitive set of scores for a particular age group. The results presented indicate whether a particular score falls in the normal range for a particular age, in a specic memory domain. This information is very valuable when planning intervention, such as mnemonic strategies when encoding is inefficient, or in order to offer the optimal learning conditions to a child found to be susceptible to interference.

REFERENCES

Barrickman, L.L., Perry, P.J., Allen, A.J., & Kuperman, S. (1995). Burotropion versus methylphenidate in the treatment of attention deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 649-657.

Bishop, J., Knights, R.M., & Stoddart, C. (1990). Rey Auditory Verbal Learning Test: Performance of English and French children aged 5 to 16. *The Clinical Neuropsychologist, 4*, 133-140.

Bleeker, M.L., Bolla-Wilson, K., Agnew, J., & Meyers, D.A. (1988). Age-related sex differences in verbal memory. *Journal of Clinical Psychology, 44*, 403-411.

Crossen, J.R., & Wiens, A.N. (1988). Residual neuropsychological deficits following head injury on the Wechsler Memory Scale-Revised. *The Clinical Neuropsychologist, 2*, 393-399.

Forrester, G., & Geffen, G. (1991). Performance measures of 7- to 15-year old children on the Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, *5*, 345-359.

Grasby, P.M., Frith, C.D., Friston, K.J., Bench, C., Frackowiak, R.S.J., & Dolan, R.J. (1993). Implicated in auditory-verbal memory function. *Brain*, *116*, 1-20.

Geffen, G., Moar, K.J., O'Hanlon, A.P., Clark, C.R., & Geffen, L.B. (1990). Performance measures of 16- to 86-year-old males and females on the Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, *4*, 45-63.

Hasher, L., & Zacks, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356-388.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1992). Mayo's older Americans normative studies: Update AVLT norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6*, 83-104.

Ivnik, R.J., Malec, J.F., Tangalos, E.G., & Petersen, R.C. (1990). The Auditory Verbal Learning Test (AVLT): Norms for ages 55 years and older. *Psychological Assessment*, *2*, 304-312.

Lange, G. (1978). Organization-related processes in children's recall. In P.A. Ornstein (Eds.), *Memory development in children* (pp. 101-128). Hillsdale, NJ: Lawrence Erlbaum.

Lezak, M.D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Mitrushina, M., Satz, P., Chervinsky A., & D'Elia, L. (1991). Performance of four age groups of normal elderly on the Rey Auditory-Verbal learning Test. *Journal of Clinical Psychology*, *47*, 351-357.

Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp.315-372). Hillsdale, NJ: Lawrence Erlbaum.

Moses, J.A. (1989). Replicated factor structure of Benton's tests of visual retention, visual construction and visual form discrimination. *International Journal of Clinical Neuropsychology*, *11*, 30-37.

Paris, S.G. (1978). The development of interference and transformation as memory operations. In P.A. Ornstein (Eds.), *Memory development in children* (pp. 129-156). Hillsdale, NJ: Lawrence Erlbaum.

Pascual-Leone, J. (1970). Mathematical model for the transition rule in Piaget's developmental strategies. *Acta Psicologica*, *32*, 301-345.

Passler, M.A., Isaac, W., & Hynd, G.W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology*, *1*, 349-370.

Powell, J.B., Cripe, L.I., & Dodrill, C.B. (1991). Assessment of brain impairment with the Rey Auditory Verbal Learning Test: A comparison with other neuropsychological measures. *Archives of Clinical Neuropsychology*, *6*, 241-249.

Query, W.T., & Berger, R.A. (1980). AVLT memory scores as a function of age among general medical, neurologic and alcoholic patients. *Journal of Clinical Psychology*, *36*, 1009-1012.

Query, W.T., & Megran, J. (1983). Age-related norms for AVLT in a male patient population. *Journal of Clinical Psychology*, *39*, 136-138.

Ryan, J.J., Geisser, M.E., Randall, D.E., & Georgemiller, R.J. (1986). Alternate form reliability and equivalency of the Rey Auditory Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology*, *8*, 611-616.

Ryan, J.J., Rosenberg, S.J., & Mittenberg, W. (1984). Factor analysis of the Rey Auditory-Verbal Learning Test. *The International Journal of Clinical Neuropsychology*, *6*, 239-341.

Savage, R.M., & Drew Gouvier, W. (1992). Rey Auditory-Verbal Learning Test: The effect of age and gender, and norms for delayed recall and story recognition trials. *Archives of Clinical Neuropsychology*, *7*, 407-414.

Schacter, D.L., Harbluk, J.L., & McLachlan, D.R. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior*, *23*, 593-611.

Shimamura, A.P. (1995). Memory and frontal lobe function. In M.S. Gazzaniga (Ed.), *The cognitive neuroscience* (pp.803-813) Cambridge, MA: MIT Press.

Simon, H.A. (1974). How big is a chunk? *Science*, *183*, 482-488.

Smith, G.E., Ivnik, R.J., Malec, J.F., Kokmen, E., Tangalos, E.G., & Kurland, L.T. (1992). Mayo's Older Americans Normative Sample (MOANS): Factor structure of a core battery. *Psychological Assessement*, *4*, 382-390.

Smith, G.E., Ivnik, R.J., Malec, J.F., & Tangalos, E.G. (1993). Factor structure of the Mayo Older Americans Normative Sample (MOANS) core battery: Replication in clinical sample. *Psychological Assessement*, *5*, 121-124.

Smith, G.E., Malec, J.F., & Ivnik, R.J. (1992). Validity of the construct of nonverbal memory: A factor analytic study in a normal elderly sample. *Journal of Clinical and Experimental Neuropsychology*, *4*, 211-221.

Talley, J.L. (1986). Memory in learning-disabled children: Digit span and the Rey Auditory Verbal Learning Test. *Archives of Clinical Neuropsychology*, *1*, 315-322.

Tzeng, O.J.L., Lee, A.T., & Wetzel, C.D. (1979). Temporal coding in verbal information processing. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 52-64.

Vakil, E., & Blachstein, H. (1993). Rey Auditory-Verbal Learning Test: Structure analysis. *Journal of Clinical Psychology*, *49*, 883-890.

Vakil, E., & Blachstein, H. (1994). A supplementary measure in the Rey AVLT for assessing incidental learning of temporal order. *Journal of Clinical Psychology*, *50*, 240-245.

Vakil, E., & Blachstein, H. (1997). Rey AVLT: Developmental norms for adults and the sensitivity of different memory measures to age. *The Clinical Neuropsychologist*, *11*, 356-369.

Vakil, E., Blachstein, H., & Hoofien, D. (1991). Automatic temporal order judgment: The effect of intentionality of retrieval on closed-head-injured patients. *Journal of Clinical and Experimental Neuropsychology*, *13*, 291-298.

Welsh, M.C., & Pennington, B.F. (1988). Assessing frontal lobe functionig in children: Views from developmental psychology. *Developmental Neuropsychology*, *4*, 199-230.

Wiens, A.N., McMinn, M.R., & Crossen, J.R. (1988). Rey Auditory-Verbal Learning Test: Development of norms for healthy young adults. *The Clinical Neuropsychologist*, *2*, 67-87.

Wolf, M.E., Ryan. J.J., & Mosnaim, A.D. (1983). Cognitive functions in tardive dyskinesia. *Psychological Medicine*, *13*, 671-674.