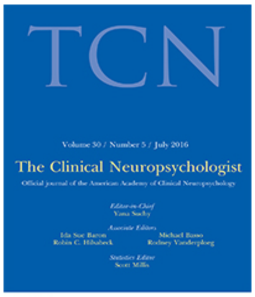

The Clinical Neuropsychologist

ISSN: 1385-4046 (Print) 1744-4144 (Online) Journal homepage: https://www.tandfonline.com/loi/ntcn20

Rey AVLT: Developmental norms for adults and the sensitivity of different memory measures to age

Eli Vakil & Haya Blachstein

To cite this article: Eli Vakil & Haya Blachstein (1997) Rey AVLT: Developmental norms for adults and the sensitivity of different memory measures to age, The Clinical Neuropsychologist, 11:4, 356-369, DOI: 10.1080/13854049708400464

To link to this article: https://doi.org/10.1080/13854049708400464

Published online: 08 Nov 2007.

Submit your article to this journal

Article views: 506

View related articles

Citing articles: 15 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ntcn20

The Clinical Neuropsychologist
1997, Vol. 11, No. 4, pp. 356-369
1385-4046/97/1104-356$12.00
© Swets & Zeitlinger

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

Eli Vakil and Haya Blachstein
Bar-Ilan University, Ramat-Gan, Israel

ABSTRACT

The Rey Auditory Verbal Learning Test (AVLT) has been found to be differentially affected by age, intelligence, and population type. However, previous studies did not specifically report which scores differentiate between age groups. Five hundred and twenty-eight males and females, ranging in age from 21 to 91 years, were administered the Hebrew version of the Rey AVLT. Resulting norms are, therefore, based upon a very large sample of males and females. In addition, findings provide a detailed description of the differential sensitivity of 22 extracted scores on six sequential age groups. Some of these scores were found to be sensitive to age, whereas others were not. The results also show a significant and consistent advantage for females over males on most of the verbal memory measures. The pattern of verbal memory change until the age of 60 years is moderate, as compared to the changes observed from the age of 60 onward. This type of analysis contributes to both the validity of the Rey AVLT, and to its usefulness as a diagnostic tool.

The Rey Auditory Verbal Learning Test (AVLT) is widely used for clinical and research purposes (Lezak, 1983). One of the reasons for its popularity is that many memory measures can be extracted from it, such as immediate and delayed recall, learning rate, recognition, proactive and retroactive interference, and primacy and recency effects (Lezak, 1983; Query & Megran, 1983; Ryan, Rosenberg, & Mittenberg, 1984; Wiens, Mcminn, & Crossen, 1988). The fact that a number of memory components are measured enhances the test's sensitivity as a diagnostic tool. The Rey AVLT has been found to be differentially affected by age, intelligence, and population type (Query & Berger, 1980; Query & Megran, 1983; Wiens et al., 1988).

When submitted to factor analysis, these different memory measures were found to represent several basic memory processes. Ryan, Rosenberg, and Mittenberg (1984) found that three of the Rey AVLT scores, together with the scores of the Paired Associates and Logical Memory subtests of the Wechsler Memory Scale (WMS), reflect a "verbal learning and memory" factor. Moses (1989) grouped three of the Rey AVLT scores with Benton's Visual Retention Test (BVRT) 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), WMS-R, and two scores from the Rey AVLT to factor analysis. A five-factor model of their battery was supported. One of the Rey AVLT scores loaded on the "learning" factor and the other on the "retention" factor. In a more recent study by Vakil and Blachstein (1993), factor analyses 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.

Until recently, one of the major limitations of the Rey AVLT was the lack of adequate norms. This was due to language and procedure changes

* Address correspondence to: Eli Vakil, Psychology Department, Bar-Ilan University, Ramat-Gan, 52900, Israel.
Accepted for publication: January 14, 1997.

from the original version (Rey, 1964), or to sampling problems such as using patient populations, excluding females or small samples (Geffen, Moar, O'Hanlon, Clark, & Geffen, 1990; Ivnik et al., 1992). Furthermore, in many studies, very few of the Rey AVLT scores were analyzed and reported (Geffen et al., 1990). For those studies that did refer to a number of Rey AVLT measures, the test itself was most often found to be sensitive to age (Bleeker, Bolla-Wilson, Agnew, & Mayers, 1988; Geffen et al., 1990; Ivnik, Malec, Tangalos, & Petersen, 1990; Mitrushina, Satz, Chervinsky, & D'Elia, 1991; Smith, Malec, & Ivnik, 1992) and gender (Bleeker et al., 1988; Geffen et al., 1990; Smith et al., 1992; but see Savage & Drew Gouvier, 1992 for different results).

However, the effect of age was not uniformly found for all the memory measures extracted from the Rey AVLT. For example, some measures, such as recognition, were found to be sensitive to age by some researchers (Bleeker et al., 1988; Geffen et al., 1990; Mitrushina et al., 1991; Rey, 1964) but not by others (Schonfield & Robertson, 1966; Wiens et al., 1988). Acquisition rate in the learning trials (Geffen et al., 1990; Mitrushina et al., 1991; Wiens et al., 1988), and proactive interference (Geffen et al., 1990) were also found to be sensitive to age. However, in these studies, degree of sensitivity to age was not found uniformly across all age groups.

Age-related memory changes have been reported on a variety of measures, including free recall, cued recall, and recognition of verbal and nonverbal material (Burke & Light, 1981; Light, 1991; Poon, 1985). However, the specific memory mechanism primarily affected by age is still a controversial issue (for review see Kausler, 1994; Light, 1991). Although some investigators suggest an impairment of the encoding processes (Erber, 1974; Gordon & Clark, 1974), others suggest that it is the retrieval mechanism that is primarily affected by age (Mitrushina et al., 1991). As mentioned above, previous studies have demonstrated that different scores of the Rey AVLT reflect different memory processes (Mitrushina et al., 1991; Vakil & Blachstein, 1993). In light of this, the Rey AVLT seems to be an ideal tool to aid in the clarification of the effect of age on the various memory processes. Most studies that have reported norms only analyzed the overall effect of age on different memory scores; however, they did not report which scores specifically differentiate between age groups.

Thus, one of the goals of the present study was to compare groups with each other on different memory scores extracted from the test, raw as well as combined, or summary scores. This will help to identify the measures most sensitive to change for each age group; it is possible that different aspects of memory are more sensitive to age at different segments of the age continuum. With regard to combined scores, we also wanted to test the claim by Geffen et al. (1990) that ratio scores are more sensitive to level of performance than are difference scores. The final purpose was to publish norms for the Hebrew version of the Rey AVLT. Using the same Hebrew version of the Rey AVLT, Vakil and Blachstein (1994) suggested executing an additional trial of the Rey AVLT in order to test memory for temporal order. Norms of this score are analyzed and reported as well. Several studies have been reported on the basis of this Hebrew version of the Rey AVLT (Blachstein, Vakil, & Hoofien, 1993; Vakil, Blachstein, & Hoofien, 1991; Vakil & Blachstein, 1993, 1994; Vakil, Hoofien, & Blachstein, 1992). The convergence of findings of the test administered in a different language and in a different culture contributes to the reliability and validity of the test.

METHOD

Participants

Five hundred and twenty-eight younger and older adults participated voluntarily in this study. The age of the sample population ranged from 21 to 91 years. The participants were divided into six groups, representing every decade, with the exception of the oldest group, which included participants between the ages of 70 to 91 years. The younger participants were volunteers who responded to advertisements placed at Bar-Ilan University (Israel) and other public places. The older participants were recruited either from among students attending a special series of courses for elderly people offered at Bar-Ilan university, or from several senior citizen community centers, serving the population in the central region of Israel. The latter were referred by social workers, who judged them to be active and independent, cooperative and communicative. All of the elderly participants were alert and oriented to time and place when tested. None of the participants, based on their report, had a history of alcohol, drug abuse, or psychiatric illness. All participants met the criteria of living in Israel for at least 10 years, and spoke Hebrew fluently. In fact, the vast majority of the participants lived in the country much longer than 10 years. Table 1 provides a more detailed description of the participants in all age groups.

Test and Procedure

The Hebrew version of the Rey AVLT was used (Vakil & Blachstein, 1993). Administration was standardized, as described by Lezak (1983). Participants were tested individually. The test consists of 15 common nouns, which were read to the subject, 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 et al., 1991; Vakil & Blachstein, 1994). 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 rewrite the words in their original order.

Table 1. Demographic Characteristics of the Sample.

Age group (years)	Gender	Age (years)	Education (years)	*n*
20–29	Male	24.74	13.39	57
	Female	24.27	13.85	60
	Both	24.50	13.62	117
30–39	Male	32.82	13.95	39
	Female	33.71	14.75	24
	Both	33.16	14.25	63
40–49	Male	44.74	12.93	42
	Female	45.09	13.89	44
	Both	44.92	13.42	86
50–59	Male	53.89	13.11	27
	Female	53.25	13.71	28
	Both	53.56	13.42	55
60–69	Male	64.77	12.44	48
	Female	64.21	12.39	67
	Both	64.44	12.41	115
70–91	Male	76.02	13.11	44
	Female	77.19	11.79	48
	Both	76.63	12.42	92
Total				528

RESULTS

Results are presented in four sections. Each section represents a different category of memory. The following is the order of test administration: learning, interference, delayed recall and recognition, and temporal order judgment. Norms are reported for the raw scores of the nine trials of the test and for an additional three scores reflecting Trial 10. Means and standard deviations for these memory scores are presented in Tables 2a and 2b, for males and females respectively, for each age group.

Twenty-two different scores were derived from the Rey AVLT (i.e., either raw or summary scores) for further analyses. These scores are frequently applied in the literature to reflect different aspects of memory (Geffen et al., 1990; Ivnik et al., 1992; Query & Megran, 1983; Ryan et al., 1984; Vakil & Blachstein, 1993, 1994; Wiens, McMinn, & Crossen, 1988). A number of the above-mentioned scores are summary scores, and they are expressed either as difference or ratio scores, in order to compare their sensitivity to age. At times, there is redundancy because the same scores might be analyzed in different combinations; however, the analysis of each measure and the comparison of the different measures is of diagnostic value.

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. In only seven variables out of the 22, the underlying assumption of

Table 2a. Males: Means and Standard Deviations of the Raw Memory Scores for Each Age Group.

	Age group (years)					
Trial	20–29 (*n* = 57)	30–39 (*n* = 39)	40–49 (*n* = 42)	50–59 (*n* = 27)	60–69 (*n* = 48)	70–91 (*n* = 44)
T1 (ListA)	7.72	7.28	7.38	6.52	6.13	5.30
	(1.51)	(2.16)	(2.41)	(1.31)	(1.88)	(1.37)
T2	10.37	9.87	9.90	9.22	8.29	7.00
	(1.99)	(2.44)	(2.54)	(1.80)	(2.10)	(1.73)
T3	12.05	11.69	11.71	10.70	10.04	8.50
	(1.85)	(2.14)	(2.14)	(1.66)	(2.01)	(1.92)
T4	13.16	12.59	12.57	11.96	10.96	9.84
	(1.44)	(1.76)	(2.00)	(1.53)	(2.26)	(1.93)
T5	13.39	13.08	12.71	12.41	11.38	10.84
	(1.33)	(1.71)	(1.73)	(1.95)	(2.35)	(2.10)
T6 (List B)	7.35	6.67	6.40	5.56	5.50	4.36
	(2.59)	(2.23)	(2.31)	(2.10)	(1.25)	(1.42)
T7 (List A)	12.19	11.49	11.52	10.48	9.21	7.70
	(1.99)	(3.04)	(2.10)	(2.71)	(3.08)	(2.64)
T8 (DR)	12.12	11.67	11.33	10.52	9.13	8.14
	(2.24)	(2.66)	(2.47)	(3.08)	(3.32)	(2.62)
T9 (RC)[a]	14.21	14.00	14.19	13.74	12.92	12.25
	(1.08)	(1.12)	(1.11)	(1.46)	(1.81)	(2.63)
T10 (AD)	26.32	27.00	34.33	35.15	39.94	45.73
	(15.62)	(17.49)	(17.87)	(15.19)	(14.34)	(14.47)
T10 (CO)	0.79	0.77	0.69	0.66	0.60	0.53
	(0.18)	(0.22)	(0.25)	(0.20)	(0.20)	(0.23)
T10 (Hits)	6.16	5.82	4.50	4.22	3.56	3.16
	(3.69)	(3.28)	(3.01)	(3.08)	(2.78)	(1.74)

Note. DR = delayed recall; RC = recognition (hit rate); AD = absolute deviation; CO = correlation score; SD between parentheses.

[a] Since 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).

Table 2b. Females: Means and Standard Deviation of the Raw Memory Scores for Each Age Group.

	Age group (years)					
Trial	20–29 (n = 60)	30–39 (n = 24)	40–49 (n = 44)	50–59 (n = 28)	60–69 (n = 67)	70–91 (n = 48)
T1 (ListA)	7.88	7.58	7.64	7.18	6.46	5.25
	(2.04)	(1.53)	(2.39)	(1.70)	(2.04)	(2.30)
T2	11.10	10.83	10.70	10.29	9.12	7.46
	(1.87)	(1.76)	(2.24)	(1.86)	(2.21)	(2.14)
T3	12.77	12.42	11.45	12.14	10.52	9.06
	(1.70)	(1.53)	(2.17)	(1.41)	(2.27)	(2.36)
T4	13.58	13.33	12.55	12.93	11.78	9.83
	(1.28)	(1.43)	(1.70)	(1.39)	(2.17)	(2.18)
T5	14.15	13.88	13.07	13.32	12.33	10.69
	(1.02)	(1.51)	(1.55)	(1.44)	(1.85)	(2.07)
T6 (List B)	7.28	7.00	6.32	6.64	5.16	4.56
	(2.48)	(1.79)	(1.94)	(2.13)	(1.94)	(1.75)
T7 (List A)	12.88	13.13	11.59	11.93	10.18	8.15
	(1.98)	(1.87)	(2.51)	(2.39)	(2.81)	(3.00)
T8 (DR)	12.87	12.75	11.70	12.11	10.72	8.08
	(2.45)	(2.25)	(2.33)	(2.23)	(2.81)	(2.73)
T9 (RC)[a]	13.35	14.79	14.05	14.32	12.97	12.10
	(0.92)	(0.51)	(1.12)	(1.28)	(2.01)	(2.17)
T10 (AD)	22.17	23.67	34.68	31.29	40.30	44.46
	(16.04)	(14.80)	(17.88)	(10.98)	(18.79)	(14.84)
T10 (CO)	0.83	0.82	0.67	0.73	0.61	0.55
	(0.19)	(0.19)	(0.26)	(0.15)	(0.25)	(0.23)
T10 (Hits)	7.05	6.08	4.55	4.82	3.85	3.48
	(3.60)	(3.15)	(3.24)	(2.42)	(2.78)	(2.09)

Note. DR = delayed recall; RC = recognition (hit rate); AD = absolute deviation; CO = correlation score; SD between parentheses.

[a] Since the distribution of this measure was found not to be normal, medians are presented for the age groups in ascending order (15, 15, 14, 15, 14, 13).

normal distribution was violated (i.e., $-1.96 < Z > +1.96$). For these seven variables non-parametric tests were applied (i.e., Kruskal-Wallis test followed up by multiple Mann-Whitney tests). For the remaining variables parametric tests were applied (ANOVA followed up by Duncan procedure).

Learning

In this section, the different learning measures extracted from the Rey AVLT are analyzed. These measures involve the recall scores of the first five learning trials expressed as: learning curve, immediate memory (Trial 1), and best learning (Trial 5). Two summary scores, reflecting other aspects of learning, are also analyzed: learning rate and total learning. Learning rate (Trial 5 score minus Trial 1 score) represents the learning ability of the participant (Mitrushina et al., 1991; Query, Randy, & Berger, 1980). Total learning (sum of the scores of Trials 1 to 5), represents the capacity to recall and accumulate words across learning trials (Crossen & Wiens, 1988; Moses, 1989; Ryan, Geisser, Randall, & Georgemiller, 1986; Ryan et al., 1984; Wolf, Ryan, & Mosnaim, 1983).

Learning curve (Trials 1 to 5)

A mixed design ANOVA was conducted to analyze the effect of age group (1 to 6), gender, and learning trials (Trials 1 to 5); age group and gender are between-subjects factors and learning trials a within-subjects factor. As can be seen in Tables 2a and 2b, the three main effects reached significance: Age group, $F(5, 516) = 55.09$, $p < .001$, that is, the younger the group the more words learned; gender, $F(1, 516) = 15.13$, $p < .001$, overall, females recalled more words than males; and learning effect, $F(4, 2064) = 1646.59$, $p < .001$, 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, $F(5, 516) = .76$, $p > .05$, indicating that the females' advantage remained constant across all age groups. The Learning Trials × Age Group interaction reached significance, $F(20, 2064) = 2.41$, $p < .001$, suggesting that the younger the group the steeper the learning rate. The learning rate of the females is steeper than that of the males, $F(4, 2064) = 2.87$, $p < .03$. Because gender did not interact with age group, the results of males and females were collapsed in the following analyses. In order to detect the source of the age-group effect, multiple comparison procedures were conducted. Mann-Whitney test was used when a nonparametric statistic was required, otherwise Duncan procedure was used as the follow-up procedure. This procedure was conducted on several learning measures extracted from the five learning trials. The groups that were found significantly different from one another in the different learning measures are indicated in Table 3.

Table 3. Age Group Comparisons (Using Multiple Comparison Procedures) for the Learning Measures and Pearson Product-Moment Correlation Scores of these with Age.

	Learning Measures						
Age group (years)	TL	TLc	T1	T5	LRd	LRr	Total
(20–29) vs. (30–39)	–	–	–	–	–	–	0
vs. (40–49)	+	–	–	+	–	–	2
vs. (50–59)	+	–	+	+	–	–	3
vs. (60–69)	+	+	+	+	–	–	4
vs. (70–91)	+	+	+	+	–	+	5
(30–39) vs. (40–49)	–	–	–	–	–	–	0
vs. (50–59)	–	–	–	–	–	–	0
vs. (60–69)	+	–	+	+	–	–	3
vs. (70–91)	+	+	+	+	–	+	5
(40–49) vs. (50–59)	–	–	–	–	–	–	0
vs. (60–69)	+	–	+	+	–	–	3
vs. (70–91)	+	–	+	+	–	+	4
(50–59) vs. (60–69)	+	–	–	+	–	–	2
vs. (70–91)	+	+	+	+	–	–	4
(60–69) vs. (70–91)	+	–	+	+	–	–	3
Total	11	4	9	11	0	3	38/90
Correlation with age for ages 20–59 ($n = 321$)	$-.24^{c}$	$-.08$	$-.13^{a}$	$-.27^{c}$	$-.08$	$.03$	
Correlation with age for ages 60–91 ($n = 207$)	$-.45^{c}$	$-.14^{a}$	$-.33^{c}$	$-.39^{c}$	$-.09$	$.13$	

Note. TL = Total learning (T1+T2+T3+T4+T5); TLc = corrected Total learning TL-(T1*5) measure used by Ivnik et al. (1992); LRd = Learning Rate – difference (T5–T1); LRr = Learning Rate – ratio (T5/T1).
$^{a}p < .05$; $^{b}p < .01$; $^{c}p < .001$.

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. The analysis of this score showed that the age groups were significantly different, $F(5, 522) = 52.94, p < .001$.

*Total learning (corrected score) (TLc) – [Total learning-(L1*5)]*

Ivnik et al. (1992) suggested an alternative measure to the total learning score which corrects for the possible baseline differences among participants in the initial trial. This score loaded on the "learning" factor in the Smith, Ivnik et al. (1992) and Smith et al. (1993) studies. The analysis of this score also showed a significant age group effect, $F(5, 522) = 4.33, p < .001$. Another score derived from the learning trials which has been used separately in previous factor analyses and found to load on the acquisition factor (Vakil & Blachstein, 1993) is *Immediate memory (Trial 1)*. This score was found to be sensitive to the effect of age, $F(5, 522) = 22.19, p < .001$.

Best learning (Trial 5)

This is another score previously used separately in factor analysis and was found to load on the retrieval factor (Vakil & Blachstein, 1993). The analysis of this score revealed a significant age group effect, $F(5, 522) = 36.40, p < .001$. The learning curve was analyzed as a continuum, using all five learning trials, as reported above. Two alternative scores are commonly applied in the literature to reflect the learning rate, either as a difference or as a ratio score of Trials 1 and 5. *Learning rate (difference score) (LRd) (Trials 5–1)*, is one of them. This score was found to be insensitive to age effect, $F(5, 522) = 1.57, p > .05$. The second alternative score is *Learning rate (ratio score) (LRr) (Trials 1/5)*. This variable was analyzed using nonparametric tests, because it did not fulfill the assumption of normal distribution. Kruskal-Wallis analysis revealed a significant age group effect, χ^2 $(5, N = 528) = 15.85, p < .01$.

Interference

In this section, the different interference measures extracted from the Rey AVLT are analyzed. These measures include the number of words recalled from the interference list (i.e., Trial 6) and the first list following the interference (i.e., Trial 7). In addition, the recall of this list as compared to the recall of the first list, enables the derivation of proactive and retroactive interference measures. Table 4 presents the groups that were found to be significantly different from each other in the different interference measures, by the multiple comparison procedures.

List B (interference list; Trial 6)

The analysis of this score revealed a significant age group effect, $F(5, 522) = 24.99, p < .001$.

List A (following the interference list; Trial 7)

An age effect was significant in the analysis of this score as well, $F(5, 522) = 42.03, p < .001$. As mentioned above, proactive and retroactive interference scores are derived from the comparison between interference list and first learning trial scores. These scores can be expressed as difference scores or as ratio scores. Both proactive interference scores, whether expressed as a difference score or a ratio score, were insensitive to age: *Proactive interference (difference score) (PId) (Trials 1-6)*, $F(5, 522) = 1.23, p > .05$, and *Proactive interference (ratio score) (PIr) (Trials 6/1)*, χ^2 $(5, N = 528) = 6.48, p > .05$. The ratio score required a nonparametric test (i.e., Kruskal-Wallis) because it violates the normal distribution assumption. On the other hand, both retroactive interference scores were sensitive to age: *Retroactive interference (difference score) (RId) (Trials 5–7)*, $F(5, 522) = 9.84, p < .001$, and *Retroactive interference (ratio score) (RIr) (Trials 7/5)*, $F(5, 522) = 16.01, p < .001$.

Delayed Recall and Recognition

This section includes the delayed recall (i.e., Trial 8) score and two measures of forgetting rate that address the relationship between delayed recall and Trial 5 scores. The recognition scores are measured in two ways: the hits rate of

Table 4. Age Group Comparisons (Using Multiple Comparison Procedures) for the Interference Measures and Pearson Product-Moment Correlation Scores of these with Age.

	Interference Measures						
Age group (years)	T6	T7	PId	PIr	RId	Rir	Total
(20–29) vs. (30–39)	–	–	–	–	–	–	0
vs. (40–49)	+	+	–	–	–	–	2
vs. (50–59)	+	+	–	–	–	–	2
vs. (60–69)	+	+	–	–	+	+	4
vs. (70–91)	+	+	–	–	+	+	4
(30–39) vs. (40–49)	–	–	–	–	–	–	0
vs. (50–59)	–	–	–	–	–	–	0
vs. (60–69)	+	+	–	–	+	+	4
vs. (70–91)	+	+	–	–	+	+	4
(40–49) vs. (50–59)	–	–	–	–	–	–	0
vs. (60–69)	+	+	–	–	+	+	4
vs. (70–91)	+	+	–	–	+	+	4
(50–59) vs. (60–69)	+	+	–	–	–	–	2
vs. (70–91)	+	+	–	–	+	+	4
(60–69) vs. (70–91)	+	+	–	–	+	+	4
Total	11	11	0	0	8	8	38/90
Correlation with age for ages 20–59 (n = 321)	–.21[c]	–.24[c]	.11[a]	–.10	.08	–.11[a]	
Correlation with age for ages 60–91 (n = 207)	–.31[c]	–.36[c]	–.07	–.08	.11	–.18[b]	

Note. PId = Proactive Interference – difference (T1–T6); PIr = Proactive Interference – ratio (T6/T1); RId = Retroactive Interference – difference (T5–T7); RId = Retroactive Interference – ratio (T7/T5).
[a] $p < .05$; [b] $p < .01$; [c] $p < .001$).

Trial 9, and a nonparametric signal detection measure that takes into account the hits and the false alarm rate as well (Geffen et al., 1990). In addition, when recognition is compared to delayed recall (i.e., Trial 8), the retrieval efficiency measure is derived either as a difference or a ratio score. Table 5 shows the groups that were found to be significantly different from each other, in the different delay measures, by the multiple comparison procedures.

Delayed recall (Trial 8)

The analysis of this score revealed a significant age group effect, $F(5, 522) = 34.67$, $p < .001$. The performance on delayed recall (i.e., Trial 8) was compared to the best learning score (i.e., Trial 5) (Ivnik et al., 1992). This score was analyzed both as difference and ratio scores. These two scores were analyzed using a nonparametric test (i.e., Kruskal-Wallis) because the normal distribution assumption was violated in these scores.

Forgetting rate (difference score) (FRd) (Trials 5–8)

This score was sensitive to age, χ^2 (5, N = 528) = 38.31, $p < .001$.

Forgetting rate (ratio score) (FRr) (Trials 8/5)

This was also sensitive to age, χ^2 (5, N = 528) =

Table 5. Age Group Comparisons (Using Multiple Comparison Procedures) for the Delayed Measures and Pearson Product Moment Correlation Scores of these with Age.

	Delayed Measures							
Age group (years)	T8	T9	FRd	FRr	REd	REr	NSD	Total
(20–29) vs. (30–39)	–	–	–	–	–	–	–	0
vs. (40–49)	+	–	–	–	+	+	–	3
vs. (50–59)	+	–	–	–	+	+	+	4
vs. (60–69)	+	+	+	+	+	+	+	7
vs. (70–91)	+	+	+	+	+	+	+	7
(30–39) vs. (40–49)	–	–	–	–	–	–	–	0
vs. (50–59)	–	–	–	–	–	–	–	0
vs. (60–69)	+	+	–	+	+	+	+	6
vs. (70–91)	+	+	+	+	+	+	+	7
(40–49) vs. (50–59)	–	–	–	–	–	–	–	0
vs. (60–69)	+	+	–	+	–	–	+	4
vs. (70–91)	+	+	+	+	+	+	+	7
(50–59) vs. (60–69)	+	+	–	–	–	–	+	3
vs. (70–91)	+	+	+	+	+	+	+	7
(60–69) vs. (70–91)	+	+	+	+	+	+	+	7
Total	11	9	6	8	9	9	10	62/105
Correlation with age for ages 20–59 ($n = 321$)	$-.20^{c}$	$-.11^{a}$	.05	−.08	$.14^{b}$	$-.19^{c}$	$-.15^{b}$	
Correlation with age for ages 60–91 ($n = 207$)	$-.37^{c}$	$-.17^{b}$	$.14^{a}$	$-.21^{b}$	$-.24^{c}$	$-.31^{c}$	−.26*c*	

Note. FRd = Forgetting Rate – difference (T5–T8); FRr = Forgetting Rate ratio (T8/T5); REd = Retrieval efficiency – difference (T9–T8); REr = Retrieval efficiency – ratio (T8/T9); NSD = Nonparametric Signal Detection measure – 0.5 (1 + Hit Rates – False Positives).
$^{a}p < .05$; $^{b}p < .01$; $^{c}p < .001$.

60.88, $p < .001$. This score (expressed as percent retention) loaded on the "retention" factor in the Smith, Ivnik et al. (1992), and Smith et al. (1993) studies. For the same reason as above the next three scores were analyzed by a nonparametric test (i.e., Kruskal-Wallis).

Recognition (Trial 9)
This score was found to be sensitive to age, χ^2 $(5, N = 528) = 104.51$, $p < .001$. In order to correct the recognition score for a positive criterion (the tendency to give a positive answer), Geffen et al. (1990) used a nonparametric signal detection measure. This measure consists of the proportion of words correctly identified from list A (hit rate (HR)) and the proportion of false positive responses (FP).

Nonparametric signal detection measure [0.5(1+HR-FP)]
This score was found to be sensitive to the age effect, χ^2 $(5, N = 528) = 120.08$, $p < .001$.

Retrieval efficiency (difference score) (REd) (Trials 9–8)
This score was also found to be sensitive to the age effect, χ^2 $(5, N = 528) = 50.65$, $p < .001$.

Retrieval efficiency (ratio score) (REr) (Trials 8/9)
This score also reached significance, $F(5, 522) = 15.07, p < .001$.

Temporal Order
A supplementary measure to the Rey AVLT for assessing temporal order was recently introduced (Vakil & Blachstein, 1994). Three alternative scores were suggested in order to measure temporal order: hits, absolute deviations, and correlations. Table 6 presents the groups that were found to be significantly different from each other, in the different temporal order measures, by the multiple comparison procedures.

All three measures of temporal order were found to be sensitive to age: *Hits*, $F(5, 522) = 18.16, p < .001$, *Absolute deviations (AD)*, $F(5, 522) = 15.81, p < .001$, and *Correlation score (CO)*, $F(5, 522) = 21.71, p < .001$.

Table 7 gives a summary of the proportional number of measures significantly differentiated between each age group for the above four categories of memory. As can be seen in this table, some age groups are more clearly differentiated than others by a number of measures. None of the measures can differentiate between the first (20–29) and the second (30–39) age groups, or between the third (40–49) and the fourth (50–59) age groups. The second (30–39) and the third (40–49) age groups as well as the second and the fourth (50–59) age group are distinguished only by the temporal order measures. However, as we move higher in the age range, even the adjacent age groups are more clearly differentiated: nine measures are sensitive to the change from the fourth (50–59) to the

Table 6. Age Group Comparisons (Using Multiple Comparison Procedures) for the Temporal Order Measures and Pearson Product-Moment Correlation Scores of these with Age.

	Temporal Order Measures			
Age group (years)	AD	CO	Hits	Total
(20–29) vs. (30–39)	–	–	–	0
vs. (40–49)	+	+	+	3
vs. (50–59)	–	+	+	2
vs. (60–69)	+	+	+	3
vs. (70–91)	+	+	+	3
(30–39) vs. (40–49)	+	+	+	3
vs. (50–59)	–	+	+	2
vs. (60–69)	+	+	+	3
vs. (70–91)	+	+	+	3
(40–49) vs. (50–59)	–	–	–	0
vs. (60–69)	+	+	–	2
vs. (70–91)	+	+	+	3
(50–59) vs. (60–69)	+	+	–	2
vs. (70–91)	+	+	+	3
(60–69) vs. (70–91)	–	+	–	1
Total	10	13	10	33/45
Correlation with age for ages 20–59 (n = 321)	.27[c]	–.25[c]	–.28[c]	
Correlation with age for ages 60–91 (n = 207)	.20[b]	–.21[b]	–.11	

Note. AD = absolute deviation; CO = correlation score.
[a] $p < .05$; [b] $p < .01$; [c] $p < .001$.

fifth (60–69) age groups, and 15 measures are sensitive to the change from the fifth to the sixth (70–91) age groups. As seen in Table 7, there is a consistent trend showing the three younger groups (i.e., 20–59) to be less distinguished from each other than the two older groups (i.e., 60–91).

Pearson Product-Moment Correlation was conducted between age and the different memory measures. This is an alternative way to assess how these memory measures are related to age. Unlike the analyses of variance conducted above, correlations take into account the exact age of each participant, rather than treating all members of each age group as the same age. The conclusion above that the three younger groups are less distinguished than the two older groups, suggests that the effect of age on most of the Rey AVLT scores is nonlinear. In order not to violate the linearity assumption, a separate set of correlations was calculated for the younger and the older age groups. These results are reported in Tables 3 to 6, along with the respective results of the multiple comparison procedures.

DISCUSSION

The norms obtained in the present study for the Hebrew version of the Rey AVLT are compatible with those found in studies that compared gender performance (Bleeker et al., 1988) and those that did not (Ivniket al., 1990; Mitrushina et al., 1991; Wiens et al., 1988). The scores of the present study differed from those reported by Geffen et al. (1990), however; that is, they were generally higher, especially for males. As in other reports on verbal memory in general, and the Rey AVLT in particular (Bleeker et al., 1988; Geffen et al., 1990), the performance of females was superior to that of males across all age groups, as indicated by lack of significant interaction in most of the measures analyzed.

Table 7. Summary of the Proportional Number of Measures SignificantlyDifferentiated Between Each Age Group for the Different Categories of Memory (Using Multiple Comparison Procedures).

Age group (years)		Measures				
		Learning	Interference	Delayed	Temporal order	Total
(20–29)	vs. (30–39)	0/6	0/6	0/7	0/3	0/22
	vs. (40–49)	2/6	2/6	3/7	3/3	10/22
	vs. (50–59)	3/6	2/6	4/7	2/3	11/22
	vs. (60–69)	4/6	4/6	7/7	3/3	18/22
	vs. (70–91)	5/6	4/6	7/7	3/3	19/22
(30–39)	vs. (40–49)	0/6	0/6	0/7	3/3	3/22
	vs. (50–59)	0/6	0/6	0/7	2/3	2/22
	vs. (60–69)	3/6	4/6	6/7	3/3	16/22
	vs. (70–91)	5/6	4/6	7/7	3/3	19/22
(40–49)	vs. (50–59)	0/6	0/6	0/7	0/3	0/22
	vs. (60–69)	3/6	4/6	4/7	2/3	13/22
	vs. (70–91)	4/6	4/6	7/7	3/3	18/22
(50–59)	vs. (60–69)	2/6	2/6	3/7	2/3	9/22
	vs. (70–91)	4/6	4/6	7/7	3/3	18/22
(60–69)	vs. (70–91)	3/6	4/6	7/7	1/3	15/22
Total		38/90	38/90	62/105	33/45	171/330

One purpose of this study was to investigate whether some memory measures extracted from the Rey AVLT are more sensitive than others to the effect of age. This was addressed by checking the correlation of a given measure with age, and by discerning the number of age groups that measure succeeded in differentiating. Both methods yielded similar results, suggesting that there is a differential degree of sensitivity of measures to age. This is relevant to measures in all four memory categories previously referred to in the Results section (i.e., learning, interference, delayed recall and recognition, and temporal order).

Among the learning measures, the Trial 5 and total learning scores were most sensitive to age. Both differentiated 11 out of 15 possible age comparisons and their correlation with age was for the younger ages $r = -.27$, and $r = -.24$, respectively, and for the older ages $r = -.39$ and $r = -.45$ respectively. In contrast, the learning rate measure correlation with age was only $r = -.08$ for the younger ages, $r = -.09$ for the older ages, and it did not differentiate age groups.

Among the interference scores, Trial 6 and 7 scores were the most sensitive. As above, these scores differentiated between 11 out of the 15 group comparisons, and their correlation with age was for the younger ages $r = -.21$, and $r = -.24$, respectively and for the older ages $r = -.31$ and $r = -.36$ respectively.

Proactive interference, whether measured as a difference or a ratio score, did not distinguish between any group. Their correlation with age for the younger ages was $r = .11$, and $r = -.10$, for the difference and ratio score, respectively and for the older ages $r = -.07$ and $r = -.08$, respectively. Trial 8 and the nonparametric signal detection measure of recognition were the most sensitive delayed scores. They distinguished among 11 and 10 of 15 group comparisons and their correlation with age for the younger ages was $r = -.20$, and $r = -.15$, respectively, and for the older ages $r = -.37$ and $r = -.26$, respectively. The forgetting rate, particularly the difference score, was the least sensitive among the delayed scores; only 6 out of 15 age group comparisons were significantly distinguishable by this score, and its correlation with age was $r = .05$ for the younger ages and $r = .14$ for the older ages.

The clear conclusion from these findings is that the age-related memory measures do not uniquely belong to a specific category of memory. Thus, there is no support to the claim that a particular memory process, such as encoding (Erber, 1974; Gordon & Clark, 1974) or retrieval (Mitrushina et al., 1991), is most sensitive to age. There is a trend indicating that the delay measures (i.e., 62/105) are more age-related than are the learning and interference measures (i.e., 38/90), but this conclusion should be viewed cautiously considering the redundancy among the scores in each category.

The three measures of temporal order were found to be sensitive to age. As seen in Table 6, 10 to 13 of 15 possible comparisons were significant with these scores. Correlation with age ranged between $r = -.25$ to $r = -.28$ for the younger ages, and $r = -.11$ to $r = -.21$ for the older ages. The similarity in the sensitivity of these scores is probably due to the fact that these scores are much more related to each other than are the different memory scores in other categories, particularly the correlation and the absolute deviation scores (Vakil & Blachstein, 1994). It is important to note that of all measures, the correlation score of temporal order judgment was the most sensitive to age. It distinguished between 13 out of 15 age comparisons. This finding lends further support to the suggestion that the temporal order judgment be added to the standard administration of the Rey AVLT (Vakil & Blachstein, 1994).

Overall, a clear trend is observed: until the age of 60 years, the changes in verbal memory are moderate, as compared to the changes observed from the age of 60 onward (Bleeker et al., 1988; Ivnik et al., 1990; Mitrushina et al., 1991). Further support for this conclusion can be drawn from the correlation pattern. The correlations of the different memory scores with age were consistently higher in the older than in the younger age groups. The rapid changes in memory performance in the older age group stresses the importance of having appropriate norms, particularly for this age group.

With regard to the question as to whether the difference or the ratio scores is more sensitive to age, our data indicate a tendency for the ratio scores to be more sensitive than the difference scores. The most pronounced difference can be found with reference to the learning rate measures. As mentioned above, none of the groups was differentiated by the difference score (r = –.08 for the younger and r = –.09 for the older), whereas three age groups were differentiated by the ratio score (r = –.03 for the younger and r = –.13 for the older). As seen in Tables 3 and 5 (but not in Table 4), there is a very little, but consistent, advantage of the ratio scores over the difference scores in the number of groups they differentiate, and in the correlation with age.

In conclusion, this Rey AVLT study used a large sample of both males and females, and extends the usefulness of the Rey AVLT as a diagnostic tool. The analyses reported in this paper provide for a better understanding and a more valid use of the variety of scores that can be extracted from the Rey AVLT.

REFERENCES

Blachstein, H., Vakil, E., & Hoofien, D. (1993). Impaired learning in closed-head injured patients: An analysis of components of the acquisition process. *Neuropsychology, 7,* 530-535.

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.

Burke, D. M., & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin, 90,* 513-546.

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.

Erber, J. T. (1974). Age differences in recognition memory. *Journal of Gerontology, 29,* 177-181.

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

Gordon, S. K., & Clark, W. C. (1974). Application of signal detection theory to prose recall and recognition in elderly and young adults. *Journal of Gerontology, 29,* 64-72.

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.

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.

Kausler, D. H. (1994). *Learning and memory in normal aging.* San Diego: Academic Press.

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

Light, L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review in Psychology, 42,* 333-376.

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.

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.

Poon, L. W. (1985). Differences in human memory with aging: Nature, causes, and clinical implications. In J. E. Birren & W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 427-462). New York: Van Nostrand Reinhold.

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.

Query, W. T., Randy, A., & 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.

Rey, A. (1964). *L'examen clinique en psychologie.* Paris: Presses Universitaires de France.

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.

Schonfield, D., & Robertson, B. A. (1966). Memory storage and aging. *Canadian Journal of Psychology, 20,* 228-236.

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 Assessment, 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 Assessment, 5,* 121-124.

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

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.

Vakil, E., Hoofien, D., & Blachstein, H. (1992). Total amount learned versus learning rate of verbal and nonverbal information, in differentiating left- from right-brain injured patients. *Archives of Clinical Neuropsychology, 7,* 111-120.

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.

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.