Journal of the International Neuropsychological Society (2004), **10**, 72–81.

DOI: 10.1017/S1355617704101094

A normative study of the CERAD neuropsychological assessment battery in the Korean elderly

DONG Y. LEE,[1] KANG U. LEE,[2] JUNG H. LEE,[2] KI W. KIM,[3] JIN H. JHOO,[4] SUNG Y. KIM,[5] JONG C. YOON,[6] SUNG I. WOO,[7] JIN HA,[8] AND JONG I. WOO,[8,9]

[1]Department of Neuropsychiatry, Kyunggi Provincial Hospital for the Elderly, Yongin, Kyunggi-do, Korea
[2]Department of Neuropsychiatry, Kangwon National University Hospital, Chunchon, Kangwon-do, Korea
[3]Department of Neuropsychiatry, Seoul National University Bundang Hospital, Sungnam, Kyunggi-do, Korea
[4]Department of Neuropsychiatry, Daejin Medical Center, Sungnam, Kyunggi-do, Korea
[5]Department of Psychiatry, Asan Medical Center, Seoul, Korea
[6]Clinical Research Institute, Seoul National University Hospital, Seoul, Korea
[7]Department of Psychiatry, Soon Chun Hyang University Hospital, Seoul, Korea
[8]Department of Neuropsychiatry, Seoul National University Hospital, Seoul, Korea
[9]Neuroscience Research Institute, Medical Research Center, Seoul National University, Seoul, Korea

(RECEIVED September 12, 2002; REVISED April 22, 2003; ACCEPTED April 22, 2003)

Abstract

This study aimed to explore the effects of age, education and gender on the performance of eight tests in the Korean version of the CERAD neuropsychological assessment battery and to provide normative information on the tests in the Korean elderly. The battery was administered to 618 healthy volunteers aged from 60 to 90. People with serious neurological, medical and psychiatric disorders, including dementia, were excluded. Multiple linear regression analyses were performed to assess the relative contribution of the demographic factors on the score of each cognitive test. Age, education, and gender were found to have significant effects on the performance of many tests in the battery. Based on these results, 4 overlapping age normative tables (60 to 74, 65 to 79, 70 to 84, and 75 to 90 years of age) with 3 educational strata (0 to 3 years, 4 to 6 years, and 7 years and more) for both genders are presented. The normative information will be useful for a clinical interpretation of the CERAD neuropsychological battery in Korean elderly as well as for comparing the performance of the battery across countries. (*JINS*, 2004, *10*, 72–81.)

Keywords: Neuropsychological battery, Norms, CERAD, Korean, Elderly, Age, Education, Gender

INTRODUCTION

The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) developed standardized clinical and neuropsychological assessment batteries for a uniform evaluation of patients with Alzheimer's disease (AD) (Morris et al., 1989). The CERAD neuropsychological assessment battery includes tests to assess the specific cognitive deficits that appear in AD, such as memory impairment, disorientation, loss of expressive and receptive language, and dyspraxia. Though it is more comprehensive than a simple cognitive screening tool, the battery is relatively brief and is easy to apply to geriatric patients. It also has been proved to have good reliability and psychopathological validity (Welsh-Bohmer & Mohs, 1997). J. Lee et al. translated the original English version of the CERAD neuropsychological assessment battery into Korean, which has proven to be equally reliable and valid (J. Lee et al., 2002).

Many studies have shown that cognitive test results are influenced by major demographic factors such as advancing age, educational levels and gender (Lezak, 1995). The degree of influence from each factor varies considerably according to individual cognitive test or domain. Therefore, in order to obtain clinically useful normative information for a cognitive test, the variables that confound the score of an individual test need to be clarified. From there, separate norms for the subjects, which are stratified according to one or more confounding variables, can be obtained.

Previous studies (Ganguli et al., 1991; Welsh et al., 1994) reported the influence of demographic variables on the cog-

Reprint requests to: Dr. Jong Inn Woo, Department of Neuropsychiatry, Seoul National University Hospital, 28 Yongon-dong, Chongno-gu, Seoul, Korea (110-744). E-mail: jiwoomd@plaza.snu.ac.kr

nitive tests in the CERAD neuropsychological battery and presented some normative data for the elderly. However, they were based mainly on well-educated western Whites (Ganguli et al., 1991; Welsh et al., 1994). Ethnic or cultural background can be a confounder for a cognitive test result (Escobar et al., 1986; Kittner et al., 1986). Moreover, a considerable number of elderly from Asian countries including Korea are poorly educated and are not normally exposed to neuropsychological tests. The more individuals in a population are exposed to similar cognitive tests, the better performance a cognitive test can become (Flynn, 1987). Therefore, it is not advisable to extrapolate normative information from highly educated, White subjects to populations with different educational and ethnic backgrounds.

This study aimed to investigate the effects of age, education, and gender on the performance of eight tests in the Korean version of the CERAD neuropsychological assessment battery, CERAD–K(N), and to provide the normative information on these tests in the Korean elderly.

METHODS

Research Participants

Six hundred eighteen healthy elderly subjects (209 male and 409 female; 60–90 years of age) were included in this study. Seven centers located in a variety of regions in Korea (one public health center, two welfare centers for the elderly, and four dementia or memory clinics) had been conducting a service program for the early detection and management of dementia for elderly in the community with the support of the Korean Association for Dementia. An announcement for this program was made through the local newspapers and posters placed on a bulletin board at each center. The study subjects were recruited from a pool of the elderly registered in this program from January 1997 to September 2000. They consisted of 94 spouses of dementia patients and 524 community volunteers. All of them lived in the community, and were not in a long-term care institution. Informed consent was obtained from each participant according to the procedures approved by the institutional review board of each center.

Each subject was examined by a psychiatrist with advanced training in neuropsychiatry and dementia research according to the protocol of the Korean version of the CERAD clinical assessment battery (J. Lee et al., 2002). The tests consisted of a standardized clinical interview on demographic information, cognitive and functional status, drug inventory, depression and medical history, a cognitive state examination including the six-item Korean version of the Short Blessed test (D. Lee et al., 1999), and a general physical and neurological examination. Reliable informants were also interviewed to acquire accurate information regarding the cognitive and functional changes and medical history of the subjects. A panel consisting of four psychiatrists with expertise in dementia research made the clinical decisions including dementia diagnosis. This panel reviewed all available raw data from the clinical evaluation.

All subjects satisfied the strict entry criteria excluding dementia and other serious medical, psychiatric, and neurological disorders that could affect mental function. A diagnosis of dementia was made according to the criteria from fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV; American Psychiatric Association, 1994). All subjects possessed adequate vision and hearing, although many wore glasses and some required a hearing aid. Individuals with minor physical abnormalities (e.g., diabetes with no serious complications, essential hypertension, mild hearing loss, etc) were not excluded. Illiterate persons were also excluded because they were unable to read the word lists in some tests. Illiteracy in this study was determined by whether or not one could read a 10-word list for the Word List Memory test in the CERAD–K(N).

Procedure

The CERAD–K(N) was administered at each site by trained research nurses, who were blinded to the psychiatrist's clinical evaluation. The battery includes the eight tests listed below, which are presented in the order of administration with the battery.

1. *Verbal Fluency: "Animal Category."* This test measures verbal production, semantic memory, and language (Rosen, 1980). It requires the subject to name as many examples of the category "animal" as possible within 1 min.
2. *Modified Korean version of the Boston Naming Test*. This test measures visual naming and presents 15 line drawings of common objects from the Korean version of the Boston Naming Test (K–BNT, Kim & Na, 1997). These items are stratified into three groups containing five items each, representing objects of a high, medium, and low frequency of occurrence in the Korean language. The maximum score is 15.
3. *Mini-Mental State Examination in the Korean version of the CERAD assessment packet (MMSE–KC)*. Mini-Mental State Examination (MMSE; Folstein et al., 1975) is a well-known screening tool for the cognitive function that measures orientation, language (repetition, naming, reading and writing), concentration, constructional praxis, and memory. Because there are a significant number of illiterate people in Korea, the reading and writing items were replaced by two items focusing on judgment in MMSE–KC (J. Lee et al., 2002; Park & Kwon, 1991). The maximum score for the test is 30.
4. *Word List Memory*. This is a free-recall memory test that assesses the learning ability for new verbal information. Three trials of a 10-item word list are presented. On each trial, the 10 words are presented in a different order. The subject is instructed to read each word aloud as it is presented. Immediately following each trial, the subject is asked to recall as many items as possible (90 s

allowed). The maximum number of correct responses is 30 for the 3 trials. All the words included in the Korean version of the Word List Memory task have a high frequency and high mental imagery. Lee et al. also considered the phonemic similarity, and semantics or word length equivalence when they translated the English version of the Word List Memory into Korean (J. Lee et al., 2002).

5. *Constructional Praxis*. This task measures visuospatial and constructional abilities. Four line drawings of figures with increasing complexity (a circle, a diamond, intersecting rectangles, and a cube) are presented to the subject for copying; 2 min are allowed for each figure. The maximum score for a correct drawing is 11 for all four figures.
6. *Word List Recall*. This test assesses the ability to recall, after a delay of a few minutes, the 10 words given in the Word List Memory task. A maximum of 90 s is allowed, and the maximum score is 10.
7. *Word List Recognition*. This test assesses the recognition of the target words presented in the Word List Memory task when presented among the 10 distractor words. In order to eliminate the possibility of a correct recognition by chance, the subject's final score is calculated as the total number of correct answers for both the 10 target words and the 10 distractor words minus 10. A score of zero is given if the result is less than zero, The maximum score of this task is 10.
8. *Constructional Recall*. This task assesses the ability to recall, after a few-minute delay, the four line drawings of the figures presented in the Constructional Praxis task. The maximum score of a correct drawing is 11 for all four figures. This task also includes the recall for the interlocking pentagons of the MMSE–KC. If a subject recalled and drew those objects, we scored on them, but did not consider this score when calculating the final score for Constructional Recall.

Statistical Analysis

A separate stepwise multiple linear regression analysis was performed to assess the relative contribution of age, education, and sex on each CERAD–K(N) cognitive test score. Age and education were entered as continuous variables and sex was coded as zero and 1 for females and males, respectively. The criteria for entry and removal of the variables were $p < .05$ and $p < .10$, respectively.

A series of $2 \times 5 \times 2$ analyses of variance (ANOVA) were also performed to determine any main effects and interactions of age (60–74 and 75–90 years), education (0–3, 4–6, 7–9, 10–12, and ≥13 years), and sex (female and male) on the tests. *Post-hoc* contrasts with a two tailed *t* test were conducted when any main effect of education was determined to be significant by ANOVA at the $p < .05$ levels. No adjustment was made for multiple comparisons.

Table 1. Demographic characteristics of the subjects

Variable	Men	Women	Total
Number	209	409	618
Age (years)	73.6 ± 6.3[a]	71.2 ± 6.0	72.0 ± 6.3
60–64	6 (2.9)[b]	48 (11.7)	54 (8.7)
65–69	61 (29.2)	143 (35.0)	204 (33.0)
70–74	58 (27.8)	103 (25.2)	161 (26.1)
75–79	45 (21.5)	71 (17.4)	116 (18.8)
80–84	27 (12.9)	29 (7.1)	56 (9.1)
85–90	12 (5.8)	15 (3.6)	27 (4.4)
Education (years)	7.8 ± 5.4	4.8 ± 4.0	5.9 ± 4.8
0	87 (17.7)	125 (30.6)	162 (26.2)
1–3	8 (3.8)	39 (9.5)	47 (7.6)
4–6	68 (32.5)	159 (38.9)	227 (36.7)
7–9	26 (12.4)	40 (9.8)	66 (10.7)
10–12	29 (13.9)	35 (8.6)	64 (10.4)
13–	41 (19.6)	11 (2.7)	52 (8.4)

[a] *M* ± *SD*
[b] Number (percent)

RESULTS

Demographic Characteristics

The demographic characteristics of 618 subjects, who completed the CERAD–K(N), are shown in Table 1. The mean age of the male subjects (73.6 years) was slightly higher than that of the female subjects [71.2 years; $t(290) = -2.28$, $p < .05$]. The mean number of years of education was also much higher in the male group (7.8 years) than in the female group [4.8 years; $t(290) = 1.97$, $p < .05$].

Effects of Age, Education, and Sex on Cognitive Tests in CERAD–K(N)

Stepwise multiple regression analyses and ANOVA's revealed that seven out of eight cognitive tests in CERAD–K(N) were significantly influenced by both age and education, as can be seen from Tables 2 and 3. The exception was Word List Recognition. Younger persons outperformed the older ones and the higher-educated ones performed better than the lower-educated ones for every test in common. For the Word List Recognition, only education significantly influenced the test performance and the higher-educated subjects performed better. However, age did not have any significant impact on the test score.

Post-hoc contrasts with a two-tailed *t* test were conducted when any main effect of education was determined to be significant by ANOVA at the $p < .05$ level. The *post-hoc* contrasts for the Verbal Fluency, Boston Naming Test, and MMSE–KC revealed a significant graded difference between the 0–3, 4–6 and 7–9 year education group. In contrast, there was no significant difference between any of the two groups among the 7–9, 10–12, and 13 or older group. A significant difference was observed only between

https://doi.org/10.1017/S1355617704101094 Published online by Cambridge University Press

Table 2. Stepwise multiple linear regression of age, education and sex on the cognitive tests in CERAD–K(N)

Test		Age	Education	Sex
Verbal Fluency	B	−0.14	0.22	—
	SE(B)	0.03	0.03	—
	β	−0.22***	0.26***	—
	R^2	4.8	8.8	—
Boston Naming Test	B	−0.12	0.21	1.38
	SE(B)	0.02	0.03	0.25
	β	−0.24***	0.33***	0.21***
	R^2	3.0	20.3	3.7
MMSE–KC	B	−0.22	0.34	1.33
	SE(B)	0.02	0.02	0.28
	β	−0.36***	0.43***	0.16***
	R^2	9.4	29.9	2.2
Word List Memory	B	−0.16	0.19	−1.00
	SE(B)	0.03	0.03	0.41
	β	−0.22***	0.20***	−0.11*
	R^2	8.0	2.5	1.0
Constructional Praxis	B	−0.11	0.23	0.43
	SE(B)	0.01	0.02	0.16
	β	−0.31***	0.50***	0.09*
	R^2	7.9	34.2	0.6
Word List Recall	B	−0.09	0.04	—
	SE(B)	0.01	0.02	—
	β	−0.28***	0.10*	—
	R^2	8.7	1.0	—
Word List Recognition	B	−0.06	—	—
	SE(B)	0.01	—	—
	β	−0.24***	—	—
	R^2	5.7	—	—
Constructional Recall	B	−0.14	0.29	0.87
	SE(B)	0.02	0.02	0.24
	β	−0.28***	0.45***	0.14***
	R^2	5.4	30.0	1.5

Note. CERAD–K(N) = Korean version of the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery. B = regression coefficient, SE(B) = standard error of B, β = standardized regression coefficient, R^2 = percent variance explained by each variable. Age and education were entered as a continuous variable, and sex was coded as 0 and 1 for female and male, respectively.
*$p < .05$; ***$p < .0001$ by stepwise multiple linear regression analysis.

the 0–3 and 4–6 year groups in the case of Word List Memory, Constructional Praxis, and Word List Recall, and no significant difference was found between any two groups among the 4–6, 7–9, 10–12, and 13 or older groups. As for Constructional Recall, graded differences were found between the 0–3, 4–6, 10–12 and 13 or older groups, but not between the 4–6 and 7–9, or 7–9 and 10–12 groups.

The gender effect was somewhat different for each test (Tables 2 and 3). For the Boston Naming test, MMSE–KC, Constructional Praxis and Constructional Recall, males performed significantly better than females. In contrast, females significantly outperformed males for Word List Memory. In the case of Verbal Fluency, Word List Recall and Word List Recognition, there was no significant association between gender and the test performance.

As shown in Table 2, with respect to the percent variance explained by each significant variable (R^2), which reflects the relative degree of influence by each variable, education explained a much greater proportion of score variances than did either age or gender for the Verbal Fluency, Boston Naming Test, MMSE–KC, Constructional Praxis and Constructional Recall. However, for the Word List Memory and Word List Recall, age had a larger contribution to the variances of scores than did either education or gender.

As shown in Table 3, ANOVA's revealed a significant interaction between education and gender for the Boston Naming test, MMSE–KC, and Constructional Praxis. The test performance for these three tests declined more rapidly in women than in men with decreasing education. For Constructional Praxis, a significant interaction was also found

https://doi.org/10.1017/S1355617704101094 Published online by Cambridge University Press

Table 3. Analyses of variance for main effects and interactions between age, education and sex on the cognitive tests in CERAD–K(N)

	Main effect		Interaction	
Test	Variable	*F*	Variable	*F*
Verbal Fluency	Age	17.49***	Age × Education	0.83
	Education	11.43***	Age × Sex	1.09
	Sex	0.01	Education × Sex	1.01
Boston Naming Test	Age	15.20***	Age × Education	2.32
	Education	19.00***	Age × Sex	1.40
	Sex	4.94*	Education × Sex	2.69*
MMSE–KC	Age	34.38***	Age × Education	2.21
	Education	39.85***	Age × Sex	3.12
	Sex	9.17***	Education × Sex	2.71*
Word List Memory	Age	14.99***	Age × Education	2.14
	Education	6.51***	Age × Sex	2.15
	Sex	6.28*	Education × Sex	0.49
Constructional Praxis	Age	19.80***	Age × Education	0.79
	Education	43.56***	Age × Sex	5.47*
	Sex	3.67*	Education × Sex	3.34*
Word List Recall	Age	24.11***	Age × Education	0.87
	Education	2.53*	Age × Sex	0.22
	Sex	0.36	Education × Sex	0.19
Word List Recognition	Age	12.44***	Age × Education	0.85
	Education	1.33	Age × Sex	0.46
	Sex	0.68	Education × Sex	0.34
Constructional Recall	Age	31.69***	Age × Education	0.19
	Education	33.37***	Age × Sex	0.09
	Sex	7.60**	Education × Sex	0.12

Note. CERAD–K(N) = Korean version of the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery. Age is categorized as '60–74 and 75–90 years'; education as '0–3, 4–6, 7–9, 10–12, and ≥13 years'; and sex as 'female and male,' respectively.
*$p < .05$; **$p < .01$, ***$p < .001$ by 2 × 5 × 2 analyses of variance (ANOVA).

between Age × Gender. The performance declined more prominently in women than in men with advancing age.

Normative Data

The above results for the effects of the demographic variables indicated that age, education, and gender need to be taken into account for an accurate interpretation of the CERAD–K(N) cognitive subtests. Therefore, the total group was divided into four overlapping age tables, separately for males and females, and three strata of educational levels (0–3, 4–6, ≥7 years). The educational strata were determined by first considering the results from the *post-hoc* contrasts between the educational groups, which indicated significant differences between the lower educational groups (i.e., 0–3 *vs.* 4–6, or 4–6 *vs.* 7–9 year group) and no significant differences between the higher educational groups (i.e., 7–9 *vs.* 10–12, or 10–12 *vs.* ≥13, or 7–9 *vs.* ≥13 year group) for most tests except Constructional Recall. The number of subjects within each cell required for stable estimation of normative value was also taken into account. The overlapping age tables have been suggested as a way to maximize the quantity of information that can be obtained from normative projects (Pauker, 1988). Overlapping cell tables with midpoint ages occurring at 5-year intervals from the 67 through to 82 years were developed (Tables 4–7). The age range from which each normative value was estimated is 14 years (i.e., ±7 years around the midpoint age) for each table (exceptionally, 15 years for the oldest group) and the ranges for the adjacent tables are overlapped. In order to interpret the test scores, the user should select the table with the closest midpoint from the subject's age. For the sake of convenience, the age range of subjects to whom each table should be applied is presented in the title of the table. Each of these ranges was derived to encompass all ages closer to one midpoint age than to the other adjacent midpoint(s). This overlapping age table method provides a broad normative base for each normative cell. The normative scores for each stratified cell are presented in the form of a mean and standard deviation, and a median and range from the 5th to the 95th percentile (Tables 4–7).

DISCUSSION

The results in this study indicate that older age is associated with lower performance on all cognitive tests in the

https://doi.org/10.1017/S1355617704101094 Published online by Cambridge University Press

Table 4. CERAD–K(N) normative data estimated from the ages 60–74[a] and used for the ages 60–69: Mean with standard deviation, and median and range from the 5th to the 95th percentile

Tests	Sex	Education (years) 0–3	4–6	≥7
Verbal Fluency	Male	13.1 ± 4.4[b]	13.2 ± 3.9	15.7 ± 4.8
		12.5 (6.0–25.0)[c]	13.0 (6.0–21.0)	15.0 (7.2–24.0)
	Female	12.9 ± 3.6	14.3 ± 3.6	15.3 ± 3.4
		13.0 (7.0–19.5)	15.0 (8.0–21.0)	14.0 (11.0–21.7)
Boston Naming Test	Male	10.6 ± 3.2	11.1 ± 2.5	11.7 ± 1.9
		10.0 (5.0–15.0)	11.0 (7.0–15.0)	12.0 (8.2–15.0)
	Female	8.1 ± 2.6	9.5 ± 2.6	11.3 ± 2.4
		8.0 (4.0–12.5)	9.0 (5.8–14.0)	12.0 (6.3–14.0)
MMSE–KC	Male	24.9 ± 2.7	27.3 ± 2.2	28.1 ± 1.7
		25.0 (18.0–30.0)	27.5 (23.1–30.0)	29.0 (24.3–30.0)
	Female	23.7 ± 3.5	26.4 ± 2.6	28.2 ± 1.5
		24.0 (16.0–29.0)	27.0 (20.9–30.0)	29.0 (25.0–30.0)
Word List Memory	Male	14.5 ± 3.7	15.1 ± 3.8	16.4 ± 4.4
		14.0 (6.0–20.0)	16.0 (9.0–22.0)	17.0 (8.2–23.0)
	Female	15.2 ± 4.2	16.8 ± 4.8	18.3 ± 3.1
		15.0 (7.0–22.0)	17.0 (9.7–23.0)	19.0 (13.0–22.0)
Constructional Praxis	Male	7.9 ± 1.5	9.5 ± 1.6	10.4 ± 1.0
		7.5 (6.0–11.0)	10.0 (7.0–11.0)	11.0 (8.0–11.0)
	Female	7.5 ± 2.0	9.7 ± 1.6	10.4 ± 1.0
		7.0 (4.0–11.0)	10.0 (7.0–11.0)	11.0 (7.3–11.0)
Word List Recall	Male	5.5 ± 1.1	5.8 ± 2.0	6.0 ± 1.9
		5.0 (2.0–8.0)	6.0 (2.0–9.0)	6.0 (2.2–9.0)
	Female	5.3 ± 1.8	6.0 ± 1.9	6.4 ± 1.8
		5.0 (2.7–8.3)	6.0 (2.0–9.0)	7.0 (3.0–8.7)
Word List Recognition	Male	9.3 ± 1.0	9.1 ± 1.4	9.0 ± 1.4
		10.0 (5.0–10.0)	10.0 (5.0–10.0)	10.0 (5.2–10.0)
	Female	9.2 ± 1.3	9.1 ± 1.6	9.3 ± 1.2
		10.0 (6.7–10.0)	10.0 (6.0–10.0)	10.0 (6.0–10.0)
Constructional Recall	Male	4.0 ± 2.5	7.0 ± 2.7	7.9 ± 2.6
		4.0 (0.0–10.0)	10.0 (2.0–11.0)	8.0 (4.0–11.0)
	Female	3.9 ± 2.2	5.4 ± 2.4	7.2 ± 2.6
		4.0 (0.0–7.0)	5.0 (2.0–10.0)	7.0 (4.0–11.0)
Age mean ± SD [sample size]	Male	67.5 ± 2.5 [20]	67.2 ± 3.3 [40]	67.7 ± 3.5 [65]
	Female	66.7 ± 3.2 [92]	66.1 ± 3.3 [137]	66.1 ± 3.7 [65]

Note. CERAD–K(N) = Korean version of the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery.
[a]Midpoint age = 67.
[b]Mean ± *SD*.
[c]Median (5–95 percentile).

CERAD–K(N). Two normative studies on the CERAD neuropsychological tests conducted in the United States (Ganguli et al., 1991; Welsh et al., 1994) also reported a similar impact of age on most cognitive tests. The CERAD–K(N) is almost equivalent to the English-American version of the CERAD neuropsychological assessment battery (J. Lee et al., 2002). Therefore, these findings suggest that, regardless of ethnic difference, the cognitive functions assessed by the CERAD neuropsychological tests gradually decline with advancing age even within the elderly. However, the cohort effect cannot be excluded because these studies, including ours, were not performed longitudinally.

With regard to education, these results suggest that a higher educational level is associated with a better performance on many of the CERAD–K(N) subtests. Then, performance on the verbal memory measures, particularly the Word List Recognition, appears to be relatively less affected by education. These findings are largely parallel to those reported in the other normative studies on the CERAD neuropsychological tests (Ganguli et al., 1991; Welsh et al., 1994). However, in contrast with Welsh et al.'s study (Welsh et al., 1994), which reported no association between education and the Verbal Fluency, this study observed that the Verbal Fluency was significantly affected by education. The apparent conflict between the two studies may be due to the range of the distribution of educational level. The range of educational level among our subjects was quite wide, from zero to 22 years, while most subjects in Welsh et al.'s study had more than 12 years of education. As indicated in our *post hoc* contrasts, the education effect appears to diminish

https://doi.org/10.1017/S1355617704101094 Published online by Cambridge University Press

Table 5. CERAD–K(N) normative data estimated from the ages 65–79[a] and used for the ages 70–74: Mean with standard deviation, and median and range from the 5th to the 95th percentile

		Education (years)		
Tests	Sex	0–3	4–6	≥7
Verbal Fluency	Male	12.6 ± 4.1[b]	13.2 ± 4.2	15.4 ± 4.6
		12.0 (6.0–22.8)[c]	14.0 (6.0–21.0)	15.0 (8.0–24.0)
	Female	12.4 ± 3.7	13.7 ± 3.6	14.4 ± 3.5
		12.0 (6.8–19.0)	14.0 (7.4–20.7)	13.5 (10.0–20.8)
Boston Naming Test	Male	9.8 ± 3.2	11.0 ± 2.6	11.5 ± 1.9
		9.5 (3.5–15.0)	11.0 (6.6–15.0)	12.0 (8.0–15.0)
	Female	7.5 ± 2.6	9.4 ± 2.6	11.0 ± 2.4
		7.0 (3.0–12.0)	9.0 (5.0–14.0)	11.0 (6.0–14.0)
MMSE–KC	Male	24.6 ± 3.3	26.9 ± 2.4	28.0 ± 1.7
		25.0 (17.5–29.1)	27.0 (23.0–30.0)	28.0 (24.1–30.0)
	Female	22.5 ± 3.9	26.2 ± 2.8	27.7 ± 1.9
		23.0 (15.0–28.0)	27.0 (20.5–30.0)	28.0 (24.2–30.0)
Word List Memory	Male	13.6 ± 3.8	15.0 ± 4.3	16.2 ± 4.2
		14.0 (6.0–11.0)	15.5 (8.0–23.0)	16.0 (8.0–23.0)
	Female	14.3 ± 4.3	16.7 ± 4.8	17.4 ± 3.2
		14.0 (6.5–21.0)	17.0 (10.0–22.5)	18.0 (12.3–22.0)
Constructional Praxis	Male	7.8 ± 1.5	9.4 ± 1.6	10.3 ± 1.1
		7.0 (6.0–11.0)	10.0 (7.0–11.0)	11.0 (8.0–11.0)
	Female	7.0 ± 2.2	9.4 ± 1.7	10.2 ± 1.2
		7.0 (3.0–11.0)	10.0 (6.0–11.0)	10.0 (7.0–11.0)
Word List Recall	Male	5.0 ± 1.9	5.6 ± 2.2	5.8 ± 2.0
		5.0 (2.0–8.0)	6.0 (2.0–9.5)	6.0 (2.0–9.0)
	Female	4.9 ± 1.8	5.8 ± 1.9	6.0 ± 2.1
		5.0 (2.0–8.0)	6.0 (2.0–8.0)	6.0 (2.0–9.0)
Word List Recognition	Male	8.8 ± 2.0	8.9 ± 1.5	8.8 ± 1.4
		10.0 (2.5–10.0)	9.5 (4.0–10.0)	10.0 (5.0–10.0)
	Female	8.8 ± 1.4	9.0 ± 1.6	9.0 ± 1.5
		9.0 (6.0–10.0)	10.0 (6.0–10.0)	10.0 (5.3–10.0)
Constructional Recall	Male	3.4 ± 2.3	6.4 ± 3.2	7.7 ± 2.6
		4.0 (0.0–8.7)	7.0 (0.6–11.0)	8.0 (4.0–11.0)
	Female	3.3 ± 2.2	5.2 ± 2.3	6.5 ± 2.7
		3.0 (0.0–7.0)	5.0 (2.0–9.0)	7.0 (2.0–11.0)
Age mean ± SD [sample size]	Male	72.9 ± 4.1 [30]	71.9 ± 4.2 [53]	71.1 ± 3.8 [81]
	Female	71.5 ± 4.0 [120]	71.1 ± 3.1 [131]	71.0 ± 3.8 [67]

Note. CERAD–K(N) = Korean version of the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery.
[a]Midpoint age = 72.
[b]Mean ± *SD*.
[c]Median (5–95 percentile).

within the highly educated group for most tests including Verbal Fluency.

The neuropsychological measures in the CERAD–K(N) were affected by gender in different ways. The performance scores on Constructional Recall were significantly higher in males, while better scores on the Word List Memory were associated with female gender. Although this was not statistically significant, females also scored better on the Word List Recall and Recognition tasks (Tables 4–7). These results are consistent with the observation that females usually perform better on verbal memory tasks and males perform better on spatial memory tasks (Lezak, 1995). With regard to the tasks with a visual recognition component such as the Boston Naming Test and Constructional Praxis, the scores of men were higher than those of women, while Welsh et al. reported no gender effects on those tasks in subjects with 12 or more years of education (Welsh et al., 1994). These differences between our results and those of Welsh et al. (1994) may be attributed to differences in the educational levels of the subjects. As shown in Table 3, a significant interaction exists between education and gender in the Boston Naming Test and Constructional Praxis: the score differences on these two tests between males and females were most prominent in the lowest education group (0–3 years), gradually diminishing in the higher education groups. As for MMSE–KC, which deals with multiple cognitive domains, men outperformed women, as did better educated subjects (Table 3). The Education × Sex inter-

https://doi.org/10.1017/S1355617704101094 Published online by Cambridge University Press

Table 6. CERAD–K(N) normative data estimated from the ages 70–84[a] and used for the ages 75–79: Mean with standard deviation, and median and range from the 5th to the 95th percentile

		Education (years)		
Tests	Sex	0–3	4–6	≥7
Verbal Fluency	Male	12.5 ± 4.2[b]	14.1 ± 4.3	14.6 ± 4.3
		11.0 (6.0–22.3)[c]	14.0 (7.0–21.0)	14.5 (7.9–24.0)
	Female	11.9 ± 3.8	13.1 ± 3.6	14.0 ± 3.5
		11.0 (6.0–19.0)	13.0 (7.0–19.0)	13.0 (10.0–22.0)
Boston Naming Test	Male	9.5 ± 3.3	10.5 ± 2.9	11.3 ± 2.0
		9.5 (3.6–15.0)	11.0 (5.0–15.0)	11.0 (8.0–15.0)
	Female	7.2 ± 2.9	9.4 ± 2.6	10.5 ± 2.5
		7.0 (3.0–13.0)	10.0 (5.0–14.0)	11.0 (6.0–14.0)
MMSE–KC	Male	24.5 ± 3.2	26.3 ± 2.7	27.9 ± 1.7
		25.0 (17.0–29.5)	27.0 (21.2–30.0)	28.0 (24.9–30.0)
	Female	21.2 ± 4.1	25.9 ± 2.9	27.0 ± 2.1
		22.0 (14.9–28.0)	26.0 (22.0–30.0)	27.0 (23.0–30.0)
Word List Memory	Male	13.2 ± 4.1	14.2 ± 4.6	15.7 ± 4.1
		13.5 (6.0–20.0)	14.0 (7.1–23.0)	15.5 (7.9–23.0)
	Female	13.7 ± 4.3	16.2 ± 4.2	17.0 ± 3.7
		14.0 (6.0–21.5)	17.0 (9.0–23.0)	18.0 (11.0–22.0)
Constructional Praxis	Male	8.0 ± 1.7	9.0 ± 1.8	10.2 ± 1.1
		7.0 (6.0–11.0)	9.0 (6.1–11.0)	11.0 (7.9–11.0)
	Female	6.4 ± 2.3	9.0 ± 1.9	9.9 ± 1.3
		6.0 (2.0–11.0)	10.0 (5.0–11.0)	10.0 (7.0–11.0)
Word List Recall	Male	4.5 ± 2.1	5.1 ± 2.3	5.4 ± 2.0
		5.0 (2.0–8.0)	5.0 (1.1–9.0)	6.0 (1.9–9.0)
	Female	4.7 ± 1.9	5.7 ± 2.0	5.9 ± 2.0
		5.0 (2.0–8.0)	6.0 (2.0–8.0)	5.0 (2.0–9.0)
Word List Recognition	Male	8.6 ± 2.2	8.5 ± 1.8	8.7 ± 1.6
		9.5 (2.4–10.0)	9.0 (4.0–10.4)	9.0 (5.9–10.0)
	Female	8.6 ± 1.6	8.7 ± 1.6	8.7 ± 1.6
		9.0 (5.5–10)	9.0 (5.0–10.0)	9.0 (5.0–10.0)
Constructional Recall	Male	3.2 ± 2.6	5.7 ± 3.3	7.4 ± 2.5
		4.0 (0.0–8.9)	5.0 (0.1–11.0)	7.5 (2.0–11.0)
	Female	2.8 ± 2.1	4.9 ± 2.3	6.2 ± 3.0
		2.8 (0.0–7.0)	5.0 (1.8–9.0)	6.0 (2.0–10.9)
Age mean ± SD [sample size]	Male	774 ± 4.3 [31]	770 ± 4.3 [43]	76.7 ± 3.8 [56]
	Female	76.6 ± 4.0 [97]	76.3 ± 3.2 [64]	76.8 ± 3.1 [42]

Note. CERAD–K(N) = Korean version of the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery.
[a]Midpoint age = 77.
[b]Mean ± *SD*.
[c]Median (5–95 percentile).

action reflects better performance by poorly educated male elderly, and might be explained by differences in their respective social roles. Male elderly, even those with little formal education, had more chances to obtain informal learning during their outside occupational activities, while poorly educated female elderly, who are usually devoted to housework, had fewer opportunities for intellectual stimulation. However, these gender effects appear to be less important in those who received at least 6 years of elementary education.

Thus, gender as well as age and educational level were found to have statistically significant effects on the performance of many subtests in CERAD–K(N). For studies such as this there are challenges in recruiting very large numbers of subjects to obtain reliable normative data for neuropsychological tests. In order to overcome this dilemma, the overlapping age tables proposed by Pauker (1988) were adopted. Using this procedure, we were able to present more stratified and more accurate normative data tables with adequate numbers of subjects for most normative cells. Normative data in each age-overlapping table were estimated from a broader age range (e.g., 65–79 for Table 5), which overlaps with an adjacent one. However, these norms can be applied to people within a narrower, non-overlapping age range (e.g., 70–74 for Table 5). The users of our norms can refer to the title of each normative table, which shows the age range from which norms were estimated, and the ranges to which they can be applied.

Table 7. CERAD–K(N) normative data estimated from the ages 75–90[a] and used for the ages 80–90: Mean with standard deviation, and median and range from the 5th to the 95th percentile

		Education (years)		
Tests	Sex	0–3	4–6	≥7
Verbal Fluency	Male	11.0 ± 3.8[b]	13.5 ± 4.2	14.0 ± 3.4
		11.0 (5.0–18.8)[c]	14.0 (6.1–21.3)	14.0 (9.6–20.4)
	Female	10.9 ± 3.6	12.2 ± 3.9	12.9 ± 3.6
		10.5 (5.0–18.5)	12.0 (6.0–20.0)	12.0 (8.00–21.5)
Boston Naming Test	Male	8.2 ± 3.0	9.9 ± 2.9	10.8 ± 2.2
		8.50 (2.3–13.0)	10.0 (4.4–14.7)	11.0 (6.2–14.0)
	Female	6.1 ± 2.8	9.3 ± 2.8	10.4 ± 2.3
		6.0 (2.0–11.0)	10.0 (4.5–14.0)	11.0 (6.0–14.0)
MMSE–KC	Male	23.7 ± 4.3	25.5 ± 2.9	27.2 ± 1.6
		24.5 (13.5–28.8)	25.0 (20.5–30.0)	27.0 (24.0–29.4)
	Female	19.5 ± 4.0	24.9 ± 3.0	26.2 ± 2.2
		19.0 (13.7–26.0)	24.0 (20.0–30.0)	26.0 (20.0–30.0)
Word List Memory	Male	13.1 ± 4.0	14.0 ± 4.8	15.4 ± 3.7
		13.5 (5.0–19.2)	14.0 (6.3–23.0)	15.0 (8.2–23.0)
	Female	12.9 ± 4.2	16.1 ± 4.6	16.4 ± 3.4
		13.0 (6.0–12.6)	17.0 (7.0–22.0)	17.0 (7.5–22.0)
Constructional Praxis	Male	8.0 ± 1.8	8.5 ± 1.8	9.9 ± 1.3
		7.0 (6.0–11.0)	8.0 (6.0–11.0)	10.0 (7.0–1.0)
	Female	5.7 ± 2.4	8.0 ± 1.8	9.6 ± 1.5
		6.0 (2.0–10.5)	8.0 (5.0–11.0)	10.0 (6.0–11.0)
Word List Recall	Male	4.4 ± 2.2	4.8 ± 2.4	4.9 ± 2.0
		4.5 (0.4–8.0)	5.0 (0.6–9.7)	5.0 (1.6–9.4)
	Female	4.5 ± 1.9	5.7 ± 2.5	5.8 ± 2.0
		5.0 (1.0–7.5)	6.0 (1.0–9.5)	6.0 (2.0–9.5)
Word List Recognition	Male	8.4 ± 2.3	8.2 ± 1.7	8.3 ± 2.0
		9.0 (2.0–10.0)	9.0 (4.0–10)	9.0 (4.0–10.0)
	Female	8.4 ± 1.8	8.6 ± 1.7	8.5 ± 2.5
		9.0 (5.0–10.0)	10.0 (5.0–10.0)	10.0 (5.0–10.0)
Constructional Recall	Male	3.1 ± 2.6	4.3 ± 2.9	7.1 ± 2.9
		2.5 (0.0–8.8)	4.0 (0.0–11.0)	7.5 (1.1–11.0)
	Female	2.2 ± 2.0	4.8 ± 2.4	5.7 ± 2.9
		2.0 (0.0–6.7)	5.0 (0.0–9.0)	5.5 (1.0–9.0)
Age mean ± SD [sample size]	Male	83.0 ± 3.6 [25]	82.4 ± 3.9 [28]	82.3 ± 4.4 [31]
	Female	82.0 ± 4.3 [72]	81.9 ± 3.9 [22]	81.7 ± 3.8 [21]

Note. CERAD–K(N) = Korean version of the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery.
[a]Midpoint age = 82.
[b]Mean ± *SD*.
[c]Median (5–95 percentile).

Even when adopting the overlapping age tables, the size of some cells in the normative tables is still quite small (e.g., 20 in the cell for males with an educational level of 0–3 years in Table 4 and 21 for females with an educational level ≥7 years in Table 7). A relatively small sample size increases the standard error and possibly reduces the stability of the estimated results. Therefore, users should be more cautious when interpreting the test scores of individuals within such cells.

The higher education stratum (i.e., ≥7 years) of our norm encompasses a very wide range of educational levels (7–22 years). As the results from the *post-hoc* contrasts between the higher educational groups (7–9, 9–12, and ≥13 year group) suggested a diminished education effect at ≥7 year for most tests in this sample, we combined these groups into one. However, the *post-hoc* contrasts may not completely exclude the possibility of differences between the ends of this stratum. Welsh et al.'s study, which was conducted on highly educated subjects, also reported a positive association of education with some of the CERAD neuropsychological tests. In this study, the mean and standard deviation for educational level of subjects with a ≥7 year educational level were 11.6 and 3.2 years, respectively. Therefore, some caution should be taken when interpreting the test results of individuals with extremely high educational levels.

These results were obtained from a healthy elderly sample. This study excluded the people with dementia and other

https://doi.org/10.1017/S1355617704101094 Published online by Cambridge University Press

major illnesses that would affect the cognitive function through a clinical examination. Elderly populations usually include a considerable proportion of patients with dementia and severe medical conditions. The prevalence rate of dementia is approximately 8.2–9.5% in Korea (D. Lee et al., 2002; Woo et al., 1998) and 6.3–10.3% in Western countries (Canadian Study of Health and Aging Working Group, 1994; Evans et al., 1989; Livingston et al., 1990). For this reason, we offered norms from a relatively healthy elderly sample since the aim was to define the range of normality which might then be helpful in detecting dementia and other problems, such as *aging-associated cognitive decline* (AACD; Levy, 1994),

In summary, both age and educational level had a considerable influence on most cognitive functions assessed by the CERAD neuropsychological battery, though education appeared to have a lesser effect on verbal memory than on the other cognitive functions. Gender also had an effect on the performance of several subtests of the CERAD neuropsychological battery, particularly in the less educated group. The normative information presented in this article might be useful for the clinical interpretation of the CERAD neuropsychological tests in the Korean elderly or the elderly who have emigrated from Korea. These data may also be useful in future cross-cultural comparisons of the CERAD battery.

ACKNOWLEDGMENTS

This work was supported by Grant No. 04-2000-045 from the Seoul National University Hospital and Biotech 2000 (Grants 97-N1-02-03-A-12 and 98-N1-02-03-A-12) of the Ministry of Science and Technology of Korea.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Canadian Study of Health and Aging Working Group. (1994). Canadian Study of Health and Aging: Study methods and prevalence of dementia. *Canadian Medical Association Journal*, *150*, 899–913.

Escobar, J.I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J., & Golding, J.M. (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity. Cultural and linguistic artifacts. *Journal of Nervous and Mental Disease*, *174*, 607–614.

Evans, D.A., Funkenstein, H., Albert, M.S., Scherr, P.A., Cook, N.R., Chown, M.J., Hebert, L.E., Hennekens, C.H., & Taylor, J.O. (1989). Prevalence of Alzheimer's disease in a community population of older persons: Higher than previously reported. *Journal of the American Medical Association*, *262*, 2551–2556.

Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.

Folstein, M.F., Folstein, S.F., & McHugh, P.R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Ganguli, M., Ratcliff, G., Huff, F.J., Belle, S., Kancel, M.J., Fischer, L., Seaberg, E.C., & Kuller, L.H. (1991). Effects of age, gender, and education on cognitive tests in a rural elderly community sample: Norms from the Monongahela valley independent elders survey. *Neuroepidemiology*, *10*, 42–52.

Kim, H.H. & Na, D.L. (1997). *Korean version–Boston Naming Test*. Seoul, Korea: Hak Ji Sa.

Kittner, S.J., White, L.R., Farmer, M.E., Wolz, M., Kaplan, E., Moes, E., Brody, J.A., & Feinleib, M. (1986). Methodological issues in screening for dementia: The problem of education adjustment. *Journal of Chronic Disease*, *39*, 163–170.

Lee, D.Y., Lee, J.H., Ju, Y.S., Lee, K.U., Kim, K.W., Jhoo, J.Y., Yoon, J.C., Ha, J., & Woo, J.I. (2002). The prevalence of dementia among the elderly in an urban population of Korea: The Seoul study. *Journal of the American Geriatric Society*, *50*, 1233–1239.

Lee, D.Y., Yoon, J.C., Lee, K.U., Jhoo, J.H., Kim, K.W., Lee, J.H., & Woo J.I. (1999). Reliability and validity of the Korean version of Short Blessed Test (SBT–K) as a dementia screening instrument. *Journal of Korean Neuropsychiatric Association*, *38*, 1297–1307.

Lee, J.H., Lee, K.U., Lee, D.Y., Kim, K.W., Jhoo, J.H., Kim, J.H., Lee, K.H., Kim, S.Y., Han, S.H., & Woo, J.I. (2002). Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD–K): Clinical and Neuropsychological Assessment Batteries. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*, 47–53.

Levy, R. (1994). Aging-associated cognitive decline. *International Psychogeriatrics*, *6*, 63–68.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Livingston, G., Sax, K., Willison, J., Blizard, B., & Mann, A. (1990). The Gospel Oak Study, Stage II: The diagnosis of dementia in the community. *Psychological Medicine*, *20*, 881–891.

Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., van Belle, G., Fillenbaum, G., Mellits, E.D., Clark, C., & the CERAD investigators. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*, 1159–1165.

Park, J.H. & Kwon, Y.C. (1991). Modification of the mini-mental state examination for use with the elderly in a non-western society. Part I: Development of Korean version of mini-mental state examination. *International Journal of Geriatric Psychiatry*, *6*, 875–882.

Pauker, J.D. (1988). Constructing overlapping cell tables to maximize the clinical usefulness of normative data: Rationale and an example from neuropsychology. *Journal of Clinical Psychology*, *44*, 930–933.

Rosen, W.G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology*, *2*, 135–146.

Welsh-Bohmer, K.A. & Mohs, R.C. (1997). Neuropsychological assessment of Alzheimer's disease. *Neurology*, *49(Suppl. 3)*, S11–S13.

Welsh, K., Butters, N., Mohs, R.C., Beekly, B.S., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Part V: A normative study of the neuropsychological battery. *Neurology*, *44*, 609–614.

Woo, J.I., Lee, J.H., Yoo, K.Y., Kim, C.Y., Kim, Y.I., & Shin, Y.S. (1998). Prevalence estimation of dementia in a rural area of Korea. *Journal of the American Geriatrics Society*, *46*, 983–987.