Age, Gender, and Education May Have Little Influence on Error Patterns in the Assessment of Set-shifting and Rule Induction Among Normal Elderly

Patricia A. Lowe and Cecil R. Reynolds

Department of Educational Psychology, Texas A&M University

Error patterns have been found to be sensitive to cognitive status, but the relationship between aging and error patterns remains unclear, and may differ as a function of gender, education, and whether a task is verbal or nonverbal. The present study examined the error patterns of normal elderly individuals on a verbal measure of set-shifting and rule induction to determine whether demographic variables, that is, age, gender, and education, influenced test performance. The sample of 109 individuals, 38 males and 71 females, ranging in age from 54 to 89 years with 6 to 19 years of education, was assessed on the Classification subtest of the Test of Verbal Conceptualization and Fluency, a verbal measure of set-shifting and rule induction. Subjects’ protocols were scored for perseverative, nonperseverative, and random errors, tabulated, and analyzed. Multivariate analysis of covariance with education as the covariate as well as other statistical tests revealed nonsignificant relationships between error scores and age, gender, and education. Years of education, however, showed a significant correlation with a reduction in random responses. Results are interpreted based on Horn’s (1978) fluid-crystallized explanations of changes in intelligence with advancing age. © 1999 National Academy of Neuropsychology. Published by Elsevier Science Ltd

The frontal lobes are highly complex structures that perform a wide variety of cognitive functions. “The multiplicity of functions involved in frontal lobe control and regulation make it unlike the parietal, temporal, and occipital lobes, which tend to be more associated with direct brain–behavior relationships” (Bigler, 1988, p. 279). Damage to the parietal, temporal, or occipital lobes often leads to specific clinical disorders associated with more specific localized lesion sites as well as signs on neuropsychological measures. In contrast, damage to the frontal lobes often lead to a wide spectrum of cognitive and behavioral changes that do not conform to specific clinical syndrome entities, and therefore, are more difficult if not impossible to assess with specific neuropsychological tests (Bigler,
Clinical neuropsychologists have sought tests of frontal-lobe damage that are specifically sensitive to frontal-lobe function or executive dysfunction, however, the search so far has proved very difficult (Bigler, 1988; Heck & Bryer, 1986; Wang, 1987).

Adults with frontal lobe lesions have been found to exhibit a wide-spectrum of cognitive and behavioral impairments. In Riccio’s et al. (1994) review, the authors reported that impairments in planning, set maintenance, and cognitive flexibility have been found. Disinhibition, perseveration (Petrides & Milner, 1982), and inability to read environmental cues to direct behavior (Passler, Isaac, & Hynd, 1985) have also been reported. “Thus, the behavior results of prefrontal damage may include impairment of sequential behavior, inability to establish and maintain a set, inability to use environmental cues” (Riccio et al., 1994, p. 216), and a tendency to repeat past mistakes (Stuss, 1992; Stuss & Benson, 1984, 1989).

Along with the difficulty encountered in assessing frontal-lobe function or executive dysfunction, there is a paucity of information with respect to the influence of demographic factors on neuropsychological test performance (Boone, Ghaffarian, Lesser, Hill-Gutierrez, & Berman, 1993; Golden, 1981; Reynolds, 1997). This seems to be the norm rather than the exception in neuropsychology, that is, research has neglected the influence of demographic variables on neuropsychological test results (e.g., Reynolds, 1997; Golden, 1981). Demographic variables, such as race or ethnicity, gender, and/or education, can have an affect on test performance, and knowledge about the impact of these variables is not only critical to the researcher, but also to the clinician’s interpretation of test results (Reynolds, 1997). Systematic effects of demographic variables have been noted on numerous tasks, such as the Coding and Digit Symbol subtests of the Wechsler scales (some of the most sensitive subtests to neurological trauma), where females have been reported to outperform males (Reynolds, 1997). Whether using a level of performance or an ipsative profile analysis (e.g., Davis, 1959; Reynolds & Gutkin, 1980), ignorance of such robust findings could be misleading not only to the researcher but also the clinician (Reynolds, 1997). For intelligence and personality tests, a body of research on the effects of demographic and other nominal variables has been accumulating (e.g., Reynolds, 1995), however, very little research has addressed the influence of demographic variables on strictly neuropsychological test results, especially among the elderly (Boone, Ghaffarian, Lesser, Hill-Gutierrez, & Berman, 1993).

Age is one variable that may account for the differential outcomes in measures of frontal-lobe performance. Age-related declines in abilities subserved by the frontal lobes have been one of the least researched areas in neuropsychology, despite the fact that some investigators have hypothesized that declines in frontal-lobe abilities may be the primary cognitive deficits associated with aging (Albert & Kaplan, 1980; Hochanadel & Kaplan, 1984; Libon et al., 1994; Veroff, 1980; Whelihan & Lesher, 1985). A few investigators have reported age-related deficits among normal elderly on several frontal-lobe measures in such domains as serial digit learning (Benton, Eslinger, & Damasio, 1981), cued response inhibition (Comalli, Wapner, & Werner, 1962), and cognitive flexibility (Heaton, 1981; Heaton, Grant, & Matthews, 1991; Loranger & Misiak, 1960; Nelson, 1976; Spreen & Strauss, 1991), however, results have not been universally replicated (Boone, Miller, Lesser, Hill, & D’Elia, 1990; Uchiyama, Mitrushina, D’Elia, Satz, & Matthews, 1994; Willis, Yeo, Thomas, & Garry, 1988). For example, on nonverbal measures of cognitive flexibility and rule induction, Heaton (1981) reported a decline in test performance in normal subjects 60 years and older in comparison to younger controls on several Wisconsin Card Sorting Test (WCST; Heaton, 1981) variables, including a greater number of total and perseverative errors. Loranger and Misiak (1960) also found age-related declines in performance on the WCST among the elderly. Subjects, aged 74 to
80, showed a significant increase in perseverative and total errors. Similarly, Spreen and Strauss (1991) found a significant decline in WCST performance among 60 nonimpaired subjects, ranging in age from 60 to 94 years, however, the authors reported no age-related declines until after age 80. Likewise, Haaland, Vranes, Goodwin, and Garry (1987) failed to document lowered scores on the WCST in a sample of healthy individuals aged 64 to 80. The subsample or group composed of individuals more than 80 years of age, unlike the Spreen and Strauss (1991) findings, demonstrated no increase in perseverative errors, but showed only an increase in total errors. In contrast, Boone et al. (1990) and Willis et al. (1988) reported no age-related declines in test performance using similar measures of cognitive flexibility. Boone et al. (1990) assessed older, nonimpaired individuals, ranging in age from 50 to 79 years, on the WCST, whereas Willis et al. (1988) assessed 154 healthy subjects, aged 65 to 79, on the Halstead Category Test (HCT; Reitan & Wolfson, 1985). These findings seem to contradict the work of Heaton (1981), Heaton et al. (1991), and Loranger and Misiak (1960) and suggest that the relationship between aging and error types or patterns remains unclear (e.g., Boone et al., 1990).

Error types or patterns, for example, perseverative errors, nonperseverative errors, and random errors have been found to be sensitive to cognitive status (e.g., Heaton et al., 1991), but the relationship between aging and error patterns as previously mentioned remains unclear (e.g., Boone et al., 1990), and may differ as a function of gender (Kaufman, 1990; Kaufman, Kaufman-Packer, McLean, & Reynolds, 1991; Kaufman, McLean, & Reynolds, 1988; Reynolds, Chastain, Kaufman, & McLean, 1987), and whether the task is verbal or nonverbal (Kaufman, 1990; Kaufman et al., 1991; Kaufman et al., 1988; Reynolds et al., 1987). Research has shown that performance by adults on nonverbal cognitive measures, such as the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) Performance IQ scale, whether differences in educational attainment are controlled or not, declines with age, beginning in the mid-30s, but is not pronounced until the 50s or 60s (e.g., Kaufman, 1990). The WCST and the HCT are two examples of nonverbal measures of set-shifting and rule induction. Heaton (1981) and Heaton’s et al. (1991) findings of significant declines in performance for subjects 60 years and older are similar to Kaufman’s (1990) findings of age-related declines on the WAIS-R Performance IQ scale. Research has also shown that performance by adults on verbal cognitive measures, such as the WAIS-R Verbal IQ scale, controlling for educational differences, improves slightly or is maintained throughout adulthood (Kaufman, 1990; Reynolds et al., 1987). Whether these findings would have some bearing on a verbal measure of set-shifting and rule induction is not known.

In addition to type of task, verbal or nonverbal, gender differences have been reported on cognitive measures (Kaufman, 1990; Kaufman et al., 1991; Kaufman et al., 1988; Reynolds et al., 1987). Research has consistently found that adult males outperform adult females on more global verbal and nonverbal cognitive measures, such as the WAIS-R Verbal and Performance IQ scales (Kaufman, 1990; Kaufman et al., 1991; Kaufman et al., 1988; Reynolds et al., 1987). Gender differences have also been noted on numerous individual cognitive tasks, such as the Coding and Digit Symbol subtests of the Wechsler scales, where females have been reported to outperform males (Kaufman et al., 1991; Reynolds, 1997). These findings indicate relationships between age or gender and type of task, verbal or nonverbal, do exist and may have some bearing on the relationship between aging and error patterns on measures of set-shifting and rule induction.

Another variable that may account for the differential outcomes in measures of set-shifting and rule induction among the elderly is education. In samples or subsamples composed of elderly individuals, older cohorts frequently have received less formal edu-
cation than their younger cohorts (Uchiyama et al., 1994). Thus, differences in education, age, or the relationship between education and age may account for differences in test score or error pattern performance. For example, education has been demonstrated to be associated with age on verbal cognitive measures, such as the WAIS-R Verbal IQ scale (Kaufman, 1990; Kaufman, Reynolds, & McLean, 1989; Reynolds et al., 1987). Kaufman et al. (1989) found when educational attainment levels were held constant (weighting for education) on the WAIS-R standardization sample of 1,480 adults, ranging in age from 20 to 74 years, the decline in Verbal IQ from age to age disappeared. However, declines across the 20- to 74-year range remained for the Performance IQ and Full Scale IQ, even after controlling for education, although these declines were not as substantial as they were before the education variable was held constant. According to Uchiyama et al. (1994), when groups evidence significant differences, the educational variable should be controlled. Few studies have examined the relationship between education and error pattern performance on measures of set-shifting and rule induction among the elderly (Beatty, 1993; Boone et al., 1993; Heaton, 1981; Heaton, 1993), and even fewer studies have held the educational variable constant when groups evidence significant differences (Heaton, 1993; Uchiyama et al., 1994). Boone et al. (1993), Heaton (1981), and Heaton (1993) found a significant relationship between education and error pattern performance among the elderly, but the studies varied as to when educational differences emerged, whereas Beatty (1993) found no significant relationship between education and error pattern performance. Additional research is needed to help clarify the relationship between education and error pattern performance among the normal elderly on measures of set-shifting and rule induction.

Education, gender, and age have been shown to influence test performance by adults on verbal and nonverbal cognitive measures, such as the WAIS-R Verbal and Performance IQ scales (e.g., Kaufman, 1990). However, it remains unclear as to whether these same variables influence error pattern performance among normal elderly on measures of set-shifting and rule induction. It also remains unclear as to whether type of task, that is, verbal or nonverbal, would have some bearing on the relationship between aging and error pattern performance. Presently, there are no published studies or measures of set-shifting and rule induction in a verbal format. The Classification subtest of the Test of Verbal Conceptualization and Fluency (TVCF) is a purported verbal measure of set-shifting and rule induction.

The TVCF developed by Reynolds and Haak (1994) is a neuropsychological measure designed to assess possible anterior (frontal-lobe) cerebral dysfunction. The TVCF is composed of four subtests, including Categorical Fluency, Letter Naming, Trails C, and the Classification subtest. The Categorical Fluency subtest is a measure of verbal fluency. The subtest assesses word retrieval by conceptual category and fluency of ideation. Subjects are asked to name as many objects as possible in a given domain, for example, “things to eat,” as quickly as possible. Geriatric subjects are allowed 60 seconds per category. Raw scores are summed across five conceptual categories. The Letter Naming subtest is also a measure of verbal fluency. The subtest assesses word retrieval by initial sound and fluency of ideation. Examinees are asked to name as many words that begin with a specific letter, for example, “s,” as quickly as possible. Geriatric subjects are allowed 60 seconds per letter. Raw scores are summed across four letters. The Trails C subtest is a measure of visual perception. The subtest assesses visual and mental tracking capabilities by shifting the sequence of Arabic numerals and linguistic representations of numbers. Subjects are required to draw a line connecting the numbers and linguistic representations of numbers in numerical order as quickly as possible. This subtest is timed. The time needed to complete the task and the number of errors made are recorded. The Classification subtest is a measure of set-shifting and rule induction in a verbal format.
The subtest assesses problem solving by the generation and discarding of categorization strategies. The Classification subtest consists of four standard cards and a response deck of 116 cards. Each card has a printed word varying in appearance from one to four times per card. Subjects are required to match each card in the response deck to one of the four standard cards based on three different sorting rules, e.g., matching the cards according to color. Examinees are required to make eight consecutive matches in a row, known as a category, before the matching principle changes. Examinees are never told the matching principle. Individuals only receive feedback as to whether their responses are right or wrong on each trial. The test ends after 116 trials.

Several scores can be generated from this procedure on the TVCF Classification subtest: Categories Completed, the number of times a subject scores eight consecutive correct items; Trials to the First Category, the number of trials required to attain the first category; Failure to Maintain a Set, the number of times a subject breaks a sequence of consecutive correct items; Perseverative Errors, the number of times a subject continues to sort the cards according to the previously successful rule; Nonperseverative Errors, the number of errors a subject makes sorting the cards along an incorrect dimension, which is not the same as the previous correct response; Random Errors, the number of errors a subject makes when none of the categories match the sort; Total Errors, the sum of perseverative, nonperseverative, and random errors; and Total Correct, the number of correct responses. These scores are objective measures and provide information on a subject’s successes as well as failures that allow interpretation of the process used to achieve a specific level of performance. The purpose of this study was to analyze the error patterns in the data on the TVCF’s Classification subtest for 54- to 89-year-old adults, according to their gender, age, and education, to assist in interpreting performance on such tasks, and the influence of those named variables on cognitive decline in normal aging.

**METHOD**

**Subjects**

Subjects for the current study consisted of 109 individuals, 38 males and 71 females. The subjects ranged in age from 54 to 89 years and reported no major learning, emotional, physical, or neuropsychological disorders. The mean age of males was 74.05 years (SD = 8.01 years) and females was 73.56 years (SD = 7.48 years). No significant gender difference in age was noted, t(107) = .32, p > .05. These individuals completed 6 to 19 years of education with males completing an average of 14.39 years (SD = 3.15 years) and females completing an average of 12.56 years (SD = 2.85 years). A significant gender difference in educational attainment was found in the sample with males completing more years of schooling than females, t(107) = 3.12, p < .05. The ethnic composition of the sample included Whites (83.8%), Blacks (9.0%), Hispanics (2.7%), and others (4.5%). Residents from five states, Idaho, Nebraska, Nevada, Oregon, and Texas, and the Canadian province of Alberta were sampled.

This sample represents the entire subsample of elderly individuals assessed to date and will be part of the standardization sample of over 1000 individuals tested to develop the norms for the TVCF. The TVCF norming sample will range in age from 9 to 90 years.

**Instrument**

The Classification subtest, one of the four subtests, of the TVCF was used. Studies addressing the TVCF’s psychometric properties are currently under investigation.
Procedure

An error pattern analysis was conducted on the Classification subtest for each of the 109 older subjects in the norming sample. Protocols were scored for perseverative, nonperseverative, and random errors. Perseverative errors occur when a subject continues to sort the cards according to the previous successful rule. Nonperseverative errors are scored when the subject sorts the cards along an incorrect dimension, and this dimension was not the same as the previous correct response but may represent a match on a future rule or a rule once removed. Random errors occur when none of the categories match the sort. The three error scores were tabulated for each protocol and used for analyses. The analyses were performed using the 1996 version of SPSS, release 7.5.

RESULTS

Education as a Covariate

Research has suggested educational attainment of adults differ as a function of age, and these differences influence neuropsychological test scores (Stratta et al., 1993; Uchiyama et al., 1994). Therefore, a Pearson product–moment correlation coefficient was calculated between age and education. A significant negative correlation between age and education was found, $r(107) = -0.20, p < .05$. This suggests cohort differences in educational attainment exist, that is, adults born in more recent years have a greater number of years of education. This finding seems logical, as advanced educational opportunities were less available to the elderly population.

Since a significant correlation between age and education was found, three Pearson product–moment correlation coefficients were then calculated between education and error pattern scores to determine if education is related to perseverative, nonperseverative, or random errors. A significant negative correlation between education and random errors was found, $r(107) = -0.21, p < .05$. This finding indicates an increase in educational attainment is associated with a decrease in random errors. This suggests that individuals, who have more years of schooling, are more likely to make one of two matching rule errors, that is, perseverative or nonperseverative errors, when they err, rather than making wild guesses or matches based on rules not relevant to the Classification subtest. Correlation coefficients between education and perseverative errors, and education and nonperseverative errors were found to be nonsignificant, $r(107) = -0.08, p > .05$, and $r(107) = .01, p > .05$, respectively.

Age and Gender Differences with Education as a Covariate

To control for educational differences and to determine the potential effects of age and gender differences on the dependent variables (i.e., perseverative, nonperseverative, and random errors), subjects were stratified by age, 54 to 69 and 70 to 89 years, and a $2 \times 2$ (Age × Gender) multivariate analysis of covariance (MANCOVA), using education as the covariate, was performed. The General Linear Model (GLM) multivariate analysis of variance procedure with education as the covariate was used. The GLM model was selected over more conventional multivariate analysis procedures because of its correction for unequal cell sizes. Prior to computing the $2 \times 2$ MANCOVA, Box’s M test of equality of covariance matrices, Bartlett’s test of sphericity, tests of multicollinearity, and an overall test of homogeneity of regression were performed. Box’s M test was not significant, $F(18, 8436) = 1.24, p > .05$, confirming homogeneity of variance-
covariance matrices. On the other hand, Bartlett’s test of sphericity was significant, $\chi^2(5) = 113.06, p = .00$, and lead to the rejection of the hypothesis that the correlation matrix is an identity matrix. There was no evidence of multicollinearity, as tolerance statistics for the dependent variables ranged from .56 to .98. Furthermore, an overall test of homogeneity of regression proved to be nonsignificant, using four separate estimates (Wilks’ Lambda, Pillai’s Trace, Hotelling’s Trace, and Roy’s Largest Root). None of the $F$ values achieved a $p$ value below .66. The MANCOVA $F$s were then computed. Means and standard deviations are depicted in Table 1.

The MANCOVA $F$s were not significant for age or gender main effects or age by gender interaction effect controlling for educational differences using four separate estimates (Wilks’ Lambda, Pillai’s Trace, Hotelling’s Trace, and Roy’s Largest Root). None of the $F$ values achieved a $p$ value below .73. Large within group variances were observed in the analyses, but these variances are comparable to the within group variances reported on similar tests (e.g., the WCST) with elderly samples (Davis et al., 1990).

**Spurious Relationships**

Second-order partial correlation coefficients were then calculated between each independent variable (i.e., age, gender, and education) and dependent variable with the other two independent variables being held constant to detect possible spurious relationships. For example, a second-order partial correlation between age and perseverative errors was calculated, while holding gender and education constant. Table 2 presents the second-order partial correlation coefficients. The second-order partial correlation coefficients ranged from $-0.16$ to $0.08$. None of the second-order partial correlation coefficients were found to be significant. The second-order partial correlation coefficients were able to explain only $0.64$ to $2.7\%$ of the variance in test performance.

**Analysis of Variance Tests**

Since no significant correlations between the independent and dependent variables were observed, three $2 \times 2 \times 2$ (Age $\times$ Gender $\times$ Education) analysis of variance tests (ANOVAs) were planned. However, the number of subjects in several cells was too small

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Means ($M$) and Standard Deviations ($SD$) of Error Types by Gender and Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Errors</td>
<td>$M$</td>
</tr>
<tr>
<td>Perseverative</td>
<td>26.92</td>
</tr>
<tr>
<td>Nonperseverative</td>
<td>22.75</td>
</tr>
</tbody>
</table>
to perform the three-way ANOVAs. This is a common problem in neuropsychological research (Reynolds, 1997). Therefore, a series of univariate F-tests were performed.

Age and gender differences. To identify possible age levels where differences in the Classification subtest performance might occur, univariate F-tests were calculated on the subtest variables for two age groups: 54 to 69 and 70 to 89 years. The GLM General Factorial Analysis of Variance procedure was selected to perform the computations because it corrects for unequal cell sizes and can be used with one or more factors. In addition, Levene’s test of equality of error variances, a test to determine whether the variances are equivalent by groups, was computed prior to the calculation of each univariate F-test. No violations of homogeneity of variance were found for any of the comparisons. Moreover, Bonferroni adjustments were planned due to the number of univariate F-tests; however, none of the univariate F-tests achieved a p-value equal to or less than .05. In addition, separate analyses by age were performed for gender due to the reported differences in years of schooling. Means, standard deviations, and ANOVA results are presented in Table 3 for the total sample and male and female subsamples.

Although slight increments in the number of perseverative errors were noted with an increase in age for the total sample, as well as for both sexes, these increases were not significant. Likewise, nonsignificant differences were found between age groups and the number of nonperseverative and random errors for the total sample and for both sexes. Even though the results were nonsignificant, it is interesting to note the increasing trends in both nonperseverative and random errors in males with an increase in age, whereas, females showed declining trends in both the number of nonperseverative and random errors with an increase in age.

Education and gender differences. To examine the relationship between education and error pattern performance, two education groups were compared on the Classification subtest variables: 6 to 14 years and 15 to 19 years. These educational levels were used due to the sample’s demographics and the need to have an adequate number of subjects per cell. The GLM Factorial Analysis of Variance procedure was used in computing the univariate F-tests. Prior to computing each univariate F-test, Levene’s test of equality of error variances was performed. No violations of homogeneity of variance were found. In addition, Bonferroni adjustments were planned. Moreover, separate analyses by education were performed for gender due to the reported differences in educational attainment. Means, standard deviations, and ANOVA results are depicted in Table 4.

Group comparisons for the total sample, and male and female subsamples revealed nonsignificant differences between the educational groups on perseverative, nonperseverative, and random errors. Although the results were not significant, it was interesting to note the declining trend in random errors for the entire sample with an increase in ed-

<table>
<thead>
<tr>
<th></th>
<th>Perseverative</th>
<th>Nonperseverative</th>
<th>Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.0850</td>
<td>−.0154</td>
<td>.0802</td>
</tr>
<tr>
<td>Gender</td>
<td>.0151</td>
<td>−.1052</td>
<td>.0589</td>
</tr>
<tr>
<td>Education</td>
<td>−.0518</td>
<td>−.0220</td>
<td>−.1640</td>
</tr>
</tbody>
</table>
### TABLE 3
Means (M) and Standard Deviations (SD) of Error Types by Age (Years)

<table>
<thead>
<tr>
<th>Errors</th>
<th>Total (n = 109)</th>
<th>Female (n = 91)</th>
<th>Male (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>54–69</td>
<td>70–89</td>
<td>54–69</td>
</tr>
<tr>
<td></td>
<td>(n = 30)</td>
<td>(n = 79)</td>
<td>(n = 18)</td>
</tr>
<tr>
<td>Perseverative</td>
<td>27.10 ± 11.74</td>
<td>29.67 ± 16.36</td>
<td>28.00 ± 11.93</td>
</tr>
</tbody>
</table>

### TABLE 4
Means (M) and Standard Deviations (SD) of Error Types by Education (Years)

<table>
<thead>
<tr>
<th>Errors</th>
<th>Total (n = 109)</th>
<th>Female (n = 91)</th>
<th>Male (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6–14</td>
<td>15–19</td>
<td>6–14</td>
</tr>
<tr>
<td></td>
<td>(n = 60)</td>
<td>(n = 49)</td>
<td>(n = 13)</td>
</tr>
<tr>
<td>Perseverative</td>
<td>30.70 ± 16.05</td>
<td>26.84 ± 14.02</td>
<td>31.21 ± 14.84</td>
</tr>
<tr>
<td>Nonperseverative</td>
<td>21.82 ± 17.32</td>
<td>24.22 ± 17.19</td>
<td>20.08 ± 14.41</td>
</tr>
<tr>
<td>Random</td>
<td>9.20 ± 10.15</td>
<td>5.92 ± 6.87</td>
<td>8.88 ± 10.42</td>
</tr>
</tbody>
</table>
ucational attainment, as well as, the declining trend in perseverative errors among females but not males with an increase in years of schooling.

**DISCUSSION**

This paper is a preliminary report on the influence of certain demographic variables on error patterns in the assessment of set-shifting and rule induction among normal elderly, and the influence of those named variables on cognitive decline in normal aging. The findings suggest that age, gender, and education have little influence on error patterns in a verbal task requiring elderly individuals to generate and discard categorization strategies, beyond the influence of education in the reduction of random responses.

**Age Differences**

No age differences were found in the current study. Overall, the Classification subtest, which is a verbal task, may require individuals to tap into their accumulated cultural knowledge base to solve new and familiar problems. Therefore, according to Horn’s (1978) fluid-crystallized theory of intelligence, both crystallized intelligence and fluid intelligence are assessed. Crystallized intelligence, as a function of accumulating cultural knowledge throughout one’s lifetime, tends to increase or is maintained across the adult years. In contrast, fluid intelligence, associated with new problem-solving and incidental learning and directly related to physiological functioning and neurological integrity, declines throughout adulthood. Subjects in the current study were educated individuals who reported no learning, physical, neurological, or emotional problems, and therefore, a significant decline in performance would not be expected with an increase in age based on Horn’s fluid-crystallized theory of intelligence, and conceptualization of the Classification subtest task as a measure of crystallized intelligence with fluid-like properties.

**Gender Differences**

Similarly, no gender differences were indicated in the current study. Research has shown that males tend to score consistently higher, 1 to 2 points, on verbal and nonverbal cognitive measures, for example, the Wechsler scales, but these differences have been shown to be of no practical significance (Kaufman, 1990). Unimportant differences often emerge as statistically significant differences due to large sample sizes. Reynolds et al. (1987) analyzed five stratification variables, gender, race, region, residence, and education level, on the WAIS-R to identify those variables that were statistically significant but trivial. Separate ANOVAs for Verbal, Performance, and Full Scale IQs and the stratification variables were calculated. ANOVA results produced significant main effects for race and education, but the variables of gender, region, and residence failed to reach significance at the .05 level. Similar results for gender have been reported on the fourth edition of the Stanford-Binet (SB: FE; Thorndike, Hagen, & Sattler, 1986).

**Education Differences**

Likewise, no education differences were found in the current study, with the exception of a significant negative correlation between years of schooling and random responses. These findings are surprising. However, research has shown that differences between the most and least educated samples decrease with an increase in age on cognitive
measures (Kaufman, 1990). In addition, considerable variability in individual differences have been found on cognitive measures for individuals with the same educational attainment (Kaufman, 1990). In the current study, large within group variability was noted and is a common finding among the normal elderly population (Davis et al., 1990). These differences represent real individual differences that are not artifacts of the testing procedure but do provide appropriate explanations for test results (Mayfield & Reynolds, 1997). Mayfield and Reynolds strongly advise researchers to treat individual differences as individual variation not error variance, and clinicians, especially in neuropsychological settings, to seek alternative approaches and hypotheses to their test results when individual variation exists. Assessment of the individual and interpretations that fit the individual assessed are what is important.

The findings of this study provide some suggestions for future research. First, inclusion of nongeriatric normal groups as controls, for example, children, adolescents, and young and middle-aged adults are needed to empirically demonstrate the increase, maintenance, or possible decline in performance with an increase in age. In the current study, the sample ranged in age from 54 to 89 years, although unlikely, a range restriction in age may have been present. A study assessing the error patterns of both geriatric and nongeriatric groups may help clarify as to whether a linear or nonlinear (e.g., curvilinear) relationship exists across the age spectrum. Second, although a hefty undertaking, a study exploring the influence of demographic variables (i.e., age, gender, education, and ethnicity) on error pattern performance among the elderly using a large systematically stratified standardization sample where performance is assessed in age increments of one year at each educational grade-level might prove beneficial. A study of this magnitude might reveal differences not found in the present study. Third, research addressing the construct validity of the TVCF is needed. Specifically, a study comparing subjects’ performance on the TVCF and a nonverbal measure of set-shifting and rule induction would be interesting if researchers could control or account for the effects of exposure to one instrument in assessing subjects’ performance on the other instrument. A fourth area in need of exploration is the performance of special populations on the TVCF. Clearly, there is fertile ground for continued research with this instrument.

REFERENCES


