Journal of the International Neuropsychological Society (2004), **10**, 401–411.

DOI: 10.1017/S1355617704103081

Cancellation test performance in African American, Hispanic, and White elderly

DESIREE A. BYRD,[1] PEGAH TOURADJI,[1] MING-XIN TANG,[1,4] AND JENNIFER J. MANLY[1,2,3]

[1]G. H. Sergievsky Center, Columbia University College of Physicians and Surgeons, New York, NY
[2]Department of Neurology, Columbia University College of Physicians and Surgeons, New York, NY
[3]Cognitive Neuroscience Division, Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University College of Physicians and Surgeons, New York, NY
[4]Division of Biostatistics, School of Public Health, Columbia University College of Physicians and Surgeons, New York, NY

(Received January 22, 2003; Revised August 5, 2003; Accepted September 22, 2003)

Abstract

Shape and letter cancellation test performance was investigated among large samples of African American, Hispanic, and White non-demented elders. Ethnic minority elders took significantly longer to complete both tasks compared to Whites. An index of task efficiency, which simultaneously measures time and accuracy, suggested that slower time by minority elders was not related to a measurable effort to achieve greater accuracy. The frequency of commission errors was greater in our sample than in previous reports, especially among ethnic minority elders. Although significant differences were observed between the ethnic groups when matched for years of education, equating for literacy level eliminated all performance differences between African Americans and Whites on both cancellation tasks. (*JINS*, 2004, *10*, 401–411.)

Keywords: Cancellation tests, Cognitive style, Ethnicity differences

INTRODUCTION

Cancellation tests, which require individuals to simultaneously attend to some stimuli (targets) while ignoring others (distractors), are commonly utilized in neuropsychological batteries for the assessment of selective attentional abilities. There is established utility for these tests in the clinical and empirical evaluation of visuospatial/attentional disorders such as hemispatial neglect (Aglioti et al., 1997; Lezak, 1995; Weintraub & Mesulam, 1987). Cancellation tests are also used to evaluate attentional abilities after traumatic brain injury, stroke, and in neurological disorders such as Alzheimer's disease (Baddeley et al., 2001; Geldmacher et al., 1995; Hills & Geldmacher, 1998; Weintraub & Mesulam, 1987, 1988). Successful performance on cancellation tests requires a number of cognitive abilities, including sustained and selective attention, visual search, psychomotor speed and fine motor coordination. Performance is also influenced by a number of stimulus and task factors. For example, there is evidence that greater perceptual similarity between distractors and target stimuli increases discrimination difficulty (Aglioti et al., 1997; Geldmacher et al., 2000; Hills & Geldmacher, 1998; Lezak, 1995). Research also suggests that increased perceptual similarity among stimuli heightens the sensitivity of cancellation tasks to detect neglect syndromes (due to greater difficulty in suppressing inhibitory processes; Duncan & Humphreys, 1989; Geldmacher, 1998). A number of other task characteristics can affect performance on cancellation tests, among them: size of the matrix (numbers of rows and columns), configuration of the matrix (random *vs.* organized arrays), font size, stimuli type (numbers, shapes, letters, lines, etc.), space between stimuli, and method of cancellation (circling *vs.* striking).

It is difficult to aggregate results from studies utilizing cancellation tests. A major complication is that many different forms of the test exist, all of which vary greatly in the task features mentioned previously (e.g., stimuli type, size of array, target type, etc). While standardized forms of cancellation tasks exist with published normative data (e.g.,

Reprint requests to: Jennifer Manly, Ph.D., G.H. Sergievsky Center, 630 West 168th Street P & S Box 16, New York, NY 10032. E-mail: jjm71@columbia.edu

Digit Vigilance; Lewis & Rennick, 1979; Heaton et al., 1991), many research laboratories and neuropsychological clinics use experimental versions of the task. Another variable factor with cancellation tests is the performance index abstracted for evaluation. Most studies have used *either* a measure of accuracy (number of correctly identified target stimuli) or time to completion (measured in seconds), in isolation, as the metric for evaluating performance (Geldmacher & Hills, 1997; Lezak, 1995; Weintraub & Mesulam, 1987). Using either of these indices alone does not adequately characterize total performance on the tests, especially if the two indices are not highly correlated. More recently, researchers have developed indices that incorporate both accuracy and speed to evaluate performance. Such indices are purported to better reflect total quality of performance and encompass more of the cognitive processes involved in completing the task in a way that is superior to examining either index alone (Geldmacher & Hills, 1997; Lockwood et al., 2001). Finally, other researchers have explored the spatial pattern of errors as the index of interest, especially when investigating neglect syndromes (Geldmacher & Hills, 1997; Geldmacher & Reidel, 1999; Marshall & Halligan, 1989; Mennemeier et al., 1998).

Like other neuropsychological measures, cancellation tests are sensitive to the influence of non-neurologic demographic factors such as age, education, gender, and to an unknown degree, cultural background. There is ample evidence to suggest that increased adult age is associated with slower and less accurate performance on cancellation tests (Della Sala et al., 1992; Filley & Cullum, 1994; Foster et al., 1995; Geldmacher & Riedel, 1999; Geldmacher et al., 2000). Few researchers have documented education effects on cancellation test performance. From available reports, increased years of education is predictive of decreased completion time (i.e., faster speed) and more accurate performance on cancellation tasks. For example, Heaton et al., (1991) reported that years of education (as well as age) was a significant predictor of performance on a cancellation task, accounting for 13% of the variance in the time to complete the test and 16% of the accuracy variance. Variability in cancellation test performance due to the ethnic background of the examinee is less well known. The participant samples of most studies using cancellation tests were exclusively White or, as was most often the case, not reported at all (Della Sala et al., 1992; Foldi et al., 1992; Geldmacher et al., 2000; Kelland & Lewis, 1996; Mennemeier et al., 1998; Weintraub & Mesulam, 1987). One potentially important source of cross-cultural variance on cancellation, as well as other neuropsychological tests, is the use of different culturally influenced cognitive strategies to complete the tests (Berry et al., 2002; Willis, 1989). Significant variability in cognitive strategy could affect scores, especially if the performance index is calculated without consideration of those particular strategies.

The popularity and clinical utility of cancellation tests in neuropsychological evaluations requires that the performance of different ethnic groups on these tests be empirically investigated and better understood. One goal of the current study was to explore multiple aspects of cancellation test performance among an ethnically diverse, non-demented elderly sample and determine whether ethnicity-based differences existed. Further, we investigated whether ethnic minority elders were more likely to utilize a cognitive style that favored accuracy over speed on these tests. If such a style is preferred by these groups, then utilizing a performance measure which accounts for this style might attenuate any ethnic differences observed with traditional indices of cancellation test performance (e.g., time and accuracy in isolation).

One aspect of African American and Hispanic cognitive test performance that has been reported in the literature is a slower speed on timed tests of attention and information processing (Diehr et al., 1998; Klineberg, 1928; Llabre & Froman, 1987; Miller et al., 1993). Such performance differences, when observed in neurologically healthy participants, are not the result of differing levels of ability, but rather a consequence of a combination of various factors such as literacy level, differential test relevance, motivation, and cognitive style. One area which demonstrates a possible cultural contribution to cognitive style on timed tests is research on perceptions of time that indicate that African American and Hispanic cultures have a unique and less rushed temporal perspective when compared to White culture (Helms, 1992; Jones, 1998; Levine, 1997; Llabre, 1991). We propose that cultural differences in time orientation may impact performance on timed tests of attention by guiding the style with which the person completes the test. If such a source of time differences is not considered when analyzing test performance, consistent and discrepant group performances risk misinterpretation as impairment. The instructions for cancellation tasks do not emphasize any particular cognitive strategy; participants are instructed to work as quickly and accurately as possible, leaving open the possibility for a test taker to unequally emphasize either one of these goals at the expense of the other. Therefore, if members of ethnic minority groups indeed display a unique time orientation that is not hurried and simultaneously place greater emphasis on accuracy, they may earn lower time scores on cancellation tests due to their cognitive style and not poor attention.

Recent research has identified *quality* of education, measured through literacy level, as having a greater impact on reducing the ethnic disparities on cognitive test scores than *years* of education (Manly et al., 2002). One of the possible mechanisms through which literacy level may affect ethnicity differences is through the development of different cognitive styles. Given the documented importance of literacy on mediating ethnicity differences, it is likely that individuals with higher literacy may demonstrate a cognitive style that favors speed over accuracy. However, the interaction between literacy level and cognitive style has not yet been investigated. Therefore, literacy level was investigated as a possible mediator of ethnic differences on cancellation test performance.

The production of commission errors (marking a distractor item) during cancellation tests has not been thoroughly examined in prior studies of these tests. In reported non-demented samples, which have been primarily White, commission errors are generally rare, with reports of only one or two per neurologically normal subjects (Geldmacher, 1998; Geldmacher & Reidel, 1999). These types of errors occur so infrequently among non-demented Whites that virtually no detailed research has been conducted on their demographic and cognitive correlates. Experimental studies have documented that the likelihood of commission errors increases as the perceptual similarity of targets and distractors increases (Amieva et al., 1999; Foldi et al., 1992). A relatively large number of such errors might suggest inattention, perceptual disturbance, decreased visual acuity, motor difficulties or even a misinterpretation of the test instructions. Another goal of the current study was to explore the frequency and correlates of commission errors. The current study will document the frequency with which commission errors are made on cancellation tasks and determine whether they are as rare among non-demented ethnic minorities as in reported samples of Whites. Further, we will identify other demographic factors that should be considered when evaluating the commonality of these errors. Finally, given the established relationship between perceptual processing and commission errors, it was anticipated that performance on measures of visuoperception would be related to the occurrence of these errors.

As mentioned above, several cognitive abilities are involved in cancellation test performance. Understanding the specific cognitive contributors to test performance can aid neuropsychologists in interpreting how performance on cancellation tests relates to other cognitive constructs. In an effort to identify the specific neuropsychological constructs, among those assessed in our test battery, that underlie each cancellation test performance index (time, omission errors, commission errors) we subjected the performance indices to regression analyses in which the usefulness of neuropsychological test scores and demographic factors were tested as predictors.

Hypotheses for the study were as follows: (1) consistent with prior research, education matched African American and Hispanic elders will obtain significantly lower scores than White elders on the time, but not the accuracy (omission errors) indices when the two variables are considered alone; (2) in our sample of elders, a quality of search index which simultaneously considers time and accuracy will not significantly differ between the groups as time differences should be eliminated after adjustment for accuracy performance; (3) equating ethnic groups on literacy level will render insignificant any observed differences among the groups on measures of test performance; (4) ethnic groups will not differ on the frequency of commission errors; and (5) performance on neuropsychological measures of visuoperception, relative to tests of other cognitive constructs, will be predictive of commission errors.

METHODS

Research Participants

The study sample was selected from participants in the Washington Heights–Inwood Columbia Aging Project (WHICAP), a community-based, epidemiological study of dementia in the ethnically diverse neighborhoods of Northern Manhattan, New York. The WHICAP study follows a random sample of elderly Medicare recipients residing in selected census tracts of Washington Heights and Inwood.

Inclusion/exclusion criteria

All potential participants were aged 65 and above. Participants were included if they self-identified their race as non-Hispanic White, Hispanic, or non-Hispanic Black/African American according to US Census Criteria (United States Bureau of the Census, 1991). Potential participants were excluded if they had a history of Parkinson's disease, stroke, head injury with loss of consciousness, alcohol abuse, or serious mental illness such as depression or schizophrenia. Further, only WHICAP participants who showed no neurological or functional signs of dementia were included in the study. This determination was made on the basis of a physician's clinical examination, which included a rating of daily functioning and was used as a gold standard for the absence of dementia, since the physician's assessment was made independent of the participant's performance on the neuropsychological battery.

Procedure

Neuropsychological battery

A full neuropsychological test battery was administered to all study participants. The measures included in the battery were selected from cognitive domains that are typically affected in dementia and have demonstrated adequate sensitivity for distinguishing dementia from normal aging (see Stern et al., 1992, for full battery description). To avoid redundancy among the predictor variables used in analyses, one measure from each cognitive domain assessed in the full battery was selected for analysis: verbal learning (Selective Reminding Test, SRT; Trials 1–5 total score; Buschke & Fuld, 1974), nonverbal short-term memory (Benton Visual Recognition test, BVRT; Benton, 1955), visuoperception (multiple choice matching of figures from the BVRT); abstraction (WAIS–R Similarities subtest, raw score), and category fluency (animals, food and clothing, mean score).

For Spanish-speaking participants, all interview questions, test instructions, and stimuli were translated into Spanish by a committee of Spanish speakers from Cuba, Puerto Rico, Spain, and the Dominican Republic, and then back-translated to ensure accuracy. The Spanish version of the battery is described in detail elsewhere (Jacobs et al., 1997). Evaluations were conducted in either English or Spanish, based on the subject's opinion of which language would yield the best performance. Examiners were balanced bilinguals, who spoke both English and Spanish daily with friends, family, and colleagues.

Cancellation tasks

The cancellation tests used in this study (the Letter and Shape Tests; Sano et al., 1984) were administered as part of the neuropsychological battery described above. These paper and pencil tasks are each one page in length with 14 rows and 10 columns of five distractor shapes or letter triads randomly interspersed with a target shape or letter triad. Participants were instructed to cross out all of the target shapes or letter triads, as quickly and accurately as they could. A maximum 4-min period was allowed. The Shape task was administered prior to the Letter task. Administration order was standardized as part of the research protocol. Traditional performance indices collected from this measure included time to completion (recorded in seconds) and number and location of omission and commission errors. For the purposes of this study, accuracy was operationally defined as the number of omission errors. Additionally, for the primary hypothesis of this study, a "quality of search" index (Q), developed by Geldmacher and colleagues (1997), was utilized. Q is the ratio of correct responses to total target number multiplied by the ratio of correct responses per unit of time. Higher Q scores reflect more efficient performance:

$$Q \text{ score} = \frac{\text{(correct responses)}}{\text{(total targets)}} \times \frac{\text{(correct responses)}}{\text{(total time)}}$$

Literacy level/occupation

For a subset of the African American and White participants, reading level was measured using the Reading Recognition subtest from the Wide Range Achievement Test–Version 3 (Wilkinson, 1993). Given that the majority of Hispanic participants were not proficient in English, they did not complete the WRAT–3 and were therefore excluded from this analysis. This test required participants to name letters and pronounce words out of context. The words are listed in order of decreasing familiarity and increasing phonological complexity. Consistent with the standard administration, a basal of five correct items and a ceiling of 10 incorrect items was used.

Participant's self-reported primary lifetime occupations were grouped into the following two categories: (1) low (unskilled/semiskilled, skilled trade or craft, and clerical/office worker), and (2) high (manager business/government, and professional/technical). Women who classified their primary occupation as housewife were excluded from the analyses of occupational effects.

Statistical Methods

Given the extent of disparate distributions of educational levels between our samples, and the established importance of education to cognitive test performance, education matched groups were created from the total eligible sample. To create groups matched on years of education, a stratified random sampling method was applied. The sample was stratified by race (White, African American, and Hispanic) and five categories of completed years of education (0–3; 4–8; 9–12; 13–15, and greater than/equal to 16 years). Within each education-by-race cell, a random sample of participants was selected using the SPSS (1998) sample function. Equal numbers of participants were selected within each education-by-race stratum. For the literacy level analysis, the same procedure was applied to the entire database (not the education matched subgroups) to create groups matched on WRAT–3 Reading Recognition scores among African American and White English speakers. The sample was stratified by ethnicity and three categories of WRAT–3 scores (0–41; 42–48; and greater than 48). Within each WRAT–3 score-by-ethnicity cell, equal numbers of participants were selected.

ANOVAs were used to compare the age and years of education of the groups. Chi-square tests were used to analyze gender compositions and the occurrence of commission errors. Separate ANCOVAs were conducted on time, accuracy (omission errors), and quality of search (Q) for the Shape and Letter tasks. Follow-up pairwise comparisons were conducted to identify the source(s) of ethnicity effects, when the overall effect of race was significant in the ANCOVA. Stepwise logistic and hierarchical multiple regression analyses were used to test hypotheses related to predictors of specific performance indices.

RESULTS

Sample Characteristics

A total of 1405 WHICAP participants (418 Whites, 454 African Americans, and 533 Hispanics) had complete neuropsychological evaluations, neurological exams, and met the inclusion criteria. Among the Hispanic participants, 86% were tested in Spanish. Groups of 234 participants from each ethnic group were selected from the total sample to create education-matched groups (sampling techniques described above). Selected participants were representative of the total sample with regards to mean age and gender distributions. Among the selected education-matched participants, the White group was significantly older than the African American and Hispanic samples [$F(2,698) = 11.0$, $p < .01$; see Table 1 for detailed demographic information] but the groups did not differ in gender ratios. Therefore, age was used as a covariate in proceeding analyses. ANOVAs demonstrated in both our total and education matched samples, that, gender had neither a significant direct, nor interactive, effect on any of the cancellation test variables. Occupational data was available for a subset of participants ($n = 414$). Analysis of cancellation test data within this subsample indicated that the influence of prior occupation on test performance did not supercede the effect of ethnic-

https://doi.org/10.1017/S1355617704103081 Published online by Cambridge University Press

Table 1. Demographic and descriptive cancellation test performance information

	Whites (n = 234)	African Americans (n = 234)	Hispanics (n = 234)
Demographics	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Age*	77.1 (7.2)	75.3 (6.1)	74.3 (6.2)
Years of education	10.9 (3.6)	10.3 (3.4)	10.5 (3.7)
Gender (% female)	65%	70%	70%
Shape Cancellation	Adjusted *M* (*SE*)	Adjusted *M* (*SE*)	Adjusted *M* (*SE*)
Shape time*	81.5 (2.4)	98.1 (2.4)	90.8 (2.4)
Shape omits	5.1 (.27)	5.6 (.27)	5.6 (.27)
Shape *Q**	.17 (.01)	.13 (.01)	.14 (.01)
Letter Cancellation			
TMX Time*	88.4 (2.5)	110.4 (2.5)	103.4 (2.5)
TMX Omits	1.7 (.17)	1.8 (.17)	2.2 (.18)
TMX *Q**	.15 (.01)	.12 (.01)	.13 (.07)

*$p < .01$

ity and therefore was not included in the primary analyses described below.

Shape Cancellation

An ANCOVA revealed significant group differences on time to complete the Shape cancellation task [$F(2,698) = 12.4$, $p < .001$; see Table 1 for adjusted group mean values]. Follow-up pairwise comparisons revealed that Hispanic and African American elders did not differ significantly from one another on the time variable but both groups took more time to complete the task than the White group. An additional ANCOVA revealed that the groups did not differ on Shape omission errors [$F(2,698) = 1.6$, $p = .20$]. Analysis of the Shape Q variable indicated significant group differences [$F(2,698) = 11.7$, $p < .001$]. Follow-up analyses revealed that Hispanic and African American elders did not differ significantly from one another on Shape Q scores but both groups obtained significantly lower scores than the White group.

Letter Cancellation

An ANCOVA revealed that African American and Hispanic elders took significantly longer to complete the Letter cancellation task than White elders [$F(2,698) = 20.6$, $p < .001$; see Table 1 for adjusted group mean values]. Follow-up pairwise comparisons revealed that Hispanic and African American elders did not differ significantly from one another on time to complete the test but both groups took longer than the White group. Analysis of Letter omission errors indicated that the ethnic groups did not differ on this variable [$F(2,698) = 2.4$, $p > .05$]. Analysis of the Letter Q variable indicates that African American and Hispanic elders obtained significantly lower Q scores than White elders [$F(2,698) = 12.7$, $p < .01$]. Follow-up pairwise comparisons revealed that Hispanic and African American elders did not differ significantly from one another on the Letter Q score.

Neuropsychological Correlates of Cancellation Time and Omission Errors

Multiple regression analyses (see Tables 2 and 3) were used to evaluate the relative contributions of (1) time to complete the cancellation tests and (2) accuracy of performance (omission errors). These analyses were completed with the entire education-matched sample (combined ethnicities). Predictor variables for the multiple regression equations were demographics (age, education, and ethnicity) and select representative neuropsychological test scores (SRT Total Learning Trials 1–5, Benton matching, Benton recognition, WAIS–R Similarities raw score, and category fluency mean score). Simultaneous entry methods were utilized for all predictor variables. The analysis for Shape time indicated that the set of predictor variables accounted for 28% of the variance with age, African American ethnicity, SRT learning score, Benton matching, Benton recognition, and category fluency emerging as significant predictors at the $p < .05$ level. The analysis of Letter time produced similar results, with an adjusted R^2 of .35 and the following statistically significant predictors: age, education, African American ethnicity, SRT learning score, Benton matching, and category fluency.

Cancellation test accuracy (omission errors) was not well predicted by demographic factors and neuropsychological test scores. The model accounted for 6% of the Shape accuracy variance, with Benton matching emerging as the only significant predictor. For Letter accuracy, the model accounted for 13% of the variance; Benton matching and Category fluency were significant predictors at the $p < .05$ level.

Literacy Level

Of the eligible English-speaking participants, 202 (126 African Americans, AA; and 76 Whites, WH) had WRAT–3 data. Groups of 51 African American and 51 White elders matched on WRAT–3 Reading score were formed through

Table 2. Multiple regression analyses of the relationship between demographic and neuropsychological variables and shape and letter time

	Shape Time $R^2 = .28$			Letter Time $R^2 = .35$		
Variable	Beta	t	p	Beta	t	p
Age	.185	5.096	.000	.133	3.817	.000
Hispanic ethnicity	.004	.101	.919	.050	1.331	.184
African American ethnicity	.088	2.264	.024	.127	3.430	.001
SRT Total	−.140	−3.101	.002	−.159	−3.680	.000
Benton Recognition	−.097	−2.410	.016	−.064	−1.652	.099
Benton Matching	−.096	−2.462	.014	−.099	−2.645	.008
WAIS–R Similarities	−.045	−1.056	.292	−.062	−1.513	.131
Category Fluency	−.179	−4.069	.000	−.263	−6.213	.000

Note. SRT = Selective Reminding Test.

the stratified random sampling procedure described above. African Americans were younger [AA: $M = 74.0$, $SD = 5.4$; WH: $M = 76.4$, $SD = 6.2$; $t(100) = 2.1$, $p < .05$] and less educated than Whites [AA: $M = 10.6$, $SD = 3.4$; WH: $M = 12.2$, $SD = 4.0$; $t(100) = 2.2$, $p < .05$] but there was no significant difference in the gender composition of the groups (approximately 66% female). All of the African Americans in this sample were native English speakers while 67% of Whites had English as their first language. All participants were proficient in English.

These literacy-matched groups were compared on cancellation test performance using ANCOVAs, with age and years of education as covariates (Table 4). Results of the analyses indicate that the literacy-matched groups of African American and White elders did not earn significantly different scores on any of the Shape or Letter cancellation indices. In fact, the literacy-matched African American elders demonstrated slightly higher, though not statistically significant, scores than Whites on the Shape efficiency (Q) score.

To further explore the influence of literacy level and ethnicity on cancellation test performance, "high" and "low" African American and White literacy groups were created by splitting each group according to ethnicity specific median WRAT–3 score values. Within each ethnic group, the high and low literacy groups were compared on cancellation test indices.

The low literacy African American group ($n = 68$) had significantly fewer years of education [$t(130) = -5.4$, $p < .01$] than the high literacy group ($n = 68$); the two groups did not differ significantly in age. ANCOVAs with education as a covariate were used to analyze cancellation test performance. No significant differences emerged between the African American literacy groups on the Shape or Letter tasks but there was a trend for the low literacy group to earn significantly lower scores on the Letter accuracy variable [$F(1,75) = 3.48$, $p = .07$]. Among Whites, the low literacy group ($n = 44$) was significantly older [$t(76) = 2.06$, $p < .05$] and had fewer years of education [$t(76) = -2.27$, $p < .05$] than the high literacy group ($n = 34$). ANCOVAs with age and education as covariates were used to analyze cancellation test performance. On the Shape task, the low literacy White group earned significantly lower scores on the Shape efficiency variable [$F(1,75) = 4.59$, $p < .05$]. There

Table 3. Multiple regression analyses of the relationship between demographic and neuropsychological variables and shape and letter accuracy

	Shape Accuracy $R^2 = .06$			Letter Accuracy $R^2 = .13$		
Variable	Beta	t	p	Beta	t	p
Age	−.025	−.596	.551	−.008	−.187	.851
Hispanic ethnicity	−.001	−.015	.988	−.010	−.227	.821
African American ethnicity	.005	.111	.912	−.055	−1.297	.195
SRT Total	−.015	−.297	.767	−.060	−1.201	.230
Benton Recognition	−.064	−1.379	.168	−.082	−1.827	.068
Benton Matching	−.106	−2.356	.019	−.154	−3.564	.000
WAIS–R Similarities	−.077	−1.573	.116	−.106	−2.261	.024
Category Fluency	−.061	−1.204	.229	−.107	−2.187	.029

Note. SRT = Selective Reminding Test.

https://doi.org/10.1017/S1355617704103081 Published online by Cambridge University Press

Table 4. Effect of ethnicity on cancellation test after matching for WRAT–3 Reading score

Test	Whites (n = 51) M (SD)	African Americans (n = 51) M (SD)	F value	p
Shape cancellation				
Shape time	71.6 (18.8)	78.4 (29.9)	1.5	.23
Shape omits	5.4 (3.6)	4.4 (3.5)	3.0	.09
Shape commits	.12 (.38)	.24 (.71)	.59	.45
Shape Q	.16 (.07)	.18 (.09)	2.0	.15
Letter cancellation				
TMX time	82.2 (21.3)	94.7 (37.5)	2.5	.11
TMX omits	.86 (1.2)	.88 (1.2)	.06	.81
TMX commits	.31 (.74)	.51 (1.1)	1.1	.30
TMX Q	.14 (.07)	.15 (.08)	1.1	.29

were no significant differences between the White literacy groups on other Shape variables or any of the Letter task indices.

Commission Errors

The proportion of elders in each ethnic group from the total, non-education matched sample (n = 1405), who made at least one commission error on each task, in addition to other descriptive information regarding commission errors, is presented in Table 5. Chi-square analyses of error frequency by ethnic group indicated that ethnic minority elders made commission errors more frequently than White elders did on both the Shape [$\chi^2(2) = 31.1$, $p < .01$] and Letter tasks [$\chi^2(2) = 27.1$, $p < .01$]. Also, the range of errors was notably larger in ethnic minority groups in that the standard deviations were 2 to 4 times larger than White values. To explore whether the observed differences between ethnic groups was related to the disparate educational levels of the groups, we repeated the analysis in the education-matched subsample and discovered a similar pattern of results (minorities made significantly more commission errors than Whites). Also, stepwise logistic regression analyses were performed in the non-education-matched sample in which age and education were entered as covariates at the first step, and ethnicity at the second step. Results revealed that after accounting for age and years of education, African American ethnicity was significantly associated with letter commission errors [$\chi^2(4) = 60.0$, $p < .001$] and Hispanic ethnicity was significantly associated with the production of shape commission errors [$\chi^2(4) = 6.7$, $p < .001$].

Analyses of demographic factors that might characterize participants who made commission errors revealed that participants who made commission errors (at least one on either task), were significantly, though only slightly, older ($p < .01$) and less educated ($p < .01$; Table 6) than participants who did not make any commission errors. There were no gender differences between the commission error groups; roughly 70% of the commission and non-commission error groups were female [$\chi^2(2) = .08$, $p > .05$]. Furthermore, few participants made commission errors on *both* tests; of the 521 participants who made at least one error on either test, only 122 (23%) made commission errors on both

Table 5. Cancellation test commission errors: frequency and descriptive information

Variable	Whites (n = 418) M (SD)	African Americans (n = 454) M (SD)	Hispanics (n = 533) M (SD)
Age*	76.6 (6.9)	76.1 (6.4)	74.1 (5.9)
Years of education*	12.5 (3.7)	9.5 (3.7)	6.5 (4.2)
Gender (% female)	65%	70%	70%
Mean Shape commission errors (SD; range)	.44 (1.7; 0–16)	10 (4.2; 0–60)	1.6 (5.4; 0–60)
Mean Letter commission errors (SD; range)	.36 (1.1; 0–10)	.96 (2.9; 0–27)	.68 (2.7; 0–31)
≥ 1 Shape commission error	16.0%	21.9%	30.4%
≥ 1 Letter commission error	18.4%	32.6%	21.8%
≥ 2 Shape commission error	6.2%	12.0%	17.7%
≥ 2 Letter commission error	8.1%	14.7%	9.1%

*$p < .01$

Table 6. Descriptive performance information: commission error group membership

Variable	≥1 commission errors ($n = 521$) M (SD)	Zero commission error ($n = 870$) M (SD)	F value	p
Demographics				
Age	76.8 (6.9)	74.8 (6.3)	25.9	.000
Education	9.0 (3.9)	10.3 (4.1)	27.6	.000
Neuropsychological test score				
SRT Learning Trials 1–5	32.5 (9.7)	38.1 (10.2)	99.9	.000
Benton Recognition Test	5.6 (2.1)	6.9 (2.0)	92.3	.000
Benton Matching	7.4 (2.1)	8.5 (1.6)	112.1	.000
WAIS–R similarities raw score	8.0 (6.3)	11.5 (6.7)	84.8	.000
Category fluency mean score	12.1 (3.9)	14.0 (4.4)	67.3	.000

tests. Likewise the bivariate correlation between the error types on the two tasks was low, though significant ($r = .23$, $p < .01$).

To explore the possible relationship between commission errors and neuropsychological abilities after age, education, and ethnicity have been controlled for, two techniques were utilized. First, ANCOVAs with age, education, and ethnicity as covariates were used to compare the group of elders who made commission errors to the group who did not on neuropsychological test scores (Table 6). Results from each of the analyses indicated that the commission error group earned significantly lower scores on each of the neuropsychological tests. To gain a more specific understanding of the relative importance of specific neuropsychological abilities to commission errors, a logistic regression was performed in which demographic variables were entered at the first step and select representative neuropsychological test scores (SRT Total Learning Trials 1–5, Benton matching, Benton recognition, WAIS–R Similarities raw score, and category fluency mean score) were entered at the second step as potential predictors of commission error group membership (at least one commission error on either test *vs.* no commission errors). After the consideration of demographic factors, results indicate that a statistically significant relationship [$\chi^2(5) = 105.8$, $p < .001$] existed between the following neuropsychological test scores and commission error group membership: SRT learning (odds ratio = .977), BVRT recognition (odds ratio = .917), and BVRT matching (odds ratio = .836).

DISCUSSION

The current study revealed several important characteristics of cancellation test performance in our large, education-matched ethnically diverse group of neurologically healthy elders. First, as predicted, there were significant, though small, ethnic differences on time to complete both forms of the cancellation test, even though the groups were matched for years of education. For both tasks, African Americans and Hispanics took more time to complete the tests than Whites. The magnitude of the time difference depended on the nature of the stimuli; groups were more similar in time on the Shape task than the Letter task. The large size of our sample ($n = 702$) and related power could explain why relatively small performance differences reached statistical significance. Secondly, the hypothesized absence of ethnicity differences in accuracy (operationally defined as the number of omission errors) was supported for both the Shape and Letter tasks. Again, an effect of task was apparent in that all participants made fewer omission errors on the Letter task than the Shape task.

A major goal of this study was to determine whether ethnic minorities, on average, relied more heavily on a cognitive style that favored accuracy over speed on timed precision measures like cancellation tests and whether such a style could be responsible for the observed difference in time between ethnic groups. We utilized a measure of efficiency developed by Geldmacher et al. (1997), which simultaneously considered time and accuracy, to test this hypothesis. Results revealed that minority elders were less efficient than White elders, even when time and accuracy were considered simultaneously, suggesting that the observed difference in time to complete the tests between the groups may not have been due to a speed/accuracy tradeoff style, as measured by our efficiency score. However, it is important to note that our measure of a speed/accuracy tradeoff style (Q score) is not perfect. For example, the Q score does not include errors of commission in its measurement of accuracy. Therefore, the impact of such a speed/accuracy tradeoff style on the timed performance of ethnic minority elders cannot be ruled out from the findings of this study. A more refined, experimental assessment of such a style would be necessary to clearly determine the presence or absence of such a style and its effect on ethnic group performance pattern. Future studies of speed/accuracy tradeoff style would do well to design tasks with experimentally manipulated instructions, carefully selected performance measures, and perhaps explicit questions to examinees regarding cognitive processes employed during their completion of the task.

Consistent with previous reports of the ability of literacy level to attenuate ethnic group differences on other neuropsychological measures (Manly, et al., 2002), matching African American and White groups on WRAT–3 Reading Recognition scores rendered group differences on our cancellation test measures insignificant. Speed differences observed in the larger, education-matched sample were reduced by roughly 35% in the literacy-matched sample. Further, there were trends in this subsample, for the African Americans to be more accurate than Whites on the Shape task. The small size of this subsample (group $n = 51$), compared to our full sample (group $n = 234$) decreased our power to detect performance differences of the magnitude observed in the larger sample. Therefore, these results must be interpreted with some caution; the time difference on the Letter task observed in the literacy-matched subsample would have likely been significant had the sample size been larger. Nonetheless, these findings offer additional support for the need for neuropsychologists to expand methods of establishing an expected performance level beyond documenting the years of education alone. The current results suggest that the effects of literacy level on ethnicity-based cognitive test differences are not limited to verbal or intelligence tests. Literacy level appears to be the best available factor for use to match African Americans and Whites (and very likely, other ethnic groups) on expected performance level in research studies. However, when the influence of literacy level on test performance was examined within ethnic group, the results indicate that the "low" and "high" literacy groups did not significantly differ in their performance pattern. Future study is needed to clarify the relationship between literacy level and test performance within ethnic groups on various tests. These results further demonstrate that the nature of performance differences between certain cultural groups is related to measurable factors that can be easily implemented in clinical and research settings. Additional research is necessary to more clearly identify how literacy might impact cognitive style and to determine whether literacy level is equally important to the performance of Hispanics.

Our study is among the first to provide a detailed examination of commission errors on cancellation tests. We found that in our sample of non-demented elders, such errors are not as rare as previously reported. Anywhere from 16 to 33% of our sample made at least one commission error on our tasks and 6 to 18% made two or more errors. While the mean number of commission errors was small (around 1), there was substantial range in the distribution of errors (0–60). More importantly, greater proportions of minority, relative to White, elders made commission errors on the task, even after education differences were accounted for (African Americans on the Letter task and Hispanics on the Shape task). Also, the upper limits for the number of errors produced by minority elders were up to 3 times higher than that for White elders. Such differences among non-demented participants suggest that the presence of commission errors on cancellation tests should not necessarily be interpreted as pathological, especially among ethnic minority elders. Participants who made these errors were slightly older and less educated than those who did not. Also, commission errors on the two tests appear to be independent of one another; commission errors on Shape and Letter tasks were only mildly correlated. Analyses of neuropsychological predictors of commission errors revealed that these errors were associated with poorer performance on a test of visuoperception, which suggest that commission errors result, in part, from difficulty perceiving and distinguishing targets and distractors, which, in the case of neurologically healthy ethnic minority elders, could be related to a history of limited experience with such stimuli.

Exploratory analyses of cognitive correlates of cancellation test time and accuracy revealed interesting relationships. Shape and Letter time were best predicted by older age, African American ethnicity and performances on tests of verbal learning, visuoperception, and category fluency. For cancellation test accuracy, visuoperception and category fluency were significant predictors but the amount of accuracy variance accounted for by these variables was small. Given the role of accurate visual perception in the ability to discriminate targets from distractors, the emergence of this ability as a predictor of test accuracy and time to complete the tests was not surprising. A relationship between category fluency and cancellation test performance likely emerged as a result of the shared timed performance factor for both tests; the category test was the only other timed test included in our battery. The relationship between performance on a verbal learning task to time to complete cancellation tasks is best understood by the attentional component to verbal learning tasks. Both cancellation tests and the learning trials of verbal learning tasks require adequate attentional skills. These results support theories that propose multiple cognitive components to successful cancellation test performance. A limitation to these analyses is that our battery was not exhaustive and therefore did not include tests of other abilities that likely contribute to cancellation tests performance, such as psychomotor speed. Finally, while no demographic factors contributed to accuracy, older age and African American ethnicity were predictive of longer time to complete the tests, a finding consistent with our cross-sectional analyses.

Results from this study offer additional evidence that task characteristics are important to understanding cancellation test performance. Our two tasks (Letter and Shape) were identical in matrix organization and ratio of targets to distractors. However, all groups took longer to complete, and were more accurate on, the Letter task. This finding is consistent with previous reports of increased processing time of letters relative to shapes (Shor, 1971). When cancellation tests are considered for inclusion in clinical or research batteries, consideration should be given to task factors and how they relate to the purpose of the test battery.

Cancellation tests are often used in assessment batteries for children. Results from this study may not generalize to the performance of ethnic minority children on cancellation tests. Future research is required to determine the presence and nature of ethnicity effects on cancellation tests in children and younger adults.

Cancellation tests are among the measures that are used in diagnostic batteries for Alzheimer's disease and attentional disorders. While not the hallmark symptom of Alzheimer's disease, it is well documented that Alzheimer's patients demonstrate significant attentional difficulties (see Perry & Hodges, 1999, for review). Perry et al. (2000) offer evidence that suggests that even early in the course of the disease, components of selective and divided attention are affected. Considering our finding of ethnicity-related performance differences among non-demented elders, diagnostic decisions for minority elders should not rely heavily on this measure, especially if literacy level is not first considered.

The current study demonstrates that there are cancellation test performance differences for elders from different ethnic groups with similar years of education. While the ethnicity effect sizes were small, they were nonetheless statistically significant. If future normative samples for cancellation tests do not include adequate numbers of minorities, normative data may not be representative and clinicians may be more likely to interpret the performance of minority elders as impaired. Until the sources of cross-cultural differences on tests such as these are better understood, when evaluating ethnic minorities with cancellation tests, neuropsychologists should (1) interpret "slower" test performances and higher number of commission errors with caution; (2) not rely heavily on these tests for diagnostic purposes; and (3) when attempting to control for ethnicity differences, match groups on literacy performance.

ACKNOWLEDGMENTS

This research was supported by NIH grants AG07232 (R. Mayeux), AG16206 (J. Manly), T32 AG000261 (Y. Stern) and the Alzheimer's Association.

REFERENCES

Aglioti, S., Smania, N., Barbieri, C., & Corbetta, M. (1997). Influence of stimulus salience and attentional demands on visual search patterns in hemispatial neglect. *Brain and Cognition*, *34*, 388–403.

Amieva, H., Lafont, S., Dartigues, F., & Fabrigoule, C. (1999). Selective attention in Alzheimer's disease: Analysis of errors in Zarro's cancellation task. *Brain and Cognition*, 40, 26–29.

Baddeley, A.D., Baddeley, H.A., Bucks, R.S., & Wilcock, G.K. (2001). Attentional control in Alzheimer's disease. *Brain*, *124*, 1492–1508.

Benton, A.L. (1955). *The Visual Retention Test*. New York: The Psychological Corporation.

Berry, J.W., Poortinga, Y.H., Segall, M.H., & Dasen, P.R. (2002). *Cross-cultural psychology: Research and applications* (2nd ed.). Cambridge, UK: Cambridge University Press.

Buschke, H. & Fuld, P.A. (1974). Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology*, *24*, 1019–1025.

Della Sala, S., Laiacona, M., Spinnler, H., & Ubezio, C. (1992). A cancellation test: Its reliability in assessing attentional deficits in Alzheimer's disease. *Psychological Medicine*, *22*, 805–901.

Diehr, M.C., Heaton, R.K., Miller, W., & Grant, I. (1998). The Paced Auditory Serial Addition Task (PASAT): Norms for age, education, and ethnicity. *Assessment*, *5*, 375–387.

Duncan, J. & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.

Filley, C.M. & Cullum, C. (1994). Attention and vigilance functions in normal aging. *Applied Neuropsychology*, *1*, 29–32.

Foldi, N.S., Jutagir, R., Davidoff, D., & Gould, T. (1992). Selective attention skills in Alzheimer's disease: Performance on graded cancellation tests varying in density and complexity. *Journal of Gerontology*, *47*, 146–153.

Foster, J.K., Behrmann, M., & Stuss, D.T. (1995). Aging and visual search: Generalized cognitive slowing or selective deficit in attention? *Aging and Cognition*, *2*, 279–299.

Geldmacher, D.S. (1998). Stimulus characteristics determine processing approach on random array letter-cancellation tasks. *Brain and Cognition*, *36*, 346–354.

Geldmacher, D.S., Doty, L., & Heilman, K.M. (1995). Letter cancellation performance in Alzheimer's disease. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, *8*, 259–263.

Geldmacher, D.S., Fritsch, T., & Riedel, T.M. (2000). Effects of stimulus properties and age on random-array letter cancellation tasks. *Aging, Neuropsychology and Cognition*, *7*, 194–204.

Geldmacher, D.S. & Hills, E.C. (1997). Effect of stimulus number, target-to-distractor ratio, and motor speed on visual spatial search quality following traumatic brain injury. *Brain Injury*, *11*, 59–66.

Geldmacher, D.S. & Riedel, T.M. (1999). Age effects on random-array letter cancellation tests. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *12*, 28–34.

Heaton, R.K., Grant, I., & Matthews, C.G. (1991). *Comprehensive norms for an Expanded Halstead-Reitan Battery: Demographic corrections, research findings, and clinical applications*. Odessa, FL: Psychological Assessment Resources.

Helms, J.E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, *47*, 1083–1101.

Hills, E.C. & Geldmacher, D.S. (1998). The effect of character and array type on visual spatial search quality following traumatic brain injury. *Brain Injury*, *12*, 69–76.

Jacobs, D.M, Sano, M., Albert, S., Schofield, P., Dooneief, G., & Stern, Y. (1997). Cross-cultural neuropsychological assessment: A comparison of randomly selected, demographically matched cohorts of English- and Spanish-speaking older adults. *Journal of Clinical and Experimental Neuropsychology*, *19*, 331–339.

Jones, J. (1998). Cultural differences in temporal perspectives. In J. McGrath (Ed.), *The social psychology of time, new perspectives* (pp. 21–38). Beverly Hills, CA: Sage Publications.

Kelland, D.Z. & Lewis, R.F. (1996). The Digit Vigilance Test: Reliability, validity, and sensitivity to diazepam. *Archives of Clinical Neuropsychology*, *11*, 339–344.

Klineberg, O. (1928). An experimental study of speed and other factors in racial differences. *Archives of Psychology*, Number 93.

Levine, R. (1997). *A geography of time: The temporal misadventures of a social psychologist, or how every culture keeps time just a little bit differently*. New York : BasicBooks.

Lewis, R.F. & Rennick, P.M. (1979). *Manual for the Repeatable Cognitive–Perceptual–Motor Battery*. Clinton Township, MI: Ronald F. Lewis.

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

Llabre, M. (1991). Time as a factor in the cognitive test performance of Latino college students. In G.D. Keller, J.R. Deneen, & R.J. Magallan (Eds.), *Assessment and access: Hispanics in higher education* (pp. 95–104). Albany: State University of New York Press.

Llabre, M. & Froman, T. (1987). Allocation of time to test items: A study of ethnic differences. *Journal of Experimental Education*, *55*, 137–140.

Lockwood, K.A., Marcotte, A.C., & Stern, C. (2001). Differentiation of attention-deficit/hyperactivity disorder subtypes: Application of neuropsychological model of attention. *Journal of Clinical and Experimental Neuropsychology*, *23*, 317–330.

Manly, J.J., Jacobs, D.M., Touradji, P., Small, S.A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, *8*, 341–348.

Marshall, J.C. & Halligan, P.W. (1989). Does the midsagital plane play any privileged role in "left" neglect? *Cognitive Neuropsychology*, 6, 403–422.

Mennemeier, M., Rapscak, S.Z., Dillon, M., & Vezey, E. (1998). A search for the optimal stimulus. *Brain and Cognition*, *37*, 439–459.

Miller, E.N., Bing, E.G., Selnes, O.A., Wesch, J., & Becker, J.T. (1993). The effects of sociodemographic factors on reaction time and speed of information processing. *Journal of Clinical and Experimental Neuropsychology*, *15*, 66.

Perry, R.J. & Hodges, J.R. (1999). Attention and executive deficits in Alzheimer's disease: A critical review. *Brain*, *122*, 383–404.

Perry, R.J., Watson, P., & Hodges, J.R. (2000). The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: Relationship to episodic and semantic memory impairment. *Neuropsychologia, 38*, 252–271.

Sano, M., Rosen, W., & Mayeux, R. (1984, August). *Attention deficits in Alzheimer's disease*. Poster session presented at the annual meeting of the American Psychological Association, Toronto, Ontario, Canada.

Shor, R. (1971). Symbol processing speed differences and symbol interference effects in a variety of concept domains. *Journal of General Psychology*, *85*, 187–205.

Stern, Y., Andrews, H., Pittman, J., Sano, M., Tatemichi, T., Lantigua, R., & Mayeux, R. (1992). Diagnosis of dementia in a heterogeneous population. Development of a neuropsychological paradigm-based diagnosis of dementia and quantified correction for the effects of education. *Archives of Neurology*, *49*, 453–460.

SPSS for Windows, Release 9.0.0 [Computer software]. (1998). Chicago: SPSS Incorporated.

United States Bureau of the Census. (1991). *Census of Population and Housing 1990. Summary tape file 1, technical documentation prepared by the Bureau of the Census*. Washington, DC: Bureau of the Census.

Weintraub, S. & Mesulam, M. (1987). Right cerebral dominance in spatial attention: Further evidence based on ipsilateral neglect. *Archives of Neurology*, *44*, 621–625.

Weintraub, S. & Mesulam, M. (1988). Visual hemispatial inattention: Stimulus parameters and exploratory strategies. *Journal of Neurology, Neurosurgery and Psychiatry*, *51*, 1481–1488.

Wilkinson, G.S. (1993). *Wide Range Achievement Test 3–Administration manual*. Wilmington, DE: Jastak Associates, Inc.

Willis, M.G. (1989). Learning styles of African American children: A review of the literature and interventions. *Journal of Black Psychology*, *16*, 47–65.