Journal of the International Neuropsychological Society (2003), **9**, 64–71.

DOI: 10.1017/S1355617703910071

Computerized reaction time battery *versus* a traditional neuropsychological battery: Detecting HIV-related impairments

RAUL GONZALEZ,[1,2] ROBERT K. HEATON,[1,2,3] DAVID J. MOORE,[1,2] SCOTT LETENDRE,[1] RONALD J. ELLIS,[1,4] TANYA WOLFSON,[1] THOMAS MARCOTTE,[1,3] MARIANA CHERNER,[1,3] JULIE RIPPETH,[1,3] IGOR GRANT,[1,3,5] AND THE HNRC GROUP*

[1]HIV Neurobehavioral Research Center and University of California, San Diego, CA
[2]San Diego State University/University of California, San Diego—Joint Doctoral Program in Clinical Psychology, CA
[3]Department of Psychiatry, University of California, San Diego, CA
[4]Department of Neurosciences, University of California, San Diego, CA
[5]VA San Diego Healthcare System, San Diego, CA
*The San Diego HIV Neurobehavioral Research Center [HNRC] group is affiliated with the University of California, San Diego, the Naval Hospital, San Diego, and the San Diego Veterans Affairs Healthcare System, and includes: Director: Igor Grant, M.D.; Co-Directors: J. Hampton Atkinson, M.D. and J. Allen McCutchan, M.D.; Center Manager: Thomas D. Marcotte, Ph.D.; Naval Hospital San Diego: Mark R. Wallace, M.D. (P.I.); Neuromedical Component: J. Allen McCutchan, M.D. (P.I.), Ronald J. Ellis, M.D., Scott Letendre, M.D., Rachel Schrier, Ph.D.; Neurobehavioral Component: Robert K. Heaton, Ph.D. (P.I.), Mariana Cherner, Ph.D., Julie Rippeth, Ph.D.; Imaging Component: Terry Jernigan, Ph.D. (P.I.), John Hesselink, M.D.; Neuropathology Component: Eliezer Masliah, M.D. (P.I.); Clinical Trials Component: J. Allen McCutchan, M.D., J. Hampton Atkinson, M.D., Ronald J. Ellis, M.D., Ph.D., Scott Letendre, M.D.; Data Management Unit: Daniel R. Masys, M.D. (P.I.), Michelle Frybarger, B.A. (Data Systems Manager); Statistics Unit: Ian Abramson, Ph.D. (P.I.), Reena Deutsch, Ph.D., Tanya Wolfson, M.A.

(RECEIVED June 8, 2001; REVISED January 3, 2002; ACCEPTED January 22, 2002)

Abstract

In recent years, interest in the use of computerized neuropsychological (NP) assessment measures has increased. However, there are limited data regarding how performance on these measures relates to performance on more traditional, clinical instruments. In the present study, 82 HIV+ men, who were all believed on clinical grounds to have neurobehavioral impairment, completed a traditional NP battery (TNB) and the California Computerized Assessment Package (CalCAP, a collection of computerized reaction time tests). Summary scores based on a TNB, as well as those based on the CalCAP, demonstrated significant associations with both degree of immunosuppression (CD4 count) and detectable viral load in cerebrospinal fluid, but not with detectable viral load in plasma. Established norms on the TNB and CalCAP batteries resulted in classifying 57% and 49% of the HIV+ sample as impaired, respectively. When using the TNB as the "gold standard," impairment classifications based on CalCAP summary scores exhibited a sensitivity of 68% and a specificity of 77%. Overall agreement on impairment classifications between batteries was low (kappa = .44). Data from this study suggest that traditional NP batteries and computerized reaction time tests do not measure the same thing, and are not interchangeable in examining HIV-related NP impairments. (*JINS*, 2003, *9*, 64–71.)

Keywords: Neuropsychological assessment, HIV, Reaction time, CalCAP, Computerized testing, Sensitivity, CSF viral load, CD4

INTRODUCTION

HIV infection is associated with increased risks for a variety of well-documented deficits in neurobehavioral performance (Grant & Martin, 1994; Heaton et al., 1995). Nevertheless, there is disagreement on the optimal approach to detect and quantify these impairments. Many researchers investigating the neurobehavioral impairments associated with HIV infection use neuropsychological (NP) batteries consisting of traditional tests, such as measures from the Halstead-Reitan Battery, various memory assessments, and one of the Wechsler Adult Intelligence Scale (WAIS) series. Much literature is available on the psychometric properties of these measures, including extensive

Dr. Erin Bigler served as action editor during the course of this review.
Reprint requests to: Raul Gonzalez, HIV Neurobehavioral Research Center (HNRC), University of California, San Diego, 150 West Washington Street, 2nd Floor, San Diego, CA 92103. E-mail: r2gonzalez@ucsd.edu

normative data. In addition, there are many studies documenting their reliability and sensitivity in detecting impairments in individuals with a variety of neurobehavioral disorders. These factors have contributed to the popularity of such measures and their acceptance among clinicians and researchers; however, traditional NP measures are not without criticism. Established NP batteries can be lengthy, time consuming, and require the use of trained psychometrists, involving commitment of many resources. It would be desirable to have briefer instruments that provide equivalent sensitivity with lower costs.

The number and availability of computerized NP measures have dramatically increased in recent years. Such measures provide rigorous standardization, easier data collection and effortless scoring, as well as incomparable precision in measuring performance speed or reaction time (RT). In addition, many conventional NP instruments now have computerized counterparts (e.g., the Wisconsin Card Sorting Test), and researchers are developing and using novel computerized measures as potentially attractive alternatives to traditional NP methods.

Studies have emerged using computerized RT tests to detect neurobehavioral impairments in HIV-infected individuals. Their findings indicate that some of these measures are capable of detecting impairments among HIV+ individuals (Dunlop et al., 1992; Martin et al., 1992; Wilkie et al., 1990). Indeed, there is evidence that reduced speed of information processing is a cardinal feature of HIV associated brain dysfunction (Heaton et al., 1995), perhaps making this an ideal disease with which to demonstrate the relative value of computerized RT procedures. Furthermore, a preliminary study by Martin et al. (1999) found that HIV-seropositive patients receiving antiretroviral therapy performed better on RT tests when compared to those not on antiretroviral therapy, suggesting that RT tests may provide efficient, sensitive measures of important treatment effects. Given the apparent technical benefits of computerized tests, and the available literature suggesting their clinical utility with HIV related disorders, it is of interest to compare their sensitivity and rates of diagnostic agreement with those of conventional NP measures in HIV disease.

To date, few have compared computerized RT measures to conventional NP measures on their ability to detect neurobehavioral impairments in an HIV+ cohort. Miller et al. (1991) used the California Computerized Assessment Package (CalCAP) and a traditional NP battery in an attempt to differentiate between medically symptomatic HIV+ individuals, asymptomatic HIV+ individuals, and seronegative controls. Scores from seven CalCAP subtests were used (Simple RT, Choice RT, Sequential RT, Lexical Discrimination, Visual Selective Attention, Response Reversal and Visual Scanning, and Form Discrimination). The traditional NP battery used was relatively brief and consisted of seven measures (Trail Making Tests A & B, WAIS–R Digit Span, Controlled Oral Word Association Test, Grooved Pegboard Test, Symbol Digit Modalities, and the Rey Auditory Verbal Learning Test). Neither battery differentiated between asymptomatic and symptomatic HIV+ individuals. However, the seronegative group differed significantly from the symptomatic group on two of the CalCAP measures and on three measures from the conventional NP battery. Sequential RT (CRT4) was identified as one of the CalCAP measures most sensitive to HIV-associated NP impairment. In addition, based on the CalCAP scores, impairment classifications (impaired *vs.* unimpaired) significantly differed between individuals in the symptomatic group and the control group. Impairment classifications based on scores from the conventional battery were not significantly different between the control group and the symptomatic group. The authors of this study concluded that computerized RT measures might be more sensitive than conventional NP measures at detecting neurobehavioral impairments associated with HIV infection.

Worth et al. (1993) reported that, of all the CalCAP measures, Choice RT (CRT3) and Sequential RT subtests (CRT4 & CRT14) best discriminated between groups of participants at different stages of AIDS dementia complex (ADC, based on the Memorial Sloan-Ketterling ADC Clinical Staging System) and a seronegative control group. That is, ADC stage I (mild) and II (moderate) groups were both significantly slower than controls on CRT3 and CRT4, and all three ADC groups (i.e., equivocal, mild, moderate) were significantly slower than controls on CRT14.

Even though the above evidence suggests that computerized RT measures are of some use in the NP assessment of HIV+ individuals, sufficient evidence is not currently available to determine their role in relation to conventional batteries. Specifically, should computerized RT measures replace conventional NP batteries, or possibly serve to augment them? One drawback of RT measures is that they may only tap a limited scope of cognitive impairment, such as attention and speed of information processing, virtually ignoring possible deficits in other cognitive domains, such as learning and memory, executive functions, verbal, and visuospatial skills. The current study investigates impairment classifications using comparable, previously established cut-points on subtests of the CalCAP and on conventional NP measures. In addition, this study examines the relationship between scores on both types of NP tests and markers of immunosuppression (CD4) and disease progression (in CSF & plasma). We hypothesized that CalCAP subtests would correlate most strongly with conventional NP measures of attention and processing speed, but less so with other NP domains (executive/abstraction, learning/memory, and motor). Due to the breadth of ability areas assessed by the conventional NP battery, we hypothesized that it would identify a larger percentage of individuals as "NP impaired," and would demonstrate stronger associations with biomarkers of HIV disease severity.

METHODS

Research Participants

We studied 82 HIV-seropositive men who were undergoing screening for possible participation in a clinical trial inves-

https://doi.org/10.1017/S1355617703910071 Published online by Cambridge University Press

tigating the effects of a medication treatment on HIV associated neurobehavioral impairment. Participants referred for screening were strongly suspected of having neurobehavioral impairments, as evidenced by self-reports and/or physician referral; therefore, the group studied should contain a relatively high percentage of NP-impaired individuals. Various *exclusion* criteria were used to eliminate prospective participants with potential confounds, including head injuries, psychiatric disorders, substance use disorders, and developmental disorders; details of these criteria are described in Heaton et al. (1995).

Participants in this sample were middle aged (mean = 40.4 years, *SD* = 7.6) and generally of average educational attainment (mean = 13.8 years, *SD* = 2.9). The sample was ethnically diverse, consisting of 59% White/Caucasian, 23% African American, 14% Hispanic, and 4% of other ethnic origin. Most participants were in advanced stages of HIV disease. Using criteria developed by the Center for Disease Control (Centers for Disease Control, 1992), 11% of participants were at stage A, 37% at stage B, and 52% had had AIDS defining illnesses (stage C). Blood levels of CD4 T-lymphocytes at the time of testing were available for 53 participants and averaged 269 (*SD* = 191). Seventy percent were classified as having AIDS, based upon clinical history and/or CD4 cell count below 200. The Log_{10} of HIV-1 viral load in blood plasma averaged 3.62 copies/mL (*SD* = 1.75; *n* = 70; 81% with detectable viral load), and the Log_{10} of HIV-1 viral load in CSF averaged 1.85 copies/mL (*SD* = 1.80; *n* = 60; 45% with detectable viral load).

Measures and Procedures

NP evaluation

All participants received NP tests administered by trained psychometrists using standardized procedures. The battery was composed of several tests considered sensitive in detecting HIV-associated NP impairment. It is a modified version of a battery recommended by the NIMH Workshop on Neuropsychological Assessment Approaches (Butters et al., 1990) and incorporates tests recommended in various clinical intervention trial protocols. The battery consisted of traditional NP tests and takes approximately two hours to complete. Participants also completed the CalCAP "Mini" battery (Miller et al., 1991). The Mini version of the CalCAP battery contains two RT tests (CRT3 & CRT4) whose previous findings concerning HIV-related NP impairment are presented in the Introduction. Furthermore, this battery represents the most concise version of the CalCAP available and can be completed in approximately ten minutes. The traditional NP battery includes the following measures, grouped by ability domains:

Traditional NP battery:

1. Abstraction/Executive function (Exec): Halstead Category Test (Halstead, 1947; Heaton et al., 1991; Reitan & Wolfson, 1993), Trail Making Test–Part B (Army Individual Test Battery, 1944; Heaton et al., 1991), Stroop Color and Word Test (Golden, 1978)
2. Attention and working memory (Att/WM): WMS–R Visual Span (Wechsler, 1987), Paced Auditory Serial Addition Task (Diehr et al., 1998; Gronwall, 1977; Gronwall & Sampson, 1974), WAIS–R Digit Span (Wechsler, 1981)
3. Speed of information processing and perceptual motor abilities (SIP/PM): Symbol Digit Modalities (Smith, 1982), WAIS–R Block Design (Wechsler, 1981), Figural Visual Scanning (Wilkie et al., 1990), Trail Making Test–Part A (Army Individual Test Battery, 1944; Heaton et al., 1991)
4. Learning and memory (Lrn/Mem): Rey Auditory Verbal Learning Test–Learning Trials & Delayed Free Recall (Rey, 1941, 1962), WMS–R Visual Reproductions–VRI & VRII (Wechsler, 1987)
5. Complex motor skills (Motor): Grooved Pegboard Test (Heaton et al., 1991; Klove, 1963), dominant and nondominant hands.

CalCAP Mini Battery:

6. Choice Reaction Time–Digits (CRT3) (Miller, 1996; Miller et al., 1991)
7. Sequential Reaction Time #1 (CRT4) (Miller, 1996; Miller et al., 1991)

All raw scores were converted to *T*-scores using the most extensive normative data available. Whenever possible, these norms were corrected for relevant demographic influences. Because we were interested in detection and classification of NP impairment, *T*-scores from all tests were converted to deficit scores (*D*-scores) (Heaton et al., 1994, 1995). Much like clinicians' ratings, this method of transforming participants' scores emphasizes deficits in NP performance while minimizing the impact of superior performance on any particular NP test when computing composite scores. Moreover, *D*-scores have been found to closely approximate clinical ratings of NP impairment associated with HIV infection (Heaton et al., 1995); therefore, *D*-scores were generated for each test in the battery. Table 1 summarizes the relationship between *T*-scores and *D*-scores. A summary

Table 1. *T*-score to deficit-score conversion

T-score	Descriptor	Deficit score assignment
≥40	Normal	0
39–35	Mild	1
34–30	Mild-moderate	2
29–25	Moderate	3
24–20	Moderate-severe	4
<20	Severe	5

https://doi.org/10.1017/S1355617703910071 Published online by Cambridge University Press

D-score was calculated for each participant by averaging the *D*-scores for all traditional NP battery tests administered. In addition, composite *D*-scores were computed for each domain of the traditional NP battery by averaging the *D*-scores for tests within each domain for each participant. Likewise, age and education corrected *T*-scores and *D*-scores were computed for the two CalCAP measures using normative data provided by Miller (1991), and a summary CalCAP *D*-score was calculated for each participant by averaging the *D*-scores of these two subtests.

On some occasions, participants were unable to complete the traditional NP battery and CalCAP on the same day. The number of days between batteries averaged 1.7 ($SD = 10.6$) with most participants (87%) receiving both tests on the same day and 96% of participants completing both batteries within 15 days.

Neuromedical evaluation

Participants underwent a standardized neuromedical assessment conducted by trained nurse-clinicians. Blood components and CSF were collected. CD4 T-lymphocyte subset enumeration was obtained by flow cytometry. Levels of HIV-1 RNA in plasma and CSF were both quantified using Roche PCR assay (Ellis et al., 1997; Mulder et al., 1994). Due to the limitations of the instruments used in assessing plasma and CSF viral load, only values greater than 200 copies/mL were deemed to accurately represent the presence of "detectable" HIV-1 virus. Neuromedical evaluations were completed on the same day as NP testing for 43% of subjects, and for 66% neuromedical and NP evaluations were obtained within 15 days. Participants receiving the neuromedical and neuropsychological evaluation on the same day did not differ significantly on demographic, neuromedical, and neuropsychological variables.

Data Analyses

Relationships between the composite *D*-scores and three biomarkers of HIV disease progression and immunosuppression (CD4, CSF viral load, and plasma viral load) were examined. Associations between *D*-scores and CD4 count were analyzed using linear regression. CSF viral load and plasma viral load were converted to dichotomous variables (undetectable/detectable) due to their severely skewed distributions; therefore, analyses employing them as dependent variables were examined using logistic regression. A participant's viral load was considered "detectable" when the value was over 200 copies/mL. Otherwise, the viral load for that individual was classified as "undetectable."

Type I error was controlled at the .05 level by applying the Bonferroni adjustment to each family of analyses. Therefore, analyses using a single overall *D*-score (either traditional or CalCAP) were considered statistically significant if the observed *p*-value was .05 or less. Analyses using the traditional NP battery domain *D*-scores were deemed statistically significant if the observed *p*-value was .01 or less. Those analyses using the individual CalCAP *D*-scores were considered significant when the observed *p*-value was .025 or less.

RESULTS

Impairment Classifications

Average raw scores and average deficit scores for all tests administered appear in Table 2. Average summary deficit scores were 0.96 ($SD = 1.03$) and 0.87 ($SD = 1.14$) for the traditional NP battery and the CalCAP, respectively. Participants were classified as NP impaired when obtaining a summary *D*-score of .5 or above. This cut-point is roughly equivalent to being mildly impaired on one half of the component measures, and was determined to provide optimal agreement with clinician ratings of NP impairment in a prior, large scale study ($N = 500$) of HIV seropositive and seronegative subjects (Heaton et al., 1995). Using a test of significant differences between correlated proportions, *D*-scores based on the traditional NP battery resulted in a trend toward significantly higher NP impairment rates when compared to *D*-scores based on the CalCAP (traditional NP battery = 57% and CalCAP = 49%; $z = 1.46$, $p = .07$). Using the traditional NP battery as the "gold standard," the CalCAP exhibited moderate specificity (77%) and sensitivity (68%), with an overall agreement of 72%. Nevertheless, the instruments demonstrated low to moderate agreement on impairment classifications (kappa = .44, $SD = 0.88$). A Receiver Operator Characteristic (ROC) curve was constructed to establish what cut-point would provide optimal sensitivity and specificity for impairment classifications based on *D*-scores from the CalCAP Mini battery in comparison to impairment classifications based on the traditional NP battery. Again, we found a cut-point of .5 or greater to provide optimal sensitivity and specificity in relation to NP impairment classifications based on the traditional NP battery.

D-scores based on the traditional NP battery domains showed low to moderate correlations with *D*-scores based on individual CalCAP tests (see Table 3). Summary scores based on each battery were moderately correlated ($r = .49$; $p < .001$). Contrary to expectations, the CalCAP measures were not primarily related to traditional measures of SIP/PM ability. Even though scores on CRT3 demonstrated the highest correlation with the SIP/PM domain, correlations with other domains (i.e., Exec, Attn/WM) were of similar, and quite modest, magnitude.

Associations with Biomarkers of HIV Disease Progression

Overall, scores based on measures from the traditional NP battery showed slightly (though not significantly) greater association with biomarkers of HIV disease progression than scores based on the CalCAP. Specifically, the summary score

https://doi.org/10.1017/S1355617703910071 Published online by Cambridge University Press

Table 2. Mean deficit and raw scores on NP tests

Test	Raw score/D-score *M* (*SD*)	Test	Raw score/D-score *M* (*SD*)
Halstead Category Test	50.41(27.83)	Trails A	31.94(17.47)
	0.91(1.36)		0.70(1.29)
Trails B	93.56(71.18)	RAVLT Learning Trials	42.16(10.28)
	0.84(1.53)		1.22(1.62)
Stroop Color and Word	133.43(40.27)	RAVLT Delayed Free Recall	7.41(3.32)
	1.91(1.98)		1.15(1.43)
WMS–R Visual Span	15.68(3.67)	WMS–R Visual Reproduction I	30.48(6.14)
	0.35(0.81)		0.38(0.93)
PASAT	105.39(36.21)	WMS–R Visual Reproduction II	24.43(8.91)
	0.82(1.46)		0.73(1.50)
WAIS–R Digit Span	14.55(3.93)	Grooved Pegboard DH	76.18(20.89)
	0.55(1.01)		1.04(1.49)
Symbol Digit Modalities	44.78(11.85)	Grooved Pegboard NDH	86.45(22.46)
	1.37(1.78)		1.22(1.52)
WAIS–R Block Design	29.00(10.45)	Choice Reaction Time 3–Digits	438.52(70.66)
	0.56(1.03)		0.98(1.64)
Figural Visual Scanning	78.04(23.18)	Sequential Reaction Time 1	595.24(122.72)
	1.19(1.61)		0.76(1.29)

Note. WMS–R = Wechsler Memory Scale–Revised; PASAT = Paced Auditory Serial Addition Task; WAIS–R = Wechsler Adult Intelligence Scale–Revised; RAVLT = Rey Auditory Verbal Learning Test; TNB = traditional neuropsychological battery; DH = dominant hand; NDH = non-dominant hand.

based on traditional NP battery tests demonstrated a significant association with CD4 ($F(1, 52) = 12.36$, $p = .001$; see Table 4) and detectable CSF viral load (chi-square = 6.76, $p = .009$; see Table 5), but was unrelated to detectable plasma viral load (chi-square < 0.01, $p = .994$). The summary *D*-score based on the CalCAP correlated with both CD4 ($F(1, 52) = 4.13$, $p = .047$) and detectable CSF viral load (chi-square = 3.97, $p = .046$) to a lesser degree than the traditional NP battery—likewise, it did not relate to detectable plasma viral load (chi-square = 0.05, $p = .829$). A nonparametric bootstrap method to test differences between dependent correlations revealed that the strength of associations between biomarkers and the two test batteries were not significantly different from each other.

DISCUSSION

This study compared a traditional neuropsychological battery and subtests from the California Computerized Assessment Package (CalCAP; a collection of computerized reaction time tests) with respect to their classification of NP impairment in a sample of individuals suspected of HIV-related NP deficits. Furthermore, we determined the degree of association between these tests and biomarkers of HIV-associated immunosuppression and disease progression. Although many studies have examined the ability of either computerized RT or traditional NP tests to detect HIV-associated NP impairments, few have directly compared the two approaches in this manner.

Our sample was suspected of experiencing NP deficits, as indicated by self-report and/or physicians' referrals. The majority of participants in this study were in advanced stages of HIV disease—over half of them were classified as CDC stage C, and most exhibited low CD4 counts. Therefore, we anticipated that the more sensitive (accurate) battery would identify a greater number of individuals as impaired, and that its impairment classifications would be more related to biomarkers of HIV-related disease severity. However, to the extent that individuals without cognitive impairment were inappropriately referred to and included in the study, a lower than 100% impairment classification would actually be desirable. Although we have no independent gold standard for correct subject classification in this study, our comparisons between the traditional NP battery and the CalCAP suggest that the traditional NP battery may be more successful at identifying participants with HIV-related NP impairment.

Our results provide some empirical support for the seemingly obvious assertion that traditional NP tests measure different cognitive domains or constructs than are assessed

Table 3. Correlations between traditional NP domains and CalCAP-mini subtests

	Exec	Attn/WM	SIP/PM	Lrn/Mem	Motor
CRT3	.36**	.37**	.38**	.18	.22*
CRT4	.43**	.24*	.35**	.41**	.40**

Note. $N = 82$, except Attn/WM ($n = 80$); CRT3 = choice reaction time 3; CRT4 = sequential reaction time 1; Exec = abstraction/executive; Attn/WM = attention/working memory; SIP/PM = speed of information processing/perceptual motor; Lrn/Mem = learning/memory.
*$p < .05$ **$p < .01$

https://doi.org/10.1017/S1355617703910071 Published online by Cambridge University Press

Table 4. Relationships between battery deficit-scores and CD4 count

TNB measures	*r*	*p*-value	CalCAP measures	*r*	*p*-value
Summary TNB	−.44	.001*	Summary CalCAP	−.27	.047*
Abstraction/Executive	−.43	.001*	CRT3	−.16	.264
SIP/Perceptual Motor	−.37	.006*	CRT4	−.27	.050
Attention/Working Memory	−.34	.015			
Learning/Memory	−.30	.028			
Complex Motor	−.40	.003*			

Note. $N = 53$, except Attn/WM ($n = 52$); TNB = traditional neuropsychological battery; SIP = speed of information processing; CalCAP = California Computerized Assessment Package; CRT3 = choice reaction time 3; CRT4 = sequential reaction time 1.
* fulfilled criteria for statistical significance presented in Method section.

by computerized reaction time tests. Agreement on impairment classification between these instruments was modest to low, as were correlations between traditional NP battery domains and CalCAP subtests. These findings are consistent with a factor analysis conducted by Miller et al. (1991) that included other traditional NP measures and the CalCAP, and which indicated that CalCAP subtests load on separate factors that do not include traditional NP battery tests. The results of that study also suggest that computerized RT tests assess different NP domains than those assessed by a traditional battery.

Another method used to compare the batteries' ability to detect the neurobehavioral impact of HIV infection was to determine their association with biological markers of HIV disease progression, including evidence of HIV viral penetration in the brain, and ostensibly potential brain injury. CSF viral load, in particular, may serve as an indirect index of viral presence in the brain, and therefore may be more related to NP function than plasma viral levels (Ellis et al., 1997). Moreover, the brain may be a "sanctuary" for viral replication that can go on relatively independent from events in the periphery.

Even though the two test batteries demonstrated low agreement in impairment classifications, both demonstrated associations with HIV-related biomarkers. Traditional NP battery-based scores demonstrated somewhat stronger associations with the biomarkers under investigation than did those based on the CalCAP, although the differences between measures of association strength failed to demonstrate statistical significance. Our results also indicate that measures of processing speed and complex motor skills from the traditional NP battery consistently demonstrate greater associations with detectable CSF viral load and CD4 count, thus suggesting that these measures may be more sensitive to HIV-associated NP impairment than other domains from the traditional NP battery or CalCAP subtests.

Our results lend additional support to the role of CSF viral load as an indirect marker of viral burden in the brain. Although both neuropsychological batteries demonstrated a significant relationship with detectable CSF viral load, neither was reliably associated with viral load in plasma. Nevertheless, it could be suggested that our findings regarding CSF viral load may be attributable to the chronicity of HIV disease resulting in an indirect "psychological" confound (e.g., depression), associated with both CSF viral load and neuropsychological functioning. However, several studies have demonstrated that neuropsychological impairment in HIV disease cannot be explained on the basis of depression (Rourke et al., 1999a, 1999b) or constitutional

Table 5. Relationships between battery *D*-scores and detectable HIV viral load markers in cerebrospinal fluid ($n = 60$)

NP measures	R^2	*p*-value	NP measures	R^2	*p*-value
Summary TNB	.08	.009*	Summary CalCAP	.05	.046*
Abstraction/Executive	.04	.077	Choice Reaction Time 3	.06	.024
SIP/Perceptual Motor	.10	.005*	Sequential Reaction Time 1	.00	.596
Attention/WM$^+$	.03	.094			
Learning/Memory	.03	.136			
Complex Motor	.09	.006*			

Note. TNB = traditional neuropsychological battery; WM = Working Memory; SIP = speed of information processing; CalCAP = California Computerized Assessment Package.
$^+$sample size was ($n - 1$) for Attn/WM domain.
*fulfilled criteria for statistical significance presented in Method section.

https://doi.org/10.1017/S1355617703910071 Published online by Cambridge University Press

symptoms (Heaton et al., 1995), thus making this possibility unlikely. Analyses conducted with the current sample did not reveal statistically significant associations between scores on the Beck Depression Inventory (BDI) and CSF viral load ($r = 0.19$, $p = 0.13$) or with overall neuropsychological performance (as assessed by the traditional battery; $r = 0.17$, $p = 0.14$). Similarly, no statistically significant relationship was observed between overall neuropsychological performance on the traditional battery and BDI scores reflecting only somatic complaints ($r = .14$, $p = 0.24$). We also investigated a small subset of our sample with estimated dates of seroconversion to HIV-positive status ($n = 35$). In this subgroup, no meaningful relationship was observed between estimated duration of HIV infection and CSF viral load ($r = 0.16$, $p = 0.36$).

Our study has several limitations, the most significant being the lack of another "gold standard" measure of brain injury, aside from NP tests. Conducting a similar study with the addition of neuroimaging data and/or information on post-mortem neuropathology would potentially yield results that are more conclusive. Also, rather than relying solely upon previously published norms for these tests, the inclusion of a demographically comparable HIV seronegative control group would have been helpful. Furthermore, it is important to note that our results reflect a comparison of a traditional NP battery with a much briefer computerized battery (CalCAP-Mini); it is possible that a longer computerized battery (e.g., the full CalCAP or CalCAP-Abbreviated) may have demonstrated greater concordance with the traditional NP battery. Nevertheless, our goal was to compare the sensitivity of the CalCAP-Mini battery to that of a traditional NP battery similar to those commonly used and recommended in HIV research.

In summary, the results of our study suggest that computerized RT tests and a traditional NP battery are both sensitive to HIV-associated NP impairment. Traditional NP tests may be somewhat more sensitive, but the clearest conclusion from our results is that the test batteries are not equivalent in what they measure; that is, they are not interchangeable. An added advantage of assessing multiple ability domains, as is done by traditional batteries, is that it allows clinicians and researchers to better understand the nature of the specific deficits observed, thus allowing informed predictions about the functional impact of an individual's NP impairment and potential difficulties in everyday functioning. Although we did not conduct such analyses in our study, traditional NP batteries lend themselves more readily to the examination of the process by which an individual approaches the problem-solving demands required on many tests. Although it is possible for computerized RT tests to incorporate such analyses into their scoring systems, as far as we know such methods are not currently available. Further research may suggest that incorporation of computerized RT tests into traditional NP batteries may be useful, and provide information on how deficits detected by such batteries differ from those detected by traditional NP tests.

ACKNOWLEDGMENTS

The HIV Neurobehavioral Research Center (HNRC) is supported by Center award P50-MH 45294 from NIMH (DG). The Clinical Intervention Trial component is supported by award RO1 MH58076 (RJE). Stipend support for the first author (RG) was provided by the San Diego Fellowship. Gratitude is expressed to Eric N. Miller, Ph.D. for providing additional support and information on the California Computerized Assessment Package (CalCAP). Preliminary results associated with this investigation were presented at the 28th annual meeting of the International Neuropsychological Society. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the United States Government.

REFERENCES

Army Individual Test Battery. (1944). *Manual of directions and scoring*. Washington, D.C.: War Department, Adjutant General's Office.

Butters, N., Grant, I., Haxby, J., Judd, L.L., Martin, A., McClelland, J., Pequegnat, W., Schacter, D., & Stover, E. (1990). Assessment of AIDS-related cognitive changes: Recommendations of the NIMH Workshop on Neuropsychological Assessment Approaches. *Journal of Clinical & Experimental Neuropsychology*, *12*, 963–978.

Centers for Disease Control. (1992). 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *Morbidity and Mortality Weekly Report*, *41*, 1–19.

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.

Dunlop, O., Bjorklund, R.A., Abdelnoor, M., & Myrvang, B. (1992). Five different tests of reaction time evaluated in HIV seropositive men. *Acta Neurologica Scandinavica*, *86*, 260–266.

Ellis, R.J., Hsia, K., Spector, S.A., Nelson, J.A., Heaton, R.K., Wallace, M.R., Abramson, I., Atkinson, J.H., Grant, I., McCutchan, J.A., & the HNRC Group. (1997). Cerebrospinal fluid human immunodeficiency virus type 1 RNA levels are elevated in neurocognitively impaired individuals with acquired immunodeficiency syndrome. *Annals of Neurology*, *42*, 679–688.

Golden, C.J. (1978). *Stroop Color and Word Test*. Chicago: Stoelting.

Grant, I. & Martin, A. (1994). *Neuropsychology of HIV infection*. New York: Oxford University Press.

Gronwall, D.M. & Sampson, H. (1974). *The psychological effects of concussion*. Auckland, New Zealand: Auckland U Press.

Gronwall, D.M. (1977). Paced auditory serial addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, *44*, 367–375.

Halstead, W.C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.

Heaton, R.K., Grant, I., Butters, N., White, D.A., Kirson, D., Atkinson, J.H., McCutchan, J.A., Taylor, M.J., Kelly, M.D., Ellis, R.J., Wolfson, T., Velin, R., Marcotte, T.D., Hesselink, J.R., Jernigan, T.L., Chandler, J., Wallace, M., Abramson, I., & the HNRC Group. (1995). The HNRC 500—Neuropsychology of HIV infection at different disease stages. *Journal of the International Neuropsychological Society*, *1*, 231–251.

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, Inc.

Heaton, R.K., Kirson, D., Velin, R.A., Grant, I., & the HNRC Group. (1994). The utility of clinical ratings for detecting cognitive change in HIV infection. In I. Grant & A. Martin (Eds.), *Neuropsychology of HIV infection* (pp. 188–206). New York: Oxford University Press.

Klove, H. (1963). Clinical neuropsychology. In F. M. Forster (Ed.), *The Medical Clinics of North America*. New York: Saunders.

Martin, A., Heyes, M.P., Salazar, A.M., Kampen, D.L., Williams, J., Law, W.A., Coats, M.E., & Markey, S.P. (1992). Progressive slowing of reaction time and increasing cerebrospinal fluid concentrations of quinolinic acid in HIV-infected individuals. *Journal of Neuropsychiatry and Clinical Neurosciences*, *4*, 270–279.

Martin, E.M., Pitrak, D.L., Novak, R.M., Pursell, K.J., & Mullane, K.M. (1999). Reaction times are faster in HIV-seropositive patients on antiretroviral therapy: A preliminary report. *Journal of Clinical and Experimental Neuropsychology*, *21*, 730–735.

Miller, E., Satz, P., & Visscher, B. (1991). Computerized and conventional neuropsychological assessment of HIV-1 infected homosexual men. *Neurology*, *41*, 1608–1616.

Miller, E.N. (1996). CalCAP: California Computerized Assessment Package (manual) (Version Revision 4.1). Los Angeles: Norland Software.

Mulder, J., McKinney, N., Christopherson, C., Sninsky, J., Greenfield, L., & Kwok, S. (1994). Rapid and simple PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma: Application to acute retroviral infection. *Journal of Clinical Microbiology*, *32*, 292–300.

Reitan, R.M. & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.

Rey, A. (1941). The psychological examination in the traumatic cases of encephalopathy. *Archives de Psychologie*, *28*, 286–340.

Rey, A. (1962). *The clinical examination in psychology*. Paris: Presses Universitaires de France.

Rourke, S.B., Halman, M.H., & Bassel, C. (1999a). Neurocognitive complaints in HIV-infection and their relationship to depressive symptoms and neuropsychological functioning. *Journal of Clinical & Experimental Neuropsychology*, *21*(6), 737–756.

Rourke, S.B., Halman, M.H., & Bassel, C. (1999b). Neuropsychiatric correlates of memory-metamemory dissociations in HIV-infection. *Journal of Clinical & Experimental Neuropsychology*, *21*(6), 757–768.

Smith, A. (1982). *Symbol Digit Modalities Test*. Los Angeles: Western Psychological Services.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–Revised (manual)*. New York: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale–Revised*. New York: The Psychological Corporation.

Wilkie, F.L., Eisdorfer, C., Morgan, R., Loewenstein, D.A., & Szapocznik, J. (1990). Cognition in early human immunodeficiency virus infection. *Archives of Neurology*, *47*(4), 433–440.

Worth, J.L., Savage, C.R., Baer, L., Esty, E.K., & Navia, B.A. (1993). Computer-based neuropsychological screening for AIDS dementia complex. *AIDS*, *7*, 677–681.