Published in final edited form as:
J Med Assoc Thai. 2014 February ; 97(0 2): S222–S233.

Practice Effect and Normative Data of an HIV-specific Neuropsychological Testing Battery among Healthy Thais

Pasiri Sithinamsuwan, MD[*], **Nicholas Hutchings, BA**[**], **Jintanat Ananworanich, MD, PhD**[***,****,*****], **Lauren Wendelken, MS**[******], **Putthachard Saengtawan, BNS**[***], **Robert Paul, PhD**[*******], **Nitiya Chomchey, MSc**[***], **James LK Fletcher, BMBCh**[***], **Thep Chalermchai, MD**[***], and **Victor Valcour, MD, PhD**[******]

[*]Division of Neurology, Department of Medicine, Phramongkutklao Hospital, Bangkok, Thailand [**]College of Physicians and Surgeons, Columbia University, New York, USA [***]SEARCH, The Thai Red Cross AIDS Research Centre, Bangkok, Thailand [****]HIV-NAT, The Thai Red Cross AIDS Research Centre, Bangkok, Thailand [*****]Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand [******]Memory and Aging Center, University of California, San Francisco, CA, USA [*******]Department of Psychology, University of Missouri in Saint Louis, Saint Louis, MO, USA

Abstract

Objective—A longitudinal cohort study was conducted in Bangkok, Thailand between 2008 and 2013 in order to determine the practice effect of serial neuropsychological testing and establish normative data among normal (HIV-uninfected) Thai volunteers.

Material and Method—The authors enrolled 511 cognitively healthy individuals (HIV-uninfected, no drug abuse or other previous/current neurological or psychological conditions) to assess baseline performance on a HIV-specific neuropsychological testing battery. Ninety-nine subjects were re-assessed at 6 and 12 months to evaluate practice effects.

Results—The mean age of the 99 subjects completing longitudinal visits was 49.2 years and 53 were male. The authors identified improved mean raw scores on most neuropsychological tests with repeated measurements; however, only change in WHO Auditory Verbal Learning Test

Correspondence to: Sithinamsuwan P, Division of Neurology, Department of Medicine, Phramongkutklao Hospital, Bangkok 10400, Thailand. Phone: 08-3236-7772, pasiri21795@yahoo.co.uk.

Competing Interests
The authors declare no conflicts of interest.

Authors' contributions
PS: contributed to study design, data collection, data analysis, literature review and manuscript preparation.
NH: contributed to data analysis, literature review, and manuscript preparation.
JA: contributed to study design, data collection, literature review and manuscript preparation.
LW: contributed to data analysis, literature review, and manuscript preparation.
PS: contributed to data collection, and manuscript preparation.
RP: contributed to study design, data collection, and manuscript preparation.
NC: contributed to data collection, and manuscript preparation.
JF: contributed to study design, data collection, and manuscript preparation.
TC: contributed to study design, data collection, and manuscript preparation.
VV: contributed to study design, data collection, data analysis, literature review and manuscript preparation.

Potential conflicts of interest
None.

(AVLT) scores (learning, attention, immediate and delayed recall tasks) met statistical significance, with larger differences seen between baseline and 6-month compared to 6 and 12 months follow-up. Older age correlated with poorer baseline raw score, and was a predictor of worse performance at 6 months and 12 months on several tasks. Level of education was associated with practice effects on several tests. No similar effects were observed with gender.

Conclusion—The authors identified improved performance after repeated measurements revealing a significant practice effect on an HIV-specific neuropsychological testing battery employed in Bangkok, Thailand. Main predictors were age and educational attainment.

Keywords

Thailand; neuropsychological test; cognition

The determination of cognitive impairment is heavily based on neuropsychological testing performance, which, in turn, requires a clear understanding of normative performance within a population in order to interpret accurately test scores. Neuropsychological evaluations must be repeated at multiple time points to assess individual or group improvement or deterioration over time. Such assessment must take into account the practice effect[1,2]. Generally, test scores can be expected to improve for cognitively healthy individuals if the same subject takes the same test multiple times, particularly if this is done within a short interval of one year or less. There are a variety of factors that may influence the magnitude of practice effects, including age and education[3]. Practice effects confound the determination of impairment, since modest improvement in neuropsychological testing may reflect learning rather than true cognitive improvement, and unchanged test scores over time may actually indicate cognitive decline offset by the practice effect. Practice effects are commonly observed on tests that have content that requires learning or memorization, such as verbal and visual memory tests. Separately, practice effects also can be seen when a skill is learned (e.g. grooved pegboard or more accurately reverse-mirror star drawing). Finally, practice effects can be observed when a test has a "solution" and once a subject has solved the problem or task, then repeat exposure to that same test will reflect the learning of the solution.

The SEARCH group, operating in collaboration with the Thai Red Cross AIDS Research Centre in Bangkok, investigates the cognitive complications of living with HIV among Thais in Bangkok. Over the course of the past five years, the authors have assessed neuropsychological test performance in HIV-infected and uninfected adults. Here, the authors analyzed scores from serial testing conducted at baseline, 6 and 12 months among 99 healthy HIV-uninfected Thai subjects, ranging from 20 to 70 years old, in order to define practice effects for an HIV-specific neuropsychological testing battery. In addition, we report normative performance in this cohort providing important data for the interpretation of neuropsychological testing in Thailand.

Material and Method

Subject selection

HIV-uninfected, healthy adult subjects were recruited for establishing normative neuropsychological data to support HIV research studies conducted at SEARCH-Thailand (ClinicalTrials.gov NCT00713752). Five hundred and eleven control subjects stratified by age and educational attainment were enrolled to develop normative Thai data and 99 of these subjects were re-assessed at 6 months and 12 months to investigate practice effects and the primary determinants of these effects for individual tests in this international HIV neuropsychological testing battery.

Enrollees were stratified by age (10-year age intervals, 20 to 79 years old) and level of education (no certificate or primary school certificate; less than high school or vocational certificate; high school diploma or vocational certificate; and Bachelor's degree or higher). Subjects were tested to be HIV-antibody negative and were excluded for traumatic head injury, illicit drug use, severe illness, fevers or meningeal signs, pre-existing neurologic disease, major depression, learning disability, and signs and symptoms of metabolic encephalopathy or delirium. In addition, study physicians excluded subjects with focal neurological deficit upon examination and those deemed unable to complete visit requirements.

Neuropsychological Testing

Neuropsychological tests were administered by nurses trained in test administration and scoring. A test manual was developed and utilized at each assessment to ensure standardized test administration parameters at each visit. The test manual included instructions for test administration and scoring of responses. Quality assurance was provided by one of two investigators with training in neuropsychological testing (RP and VV), and included oversight of test administration and verification of scoring procedures. The testing battery was previously designed to minimize cultural influences for use in international HIV studies[4]. The neuropsychological battery consisted of the following tests: Timed Gait, Finger Tapping for both dominant and non-dominant hands; the WHO Auditory Verbal Learning Test (AVLT), Color Trails Tests 1 and 2, the Escala de Inteligencia de Wechsler para Adultos (EIWA) Block Design task, the EIWA Digit Symbol Modalities Test, the Grooved Pegboard test for both dominant and non-dominant hands, the Trail Making Test A, the Brief Visual Memory Retention Test-Revised (BVMT-R) and verbal fluency animals and first names. The battery was conducted in the same order for all subjects at all time points; alternate forms of the tests were not used. Test administrators ensured a comparable testing environment at each administration of the battery.

Several measures within these tests were studied for practice effects. Within the AVLT, the sum of trials 1–5 (AVLT 1–5) was used to measure verbal learning efficiency. Here, the participant is read the same 15 words 5 times, and each time is asked to recall as many words as possible. We estimated practice effects for the learning trials (sum of trials 1–5), as well as practice effects during trial 6 (AVLT 6), where the subject is read an interference list of 15 new words and must recall as many as possible. Trial 6 is generally thought to measure attention. Then, in trial 7 the subject is asked to recall the first list (AVLT Trial 7, immediate recall), and after a 30-minute delay, the subject is again asked to recall the original list of words (AVLT Trial 8, delayed recall). For the BVMT-R, we report practice effects for the sum of learning trials 1- 4 (BVMT-R Total) and the delayed recall trial. For other tests, we examined scores for each of the measures at each time point

Statistical analysis

Data were reported in number, percentage, mean, median and standard deviation. Multiple comparisons (ANOVA for parametric variables) and paired t-test /multiple related samples (for non-parametric variables) were used to determine difference among groups and within group at each time point. The authors evaluated correlations of age and Neurological scores using Pearson's test. Practice effects were estimated as a change in score between one testing session and the next, comparing group means. Regression models were used to determine predictors of change in test score. A p-value<0.05 was considered statistically significant. Data were analyzed by Stata version 13.

Results

Baseline data from the full cohort were used to create normative lookup tables, available in Supplement 1. Of the 99 subjects who completed three longitudinal visits, the mean (SD) age was 49.2 (13.8) years (range 20–69), and 53 (53.5%) were male (see Table 1). Thirty subjects (30.3%) had primary school certificate or less, 27 (27.3%) had less than high school or vocational certificate, 28 (28.3%) had a high school diploma or vocational certificate, and 14 (14.1%) had a Bachelor's degree or above.

Visual inspection of mean raw scores across the longitudinal cohort suggested higher scores over time in nearly all tests (Table 2). Differences in mean scores across time points met statistical significance only for portions of the AVLT – specifically, learning (sum of Trials 1–5), attention (Trial 6) immediate (Trial 7) and delayed recall (Trial 8). Examination of contrasts within these models noted that the greatest improvement was generally seen between baseline and 6 months with the following mean change: learning: 3.67 (p<0.001); immediate: 0.72 (p=0.023); delayed recall: 0.75 (p=0.018). Some additional improvement was also seen between 6 months and 12 months for the learning score: 1.74 (p=0.012) and a small but significant decline was seen in attention (Trial 6): −0.62 (p=0.004).

As expected, age had a significant influence on neuropsychological testing performance. At baseline, older age correlated to poorer neuropsychological test performance on all scores except on the test of AVLT learning (trial 6) and verbal fluency for first names and animals (Table 3). Age was also a significant and negative predictor of change in test scores on BVMT-R Total, Color Trails 1, Color Trails 2 and EIWA Block Design at 6 months compared to baseline, and on AVLT attention (Trial 6), immediate recall (Trial 7) and delayed recall (Trial 8), Color Trails 2, and Grooved pegboard for both dominant and non-dominant hands at 12 months compared to baseline (Table 4).

Level of educational attainment was associated with neuropsychological test performance on many, but not all, measures when comparing median scores (Table 5). Associations were

NIH-PA Author Manuscript

NIH-PA Author Manuscript

seen particularly for tests of psychomotor speed, executive functioning, and verbal fluency, where higher educational attainment was associated with better performance. Education also influenced the practice effect on some tasks, with significant differences in change in test scores seen between educational groups across the three testing sessions (Table 6). Although the authors did not find any significant differences in practice effect between genders, the pattern of change was different between males and females. Both groups showed improvement in scores on AVLT subtests, while only females showed significant improvement in Trails A at 6 and 12 months.

Discussion

To investigate cognitive most accurately performance by neuropsychological testing, it is important to include control subjects from a similar cultural and demographic background to allow appropriate adjustment. However, few normative data are available for neuropsychological testing in Thailand, and even fewer data provide insight into anticipated changes due to practice effect among healthy individuals as would be needed to interpret scores from longitudinal studies. This report provides such data for a standardized HIV-specific testing battery used extensively in research studies within Bangkok. These data are not limited to use in research, however, since normative data and anticipated practice effects may be useful in monitoring treatment effects for HIV-infected Thais in a clinical setting. The relevance of these findings in clinical settings is attenuated by the small magnitude of measured practice effects.

Although performances on nearly all tests improved over time, performance on only about half of these tests met statistical significance for practice effect across the three time points. Among these, most changed significantly between baseline and 6 months as anticipated, confirming that the largest practice effects are noted with the first repeat testing. Moreover, the magnitude of change in test scores was universally larger at 6 months versus baseline, compared to 6 versus 12 months, again demonstrating that the majority of learning occurred between the first and second testing sessions. Our finding that significant changes were still noted to occur in some tasks between 6 and 12 month time points provides evidence that learning continues to occur beyond the second testing session.

The primary factors that predicted baseline scores and change in scores over time were age and education. The authors found little effect of gender. Although practice effect patterns were slightly different between males and females, there were no significant differences between the groups. Older individuals and those with less education tended to have lower scores at baseline, and both age and education influenced the practice effect across a range of cognitive domains. Tests of psychomotor speed were not influenced by education, though age did influence change in test scores on the grooved pegboard test. Some tests that were influenced by age and education at 6 months were less influenced at 12 months, and vice versa, suggesting that number of sessions and interval between testing sessions moderates these effects.

In the U.S., researchers have shown that the magnitude of practice effect varies with type of test and interval between testing sessions[3]. The effect is anticipated to be stronger with shorter intervals between testing sessions. For some tests, participants may develop a strategy during early testing sessions which subsequently improves their performance[5]. In order to control for the practice effect, a number of statistical and practical methods have been proposed[6]. For memory-based tests, the effect can be ameliorated to some extent by the use of alternate forms of the test, where the questions or tasks are equivalent but not identical to the previous form administered[7]. It is important to note that in order for an alternative form of a test to be considered valid, it must first be shown to be comparable to the original form via normative testing.

Alternatively, some researchers have proposed using a “dual baseline” approach, wherein the same test is administered twice within a short period of time, thereby capturing the largest degree of learning at the baseline visit[8,9]. This approach relies on the assumption that the greatest change occurs between the first and second test sessions, and the second score is used as a practice-effect-adjusted score against which further testing may be compared. Research suggests that this may be a less effective strategy, because test scores do not appear to plateau after the second administration, and as noted in our own analysis here, practice effect continues to be seen with further assessments[1].

If the same test form is used at multiple time points, the final scores can be adjusted to compensate for practice effect through statistical methods[10,11]. This approach relies on complex modeling and enables the researcher to take into account other factors that can influence the practice effect such as age, test-retest interval, and the style of test in question. However, this statistical approach is less practical for a clinical setting. It is also dependent on very clean and robust normative data, demographically and culturally matched to the study cohort, in which the practice effect has been determined.

Conclusion

In summary, this work demonstrates practice effects on neuropsychological test performance among Thais. Our findings are consistent with the literature from U.S. populations where age and education influence the practice effect, as well as overall performance. Many of these tests were influenced by age and education, further highlighting the need for culturally appropriate norms to allow accurate interpretation of data. The tables included in this manuscript and the normative data provided in Supplement 1 provide precisely this essential information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to thank the study subjects for their participation and Isabel Allen for assistance with the statistical analyses.

The authors also thank all our contributors from SEARCH 001 & 002 protocol teams. From UCSF: Carol Liu, Elijah Mun, Stephanie Chiao, Lauren Wendelken, and Nicholas Hutchings; From SEARCH: Nittaya Phanuphak, Eugene Kroon, Duanghathai Suttichom, Somprartthana Rattanamanee, Somporn Tipsuk, Ratchapong Kanaprach, and Wichitra Apateerapong; From Phramongkutklao Hospital: Dr. Samart Nidhinansana, and Dr. Parnsiri Chairungsaris.

Funding

This work was supported by the United States National Institutes of Health grants R01-NS061696 (VV) and R21 MH072388 (VV).

References

1. Johnson BF, Hoch K, Johnson J. Variability in psychometric test scores: the importance of the practice effect in patient study design. Prog Neuropsychopharmacol Biol Psychiatry. 1991; 15:625–635. [PubMed: 1956991]
2. Basso MR, Carona FD, Lowery N, Axelrod BN. Practice effects on the WAIS-III across 3- and 6-month intervals. Clin Neuropsychol. 2002; 16:57–63. [PubMed: 11992227]
3. Dikmen SS, Heaton RK, Grant I, Temkin NR. Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery. J Int Neuropsychol Soc. 1999; 5:346–356. [PubMed: 10349297]
4. Maj M, Satz P, Janssen R, Zaudig M, Starace F, D'Elia L, et al. WHO Neuropsychiatric AIDS study, cross-sectional phase II. Neuropsychological and neurological findings. Arch Gen Psychiatry. 1994; 51:51–61. [PubMed: 8279929]
5. Basso MR, Bornstein RA, Lang JM. Practice effects on commonly used measures of executive function across twelve months. Clin Neuropsychol. 1999; 13:283–292. [PubMed: 10726600]
6. Sawrie SM, Marson DC, Boothe AL, Harrell LE. A method for assessing clinically relevant individual cognitive change in older adult populations. J Gerontol B Psychol Sci Soc Sci. 1999; 54:116–124.
7. Benedict RH, Zgaljardic DJ. Practice effects during repeated administrations of memory tests with and without alternate forms. J Clin Exp Neuropsychol. 1998; 20:339–352. [PubMed: 9845161]
8. Collie A, Maruff P, Darby DG, McStephen M. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. J Int Neuropsychol Soc. 2003; 9:419–428. [PubMed: 12666766]
9. Bartels C, Wegrzyn M, Wiedl A, Ackermann V, Ehrenreich H. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. BMC Neurosci. 2010; 11:118. [PubMed: 20846444]
10. Temkin NR, Heaton RK, Grant I, Dikmen SS. Detecting significant change in neuropsychological test performance: a comparison of four models. J Int Neuropsychol Soc. 1999; 5:357–369. [PubMed: 10349298]
11. Woods SP, Childers M, Ellis RJ, Guaman S, Grant I, Heaton RK. A battery approach for measuring neuropsychological change. Arch Clin Neuropsychol. 2006; 21:83–89. [PubMed: 16169705]

Table 1

Demographic characteristics of 99 patients with longitudinal assessment

Variables	Total n = 99
Age (years), mean (SD)	49.21 (13.8)(range 20.02–69.92 years)
Male	53 (53.5%)
Education	
Group 1: No certificate or primary school certificate	30 (30.3%)
Group 2: Less than high school diploma or vocational certificate	27 (27.3%)
Group 3: High school diploma or vocational certificate	28 (28.3%)
Group 4: Bachelor's degree or higher	14 (14.1%)
Employed	73 (73.7%)
Depression score, mean(SD)(by Thai depression inventory)	7.86 (6.0) (range 0–31)
No depression	95 (96%)
Mild depression	2 (2%)
Moderate depression	1 (1%)
Major and severe depression	0 (0%)

Table 2

Neuropsychological test scores across the longitudinal cohort at baseline, 6 months, and 12 months (n = 99)

Visit	Baseline			6-month			12-month			*p*-value[b]
Neuropsychological test	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**	
Finger Tapping - dominant hand	50.56	6.8	50.8	49.82	6.8	49.5	49.89	7.3	50.6	0.718
Finger Tapping - non-dominant hand	46.07	6.0	46.1	46.31	6.9	45.6	45.99	6.2	45.6	0.937
AVLT sum of Trials1-5 (learning)	53.08	6.9	52.0	56.75	6.2	58.0	58.48	7.2	59.0	<0.001*
AVLT-Trial 6 (attention)[a]	6.36	1.8	6.0	6.50	1.7	6.0	5.88	1.7	6.0	0.031*
AVLT-Trial 7 (immediate recall)[a]	11.22	2.3	12.0	11.94	2.1	12.0	12.19	2.0	12.0	0.004*
AVLT-Trial 8 (delayed recall)[a]	11.35	2.3	12.0	12.10	1.9	12.0	12.30	2.3	13.0	0.007*
BVMT total of trials 1–4	21.66	7.5	22.0	22.50	7.0	23.0	20.80	7.3	22.0	0.264
BVMT – delayed recall	8.93	2.7	10.0	8.76	2.6	9.0	8.18	2.5	9.0	0.021*
Color Trails 1[a]	54.26	23.1	48.1	50.24	23.3	44.8	48.12	20.1	42.4	0.144
Color Trails 2	110.45	45.9	100.3	102.85	38.1	96.3	99.36	37.1	92.2	0.146
EIWA Digit Symbol	49.83	16.6	52.0	52.73	15.5	52.0	52.87	16.4	55.0	0.329
Grooved Pegboard - dominant hand[a]	69.47	17.3	66.8	68.27	17.4	65.3	67.09	13.8	64.1	0.592
Grooved Pegboard - non-dominant hand[a]	77.27	15.8	73.5	78.82	23.0	72.9	78.06	19.3	73.6	0.857
Timed Gait[a]	12.42	3.3	12.0	12.19	2.9	11.9	11.97	1.5	11.8	0.498
Verbal Fluency – first names	19.60	5.7	19.0	20.12	5.8	21.0	20.3	5.6	20.0	0.659
Trails A	37.07	14.4	34.8	33.78	12.1	31.2	35.74	13.9	32.0	0.226
EIWA Block Design	33.31	7.5	33.0	33.33	8.2	32.0	33.76	8.6	34.0	0.910
Verbal Fluency - animals	22.19	5.2	22.0	22.72	6.0	22.0	22.69	5.3	23.0	0.752

Table 3

Correlation of age and baseline neuropsychological test score

Neuropsychological test	r	*p*-value[b]
Finger Tapping dominant hand	−0.477	<0.001*
Finger Tapping non-dominant hand	−0.510	<0.001*
AVLT sum of trials 1–5	−0.309	0.002*
AVLT Trial 6[a]	−0.088	0.387
AVLT Trial 7[a]	−0.325	0.001*
AVLT Trial 8[a]	−0.332	0.001*
BVMT total of trials 1–4	−0.517	<0.001*
BVMT – delayed recall	−0.460	<0.001*
Color Trails 1[a]	0.441	<0.001*
Color Trails 2	0.512	<0.001*
EIWA Digit Symbol	−0.470	<0.001*
Grooved pegboard dominant hand[a]	0.368	<0.001*
Grooved pegboard non-dominant hand[a]	0.491	<0.001*
Timed Gait[a]	0.254	0.011*
Verbal Fluency – first names	−0.042	0.682
Trails A	0.508	<0.001*

Table 4

Effect of age on practice effect

Neuropsychological test	**Change coefficient**	**Standard error**	***p*-value[b]**
6 month vs. baseline			
BVMT Total	−0.195	0.075	0.011*
Color Trails 1[a]	0.432	0.230	0.064
Color Trails 2	1.053	0.417	0.014*
EIWA Block Design	−0.164	0.096	0.092
12 month vs. baseline			
AVLT-6[a]	−0.050	0.019	0.011*
AVLT-7[a]	−0.039	0.020	0.057
AVLT-8[a]	−0.045	0.019	0.023*
Color Trails 2	0.388	0.205	0.063
Grooved pegboard – dominant hand[a]	0.214	0.103	0.041*
Grooved pegboard – non-dominant hand[a]	0.357	0.148	0.019*

Table 5

Baseline NP test score among 4 different certifications. Group 1: no certificate or primary school certificate; Group 2: less than high school diploma or vocational certificate; Group 3: high school diploma or vocational certificate; Group 4: Bachelor's degree or higher

Certification Group	**Group 1**	**Group 2**	**Group 3**	**Group 4**	
	(n = 30)	**(n = 27)**	**(n = 28)**	**(n = 14)**	***p*-value[b]**
Neuropsychological test	**Median score**	**Median score**	**Median score**	**Median score**	
Finger Tapping dominant hand	51.8	49.6	51.9	47.3	0.475
Finger Tapping non-dominant hand	46.8	47.6	46.3	41.4	0.043*
AVLT sum of trials 1–5	50.5	52.0	55.0	53.5	0.345
AVLT-6[a]	5.0	7.0	6.5	6.5	0.006*
AVLT-7[a]	12.0	11.0	12.0	11.5	0.624
AVLT-8[a]	12.0	11.0	12.0	12.0	0.924
AVLT-9[a]	15.0	15.0	14.0	14.0	0.474
BVMT Total	18.5	26.0	24.0	24.5	0.051
BVMT – delayed recall	7.5	10.0	10.0	10.0	0.021*
Color Trails 1[a]	56.9	43.3	47.7	47.1	0.032*
Color Trails 2	126.9	92.1	88.9	88.2	0.042*
EIWA Digit Symbol	34.0	55.0	53.5	59.5	0.001*
Grooved pegboard dominant hand[a]	73.8	65.5	64.4	63.5	0.386
Grooved pegboard non-dominant hand[a]	81.3	68.9	72.3	73.2	0.043*
Timed Gait[a]	11.8	12.1	12.2	12.0	0.900
Verbal Fluency – first names	17.5	18.0	21.0	22.0	0.001*
Trails A	39.9	27.1	31.85	33.2	0.024*
EIWA Block Design	28.0	36.0	32.5	34.0	0.018*

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 6

Change in neuropsychological test scores between 4 different certification groups where significant associations were noted. Group 1: no certificate or primary school certificate; Group 2: less than high school diploma or vocational certificate; Group 3: high school diploma or vocational certificate; Group 4: Bachelor's degree or higher

	6-month vs. Baseline					12-month vs. Baseline				
	Mean Change				*p*-value	Mean Change				*p*-value
Neuropsychological test	**Group 1**	**Group 2**	**Group 3**	**Group 4**		**Group 1**	**Group 2**	**Group 3**	**Group 4**	
AVLT-6[a]	0.93	−0.19	−0.39	0.57	0.151	0.31	−0.78	−0.79	−0.93	0.049*
AVLT-7[a]	−0.21	1.67	1.07	0.50	0.119	0.76	2.00	0.75	0.14	0.038*
BVMT – delayed recall	−0.13	0.14	−0.07	−1.07	0.021*	−0.37	−0.59	−0.79	−1.79	0.009*
Color Trails 2	−26.41	1.63	−12.07	23.36	0.041*	−13.48	−10.59	−12.36	−2.93	0.645
EIWA Digit Symbol	10.31	1.59	4.11	−14.93	0.011*	1.69	4.59	3.43	1.93	0.673
Verbal Fluency – first names	4.03	2.48	−2.14	−3.93	0.001*	0.34	1.33	0.25	1.50	0.707
Trails A	−9.72	−0.04	−5.61	5.86	0.044*	−1.31	1.81	−4.54	−1.43	0.225
Verbal Fluency – animals	2.03	1.63	0.71	−4.00	0.040*	0.72	0.52	0.89	−0.71	0.743