Validity of a novel computerized cognitive battery for mild cognitive impairment

Tzvi Dwolatzky¹, Victor Whitehead², Glen M Doniger³, Ely S Simon³, Avraham Schweiger⁴, Dena Jaffe⁵ and Howard Chertkow*²

Address: ¹Memory Disorders Clinic and Dept. of Geriatrics, Shaare Zedek Medical Center, Jerusalem, Israel, ²Department of Clinical Neurosciences, McGill University, Montreal, Canada, ³Department of Clinical Science, NeuroTrax Corporation, New York, NY USA, ⁴Department of Behavioral Sciences, Academic College of Tel Aviv, Tel Aviv, Israel and ⁵School of Public Health, Hebrew University – Hadassah Medical School, Jerusalem, Israel

Email: Tzvi Dwolatzky - tzvidov@bgumail.bgu.ac.il; Victor Whitehead - victorw@cam.org; Glen M Doniger - glen@neurotrax.com; Ely S Simon - ely@neurotrax.com; Avraham Schweiger - schweige@mta.ac.il; Dena Jaffe - denaj@neurotrax.com; Howard Chertkow* - howard.chertkow@mcgill.ca

* Corresponding author

Abstract

Background: The NeuroTrax Mindstreams computerized cognitive assessment system was designed for widespread clinical and research use in detecting mild cognitive impairment (MCI). However, the capability of Mindstreams tests to discriminate elderly with MCI from those who are cognitively healthy has yet to be evaluated. Moreover, the comparability between these tests and traditional neuropsychological tests in detecting MCI has not been examined.

Methods: A 2-center study was designed to assess discriminant validity of tests in the Mindstreams Mild Impairment Battery. Participants were 30 individuals diagnosed with MCI, 29 with mild Alzheimer’s disease (AD), and 39 healthy elderly. Testing was with the Mindstreams battery and traditional neuropsychological tests. Receiver operating characteristic (ROC) analysis was used to examine the ability of Mindstreams and traditional measures to discriminate those with MCI from cognitively healthy elderly. Between-group comparisons were made (Mann-Whitney U test) between MCI and healthy elderly and between MCI and mild AD groups.

Results: Mindstreams outcome parameters across multiple cognitive domains significantly discriminated among MCI and healthy elderly with considerable effect sizes ($p < 0.05$). Measures of memory, executive function, visual spatial skills, and verbal fluency discriminated best, and discriminability was at least comparable to that of traditional neuropsychological tests in these domains.

Conclusions: Mindstreams tests are effective in detecting MCI, providing a comprehensive profile of cognitive function. Further, the enhanced precision and ease of use of these computerized tests make the NeuroTrax system a valuable clinical tool in the identification of elderly at high risk for dementia.
Background

Mild cognitive impairment (MCI) is the term applied to a condition in which elderly individuals who have a subjective cognitive complaint have objective memory impairment in the absence of functional disability [1–3]. Its importance arises from the observation that it often constitutes the clinical state between normal cognition and dementia in the elderly [4]. Approximately 12–15% of MCI subjects per year convert to clinical dementia with functional disability [4,5]. For this reason, much interest has centered on the development of standardized techniques for quantification of cognitive deficits in MCI and potential therapeutic interventions for treatment of these high-risk individuals [6].

While memory impairment is the hallmark of MCI [3], multiple cognitive domains are compromised in the majority of MCI individuals [7–9], as in those with mild Alzheimer's disease (AD). Therefore, sufficiently broad and sensitive instruments are key in the effective diagnosis of these two groups. While traditional neuropsychological tests have been shown to discriminate among individuals with MCI and cognitively healthy elderly [10–12], there is no standard, comprehensive neuropsychological battery that is suitable for the screening and follow-up of mild impairments in routine clinical care. Moreover, while paper-based neuropsychological batteries have been applied in research settings, they are generally impractical for widespread clinical use due to high cost and extended administration time.

Computerized cognitive testing has the potential to effectively address the limitations posed by traditional paper-based measures. Technical innovations for accurate measurement of reaction time as well as frequency of errors enhance overall sensitivity, and on-line adjustment of level of difficulty may minimize ceiling or floor effects. Standardized batteries with alternate forms allow for accurate follow-up of patients over time. A computerized testing session can be of shorter duration and is less expensive than a paper-based session. Further, computerized testing can be made widely available via the Internet and easily administered so that high quality testing is on hand to supplement clinical evaluation in routine patient care settings. Indeed computerized tests have been developed to assist researchers and psychologists in screening for cognitive dysfunction [13,14]. We chose to evaluate Mindstreams™ (NeuroTrax Corp, NY), a new commercially available computerized testing system for comprehensive clinical assessment of cognitive impairment, designed primarily for use in the elderly. Specifically, we examine the ability of Mindstreams tests to discriminate individuals with MCI from among cognitively healthy elderly. The present study is the first to assess discriminant validity of the computerized tests as compared with that of traditional neuropsychological tests.

Methods

Participants

Participants were 98 elderly individuals assessed at two tertiary care memory clinics (Bloomfield Centre for Research in Aging, McGill-Jewish General Hospital, Montreal, Canada; Memory Disorders Clinic, Shaare Zedek Medical Center, Jerusalem, Israel). Participants were diagnosed by consensus of evaluation teams led by dementia experts at each of the sites and were diagnosed with mild cognitive impairment (MCI), mild Alzheimer's disease (AD), or as cognitively healthy. Diagnosis of MCI followed Petersen et al. [15] and included the following features: (1) a complaint of defective memory; (2) normal activities of daily living; (3) a memory deficit documented on mental status evaluation and supported by abnormalities on neuropsychological testing; and (4) absence of dementia. These criteria define the subtype of MCI known as 'MCI-amnestic' [3]. Diagnosis of mild AD was according to the Diagnostic and Statistical Manual, 4th ed. (DSM IV). Healthy elderly had no cognitive complaints and were volunteers for research testing. Each diagnostic group was taken to be representative of a distinct population defined by the criteria outlined above. Ethics Committee approval in compliance with the Declaration of Helsinki was obtained at both testing sites, and informed consent was obtained from all participants.

Table 1: Demographic and clinical characteristics for each of the three diagnostic groups.

| Characteristic                  | Healthy Elderly N = 39 | MCI N = 30 | Mild AD N = 29 |
|--------------------------------|------------------------|------------|---------------|
| Age, mean years (SD)           | 73.41 (8.00)           | 77.15 (6.43) | 80.55 (4.91)  |
| Education, mean years (SD)     | 14.95 (3.50)           | 13.07 (2.86) | 11.31 (2.85)  |
| Gender, N female               |                         |            |               |
| Female                         | 26                     | 13         | 16            |
| Computer Experience, N no      |                         |            |               |
| No                             | 10                     | 24         | 28            |
| Handedness, N left-handed      | 0                      | 1          | 2             |
| MMSE, mean score (SD)          | 29.03 (1.11)           | 27.63 (1.54) | 24.17 (3.25)  |
| ADAS-cog, mean score (SD)      | 8.01 (3.58)            | 14.06 (5.23) | 23.51 (8.33)  |

SD: standard deviation
Participants were excluded if there was a prior history of major psychiatric disorder, major depression, or any neurological disorder. All participants were native English speakers. Demographic and clinical characteristics for each diagnostic group are presented in Table 1. The MCI group was older ($t_{57} = 2.091, p = 0.040$) and had fewer years of education ($t_{67} = 2.391, p = 0.020$) relative to the healthy elderly group. Conversely, the MCI group was younger ($t_{57} = 2.276, p = 0.027$) and more highly educated ($t_{67} = 2.359, p = 0.022$) than the mild AD group. The MCI and healthy elderly groups were comparable in terms of gender and handedness (gender: $\chi^2[1, N = 69] = 3.757, p = 0.053$; handedness: $\chi^2[1, N = 69] = 1.319, p = 0.251$) as were the MCI and mild AD groups (gender: $\chi^2[1, N = 59] = 0.827, p = 0.363$; handedness: $\chi^2[1, N = 59] = 0.388, p = 0.533$). Fewer MCI participants had prior computer experience relative to the healthy elderly group ($\chi^2[1, N = 69] = 20.046, p = 0.000$), but more participants with MCI had computer experience relative to those with mild AD ($\chi^2[1, N = 59] = 3.863, p = 0.049$). Between-site differences across diagnostic groups were found for years of education ($t_{96} = 4.315, p = 0.000$), gender ($\chi^2[1, N = 98] = 7.936, p = 0.005$), and computer experience ($\chi^2[1, N = 98] = 12.001, p = 0.001$), but not for age ($t_{96} = 0.250, p = 0.802$) or handedness ($\chi^2[1, N = 98] = 0.140, p = 0.709$). Consistent with the expert consensus diagnoses, MCI participants performed more poorly than healthy elderly on the Mini-Mental State Examination (MMSE [16]; $U = 279, p = 0.000$) and on the Alzheimer's Disease Assessment Scale, cognitive subscale (ADAS-cog [17]; $U = 190.500, p = 0.000$), but the MCI group performed better than the mild AD group on these measures (MMSE: $U = 570.500, p = 0.000$; ADAS-cog: $U = 588.000, p = 0.000$).

### Procedure

All participants were administered the MMSE and the ADAS-cog and completed the Mindstreams™ (NeuroTrax Corp., NY) battery designed to detect mild cognitive impairment (the “Mild Impairment Battery”). Participants at the McGill-Jewish General Hospital site (20 with MCI, 15 healthy elderly, and 19 with mild AD) underwent a neuropsychological battery in addition, comprising standardized tests of memory, executive function, visual spatial skills, verbal fluency, and verbal fluency. Memory tests included the Logical Memory subtest of the Wechsler Memory Scale, 3rd Edition (WMS-III) and the Rey Auditory Verbal Learning Test (RAVLT), Version 1. Tests of executive function included the Clock Drawing Test, the Trail Making Test (Part A), the Digit Symbol subtest of the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III), and the Mental Control subtest of the WMS-III. Visual spatial skills were assessed with the Block Design subtest of the WAIS-III. Tests of verbal fluency included the Boston Naming test, the Controlled Oral Word Association (COWA) test, and the Similarities subscale of the WAIS-

III. Paper-based tests at both sites and Mindstreams tests at the Shaare Zedek site and were administered by research assistants blind to diagnosis. Mindstreams tests at the McGill-Jewish General Hospital site were administered by a research assistant aware of the diagnoses, but uninvolved in their determination. Notably, research assistants at both sites were trained not to assist participants during actual testing, but rather to ensure that they sufficiently understood the instructions prior to each Mindstreams test. Paper-based tests were administered first, followed by the Mindstreams Mild Impairment Battery. Both Mindstreams and paper-based tests were administered in English.

### Mindstreams Computerized Cognitive Testing

A detailed treatment of the NeuroTrax system, including the computerized tests, data processing, and usability considerations appears in a supplementary document (Additional File 1). In brief, Mindstreams consists of custom software that resides on the local testing computer and serves as a platform for interactive cognitive tests that produce precise accuracy and reaction time (milliseconds timescale) data. Tests are adaptive, in that the level of difficulty is adjusted accordingly depending upon performance. This feature increases sensitivity and minimizes the prevalence of ceiling effects. Feedback is provided in the practice sessions that precede each test, but not during the actual tests. Web-based administrative features allow for secure entry and storage of patient demographic data. Once tests are run on the local computer, data are automatically uploaded to a central sever, where calculation of outcome parameters from raw single-trial data and report generation occur.

The Mindstreams Mild Impairment Battery (administration time: 45 minutes) samples a wide range of cognitive domains, including memory (verbal and non-verbal), executive function, visual spatial skills, verbal fluency, attention, information processing, and motor skills (see Table 2). The tests that comprise this battery were designed for use with the elderly. All responses were made with the mouse or with the number pad on the keyboard (intuitively similar to the telephone keypad). Participants were familiarized with these input devices at the beginning of the battery, and practice sessions prior to the individual tests prepared them for the specific types of responses required for each test. Outcome parameters varied with each test, as in Table 2. Given the speed-accuracy tradeoff, (e.g., [18]) a performance index (computed as [accuracy/reaction time]*100) was computed for timed Mindstreams tests in an attempt to capture performance both in terms of accuracy and reaction time (RT). Tests were run in the same fixed order for all participants.
Following are brief descriptions of the *Mindstreams* tests included in the Mild Impairment Battery:

**Verbal Memory**
Ten pairs of words are presented, followed by a recognition test in which one member (the target) of a previously presented pair appears together with a list of four candidates for the other member of the pair. Participants must indicate which word of the four alternatives was paired with the target when presented previously. Four consecutive repetitions of the recognition test are administered during the ‘learning’ phase. An additional recognition test is administered following a delay of approximately 10 minutes.

**Non-Verbal Memory**
Eight pictures of simple geometric objects are presented, followed by a recognition test in which four versions of each object are presented, each oriented in a different direction. Participants are required to remember the orientations of the originally presented objects. Four consecutive repetitions of the recognition test are administered.

---

Table 2: Validity of *Mindstreams* tests in discriminating participants with mild cognitive impairment (MCI; \(N = 30\)) from cognitively healthy elderly (\(N = 39\)).

| Cognitive Domain | Mindstreams Test (outcome parameter) | AUC   | SE   | Asymptotic Significance | 95% CI       |
|------------------|--------------------------------------|-------|------|-------------------------|-------------|
| MEMORY           | Verbal Memory (accuracy, first repetition trial) | 0.779 | 0.057 | 0.000 | 0.668–0.890 |
|                  | Verbal Memory (accuracy, second repetition trial) | 0.838 | 0.05  | 0.000 | 0.741–0.936 |
|                  | Verbal Memory (accuracy, third repetition trial) | 0.752 | 0.063 | 0.000 | 0.629–0.876 |
|                  | Verbal Memory Accuracy (accuracy, final repetition trial) | 0.783 | 0.062 | 0.000 | 0.662–0.904 |
|                  | Verbal Memory (accuracy, all repetition trials) | 0.859 | 0.048 | 0.000 | 0.765–0.953 |
|                  | Verbal Memory (accuracy, delayed recognition) | 0.771 | 0.061 | 0.000 | 0.651–0.890 |
|                  | Non-Verbal Memory (accuracy, first repetition trial) | 0.654 | 0.067 | 0.033 | 0.522–0.786 |
|                  | Non-Verbal Memory (accuracy, second repetition trial) | 0.698 | 0.069 | 0.006 | 0.563–0.833 |
|                  | Non-Verbal Memory (accuracy, third repetition trial) | 0.712 | 0.065 | 0.003 | 0.585–0.839 |
|                  | Non-Verbal Memory Accuracy (accuracy, final repetition trial) | 0.773 | 0.061 | 0.000 | 0.654–0.892 |
|                  | Non-Verbal Memory (accuracy, all repetition trials) | 0.756 | 0.063 | 0.000 | 0.633–0.879 |
|                  | Non-Verbal Memory (accuracy, delayed recognition) | 0.728 | 0.065 | 0.002 | 0.601–0.854 |
| EXECUTIVE FUNCTION| Go-NoGo (accuracy) | 0.736 | 0.065 | 0.001 | 0.608–0.863 |
|                  | Go-NoGo (performance index) | 0.810 | 0.056 | 0.000 | 0.701–0.920 |
|                  | Problem Solving (accuracy) | 0.768 | 0.059 | 0.000 | 0.653–0.884 |
|                  | Stroop Interference (performance index) | 0.703 | 0.077 | 0.023 | 0.551–0.854 |
|                  | Catch Game (accuracy) | 0.696 | 0.075 | 0.016 | 0.548–0.843 |
| VISUAL SPATIAL   | Visual Spatial Imagery (accuracy) | 0.765 | 0.060 | 0.000 | 0.648–0.881 |
| VERBAL           | Verbal Function (accuracy, naming) | 0.716 | 0.067 | 0.005 | 0.584–0.848 |
|                  | Verbal Function (accuracy, rhyming) | 0.824 | 0.051 | 0.000 | 0.724–0.923 |
| ATTENTION        | Go-NoGo (RT) | 0.771 | 0.063 | 0.000 | 0.648–0.893 |
|                  | Go-NoGo (standard deviation of RT) | 0.706 | 0.066 | 0.006 | 0.576–0.835 |
|                  | Choice Reaction Time (performance index) | 0.490 | 0.084 | 0.894 | 0.326–0.653 |
| INFORMATION PROCESSING| Staged Information Processing, Low Load (performance index) | 0.579 | 0.083 | 0.345 | 0.416–0.743 |
|                  | Staged Information Processing, Medium Load (performance index) | 0.783 | 0.070 | 0.001 | 0.646–0.920 |
|                  | Staged Information Processing, High Load (performance index) | 0.688 | 0.088 | 0.037 | 0.515–0.860 |
| MOTOR SKILLS     | Finger Tapping (inter-tap interval) | 0.595 | 0.076 | 0.215 | 0.446–0.744 |
|                  | Finger Tapping (standard deviation of inter-tap interval) | 0.625 | 0.080 | 0.101 | 0.469–0.782 |
|                  | Catch Game (time to first move) | 0.521 | 0.082 | 0.797 | 0.360–0.682 |

*Mindstreams* outcome parameters with insufficient data (<13 data points per group) due to failed practice sessions (see Methods) were excluded. AUC = area under the curve SE = standard error CI = confidence interval RT = reaction time

Mindstreams tests with insufficient data (<13 data points per group) due to failed practice sessions (see Methods) were excluded.
during the ‘learning’ phase of the test. An additional recognition test is administered following a delay of approximately 10 minutes.

**Go-NoGo test**
A series of large colored stimuli are presented at pseudorandom intervals. Participants are instructed to respond as quickly as possible by pressing a mouse button if the color of the stimulus is any color except red, for which no response is to be made.

**Problem Solving**
Pictorial puzzles of gradually increasing difficulty are presented. Each puzzle consists of a $2 \times 2$ array containing three black-and-white geometric forms with a certain spatial relationship among them and a missing form. Participants must choose the best fit for the fourth (missing) form from among six possible alternatives.

**Mindstreams Stroop test**
The Stroop is a well-established test of response inhibition [19]. The *Mindstreams* Stroop test consists of three phases. Participants are presented with a pair of large colored squares, one on the left and the other on the right side of the screen. In each phase, participants are instructed to choose as quickly as possible which of the two squares is a particular color by pressing either the left or right mouse button, depending upon which of the two squares is the correct color. First, participants are presented with a general word in colored letters. In the next phase (termed the Choice Reaction Time test), participants are presented with a word that names a color in white letters. In the final phase (the Stroop phase), participants are presented with a word that names a color, but the letters of the word are in a color other than that named by the word. The instructions for the final phase are to choose the color of the letters, and not the color named by the word.

**Verbal Function**
Pictures of common objects of low and high familiarity are presented. Participants are instructed to select the name of the picture from four choices. In a related test, participants are instructed to select the word that best rhymes with the name of the picture.

**Visual Spatial Imagery**
Computer-generated scenes containing a red pillar are presented. Participants are instructed to imagine viewing the scene from the vantage point of the red pillar. Four alternative views of the scene are presented as choices.

**Staged Information Processing test**
This test comprises three levels of information processing load: single digits, two-digit arithmetic problems (e.g., 5-1), and three-digit arithmetic problems (e.g., 3+2-1). For each of the three levels, stimuli are presented at three different fixed rates, incrementally increasing as testing continues. Participants are instructed to respond as quickly as possible by pressing the left mouse button if the digit or result is less than or equal to 4 and the right mouse button if it is greater than 4.

**Finger Tapping**
Participants are instructed to tap on the mouse button for 12 seconds with their dominant hand. This task is repeated twice.

**Catch Game**
The Catch game is a novel motor screen that assesses cognitive domains distinct from those in other *Mindstreams* tests. Participants must “catch” a rectangular white object falling vertically from the top of the screen before it reaches the bottom of the screen. Mouse button presses move a rectangular green “paddle” horizontally so that it can be positioned directly in the path of the falling object. The test requires hand-eye coordination, scanning and rapid responses.

**Data Analysis**
*Mindstreams* data were uploaded to the NeuroTrax central server, where automatic data processing occurred, during which aggregate outcome parameters were computed from the raw single-trial data (Additional File 1). Outcome parameters were calculated using custom software that was blind to diagnosis or testing site, and results were relayed to each of the sites for review and analysis. Outcome parameters were computed for each test only when performance on the preceding practice session exceeded a predetermined minimum accuracy. The actual test was not given when practice session performance was below this cutpoint. Given this source of ‘missing’ data, a minimum of 13 data points was deemed acceptable for inclusion of a group in statistical analyses.

All statistics were computed with SPSS statistical software (SPSS, Chicago, IL). Two-tailed statistics were used throughout, and $p < 0.05$ was considered significant. Receiver operating characteristic (ROC) analysis was used to evaluate the ability of *Mindstreams* outcome parameters and traditional neuropsychological tests to discriminate participants with MCI from cognitively healthy elderly. Area under the curve (AUC), an index of effect size, was the primary result of the ROC analysis. For each measure, the AUC indicated the probability that a randomly selected individual with MCI would perform more poorly than a randomly selected cognitively healthy individual. An AUC of 0.50 indicated no better than chance discrimination, and an AUC of 1.00 indicated perfect discrimination. If the 95% confidence interval around an AUC included 0.50, the measure was unable to discriminate
among MCI and healthy elderly at a significance level of $p < 0.05$. Separate between-group comparisons were made on Mindstreams outcome parameters between MCI and cognitively healthy and between MCI and mild AD. Given heterogeneous variances across these pairs of groups for numerous outcome parameters (Brown-Forsythe test, $p > 0.05$), the non-parametric Mann-Whitney U was used to make the comparisons.

Results

**Discriminant Validity of Mindstreams Outcome Parameters: Effect Sizes**

Results of an ROC analysis measuring the ability of Mindstreams outcome parameters to discriminate MCI from cognitively healthy elderly are presented in Table 2, subdivided by cognitive domain. All Mindstreams memory outcome parameters discriminated significantly, with a maximum AUC of 0.859 for mean accuracy across all recognition test repetitions in the ‘learning’ phase of the Verbal Memory test. Similarly, all executive function outcome parameters discriminated significantly, with a maximum AUC of 0.810 for the Go-NoGo performance index. All outcome parameters measuring visual spatial skills and verbal fluency discriminated significantly, with a maximum AUC of 0.824 for accuracy on the rhyming portion of the Verbal Function test. Significant discriminability was found for attention outcome parameters from the Go-NoGo test, with an AUC of 0.771 for Go-NoGo RT and 0.706 for Go-NoGo standard deviation of RT. The Choice Reaction Time performance index did not discriminate significantly among MCI and cognitively healthy elderly. Medium- (AUC = 0.783) and high-load (AUC = 0.688) information processing outcome parameters discriminated significantly, but the low-load parameter did not. All motor skills outcome parameters did not discriminate significantly.

**Effect Sizes of Mindstreams Outcome Parameters Relative to Traditional Neuropsychological Tests**

Table 3 presents ROC analysis results for the subset of MCI and healthy elderly participants who received a battery of standardized neuropsychological tests in addition to Mindstreams testing. For each cognitive domain, the ability of Mindstreams outcome parameters to discriminate MCI from cognitively healthy elderly was compared with paper-based neuropsychological tests designed to tap the same domain. Outcome parameters measuring attention, information processing, and motor skills were excluded from this analysis for lack of corresponding traditional tests in these cognitive domains.

As above (Table 2), the Mindstreams memory outcome parameter that best discriminated MCI from cognitively healthy elderly was accuracy across the ‘learning’ phase of the Verbal Memory test (AUC = 0.894; Table 3); the best traditional memory test was the WMS-III Logical Memory II subtest (AUC = 0.885). Also as above, the Mindstreams executive function outcome parameter that discriminated best was the Go-NoGo performance index (AUC = 0.840); the best traditional executive function test was the WAIS-III Digit Symbol subtest (AUC = 0.729; Figure 1A). The Mindstreams visual spatial outcome parameter discriminated significantly (AUC = 0.778), but the corresponding WAIS-III Block Design subtest did not (Figure 1B). The Mindstreams verbal outcome parameter that discriminated best was accuracy on the naming portion of the Verbal Function test (AUC = 0.837); the best traditional verbal test was the COWA FS test (AUC = 0.768).

**Discriminant Validity of Mindstreams Outcome Parameters: Group Differences**

Mindstreams outcome parameters discriminated MCI from cognitively healthy and from mild AD (e.g., Figure 2). Descriptive statistics for outcome parameters in each cognitive domain are presented in Table 4, subdivided by diagnostic group. The results of two Mann-Whitney U tests are shown for each parameter, one comparing MCI and healthy elderly participants and the other comparing MCI and mild AD participants. Results for the MCI/healthy elderly comparison are similar to those in Table 2. For cognitive domains with sufficient data for conclusive results, significant differences between MCI and mild AD participants were found for memory, visual spatial, and verbal outcome parameters. Results were mixed for attention outcome parameters, such that timed Go-NoGo parameters did not significantly discriminate among MCI and mild AD, but the performance index from the Choice Reaction Time test did.

**Discussion**

Cognitive assessment is essential to the effective care and treatment of the elderly. Given that the number of elderly is predicted to increase steeply as the baby boomer generation ages [20], there is an urgent need for standardized cognitive assessment tools that deliver high quality information and are practical for routine clinical use. Traditional paper-based neuropsychological testing is seriously limited by the formidable cost in time and money. We therefore evaluated a novel computerized cognitive testing system, the Mindstreams system (NeuroTrax Corp., NY), which was designed for widespread clinical application in the detection of MCI and mild dementia.

The present study evaluated the discriminant validity of Mindstreams tests in distinguishing individuals with MCI from healthy elderly. Outcome parameters across multiple cognitive domains significantly discriminated among MCI and healthy elderly with considerable effect sizes (Table 2). Particularly strong results were obtained for outcome parameters assessing memory, executive
Discriminant Validity of Executive Function and Visual Spatial Tests. Receiver operating characteristic (ROC) curves demonstrating the ability of Mindstreams tests to discriminate individuals with mild cognitive impairment (MCI) from among cognitively healthy elderly as compared to traditional neuropsychological tests (MCI: N = 20; Healthy Elderly: N = 15). In (A), discriminability of the Mindstreams Go-NoGo performance index is compared with that of the Wechsler Adult Intelligence Scale, 3rd Edition (WAIS-III) Digit Symbol subtest. In (B), discriminability of accuracy on the Mindstreams Visual Spatial Imagery test is compared with that of the WAIS-III Block Design subtest. Discriminant validity of the Mindstreams tests was better than traditional paper-based tests designed to tap similar cognitive domains. AUC=area under the curve.
function, visual spatial skills, and verbal function. Further, effect sizes of the computerized tests in these domains were at least comparable to neuropsychological tests designed to assess the same domains (Table 3; Figure 1).

Results were mixed for *Mindstreams* attention and information processing outcome parameters, and those assessing motor skills did not discriminate among MCI and healthy elderly (Table 2).

The current findings are consistent with those of studies designed to identify traditional neuropsychological tests that predict conversion to dementia. Many such studies have found standard tests of verbal- and non-verbal
Computerized tests other than Mindstreams have been employed to discriminate MCI from cognitively healthy elderly. Indeed the paired associates learning (PAL) test of the Cambridge Neuropsychological Test Automated Battery (CANTAB) has been shown sensitive to cognitive decline [26]. While demonstrating the general utility of computerized cognitive testing, the CANTAB-PAL is limited in scope, difficult to use, and requires specialized equipment. A brief set of three tests developed by CogState Ltd. and administered serially four times in 3 hours has recently been shown to discriminate among MCI and cognitively healthy elderly on the basis of learning performance [7]. However, the CogState tests fail to provide a comprehensive cognitive profile, consisting exclusively of reaction time tests. Finally, MicroCog [27], a multi-domain computerized battery, showed good discriminability among participants with mild dementia and cognitively healthy elderly in an initial validity study [28]. However, MicroCog has not been widely used clinically, likely because it tests only selected cognitive domains and must be administered by a trained psychologist [29].

It is important to note that the results reported in the present study are preliminary. Population based studies with longitudinal follow-up, pathological confirmation of diagnosis, and comparison with a wider array of traditional tests are required to fully establish the validity of the Mindstreams tests in MCI detection. Further, given the between-group differences in age and years of education in the present study, future studies must collect normative data on Mindstreams tests so that performance can be standardized according to age and years of education. Given the between-group difference in computer experience in the present study, subsequent studies will collect more detailed information on participants' facility with the computer in general and with each of the Mindstreams tests in particular. However, the absence of between-group differences on Mindstreams motor skills tests in the current study, those most dependent upon facility with the computer, suggests that differential computer experience did not confound the results. Finally, future work might incorporate test data in the event of a failed practice session. As such data was labeled 'missing' in the present study, the reported results likely underestimate the true discriminant validity of the Mindstreams tests.

An important limitation imposed upon the present study and all studies of MCI arises from lack of consensus regarding the clinical definition of MCI [30,31]. Our MCI participants were selected according to the standard definition in the field [4], but these criteria for 'MCI-amnestic' [3] require only memory impairment. Consistent with the present results, individuals classified as 'MCI-amnestic' are often impaired in other cognitive domains [7–9]. A more clinically valid classification of this pre-dementia...
state may be Aging-Associated Cognitive Decline (AACD; [32]), which has clearly defined diagnostic criteria and requires impairment in multiple cognitive domains [30,31]. Indeed AACD has recently been validated as a predictor of conversion to dementia [33,34].

Computerized testing has been criticized relative to paper-based testing in terms of technical limitations and appropriateness for clinical use [35]. Perhaps the most pervasive technical limitation is measurement error that varies depending upon the refresh rate of the monitor, the sampling rate of the input device, operating system activities, and the data acquisition software. Mindstreams, which runs under Microsoft Windows, utilizes the DirectX library to minimize imprecision due to operating system activities and data acquisition software to sub-millisecond levels. The remaining sources of error are hardware-dependent and typically result in imprecision on the order of less than 20 milliseconds, still far better than human measurement error. Computerized assessment has also been criticized on the grounds that testing is not customizable for the individual participant. While paper-

Table 4: Discriminant validity of Mindstreams computerized tests: summary statistics and group differences (Mann-Whitney U).

| Cognitive Domain | Mindstreams Test (outcome parameter) | Healthy Elderly (N = 39) | MCI (N = 30) | Mild AD (N = 29) | MCI / Healthy Elderly p-value | MCI / Mild AD p-value |
|------------------|-------------------------------------|--------------------------|--------------|------------------|-------------------------------|------------------------|
| MEMORY           | Verbal Memory (accuracy, first repetition trial) | 71.35 (18.88) | 48.21 (22.94) | 24.00 (20.41) | 0.000                          | 0.000                  |
|                  | Verbal Memory (accuracy, second repetition trial) | 87.37 (13.09) | 60.71 (25.08) | 23.60 (23.96) | 0.000                          | 0.000                  |
|                  | Verbal Memory (accuracy, third repetition trial) | 94.47 (7.60) | 71.43 (29.9) | 20.80 (22.16) | 0.000                          | 0.000                  |
|                  | Verbal Memory Accuracy (accuracy, final repetition trial) | 97.37 (5.54) | 73.21 (29.07) | 22.00 (23.8) | 0.000                          | 0.000                  |
|                  | Verbal Memory (accuracy, all repetition trials) | 87.30 (12.17) | 62.50 (30.26) | 18.40 (19.30) | 0.000                          | 0.000                  |
|                  | Non-Verbal Memory (accuracy, first repetition trial) | 49.92 (21.25) | 38.71 (15.61) | 27.96 (16.35) | 0.027                          | 0.017                  |
|                  | Non-Verbal Memory (accuracy, second repetition trial) | 63.11 (20.63) | 46.68 (25.91) | 30.13 (13.95) | 0.006                          | 0.014                  |
|                  | Non-Verbal Memory (accuracy, third repetition trial) | 74.89 (23.65) | 55.18 (27.44) | 30.74 (17.98) | 0.003                          | 0.002                  |
|                  | Non-Verbal Memory Accuracy (accuracy, final repetition trial) | 82.84 (18.19) | 55.64 (28.36) | 28.92 (15.33) | 0.000                          | 0.000                  |
|                  | Non-Verbal Memory (accuracy, all repetition trials) | 67.79 (17.17) | 49.21 (20.36) | 28.96 (11.24) | 0.000                          | 0.000                  |
|                  | Non-Verbal Memory (accuracy, delayed recognition) | 74.92 (21.92) | 55.14 (25.00) | 37.68 (17.65) | 0.002                          | 0.012                  |
| EXECUTIVE        | Go-NoGo (accuracy) | 94.28 (6.59) | 84.89 (18.93) | 81.90 (14.19) | 0.001                          | 0.238                  |
| FUNCTION         | Go-NoGo (performance index) | 20.55 (2.63) | 16.47 (3.75) | 13.68 (4.77) | 0.000                          | 0.036                  |
|                  | Problem Solving (accuracy) | 73.45 (15.27) | 61.36 (13.09) | 55.00 (15.62) | 0.000                          | 0.161                  |
|                  | Stroop Interference (performance index) | 11.70 (6.40) | 7.16 (5.38) | ---           | 0.023                          | ---                    |
|                  | Catch Game (accuracy) | 59.67 (14.91) | 46.24 (19.80) | ---           | 0.016                          | ---                    |
| VISUAL SPATIAL   | Visual Spatial Imagery (accuracy) | 60.16 (15.61) | 44.37 (15.53) | 35.56 (15.43) | 0.000                          | 0.015                  |
| VERBAL           | Verbal Function (accuracy, naming) | 92.33 (8.79) | 83.22 (14.05) | 66.96 (18.37) | 0.004                          | 0.002                  |
|                  | Verbal Function (accuracy, rhyming) | 80.38 (16.97) | 52.21 (29.36) | 37.79 (22.79) | 0.000                          | 0.02                    |
| ATTENTION        | Go-NoGo (RT) | 464.39 (56.66) | 556.04 (102.82) | 666.38 (243.81) | 0.000                          | 0.244                  |
|                  | Go-NoGo (standard deviation of RT) | 119.17 (46.40) | 71.96 (89.02) | 242.81 (160.75) | 0.006                          | 0.149                  |
|                  | Choice Reaction Time (performance index) | 15.84 (3.84) | 15.52 (5.51) | 9.34 (4.41) | 0.894                          | 0.004                  |
| INFORMATION      | Staged Information Processing, Low Load (performance index) | 15.21 (1.91) | 14.52 (2.28) | ---           | 0.345                          | ---                    |
| PROCESSING       | Staged Information Processing, Medium Load (performance index) | 8.83 (1.55) | 7.04 (1.75) | ---           | 0.001                          | ---                    |
|                  | Staged Information Processing, High Load (performance index) | 5.55 (1.16) | 4.71 (1.51) | ---           | 0.037                          | ---                    |
| MOTOR SKILLS     | Finger Tapping (inter-tap interval) | 221.10 (42.71) | 242.83 (64.88) | 279.00 (66.09) | 0.215                          | 0.055                  |
|                  | Finger Tapping (standard deviation of inter-tap interval) | 38.38 (28.82) | 58.35 (49.46) | 66.62 (78.09) | 0.101                          | 0.888                  |
|                  | Catch Game (time to first move) | 774.91 (158.10) | 783.14 (169.16) | --- | 0.797                          | ---                    |

Summary statistics presented as mean (standard deviation). Accuracies are given as percent correct. Timed outcome parameters are given in milliseconds. ‘---’ indicates insufficient data (<13 data points per group) due to failed practice sessions (see Methods). MCI = mild cognitive impairment AD = Alzheimer’s disease RT = reaction time
based tests are indeed more flexible, the inherent lack of uniformity confounds the valid comparison of test results across participants. Further, Mindstreams testing batteries can be customized to suit specific clinical needs. Batteries can be constructed to include only relevant tests, and stimulus presentation parameters can be altered as appropriate for a particular clinical population.

We found Mindstreams tests straightforward to administer and easy for even the mild AD participants to learn. Administration time for the comprehensive testing battery used in this study (45 minutes) was appropriate, and participants were pleased with the positive feedback that the system provided throughout the session. The automatic uploading and scoring of the data streamlined the entire data collection process, and, in our view, these features may lead to widespread adoption of computerized cognitive testing.

The present study is evaluative in that it serves to guide future studies in determining the optimal set of Mindstreams tests and outcome parameters for differentiating among various patient groups. For example, not all information processing outcome parameters discriminated equally among MCI and cognitively healthy elderly (Table 2). It appears that the level of difficulty associated with the 2-digit arithmetic (i.e., medium load) portion of the Information Processing test discriminated best, while that associated with the single digit (i.e., low load) portion of the test was ineffectual in discriminating. This suggests that level of difficulty is an important consideration in selecting the Mindstreams parameters that best discriminate among groups. Similarly, the mixed pattern of results for attention outcome parameters (i.e., Choice Reaction Time did not discriminate, but Go-NoGo timed outcome parameters did discriminate; Table 2) can be accounted for by inter-task differences in level of difficulty. These observations may guide both clinical research on existing Mindstreams tests and future test development.

Conclusions

The present preliminary study demonstrates the ability of Mindstreams computerized cognitive tests to discriminate individuals with MCI from cognitively healthy elderly. Mindstreams measures of memory, executive function, visual spatial skills, and verbal fluency discriminated best, and discriminability was at least comparable to that of traditional neuropsychological tests in these domains. Our findings and experience with the NeuroTrax system underscore the utility of this novel clinical tool in the diagnosis of MCI and mild dementia in circumstances where full neuropsychological evaluation is unavailable or impractical. Guided by the present results, further work is necessary to examine the suitability of Mindstreams tests for additional clinical and non-clinical validation cohorts and for longitudinal use.

Competing Interests

T Dwolatzky: None Declared

V Whitehead: None Declared

GM Doniger: Employee of NeuroTrax Corp.

ES Simon: Employee of NeuroTrax Corp.

A Schweiger: Consultant for NeuroTrax Corp.

D Jaffe: Consultant for NeuroTrax Corp.

H Chertkow: None Declared

This work was partially supported by a grant from NeuroTrax Corporation to the participating clinical institutions. There were no restrictions placed upon the authors regarding the use or publication of the results reported herein. NeuroTrax served to manage collaboration among the institutions where the work was carried out.

Authors’ Contributions

TD and HC contributed to study design and oversaw data collection at each of the two sites. Data analysis was performed by GMD and VW, guided by contributions from DJ, ES, AS and HC. Manuscript preparation was handled primarily by GMD and ES, with contributions from HC, AS, and TD. All authors read and approved the final manuscript.

Additional material

Additional File 1
Mindstreams Cognitive Health Assessment, Version 2.1 (May 2003). A detailed treatment of the NeuroTrax system, including the computerized tests, data processing, and usability considerations. Click here for file [http://www.biomedcentral.com/content-supplementary/1471-2318-3-4-S1.pdf]

Acknowledgements

The authors wish to thank Dr. Ted Miller for contributing to a pilot study leading to the present work. We owe gratitude to Dr. Natalie Phillips for input relating to data analysis and for manuscript review. Shimon Amit and Judy Simon provided excellent technical support throughout the study.

References

1. Kabani NJ, Sled JG, Shuper A and Chertkow H: Regional magnetization transfer ratio changes in mild cognitive impairment. Magn Reson Med 2002, 47:143-148.
2. Massoud F, Chertkow H, Whitehead V, Overbury O and Bergman H: Word-reading thresholds in Alzheimer disease and mild memory loss: a pilot study. Alzheimer Dis Assoc Disord 2002, 16:31-39.
3. Petersen RC, Doyley R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L and Winblad B: Current concepts in mild cognitive impairment. Arch Neurol 2001, 58:1985-1992.
4. Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL and DeKosky ST: Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001, 56:133-142.
5. Flicker C, Ferris SH and Reisberg B: Mild cognitive impairment in the elderly: predictors of dementia. Neurology 1991, 41:1006-1009.
6. Burns A and Zaudig M: Mild cognitive impairment in older people. Lancet 2000, 356:1963-1965.
7. Darby D, Maruff P, Collie A and Mcstephen M: Mild cognitive impairment can be detected by multiple assessments in a single day. Neurology 2002, 59:1042-1046.
8. Olichney JM, Morris SK, Ochoa C, Salmon DP, Thal LJ, Kurz A and Iragui VJ: Abnormal event related potentials in mild cognitive impairment and incipient Alzheimer’s disease. J Neurol Neurosurg Psychiatry 2002, 73:377-384.
9. Xu G, Meyer JS, Thornby J, Chowdhury M and Quach HH: MicroCog: Assessment of Cognitive Functioning. [2.1] San Antonio, TX: The Psychological Corporation 1999.
10. Albert MS, Moss MB, Tanzi R and Jones K: Preclinical prediction of AD using neuropsychological tests. J Int Neuropsychol Soc 2001, 7:631-639.
11. Kawas CH, Corrada MM, Brookmeyer R, Morrison A, Resnick SM, Zonderman AB and Arendt D: Visual memory predicts Alzheimer’s disease more than a decade before diagnosis. Neurology 2003, 60:1089-1093.
12. Mapstone M, Steffenella TM and Duffy CJ: A visuospatial variant of mild cognitive impairment: getting lost between aging and AD. Neurology 2003, 60:802-808.
13. White RF, James KE, Vasterling JJ, Letz R, Marans M, Delaney RR, Krenkel P, Rose F and Kraemer HC: Neuropsychological screening for cognitive impairment using computer-assisted tasks. Assessment 2003, 10:86-101.
14. Swanson R, Hodges JR, Galton CJ, Semple J, Michael A, Dunn BD, Iddon JL, Robbins TW and Sahakian BJ: Early detection and differential diagnosis of Alzheimer’s disease and depression with neuropsychological tasks. Dement Geriatr Cogn Disord 2001, 12:265-280.
15. Petersen RC, Smith GE, Waring SC, Innik RJ, Tangalos EG and Korkman M: Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999, 56:303-308.
16. Folstein MF, Folstein SC and McHugh PR: "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975, 12:189-198.
17. Mohs RC, Knapman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bielaskas L, Geldmacher D and Clark C et al.: Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer’s Disease Assessment Scale that broaden its scope. The Alzheimer’s Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997, 11(Suppl 2):S53-S21.
18. Cavanaugh JH: Speed-accuracy tradeoff during response preparation. Res Q Exerc Sport 1990, 61:331-337.
19. MacLeod CM: Half a century of research on the Stroop effect: an integrative review. Psychol Bull 1991, 109:163-203.
20. Population and Life Tables, Population Projections of the United States by Age, Race, Sex, Hispanic Origin and Nativity: 1999 to 2100 [http://www.census.gov/population/projections/nation/detail/]
21. Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST and Ganguli M: Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. Neurology 2000, 55:1847-1853.
22. Reid W, Broe G, Creasey H, Grayson D, McCusker E, Bennett H, Longley W and Sulway MR: Age at onset and pattern of neuropsychological impairment in mild early-stage Alzheimer disease. A study of a community-based population. Arch Neurol 1996, 53:1056-1061.
23. Tierney MC, Salz PJ, Snow WG, Fisher RH, Nores A, Nagon D, Dunn E and George-Hyslop PH: Prediction of probable Alzheimer’s disease in memory-impaired patients: a prospective longitudinal study. Neurology 1996, 46:661-665.
24. Jacobs DM, Sano M, Donoifee G, Marder K, Bell KL and Stern Y: Neuropsychological detection and characterization of preclinical Alzheimer’s disease. Neurology 1995, 45:957-962.
25. Hannonen T, Hallikainen M, Koivisto K, Helkala EL, Reinkainen KJ, Soininen H, Mykkanen L, Laasko M, Pyorala K and Riekkinen P: Sr A follow-up study of age-associated memory impairment: neuropsychological predictors of dementia. J Am Geriatr Soc 1995, 43:1007-1015.
26. Fowler KS, Saling MM, Conway EL, Semple JM and Louis WJ: Paired associate performance in the early detection of DAT. J Int Neuropsychol Soc 2002, 8:S8-71.
27. Powell DH, Kaplan EF, Whitha D, Weintraub S, Catlin R and Funkenstein HH: MicroCog: Assessment of Cognitive Functioning. [2.1] San Antonio, TX: The Psychological Corporation 1993.
28. Green RC, Green J, Harrison JM and Kutner MH: Screening for cognitive impairment in older individuals. Validation study of a computer-based test. Arch Neurol 1994, 51:779-786.
29. Elwood RM: MicroCog: assessment of cognitive functioning. Neuropsychol Rev 2001, 11:89-100.
30. Ritchie K, Artero S and Touchon J: Classification criteria for mild cognitive impairment: a population-based validation study. Neurology 2001, 56:437-442.
31. Bischkopf J, Busse A and Angermeyer MC: Mild cognitive impairment – a review of prevalence, incidence and outcome according to current approaches. Acta Psychiatr Scand 2002, 106:403-414.
32. Levy R: Aging-associated cognitive decline. Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization. Int Psychogeriatr 1994, 6:63-68.
33. Daly E, Zaitchik D, Copeland M, Schmahman J, Gunther J and Albert M: Predicting conversion to Alzheimer disease using standardized clinical information. Arch Neurol 2000, 57:675-680.
34. Bozoki A, Giordani B, Heidebrink JL, Berent S and Foster NL: Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss. Arch Neurol 2001, 58:411-416.
35. Schatz P and Browndyke J: Applications of computer-based neuropsychological assessment. J Head Trauma Rehabil 2002, 17:395-410.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2318/3/4/prepub

Publish with BioMed Central and every scientist can read your work free of charge
"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."
Sir Paul Nurse, Cancer Research UK
Your research papers will be:
· available free of charge to the entire biomedical community
· peer reviewed and published immediately upon acceptance
· cited in PubMed and archived on PubMed Central
· yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp