

Two Novel Psychomotor Tasks in Idiopathic Normal Pressure Hydrocephalus

Maria A. Rossetti,^{1,2} Irene Piryatinsky,³ Fayeza S. Ahmed,^{4,5} Petra M. Klinge,³ Norman R. Relkin,⁵ Stephen Salloway,³ Lisa D. Ravdin,⁵ Einat Brenner,³ Paul F. Malloy,³ Bonnie E. Levin,^{1,6} Michael Broggi,³ Rebecca Gavett,³ James S. Maniscalco,^{5,7,8} AND Heather Katzen⁶

¹Department of Psychology, University of Miami, Coral Gables, Florida

²Michael E. DeBakey VA Medical Center, Houston, Texas

³Memory and Aging Program, Brown University, Providence, Rhode Island

⁴Department of Psychology, Eastern Illinois University, Charleston, Illinois

⁵Department of Neurology, Weill Cornell Medical College, New York, New York

⁶Department of Neurology, University of Miami Miller School of Medicine, Miami, Florida

⁷Department of Psychology, CUNY Queens College, Flushing, New York

⁸Department of Psychology, CUNY Graduate Center, New York, New York

(RECEIVED October 15, 2014; FINAL REVISION September 20, 2015; ACCEPTED October 14, 2015; FIRST PUBLISHED ONLINE January 28, 2016)

Abstract

Objective: Idiopathic normal pressure hydrocephalus (INPH) is a neurological disorder presenting with gait, cognitive, and bladder symptoms in the context of ventricular enlargement. Although gait is the primary indicator for treatment candidacy and outcome, additional monitoring tools are needed. Line Tracing Test (LTT) and Serial Dotting Test (SDT), two psychomotor tasks, have been introduced as potential outcome measures but have not been widely studied. This preliminary study examined whether LTT and SDT are sensitive to motor dysfunction in INPH and determined if accuracy and time are important aspects of performance. **Methods:** Eighty-four INPH subjects and 36 healthy older adults were administered LTT and SDT. Novel error scoring procedures were developed to make scoring practical and efficient; interclass correlation showed good reliability of scoring procedures for both tasks (0.997; $p < .001$). **Results:** The INPH group demonstrated slower performance on SDT ($p < .001$) and made a greater number of errors on both tasks ($p < .001$). Combined Time/Error scores revealed poorer performance in the INPH group for original-LTT ($p < .001$), modified-LTT ($p \leq .001$) and SDT ($p < .001$).

Conclusions: These findings indicate LTT and SDT may prove useful for monitoring psychomotor skills in INPH. While completion time reflects impaired processing speed, reduced accuracy may suggest planning and self-monitoring difficulties, aspects of executive functioning known to be compromised in INPH. This is the first study to underscore the importance of performance accuracy in INPH and introduce practical/reliable error scoring for these tasks. Future work will establish reliability and validity of these measures and determine their utility as outcome tools. (*JINS*, 2016, 22, 341–349)

Keywords: Normal pressure hydrocephalus, Cognition, Dementia, Neuropsychology, Speed of processing, Psychomotor speed

INTRODUCTION

Idiopathic normal pressure hydrocephalus (INPH) is a progressive neurological disorder in which ventricular enlargement occurs in the context of minimal cerebral atrophy and no macroscopic evidence of a cerebral spinal fluid (CSF) obstruction (Adams, Fisher, Hakim, Ojemann, & Sweet, 1965; Kitagaki et al., 1998; Shprecher, Schwalb, &

Kurlan, 2008). Clinical presentation of INPH involves the symptom triad of gait disturbance, cognitive impairment, and urinary incontinence (Marmarou, Bergsneider, Klinge, Relkin, & Black, 2005; Savolainen, Hurskainen, Paljärvi, Alafuzoff, & Vapalahti, 2002). Cognitive dysfunction in INPH can range from subtle mental status changes to dementia (Iddon et al., 1999). The pattern of cognitive decline is characterized as frontal network dysfunction, evidenced by declines in attention, executive functioning, and psychomotor speed (Klinge et al., 2001). Impaired executive functioning is observed on neuropsychological exam, with poor performance on tasks sensitive to frontal lobe or frontostriatal

Correspondence and reprint requests to: Heather Katzen, Department of Neurology, University of Miami Leonard M. Miller School of Medicine, 1120 NW 14th Street, 13th Floor, Miami, FL 33136. E-mail: hkatzen@med.miami.edu

dysfunction (Devito et al., 2005; Iddon et al., 1999; Klinge et al., 2001). Screening tests such as the Mini Mental Status Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) are not typically sensitive to the types of deficits observed in INPH, especially early in the disease course, and are fairly insensitive to changes over time. Thus, detailed neuropsychological testing is needed for diagnostic purposes, to help determine whether intervention is warranted and to monitor response to treatment.

INPH is most commonly treated via shunt placement (Kahlon, Sjunnesson, & Rehncrona, 2007; Pujari et al., 2008; Thomas et al., 2005). Candidacy for shunt is determined by careful diagnostic work-up, which often includes a lumbar puncture to establish whether symptoms improve after CSF removal (Bergsneider, Black, Klinge, Marmarou, & Relkin, 2005; Foss, Eide, & Finset, 2006; McGirt et al., 2005). Gait has been shown to improve most frequently following surgery (Hellström et al., 2008). While improvements in higher order cognitive functions have been observed, there is a great deal of variability in the pattern and course of cognitive recovery (Katzen et al., 2011).

To our knowledge, no systematic neuropsychological tests have been shown to reliably identify candidates for shunt or monitor response to treatment. There is a clear need for additional measures that are both brief and practical to efficiently manage INPH patients. Assessment of upper extremity motor (UEM) skills in INPH is particularly important given that some patients are wheelchair dependent, prohibiting gait assessment, and others have orthopedic issues that interfere with gait evaluation.

Two experimental measures of psychomotor function, the Line Tracing Test (LTT) (Schomerus, Weissenborn, Hamster, Rückert, & Hecker, 1999; Wechsler, 1981; Weissenborn, Ennen, Schomerus, Rückert, & Hecker, 2001) and the Serial Dotting Test (SDT) (Schomerus et al., 1999; Wechsler, 1981; Weissenborn et al., 2001) have potential as reliable and practical outcome assessments in INPH (Klinge et al., 2001). These measures are different from traditional psychomotor tasks typically used in a neuropsychological examination since they allow for assessment at all skill levels. Many of the more traditional motor tasks at our disposal are too difficult for INPH patients, and in many instances accurate assessment is not possible due to a floor effect. While LTT and SDT have not been widely studied in INPH, psychometric studies have shown that these tasks are suitable to assess psychomotor skills in other populations with significant motor impairment, such as hepatic encephalopathy (Weissenborn, 2013; Biller & Ferro, 2014).

LTT and SDT were first introduced for use in INPH by Klinge and colleagues, who reported improvement in performance on these measures 1 week after shunt placement; these changes were also sustained seven months post-shunt (Klinge et al., 2002). For those readers who are unfamiliar with the test, as a way of brief introduction, LTT and SDT, are measures of upper extremity dexterity, that have been reported by Klinge et al. (2002) as sensitive measures to shunt response in INPH patients. Tsakanikas, Katzen, Ravdin, and Relkin (2009), further demonstrated that LTT and SDT identified INPH responders to CSF drainage (tap-test). Post-tap changes on

these tasks correlated with shunt outcome. Collectively, these findings suggest that LTT and SDT appear to be sensitive measures for evaluating early shunt response in INPH.

Despite the promising preliminary data, LTT and SDT have not yet been fully adopted into standard clinical or research INPH protocols. One of the primary obstacles has been the cumbersome scoring methodology for LTT, which is time intensive and introduces the potential for a great deal of error due to the subjective nature of the scoring decisions required by raters. For this reason, many investigators have used only completion time as a potential measure to determine change and have not examined performance accuracy, which may be a critical component of performance (Tsakanikas et al., 2009). Accuracy is important because an individual who completes the task quickly may compromise accuracy for speed, whereas another individual might perform the task slowly to achieve greater precision. To further explore the importance of accuracy, in addition to speed, we developed an alternative LTT error scoring methodology. We also introduced an error scoring system for SDT, which previously was scored only for completion time. Both new scoring methods have been introduced in our hydrocephalus clinics and have been valuable for measuring both response to drainage and shunt outcome in individual patients. The overall goal of the present study was to determine the utility and reliability of the original and revised error scoring procedures and examine whether these newly devised LTT and SDT error scoring variables may be sensitive to UEM dysfunction in INPH. This preliminary study will help determine whether LTT and SDT should be further investigated and developed as outcome assessment tools in INPH as well as whether the adapted scoring methodology adds useful information in INPH.

METHODS

Participants: Eighty-four INPH subjects were recruited from hydrocephalus programs at three neurological centers [Weill Cornell Memory Disorders Program (WCMC), $n = 25$; University of Miami Department of Neurology (UM), $n = 24$; and Butler Hospital Memory and Aging Program (MAP), $n = 35$]. A diagnosis of INPH was made by the treating neurologist, based upon neurological and neuropsychological evaluation. Inclusion/exclusion criteria were: (1) Enlarged ventricles out of proportion to sulcal atrophy on computed tomography or magnetic resonance imaging (Evan's index of at least 0.3; Shprecher et al., 2008), (2) Gait disturbance with either urinary incontinence and/or cognitive impairment, (3) No evidence of a known cause for hydrocephalus, (4) No history of alcohol abuse, significant psychiatric diagnosis or clinically significant hearing/visual loss, (5) No history of large and medium cortical strokes or neurologic disease other than INPH, and (6) Impaired vision interfering with completion of cognitive assessment. All INPH patients were pre-shunt placement at the time of testing. Of the INPH participants, 2 were wheelchair bound and 15 required the use of an assistive device for ambulation (cane or walker).

The remaining 67 participants were able to ambulate independently for the purposes of the gait evaluation.

Thirty-six healthy older adults were recruited from two of the sites as a comparison group and included INPH caregivers and community-dwelling healthy older adults [Weill Cornell Memory Disorders Program (WCMC), $n = 25$; Butler Hospital Memory and Aging Program (MAP), $n = 11$]. Exclusion criteria for the comparison group included current or past history of alcoholism, drug use, mental illness, neurologic diagnosis, brain injury, or cognitive impairment (MMSE < 25). Informed consent was obtained from all participants. Proxy consent was obtained in any INPH patient who was deemed unable to provide informed consent due to the severity of their cognitive difficulties. This study was approved by the Institutional Review Board at each center.

Procedures

All participants underwent comprehensive neuropsychological assessment, including measures of attention, executive functioning, construction, visuospatial skills, learning and memory, and motor/psychomotor skills, as well as brief mood and behavior screening as part of a clinical work-up at each center. Neuropsychological measures were administered and completed using standard instructions outlined in their respective test manuals; the full battery varied across centers.

In addition to the traditional neuropsychological tasks, two experimental measures of psychomotor functions, SDT and LTT, were administered to each participant (Schomerus et al., 1999; Weissenborn et al., 2001). Modifications were made to the established administration and scoring procedures to adapt these measures for use with individuals with INPH.

First, we created a modified scoring methodology for LTT and introduced an error scoring procedure for SDT, which was previously only scored for completion time. Second, a felt-tip red pen was used rather than a soft-pencil to increase response precision and improve scoring through better visualization of the participants' data. Lastly, during both tasks, the stimuli were taped to the table (top and bottom) to prevent movement and to ensure that the same orientation was used throughout. All administration procedures and error scoring methods are described in detail below.

Line Tracing Administration and Scoring

LTT (Figure 1a) requires the participant to draw a line inside two given lines as quickly as possible without touching/crossing the given boundary lines. There are four alternate forms. All forms are mirror images, and the same scoring template and total number of possible errors applies. WCMC and UM used LTT Form 1 for all participants and MAP counterbalanced the forms. Participants are given a red felt-tip marker and asked to complete the task without lifting the pen. A short sample trial is demonstrated by the examiner and completed by the participant before proceeding to the test stimuli. Completion time is recorded (seconds) as well as the number of errors made using two different error scoring methods (original and modified), described below.

The original LTT (LTT-O) error scoring method uses a scoring template (Figure 1b) that is placed over the stimulus and divides the page into small individual sections of equal size. Error points are assigned for each segment depending on whether the drawn line is within the boundary line (0 points), touching the boundary line (1 point), outside the borders of

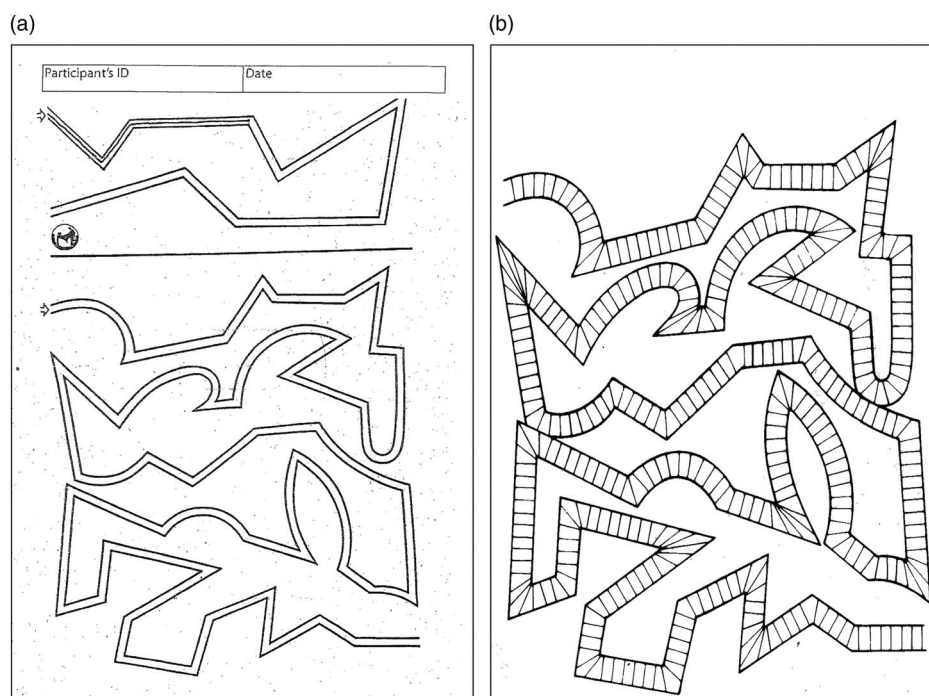


Fig. 1. a: The Line Tracing Test (LTT). **b:** Original LTT scoring template.

the boundary line (2 points) or outside the template border (3 points). A segment is marked as outside the template border if the drawn line falls outside the boundary line. There are 365 segments and error scores can range from 0 to 1095. This scoring procedure takes approximately 10 min.

The modified LTT scoring method (LTT-M) was developed in an attempt to simplify the procedure and create a scoring method that is practical, demonstrates strong inter-rater reliability and can be easily implemented in the clinic setting. The LTT-M scoring method does not use a template or individual segments. Instead, error points are assigned each time the drawn line touches (1 point) or crosses the boundary line (2 points). There is no maximum score, since segments are not used. Total scores of INPH participants using this scoring method typically range from 40 to 70. This scoring procedure takes approximately 2 to 4 min.

In addition to the individual time (seconds) and error scores, combined scores (time+errors) were calculated for LTT-O and LTT-M scoring methods to evaluate overall performance. Raw scores were converted to Z-scores by standardizing all data points to the mean and standard deviation of the control group. These scores reflect the sum of the Z-score for time and Z-score for error. Higher scores represent poorer performance.

Serial Dotting Administration and Scoring

SDT (Figure 2a) requires the participant to place a dot in the center of 100 circles (1 cm diameter) arranged in a 10×10 array. Participants are instructed to work as quickly and accurately as possible. A short sample trial is demonstrated by the examiner and completed by the participant before proceeding to the test stimuli. Original scoring of SDT included only completion time. We developed error scoring

for this task that involves placing a scoring template over the stimulus (Figure 2b). Error points (1, 2, or 3) are assigned for each circle based on how far the marked dot diverges from the center. Error scores range from 0 to 300.

In addition to the individual time (seconds) and error scores, a combined score (time+errors) was calculated for SDT to evaluate overall performance. Raw scores were converted to Z-scores by standardizing all data points to the mean and standard deviation of the control group. The combined scores reflect the sum of the Z-score for time and Z-score for error. Higher scores represent poorer performance.

Double Scoring Procedure for LTT and SDT

To ensure accuracy, each LTT and SDT protocol was scored by two raters trained in scoring methods by the lead neuropsychologist at each center (H.K. and I.P.). In cases where the second rater's score was within 15% of the initial score, the initial score was used for data analysis. If the two scores were more than 15% discrepant, the protocol was scored by another independent rater. The score that fell closest (and within 15%) of the independent rater was used for analysis. For all scoring procedures, 13% (across each test and scoring method) required a third rater. If a greater than 15% discrepancy remained after the task was scored by three raters, all raters and the site neuropsychologist reviewed the data and consensus was reached.

Statistical Analyses

All variables were analyzed for normality using the Shapiro-Wilk test. Mann-Whitney *U* test was used to examine group differences in demographics and gait outcomes for non-normally distributed variables. Data transformations for

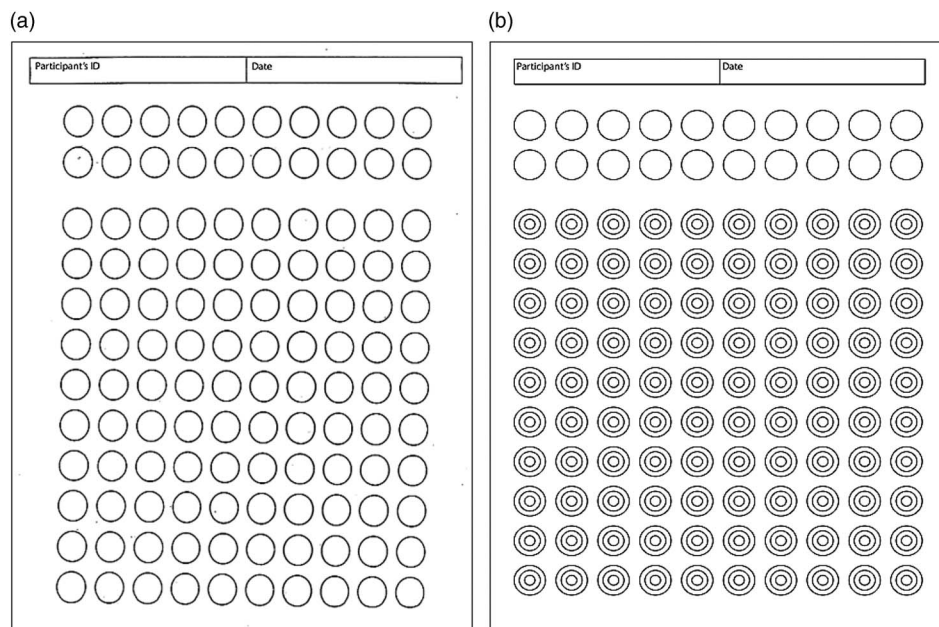


Fig. 2. a: The Serial Dotting Test (SDT). **b:** SDT scoring template.

positively skewed data (square root, log, or inversion) as well as removal of extreme outliers were used when necessary in order for all STD and LTT variables to reach normality. Independent sample *t* tests were used to examine group differences and a Bonferroni correction was used to control for multiple comparisons and reduce the chances of a Type I error. The critical *p* value for the main analyses was set to alpha of .05 and with eight comparisons; the adjusted alpha value was $p < .006$. Cohen's *d* was used to calculate effect sizes for the independent sample *t* tests of the outcome measures. Intraclass correlation coefficients were calculated to assess interrater reliability for both measures. Analyses of covariance (ANCOVA) were used to assess group differences in test performance (time, errors, and time + errors) while controlling for age, gender, and education. Partial eta squared (η^2) was used to calculate effect sizes for the ANCOVAs. Partial correlations were used to assess the relationship between LTT and SDT with both speed of walking and number of steps during 10-meter walk for both groups, gait time (10 meter walk), and mean number of steps for both groups.

RESULTS

Demographic Information

Demographic variables were not normally distributed. Mann-Whitney *U* tests revealed that the INPH group was older ($U = 1,832$; $Z = 2.11$; $p = .035$) and had fewer years of education ($U = 965.5$; $Z = -2.91$; $p = .004$) than the

healthy elderly adult comparison group. The comparison group was comprised of a greater number of females. ANCOVA analyses were performed controlling for these demographic differences between groups. No significant group differences in estimated verbal IQ were observed. The majority of participants for which racial information was available ($n = 80$) identified themselves as white, non-Hispanic across both groups ($n = 68$ or 85%). The comparison group performed significantly better on a cognitive screen ($U = 392.5$; $Z = -6.26$; $p < .001$). Means and standard deviations of all demographic variables are shown in Table 1.

Interrater Reliability

Intraclass correlation coefficients (ICC) were calculated to determine interrater reliability of error scores derived by two independent raters for both LTT scoring methods and SDT. Estimated reliability for LTT-O was 0.994, $p < .001$, 95% confidence interval (CI), [0.991, 0.996]. For the LTT-M scoring procedure, the estimated reliability was 0.997, $p < .001$, 95% CI, [0.996, 0.998]. The ICC for SDT error scores was 0.997, $p < .001$, 95% CI, [0.996, 0.998].

Testing Results

Gait assessment

Number of seconds and number of steps required to walk 10 meters and the number of steps required to turn 180 degrees were recorded for each participant (average of two trials).

Table 1. Demographic information and gait assessment

	INPH ^a					Controls ^b				
	N (Percent)	Min	Max	Mean	SD	N (Percent)	Min	Max	Mean	SD
Age*		61	89	75.29	6.78		62	87	72.37	7.44
Gender										
Male	51 (61)					9(26)				
Female	33 (39)					26(74)				
Education**		8	20	14.01	3.06		9	20	15.79	2.70
Handedness										
Right	74 (88)					31(89)				
Left	10 (12)					4 (11)				
MMSE ^c ***		10	30	24.44	4.76		26	30	29.09	1.17
NAART ^d		89.54	128	113.73	9.03		97.00	128.70	115.39	6.99
Gait										
Steps ^e ***		6	52	19.24	11.22		5	18.50	9.10	3.76
Time ^f ***		5	53	15.29	8.91		4	8.50	6.11	1.14
Turn ^g ***		2	16	5.68	2.32		1	5	2.71	0.76

^a*N* = 84.

^b*N* = 36.

^cMini Mental Status Exam Total Score, *N* = 82 (INPH), *N* = 35 (Controls).

^dNorth American Adult Reading Test Total Verbal IQ Estimate, *N* = 54 (INPH), *N* = 32 (Controls).

^eMean steps to walk 10 meters, *N* = 77 (INPH), *N* = 35 (Controls).

^fMean time (seconds) to walk 10 meters, *N* = 77 (INPH), *N* = 35 (Controls).

^gMean number of steps to turn 180 degrees, *N* = 70 (INPH), *N* = 35 (Controls).

*Statistically Significant group difference, $p < .05$.

**Statistically Significant group difference, $p < .01$.

***Statistically Significant group difference, $p < .001$.

Table 2. Line Tracing and Serial Dotting Tests

	INPH			Controls			<i>p</i> Value	Cohen's <i>d</i>
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>		
Line Tracing Completion Time	75	94.59	36.99	36	89.17	29.18	.553	—
Line Tracing-O Error Scorescore	74	160.57	72.33	27	66.89	23.62	<.001	-1.7
Line Tracing-M Error Scorescore	76	67.05	21.66	34	43.00	16.79	<.001	-1.2
Serial Dotting Completion Time	77	83.88	30.55	36	58.34	14.81	<.001	-1.1
Serial Dotting Error Scorescore	81	51.74	27.33	36	26.86	22.13	<.001	-1.0

Mann-Whitney *U* test revealed better performance on all three gait measures in healthy compared to INPH participants: number of seconds to walk 10 meters ($U = 2,599$; $Z = 7.86$; $p < .001$); number of steps to walk 10 meters ($U = 2,241$; $z = 5.61$; $p < .001$); number of steps to turn 180 degrees ($U = 2,333$; $z = 7.56$; $p < .001$).

Line Tracing Test

Independent samples *t* test revealed no group differences in completion time on the LTT. The INPH group made more errors than healthy adults using both methods for LTT error calculation (see Table 2). These differences remained after controlling for age, gender, and education: (LTT-O, $F(1,95) = 30.13$; $p < .001$; $\eta_p^2 = .24$); (LTT-M, $F(1,103) = 21.30$; $p < .001$; $\eta_p^2 = .17$).

Serial Dotting Test

Independent samples *t* test revealed that the INPH group exhibited a significantly longer completion time and made more errors on the SDT (see Table 2). These differences remained significant after controlling for age, gender, and education (SDT Time: $F(1,106) = 15.90$; $p = .001$; $\eta_p^2 = .13$); (SDT Errors: $F(1,110) = 24.58$; $p < .001$; $\eta_p^2 = .18$).

Combined time and error scores

Results of the ANCOVA comparing combined time and error scores revealed that the INPH group demonstrated poorer performance compared to the healthy older adults on LTT-O Time + Errors, LTT-M Time + Errors, and SDT Time + Errors

after controlling for age, gender, and education. Results of the Time and Error Score comparisons are shown in Table 3.

Correlations

LTT-O and LTT-M were highly correlated ($r = 0.86$). Timed gait was moderately correlated with LTT-O Error, LTT-O Time + Error, SDT- Time, SDT Error, SDT Time + Error, and weakly correlated with LTT-M Error and LTT-M Time + Error. Gait mean steps was weakly correlated with LTT-O Error, LTT-O Time + Error, LTT-M Error, SDT Time, SDT Error, and SDT Time + Error (Table 4).

DISCUSSION

Our data indicate that LTT and SDT, two novel measures of psychomotor function, may be useful in examining upper extremity motor impairments observed in INPH. This is the first study that systematically evaluated both speed and accuracy to determine which aspects of performance best differentiate INPH participants from healthy older adults. Our findings indicate that accuracy may be more important than speed in evaluating performance and should not be overlooked. In fact, LTT time to completion was not different between INPH and the comparison group, highlighting the importance of examining accuracy when evaluating psychomotor skills in INPH.

A major contribution of this study is the development of the revised LTT error scoring method, and the introduction of an accuracy scoring method for SDT. Previously, researchers and clinicians who used these tasks in INPH have focused

Table 3. Combined Time and Error score

	INPH			Controls			<i>p</i> Value	η_p^2
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>		
Line Tracing-O Time + Errors	67	4.28	3.11	27	.21	1.19	<.001	.23
Line Tracing-M Time + Errors	69	1.62	1.79	33	0.29	1.21	.001	.11
Serial Dotting Time + Errors	75	2.85	2.46	35	0	.91	<.001	.20

Table 4. Partial Correlations correlations controlling for age, education, and gender

Variables	1	2	3	4	5	6	7	8	9	10
1.LTT ^a -Time	—									
2.LTT ^a -O Error	-.18	—								
3.LTT ^a -O Time + Error	.38***	.84***	—							
4.LTT ^a -M Error	-.15	.86***	.73***	—						
5.LTT ^a -M Time + Error	.89***	.22*	.70***	.32**	—					
6.SDT ^b -Time	.43***	.36***	.58***	.17	.49***	—				
7.SDT ^b -Error	.13	.59***	.63***	.48***	.34***	.31**	—			
8.SDT ^b - Time + Error	.37***	.57***	.74***	.36***	.52***	.86***	.75***	—		
9.Gait-Time	.17	.42***	.49***	.28**	.29**	.37***	.51***	.53***	—	
10.Gait- Mean Steps	.04	.29**	.29**	.21*	.14	.29**	.27**	.34**	.71***	—

^aLTT-Line Tracing Test.^bSDT-Serial Dotting Test* $p < .05$.** $p < .01$.*** $p < .001$.

primarily on time to completion; therefore, less is known about accuracy. This may have been due to the fact that the original LTT error scoring was cumbersome and a scoring paradigm for SDT accuracy had never been developed. While SDT time and errors appear to be correlated, our findings indicate that it was the errors for LTT that differentiated groups, highlighting the importance of accuracy scoring for this measure. Our data suggest that both the original and revised scoring methods demonstrate excellent inter-rater reliability and are highly correlated. The LTT-M is less cumbersome, and more efficient, and therefore, is more likely to be adopted by clinical and research centers that evaluate INPH patients. The combined scores, which examine both time and accuracy, should be further developed as a potential outcome measure to determine change in this population. Work is also needed to examine psychometric properties of these novel assessment tools to establish validity and reliability.

While many INPH patients have motor dysfunction that interferes with their performance on LTT and SDT, another factor that may contribute specifically to reduced accuracy is impairment in frontal executive skills in this population. The multiple demands of LTT and SDT introduce an executive component to the task, where self-monitoring is necessary to prioritize both accuracy and speed. This may be more problematic in INPH, as this group is known to have dysfunction of frontal networks. Future studies will examine the correlations between other frontal executive measures in the neuropsychological battery with and without a motor component (i.e., Trail Making Test B, phonemic fluency, working memory, motor programming tasks) and LTT and SDT speed and accuracy. This was not possible in the present study since each of the sites used different executive measures and did not use a uniform testing protocol across centers.

A question may arise regarding the use of experimental motor tests when there are several well-established neuropsychological measures of motor and psychomotor function with strong psychometric properties. In our experience,

standard motor tasks (i.e., Pegboard, Finger Tapping Test, Trail Making Test) do not allow for thorough assessment of deficits. In fact, some patients are unable to perform these tasks at all, making it impossible to establish a baseline performance and assess change over time. One study of 185 INPH patients that used the Trail Making Test reported that only 70 of the participants were able to complete the task at baseline and only 13 additional participants completed the task after successful shunt placement (Solana, Sahuquillo, Junqué, Quintana, & Poca, 2012). Furthermore, some patients who are able to complete these traditional tasks demonstrate a floor effect, where even dramatic improvement in raw scores will not translate to improvement in standard scores, thus making it difficult to quantify change. SDT and LTT also require less administration time than some of the traditional motor tasks. For these reasons, SDT and LTT provide valuable data in INPH. These tasks have been used in patients with Minimal Hepatic Encephalopathy (Biller & Ferro, 2014; Weissenborn, 2013) and may also be suitable for assessing psychomotor skills in other conditions that exhibit significant motor impairment.

A limitation of the current study is that the INPH and comparison group were not matched for age, education and cognitive status. Additional work will focus on collecting data on a well-matched disease comparison group. Another limitation is that one center used multiple versions of the LTT task, whereas the other two centers used only one form. This procedure had been in place before the established collaboration; the protocols have since been standardized. Another consideration is that our reported reliability may be higher than what will be observed when these tools are implemented in a clinic setting since raters at our centers underwent rigorous training. In terms of reliability, it will be important to examine inter-rater as well as test-retest reliability in a control population to better understand pre- and post-surgical changes in patients with INPH.

It is notable that when these tests are used at bedside or in a clinic setting, they do not have to be scored in a tedious

manner to be valuable. Quick hand scoring can add useful information regarding differential diagnosis and treatment outcome. While additional work is still needed to establish the psychometric properties of these measures, the revised scoring tools will make implementation into a clinic setting more practical. When LTT and SDT are implemented for research purposes, we recommend reliability training of raters, implementation of double scoring procedures until raters are proficient, and intermittent reliability checks across sites.

This is the first study examining LTT and SDT performance in INPH patients. Overall, the current data suggest that these measures show great promise and should be further developed as assessment tools to investigate outcome in INPH. This is the first step in investigating a more objective and practical scoring criteria for these tasks. Of course, the results should be interpreted with caution given relative small effect size, however, they should not be dismissed and further build upon by future investigations. Additional validation work is necessary to establish the psychometric properties of these tasks and to examine the utility for monitoring treatment outcomes. More specifically, it will be important to examine INPH patients' performance following spinal tap and shunt surgery, correlate the findings with cognitive functions, and establish test-retest reliability. Future studies should compare performance of INPH patients with other disease groups that are differentiated from INPH including Mild Cognitive Impairment, Alzheimer's disease, Parkinson's disease, and subcortical ischemic disease.

ACKNOWLEDGMENTS

We thank Ariel Brent, Mallorie Gonzalez, Alice Mathew, and Jeffrey Ruiz for serving as raters to double score LTT and SDT protocols and Ania Mikos, Ph.D. for helping create the Brown database. We also thank Professor Karin Weissenborn (Department of Neurology, Medical School Hannover, Germany) for the Line Tracing and Serial Dotting forms, which are part of the Psychometric Hepatic Encephalopathy Test (Copyrights: Medical School Hannover, Germany). *Conflicts of Interest:* The authors declare that they have no conflicts of interests related to this study to report. *Funding:* This work was funded in part by a grant from NINDS (K23 NS045051; PI: Dr. Katzen).

REFERENCES

Adams, R.D., Fisher, C.M., Hakim, S., Ojemann, R.G., & Sweet, W.H. (1965). Symptomatic occult hydrocephalus with normal cerebrospinal-fluid pressure: A treatable syndrome. *New England Journal of Medicine*, 273(3), 117–126.

Bergsneider, M., Black, P.M., Klinge, P., Marmarou, A., & Relkin, N. (2005). Surgical management of idiopathic normal-pressure hydrocephalus. *Neurosurgery*, 57(3), S2–S29.

Biller, J., & Ferro, J.M. (2014). *Portosystemic encephalopathy. Neurologic aspects of systemic disease part II: Handbook of clinical neurology* (Vol. 120, p 661). Amsterdam: Elsevier.

Devito, E.E., Pickard, J.D., Salmond, C.H., Iddon, J.L., Loveday, C., & Sahakian, B.J. (2005). The neuropsychology of normal pressure hydrocephalus (NPH). *British Journal of Neurosurgery*, 19(3), 217–224.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Minimal state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.

Foss, T., Eide, P.K., & Finset, A. (2006). Intracranial pressure parameters in idiopathic normal pressure hydrocephalus patients with or without improvement of cognitive function after shunt treatment. *Dementia and Geriatric Cognitive Disorders*, 23(1), 47–54.

Hellström, P., Edsbacke, M., Blomsterwall, E., Archer, T., Tisell, M., Tullberg, M., & Wikkelsø, C. (2008). Neuropsychological effects of shunt treatment in idiopathic normal pressure hydrocephalus. *Neurosurgery*, 63(3), 527–536.

Iddon, J.L., Pickard, J.D., Cross, J.J.L., Griffiths, P.D., Czosnyka, M., & Sahakian, B.J. (1999). Specific patterns of cognitive impairment in patients with idiopathic normal pressure hydrocephalus and Alzheimer's disease: A pilot study. *Journal of Neurology, Neurosurgery, & Psychiatry*, 67(6), 723–732.

Kahlon, B., Sjunnesson, J., & Rehncrona, S. (2007). Long-term outcome in patients with suspected normal pressure hydrocephalus. *Neurosurgery*, 60(2), 327–332.

Katzen, H., Ravdin, L. D., Assuras, S., Heros, R., Kaplitt, M., Schwartz, T.H., ... Relkin, N.R. (2011). Postshunt cognitive and functional improvement in idiopathic normal pressure hydrocephalus. *Neurosurgery*, 68(2), 416.

Kitagaki, H., Mori, E., Ishii, K., Yamaji, S., Hirono, N., & Imamura, T. (1998). CSF spaces in idiopathic normal pressure hydrocephalus: Morphology and volumetry. *AJNR American Journal of Neuroradiology*, 19(7), 1277–1284.

Klinge, P., Rückert, N., Weißenborn, K., Dörner, L., Samii, M., & Brinker, T. (2001). *A practical neuropsychological test battery for measurement of cognitive and motor function in patients with Normal-pressure hydrocephalus (NPH)*. Poster presented at the Third International Hydrocephalus Workshop, Kos, Greece.

Klinge, P., Rückert, N., Schuhmann, M., Dörner, L., Brinker, T., & Samii, M. (2002). *Neuropsychological testing to improve surgical management of patients with chronic hydrocephalus after shunt treatment* (pp. 51–53). Vienna: Springer.

Marmarou, A., Bergsneider, M., Klinge, P., Relkin, N., & Black, P.M. (2005). The value of supplemental prognostic tests for the preoperative assessment of idiopathic normal-pressure hydrocephalus. *Neurosurgery*, 57(3), S2–S17.

McGirt, M.J., Woodworth, G., Coon, A.L., Thomas, G., Williams, M.A., & Rigamonti, D. (2005). Diagnosis, treatment, and analysis of long-term outcomes in idiopathic normal-pressure hydrocephalus. *Neurosurgery*, 57(4), 699–705.

Pujari, S., Kharkar, S., Metellus, P., Shuck, J., Williams, M.A., & Rigamonti, D. (2008). Normal pressure hydrocephalus: Long-term outcome after shunt surgery. *Journal of Neurology, Neurosurgery, & Psychiatry*, 79(11), 1282–1286.

Savolainen, S., Hurskainen, H., Paljärvi, L., Alafuzoff, I., & Vapalahti, M. (2002). Five-year outcome of normal pressure hydrocephalus with or without a shunt: Predictive value of the clinical signs, neuropsychological evaluation and infusion test. *Acta Neurochirurgica*, 144(6), 515–523.

Schomerus, H., Weissenborn, K., Hamster, W., Rückert, N., & Hecker, H. (1999). *PSE syndrome test manual*. Frankfurt: Swets Test Services.

- Shprecher, D., Schwalb, J., & Kurlan, R. (2008). Normal pressure hydrocephalus: Diagnosis and treatment. *Current Neurology and Neuroscience Reports*, 8(5), 371–376.
- Solana, E., Sahuquillo, J., Junqué, C., Quintana, M., & Poca, M.A. (2012). Cognitive disturbances and neuropsychological changes after surgical treatment in a cohort of 185 patients with idiopathic normal pressure hydrocephalus. *Archives of Clinical Neuropsychology*, 27(3), 304–317.
- Thomas, G., McGirt, M.J., Woodworth, G., Heidler, J., Rigamonti, D., Hillis, A.E., & Williams, M.A. (2005). Baseline neuropsychological profile and cognitive response to cerebrospinal fluid shunting for idiopathic normal pressure hydrocephalus. *Dementia and Geriatric Cognitive Disorders*, 20(2-3), 163–168.
- Tsakanikas, D., Katzen, H., Ravdin, L.D., & Relkin, N.R. (2009). Upper extremity motor measures of Tap Test response in normal pressure hydrocephalus. *Clinical Neurology and Neurosurgery*, 111(9), 752–757.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: The Psychological Corporation.
- Weissenborn, K. (2013). Psychometric tests for diagnosing minimal hepatic encephalopathy. *Metabolic Brain Disease*, 28(2), 227–229.
- Weissenborn, K., Ennen, J.C., Schomerus, H., Rückert, N., & Hecker, H. (2001). Neuropsychological characterization of hepatic encephalopathy. *Journal of Hepatology*, 34(5), 768–773.