Evaluating the Measurement Properties of the ScanCourse, a Dual-Task Assessment of Visual Scanning

Paige Lund, Caitlyn Moir, Lisa Kristalovich, W. Ben Mortenson

Importance: The ScanCourse is used by occupational therapists to evaluate visual scanning ability during locomotion. Its measurement properties have not been examined.

Objective: To assess the interrater reliability, test–retest reliability, and construct validity of the ScanCourse.

Design: This study involved data collection at two time points. To assess test–retest reliability, the ScanCourse was administered twice within a 2-week period. To assess interrater reliability, a second rater was present for one session. To assess level of agreement, a Bland–Altman plot was created. To assess absolute reliability, the standard error of measurement was calculated. To evaluate construct validity, the results of the ScanCourse were compared with results of the Bells Test and Trail Making Test A and B.

Setting: Rehabilitation hospital.

Participants: Forty-one patients with neurological impairments.

Outcomes and Measures: The ScanCourse (participants identify numbered cards placed on both sides of a hallway at various heights during locomotion).

Results: The ScanCourse was found to have excellent interrater reliability (intraclass correlation coefficient [ICC] [1,1] = .998; 95% confidence interval [CI] [.996–.999]), test–retest reliability (ICC [1,1] = .912; 95% CI [.811–.959]), a high level of agreement, and a low standard error of measurement (.503), and it was found to be significantly correlated with Trails A ($r_s = -.436$, $p = .009$) and B ($r_s = -.364$, $p = .029$).

Conclusions and Relevance: The assessment was found to have strong measurement properties, and it is therefore an appropriate tool for assessing dual-task visual scanning among those with neurological impairments.

What This Article Adds: This research demonstrates that the ScanCourse is reliable between raters and over time and that scores on the measure vary as anticipated with scores on a related measure, which provides evidence of its validity. These findings support its use in practice.

In rehabilitation settings for people with possible or confirmed neurological conditions, occupational therapy practitioners must consider multiple areas of function. Neurological conditions can result in a wide variety of impairments that affect daily activities, including the ability to drive. Cognitive functions of attention, working memory, visual–motor coordination, motor and mental speed, and visual scanning significantly contribute to predicting driving status of people after neurorehabilitation (Perumparaichallai et al., 2014). In addition, measures of executive functioning, visual search, and psychomotor speed accurately predict road test performance (Nieuwoehner et al., 2012). Proper assessment can allow therapists to determine whether an individual can likely complete the activity, and it can also help to guide treatment before *in situ* assessment and training, which is especially important in safety-sensitive tasks such as driving (Nieuwoehner et al., 2012).

Visual attention refers to one's ability to carefully observe objects to determine information about their features and their relationship to themselves and other objects in the environment (Warren, 2006). One product of visual attention is

Citation: Lund, P., Moir, C., Kristalovich, L., & Mortenson, W. B. (2020). Evaluating the measurement properties of the ScanCourse, a dual-task assessment of visual scanning. *American Journal of Occupational Therapy*, 74, 7401185040. <https://doi.org/10.5014/ajot.2020.032052>

visual scanning. *Visual scanning* consists of the ocular strategies used to obtain visual input; it is necessary to properly process and interpret visual information (Hutman, 2013).

The Brain Injury Visual Assessment Battery for Adults (BIVABA; Warren, 2006) consists of multiple visual assessments that measure acuity, contrast sensitivity, oculomotor function, visual field, and visual search. The ScanCourse is a component of the BIVABA and assesses a person's visual attention and scanning capacity while ambulating (Warren, 2006). The ScanCourse is administered by placing numbered cards on both sides of a hallway at varied heights and asking the test takers to verbalize or point out which cards they see as they mobilize down the hallway.

The ScanCourse has been recommended to assess visual scanning abilities in driver rehabilitation, because it is easy to access, administer, and score and combines visual search with mobility, a feature not seen in paper-based assessments (Warren, 2006). This combination of functional mobility and visual scanning classifies the ScanCourse as a dual-task assessment. Dual tasking is important in many aspects of daily life, including observing surroundings while crossing the street, conversing while walking, and driving (Yang et al., 2016).

Research on the BIVABA is limited, especially for the ScanCourse component. On the basis of our review of the literature, which included CINAHL and MEDLINE database searches as well as consultation with the developer of the ScanCourse, we were unable to identify any available studies that examined the measurement properties of the ScanCourse. Therefore, the objective of this study was to determine the measurement properties of the ScanCourse, including interrater reliability among occupational therapists using the ScanCourse with neurological clients, test-retest reliability of the ScanCourse in people with neurological impairments, and construct validity by comparing scores on related measures. We hypothesized that with the provision of standardized administration guidelines, the ScanCourse would have an intraclass correlation coefficient (ICC) greater than .8 for interrater and test-retest reliability and would be moderately correlated with measures of visual scanning, the Bells Test, and the Trail Making Test Parts A and B (between .3 and .6). Moderate correlations were expected because all three assessments measure visual attention and scanning; however, the ScanCourse differs from the other two because it is a divided attention task that involves movements while scanning, and it is not paper based. Dual tasks involving divided attention are typically more difficult as a result of task interference (Alavash et al., 2015).

Method

Study Design

This measurement study involved data collection by multiple raters. The goal was for each study participant to be evaluated at two time points and on one of those occasions by two therapists. Ethical approval was obtained from the local university and institutional ethics boards at each site before beginning the study. Researchers used a convenience sample from three hospital-based rehabilitation centers in the same geographic region. Data were collected over approximately 4 mo from February 2017 through May 2017.

Participants and Recruitment

Inpatients and outpatients with potential problems with visual scanning, as identified by their therapists, were considered for the study. Inclusion criteria included the ability to independently mobilize, by walking or wheelchair; either verbalize or point to their observations; and make consent clear. Individuals who were unable to verbalize or otherwise communicate their observations, or who were unable to provide consent, were excluded from the study.

Forty-one participants were recruited from the active caseload of occupational therapists working in neurorehabilitation at inpatient and outpatient facilities of three hospitals. Participants were recruited by their treating therapist, who planned to administer the tool as part of their normal practice. Each participant was provided with a letter of initial contact and a consent form. Participants were informed that there was no obligation for them to participate and that they could withdraw from the study at any time. Informed consent and demographic information were obtained before administration of the test.

Procedure

Occupational therapists administered the ScanCourse using a standardized administration protocol that was developed by the research team. The protocol included a script to read to the participant before completing the ScanCourse, reminders to not prompt the participant with visual scanning suggestions, and instructions on how to complete and score the ScanCourse. Occupational therapists were also provided with a standardized recording form, which contained the script, order of numbers in a table format for easy marking, and a checklist of visual scanning techniques potentially used by participants. The therapists' experience levels ranged from <1 yr to >10 yr. Therapists assessed their own clients; when acting as a second rater, they assessed other clients. One to two therapists were involved at a given time.

Researchers set up a partially standardized ScanCourse at each site. Numbers 1–20 were printed at 1.25 in. high using Arial font on a 3.25-in. × 4.25-in. card, which was laminated. The numbers were placed with the middle of the card at wall heights of 71 in., 42 in., and 20 in. Fifty percent of the numbers were on the left side of the hallway, and 50% were on the right side of the hallway. Hallway width and length were not standardized because of environmental constraints and therefore varied from 7 ft to 8 ft and 72 ft to 100 ft, respectively.

The Bells Test (Gauthier et al., 1989) and Trail Making Test Parts A and B (Trails A and Trails B; Marvin, 2012) were used to assess the level of convergent construct validity. The Bells Test is a cancellation test that measures visual neglect by asking individuals to cross out 35 bells embedded within 280 distractors on an 11-in. × 8.5-in. page (Zeltzer & Menon, 2011). The Bells Test demonstrates extreme group validity in that there is a statistically significant difference between scores of individuals with and without neglect (Gauthier et al., 1989; Vanier et al., 1990).

The Trails A and Trails B (Marvin, 2012) involve a variety of skills, including letter and number recognition, visual scanning, and motor function and primarily measure individual differences in speed and fluid cognitive abilities (Salthouse, 2011). Trails A involves connecting randomly distributed numbers in numerical sequence. Trails B involves connecting randomly distributed letters and numbers, alternating between a letter and a number in ascending order (i.e., 1, A, 2). The Trail Making Test demonstrates test–retest reliability of $r = .78$ for Trails A and $r = .67$ for Trails B as well as predictive validity for driving test outcomes for Trails B (Goldstein & Watson, 1989; Marshall et al., 2007).

Data Collection

To measure interrater reliability, during one of the two ScanCourse administrations, two raters observed participants and scored them on the standardized recording form provided. The second rater was another occupational therapist or an occupational therapy student who had previous experience administering the ScanCourse. Occupational therapy students were on first- or second-year clinical placements and were trained by their supervising therapist before administering the assessment.

To measure test–retest reliability, level of agreement, and standard error of measurement (SEM), the same occupational therapist administered the ScanCourse twice to the same participant within a 3- to 14-day period. Occupational therapists were instructed not to give any training, education, or feedback between the sessions. These measures were put in place to minimize a learning effect for the participants.

To examine the level of convergent construct validity between the ScanCourse and the Bells Test or Trails A and B, occupational therapists administered these assessments at baseline. These tests were chosen because they are used to determine potential difficulty with visual scanning tasks, such as driving (Vrkljan et al., 2011).

Data Analysis

Statistical analysis was completed using IBM SPSS Statistics (Version 23.0; IBM Corp., Armonk, NY). To determine the level of interrater reliability among occupational therapists using the ScanCourse with neurological clients and test–

retest reliability among neurological clients, ICCs (1,1) were calculated. We chose ICC (1,1) (i.e., one-way random effects model) because it is used in cases in which participants are not assessed by the same set of raters (Koo & Lee, 2016). In our study, participants were evaluated at multiple sites at which different therapists were used. A Bland–Altman plot was created to demonstrate the level of agreement between the difference in scores for ScanCourse Trial 2 and Trial 1 and the mean scores between these two trials (Figure 1). To evaluate absolute reliability (i.e., the degree of variability from trial to trial; Atkinson & Nevill, 1998), the SEM was calculated using the formula

$$S_{\text{measurement}} = S_{\text{test}} \sqrt{1 - r_{\text{test}}}.$$

Data were skewed because most participants scored high on the ScanCourse and other measures. Because data were found to be nonnormally distributed, Spearman's correlations were used instead of Pearson's correlations for construct validity testing, which compared ScanCourse Trial 1 scores with Bells Test, Trails A, and Trails B scores (Mukaka, 2012).

Results

Demographics

The demographic characteristics of the study population are shown in Table 1. The majority of participants were male and spoke English as a first language, and the most common diagnoses were stroke, brain injury, and spinal cord injury. The mean scores for the measures used for this assessment are reported in Table 2. The mean score for the ScanCourse on Trial 2 was slightly higher than the mean score for Trial 1. Note that the number of participants vary for each category because not all participants completed all components of the study as a result of time constraints or limitations of service delivery, for example, outpatients who were seen less frequently than the 2-wk time period required for readministration.

Figure 1. Mean ScanCourse scores between Trial 1 and 2.

Note. Mean ScanCourse scores between T1 and T2 on x axis; difference between scores on y axis. Solid line: mean of the difference in scores. Dotted lines: upper and lower limits of a 95% confidence interval, or limits of agreement defined as the mean difference $\pm$ 1.96 standard deviation. Upper limit, 4.61; lower limit, -2.04. T1 = Trial 1; T2 = Trial 2.

Table 1. Participant Demographics

Characteristic	<i>M</i> ($\pm$ <i>SD</i>) or <i>n</i> (%)	Total Participants Assessed
Age, yr	54.1 ($\pm$ 18.2)	41
Male	33 (82.5)	40
Most common diagnoses		40
Stroke	24 (60.0)	
Brain injury	6 (15.0)	
Spinal cord injury	3 (7.5)	
Most common method of mobility:		
Walking without aid	26 (70.0)	37
English as a first language	28 (70.3)	40

Reliability

For the participants assessed for interrater reliability ($N = 36$), the ICC was .998 (95% confidence interval [CI] [.996–.999]). The ICC for test–retest reliability ($N = 28$) was .912 (95% CI [.811–.959]). The SEM was .503.

Level of Agreement

Figure 1 displays the Bland–Altman plot.

Spearman's correlation coefficient was calculated to determine the level of correlation

of the ScanCourse to related measures. The ScanCourse was found to be significantly correlated to Trails A ($N = 35$; $r_s = -.436$; $p = .009$) and Trails B ($N = 36$; $r_s = -.364$; $p = .029$). It was not significantly correlated to the Bells Test ($N = 41$; $r_s = .140$; $p = .383$).

Discussion

This study determined the interrater reliability, test–retest reliability, and construct validity of the ScanCourse compared with the Bells Test and Trails A and B. There was good support for our first two hypotheses, that the ScanCourse would have an ICC greater than .8 for interrater and test–retest reliability. These values are sufficient for individual comparisons ($>.90$; Aaronson et al., 2002). The Bland–Altman plot shows a slight increase over time, which could be due to a learning effect of the assessment or to improvement in the participants' neurological impairment over time. The SEM indicates that a 1-point change in score would not be considered an error (i.e., a 1-point difference is considered to be statistically significant). However, further research is needed to determine which values represent clinically significant changes.

We found support for only a portion of our final hypothesis: The ScanCourse was significantly correlated to the Trails A and B, but not to the Bells Test. These results suggest that the Trail Making Test and the ScanCourse measure different skills than the Bells Test. For example, the Bells Test is a static test of near space and visual neglect, whereas the Trail Making Test and the ScanCourse involve following a path of targets. The ScanCourse and Trail Making Test also both involve recognition of numbers versus the symbols used in the Bells Test.

The correlation between Trails B and the ScanCourse provides support for the potential use of the ScanCourse in driver rehabilitation. Multiple studies support the use of the Trails B in predicting fitness to drive (Classen et al., 2013; Devos et al., 2011; Gibbons et al., 2017; Roy & Molnar, 2013; Vrkljan et al., 2011). Previous studies have found that participants who completed the Trail Making Test B in more than 90 s required further on-road assessment and that the Trail Making Test B was the best predictor of driving ability compared with the Montreal Cognitive Assessment, Motor-Free Visual Perceptual Test, and clock drawing test (Devos et al., 2011; Gibbons

et al., 2017). Therefore, an area for future research would be to examine the use of the ScanCourse in predicting fitness to drive, especially because it involves dual tasking, which is critical for driving (Baldwin & Schieber, 1995).

The study had two main limitations. We were unable to standardize hallway width, length, color, or amount of clutter because of

Table 2. Mean Scores of Study Measures

Measure (Scoring, Min–Max or Time)	<i>n</i> (%)	Total Participants Assessed
ScanCourse Trial 1 (0–20)	16.2 (3.81)	41
ScanCourse Trial 2 (0–20)	17.4 (3.39)	28
Bells Test (0–35)	33.2 (2.77)	41
Trail Making Test Part A, s	44.3 (24.0)	35
Trail Making Test Part B, s	128.9 (96.3)	36

constraints of the facilities used. Data were collected at the site level; therefore, the test–retest reliability or interrater reliability for each clinician or pair cannot be calculated.

Implications for Occupational Therapy Practice

This study has the following implications for occupational therapy practice:

- When administered using a standardized protocol, the ScanCourse demonstrates high interrater and test–retest reliability.
- Because the ScanCourse was not significantly correlated to the Bells Test, it should not be used to identify near space visual neglect.
- Because the ScanCourse is significantly correlated to the Trails A and B, which are currently used to predict driving capability (Marshall et al., 2007), the ScanCourse warrants additional study to determine how well it predicts driving capability.

Conclusion

The results of this study provide support for the use of the ScanCourse as a visual scanning assessment in clients with neurological impairment; it was found to have excellent interrater and test–retest reliability and to be significantly correlated to Trails A and B. The correlation between scores on the ScanCourse and Trail Making Test B may indicate that the ScanCourse may be appropriate for use in a driver rehabilitation setting. ■

References

- Alavash, M., Hilgetag, C. C., Thiel, C. M., & Giebing, C. (2015). Persistency and flexibility of complex brain networks underlie dual-task interference. *Human Brain Mapping*, 36, 3542–3562. <https://doi.org/10.1002/hmb.22861>
- Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26, 217–238. <https://doi.org/10.2165/00007256-199826040-00002>
- Aaronson, N., Alonso, J., Burnam, A., Lohr, K. N., Patrick, D. L., Perrin, E., & Stein, R. E.; Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*, 11, 193–205. <https://doi.org/10.1023/A:1015291021312>
- Baldwin, C. L., & Schieber, F. (1995). Dual task assessment of age differences in mental workload with implications for driving. *Proceedings of the Human Factors and Ergonomics Society 29th Annual Meeting*, 39, 167–171. <https://doi.org/10.1177/154193129503900210>
- Classen, S., Wang, Y., Crizzle, A. M., Winter, S. M., & Lanford, D. N. (2013). Predicting older driver on-road performance by means of the Useful Field of View and Trail Making Test Part B. *American Journal of Occupational Therapy*, 67, 574–582. <https://doi.org/10.5014/ajot.2013.008136>
- Devos, H., Akinwuntan, A. E., Nieuwboer, A., Truijen, S., Tant, M., & De Weerd, W. (2011). Screening for fitness to drive after stroke: A systematic review and meta-analysis. *Neurology*, 76, 747–756. <https://doi.org/10.1212/WNL.0b013e31820d6300>
- Gauthier, L., Dehaut, F., & Joannette, Y. (1989). The Bells Test: A quantitative and qualitative test for visual neglect. *International Journal of Clinical Neuropsychology*, 11, 49–54.
- Gibbons, C., Smith, N., Middleton, R., Clack, J., Weaver, B., Dubois, S., & Bédard, M. (2017). Using serial trichotomization with common cognitive tests to screen for fitness to drive. *American Journal of Occupational Therapy*, 71, 7102260010. <https://doi.org/10.5014/ajot.2017.019695>
- Goldstein, G., & Watson, J. R. (1989). Test–retest reliability of the Halstead–Reitan Neuropsychological Battery and the WAIS in the neuropsychiatric population. *Clinical Neuropsychology*, 3, 265–273. <https://doi.org/10.1080/13854048908404088>
- Hutman, H. (2013). Visual scanning. In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders* (pp. 3300–3305). New York: Springer.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Marshall, S. C., Molnar, F., Man-Son-Hing, M., Blair, R., Brosseau, L., Finestone, H. M., . . . Wilson, K. G. (2007). Predictors of driving ability following stroke: A systematic review. *Topics in Stroke Rehabilitation*, 14, 98–114. <https://doi.org/10.1310/tsr1401-98>
- Marvin, K. (2012). *Trail Making Test (TMT)*. Retrieved from <https://www.strokeengine.ca/assess/tmt/>
- Mukaka, M. M. (2012). Statistics Corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24, 69–71.
- Niewoehner, P. M., Henderson, R. R., Dalchow, J., Beardsley, T. L., Stern, R. A., & Carr, D. B. (2012). Predicting road test performance in adults with cognitive or visual impairment referred to a Veterans Affairs Medical Center driving clinic. *Journal of the American Geriatrics Society*, 60, 2070–2074. <https://doi.org/10.1111/j.1532-5415.2012.04201.x>

- Perumparaichallai, R. K., Husk, K. L., Myles, S. M., & Klonoff, P. S. (2014). The relationship of neuropsychological variables to driving status following holistic neurorehabilitation. *Frontiers in Neurology*, 5, 56. <https://doi.org/10.3389/fneur.2014.00056>
- Roy, M., & Molnar, F. (2013). Systematic review of the evidence for Trails B cut-off scores in assessing fitness-to-drive. *Canadian Geriatrics Journal*, 16, 120–142. <https://doi.org/10.5770/cgj.16.76>
- Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence*, 39, 222–232. <https://doi.org/10.1016/j.intell.2011.03.001>
- Vanier, M., Gauthier, L., Lambert, J., Pepin, E. P., Robillard, A., Dubouloz, C. J., . . . Joannette, Y. (1990). Evaluation of left visuospatial neglect: Norms and discrimination power of two tests. *Neuropsychology*, 4, 87–96. <https://doi.org/10.1037/0894-4105.4.2.87>
- Vrkljan, B. H., McGrath, C. E., & Letts, L. J. (2011). Assessment tools for evaluating fitness to drive: A critical appraisal of evidence. *Canadian Journal of Occupational Therapy*, 78, 80–96. <https://doi.org/10.2182/cjot.2011.78.2.3>
- Warren, M. (2006). *Brain Injury Visual Assessment Battery for Adults test manual*. Birmingham, AL: visAbilities Rehab Services.
- Yang, L., He, C., & Pang, M. Y. C. (2016). Reliability and validity of dual-task mobility assessments in people with chronic stroke. *PLoS One*, 11, e0147833. <https://doi.org/10.1371/journal.pone.0147833>
- Zeltzer, L., & Menon, A. (2011). Bells Test. Retrieved from <https://www.strokengine.ca/assess/bt/>

Paige Lund, BSc, MOT, is Occupational Therapist, Family Resource Association, Parksville, British Columbia, Canada.

Caitlyn Moir, BA, MOT, is Occupational Therapist, Vancouver Coastal Health Authority, Vancouver, British Columbia, Canada.

Lisa Kristalovich, BMR (OT), MRSc, is Clinical Faculty, Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, British Columbia, Canada, and Occupational Therapist, Vancouver Coastal Health Authority, Vancouver, British Columbia, Canada.

W. Ben Mortenson, BScOT, MSc, PhD, is Associate Professor, Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, British Columbia, Canada; Principal Investigator, Rehabilitation Research Program, Vancouver Coastal Health Research Institute (VCHRI), Vancouver, British Columbia, Canada; and Principal Investigator, International Collaboration on Repair Discoveries, VCHRI, Vancouver, British Columbia, Canada; ben.mortenson@ubc.ca

Acknowledgments

The authors thank the occupational therapists who contributed to the data collection. W. Ben Mortenson was supported by a New Investigator Award from the Canadian Institutes of Health Research. This is an unfunded study completed in fulfillment of a Master of Occupational Therapy degree.