Tablet-Based Trail Making Tests Over the Lifespan: Validating Target Ranges on an In-Clinic Platform

David Joffe[a], D S Oakley[a], and F. X. Palermo[b]
[a] WAVi Research, Boulder CO [b] University of Colorado School of Medicine, Aurora, CO

Abstract

Background: Trail Making tests (TMT) are commonly-used neuropsychological assessments that test a wide variety of cognitive processes. These tests can provide information regarding age- and trauma-related changes in brain function in addition to disease-related cognitive decline. Age-related targets can provide useful guides for practitioners, trainers, and patients seeking to optimize and track brain function over time.

Objective: To validate a method for presenting age-stratified target ranges for completion times of tablet-based (touch screen) Trail Making A and B tests (TMTA, TMTB), and physical reaction time (PRT) concurrent with an audio P300 test.

Participants: One thousand nine hundred and fifty subjects aged 13-90.

Methods: TMT and PRT were measured as part of a health screening exam for studies through Colorado University, Children's Hospital Colorado, Boone Heart Institute, WAVi Co., and various clinics alongside other clinical evaluations such as EEG and audio P300.

Group Results: TMT completion times in tablet-based, in-clinic testing show similar age-related changes seen in previous paper-based research, with TMTB showing the best agreement. PRT, on the other hand shows less age-related variation.

Conclusion: A target reference of TMT completion times and PRT is a useful tool to compares trends with end points of high functioning people on both ends of the age range.

Keywords

((Electroencephalogram (EEG), P300, Event Related Potential (ERP), Brainwave, Amplitude, Trail Making Test, Physical Reaction Time))

Introduction

Trail making tests (TMT) have been widely used in neuropsychological assessments.[1] [2] [3] In TMTA the participant draws lines to connect circled numbers in a numerical sequence (i.e., 1-2-3, etc.), where in TMTB the participant draws lines to connect circled numbers and letters in an alternating numeric and alphabetic sequence (i.e., 1-A-2-B, etc.). Both the time to complete the test and the number of errors are then recorded. These simple tests have been hypothesized to reflect a wide variety of cognitive processes including attention, visual search and scanning, sequencing and shifting, psychomotor speed, abstraction, flexibility, ability to execute and modify a plan of action, and ability to maintain two trains of thought simultaneously. [4] [5] [6]

A number of studies have examined the patterns of relations between trail-making ((and health, where TMTB time has been cited as one of the strongest predictors of MCI out of numerous other tests Frank reference)). [7] TMT completion times have been seen to decline with concussion, trauma, and age, where error rates may be less susceptible to age differences than time to completion. [8] [9] [10] [11]

In classifying individuals into diagnostic categories, significant group differences are seen in time to completion between control, mild cognitive impairment, and Alzheimer's, particularly for TMTB, but as with many tests the between-person variance is large enough to suggest these tests are best suited to be used in combination with other assessments and for longitudinal tracking.

Physical reaction time also represents a straightforward measurement that correlates with a host of conditions, including concussion, executive-function tasks have shown some sensitivity in predicting functional decline and mortality, detecting persons at risk for AD, psychopathology, and in high-functioning older adults PRT as a measure of successful independent functioning and medication compliance. [12] [13] [14] [15] [16] [17] [18] [19]

The WAVi platform incorporates various EEG tests, including audio P300 ERP, and visual Flanker ERP, with and a tablet-based TMT. The system is designed to maximize information and minimize testing times in-clinic. The objective here is to validate the TMT measures across the lifespan for these metrics measured in real life clinical settings against published trends in research as well as to establish an in-clinic trend for PRT.

Methods

Subjects

The subjects for this study were comprised of 1950 subjects from previous or ongoing studies, 1417 of those completing the TMT portion of the test. They are not intended to represent a normal control for a general population, rather to provide a target reference. One of the goals of this study is to compare in-vivo data with historical research to test the validity of large-scale screening. This study does have 3 control groups (13-16 years of age, 17-23, and 81-90) and these will anchor the resulting age-matched curves as discussed below. It may be the case that these controls perform differently from a found in a normal population, where 2 of these control groups were taken from elite club, High School, or NCAA athletic teams while those in the oldest age range were volunteers living independently, still interested in brain science, and still interested in their brain performance. Each group will be discussed individually, but because of the suspected other-than-normal performance, we will focus on age trends and refer to this reference group as a target reference, with end points as discussed, rather than a normal reference.

It is important to note that while there may be male/female differences, this is not the focus of this paper which is to compare trends to literature to establish a reference target.

All studies were approved by appropriate IRB's and written informed consent was obtained from the participants before study intake.

Ages 8-12

48 subjects aged 8-12 were taken from three previous studies: a study that followed athletes over the course of their sports seasons in Texas and Washington, control subjects measured as part a beta test to explore the outcome of an educational/wellbeing intervention program in an economically-challenged school, [20] and wards accompanying WAVi study volunteers discussed below.

Ages 13-16

This control group comprises 83 subjects from a previous study following 94 athletes aged 13-16 over the course of their sports seasons and at 4 different sites. These subjects are participants in youth soccer and youth basketball representing all players from single teams representing all players from single teams. [21]

In these previous studies, these subjects were controls against which pre-contact, post-concussion and return-to-play groups could be compared. Here P300 voltages, along with reaction time and Trail Making measurements, were assessed during the course of other pre-contact clinical evaluations administered by sports medicine staff. To follow the objectives of this study (as well as the above-mentioned studies) which

involves real clinical settings, and because the primary marker being studied is nonspecific, our exclusion criteria are minimal. The "control" group, therefore, is a reference group taken from all players participating on these teams with no exclusions, except that those who had lower than 80% yield on the audio P300 protocol that were excluded due to artifact. Of these 83 subjects, 66 returned and completed a valid post-season second test which will be used to discuss test-retest variability.

Ages 17-23

This second control group is taken from a previous study that followed 364 athletes aged 17-23 over the course of up to 4 sports seasons and at 5 different sites. These subjects are participants in NCAA Div. 1 men's football (172 players, representing all players from a single team), woman's soccer (29 NCAA Div. 1, representing all players from a single team), men's high school football (142 players, representing all seniors from a single team), and semipro men's ice hockey (20 players, representing all players from a single team). [9]

In these previous studies, these subjects were controls against which pre-contact, post-concussion and return-to-play groups could be compared. Here P300 voltages, along with reaction time and Trail Making measurements, were assessed during the course of other pre-contact clinical evaluations administered by sports medicine staff on 356 of the 364 players tracked. To follow the objectives of this study (as well as the above-mentioned studies) which involves real clinical settings, and because the primary marker being studied is nonspecific, our exclusion criteria are minimal. The "control" group, therefore, is a reference group taken from all players participating on these teams and exclusions are limited to the players who fell asleep during the first-year test and passing the artifact criteria discussed above, leaving a total of 304 players comprising the baseline reference group of Table I. Of these subjects, 70 returned injury free to completed a valid second test, which will be used to discuss test-retest variability for P300 amplitude, with a subset of 38 of these to test P300 latency variation (because of a change in protocol as noted in the study).[9]

Ages 24-30

128 individuals aged 24-30 were measured in clinic at baseline where some were to be tracked over the course of various interventions. Subjects include patients who visited the Boone Heart Institute Colorado for a combined preventative cardiology and EEG/ERP evaluation from June 2014 through June 2017. Only first-time patients receiving an initial evaluation were included in the sample, which was also used for a preventative cardiology study. [6] Because this is a target reference study, the exclusion criteria are minimal, the criteria being those who were taking beta-blockers or psychiatric medication and those who

had lower than 80% yield on evoked responses due to artifact.

Also included were subjects from Natural Bio Health (NBH) Texas for a first-time preventative wellness exam, evaluated from 2017 through 2018; and a random sampling of subjects measured for demonstration purposes at 5 medical conferences.

The remaining subjects were volunteers who were known to or associated with the study team and wanted to become pro-active in their brain health. In general these reference subjects were well educated and wanted to use WAVi to compare pre-intervention to post-interventions where interventions typically included some form of lifestyle change. To follow the objectives of this study, which involves real clinical settings, and because the primary marker being studied is nonspecific, our exclusion criteria are minimal and all volunteers in this age group were analyzed for the purposes of this study.

Ages 31-40
204 individuals aged 31-40 were tracked over the course of various interventions and the above-mentioned clinics, conferences, and volunteers.

Ages 41-50
297 individuals aged 31-40 were tracked over the course of various interventions and the above-mentioned clinics, conferences, and volunteers.

Ages 51-60
377 individuals aged 31-40 were tracked over the course of various interventions and the above-mentioned clinics, conferences, and volunteers.

Ages 61-70
231 individuals aged 31-40 were tracked over the course of various interventions and the above-mentioned clinics, conferences, and volunteers.

Ages 71-80
56 individuals aged 31-40 were tracked over the course of various interventions and the above-mentioned clinics, conferences, and volunteers.

Ages 81-90

Our third control group comprises 42 people taken as volunteers, discussed above. This group were living independently, had not been diagnosed with dementia, and were by definition a population who had experienced what could be called successful cognitive aging. They provide an end point against the 20-year old athletes for our target reference.

Table I Profile of Assessments.		
Age (yrs)	***# Assessments (PRT/TMT)***	***Subject Profile***
8-12	73/66	Youth Sports and Educational studies: *initial assessment upon enrollment into study*.
13-16	79/58	Youth Sports study: *initial pre-contact assessment upon enrollment into study*.
17-23	314/247	High School and University Soccer, Hockey, and US Football study: i*nitial pre-contact assessment upon enrollment into study*.
24-30	143/108	Preventative Cardiology and Healthy Aging studies: i*nitial pre-intervention assessment upon enrollment into study*.
31-40	228/187	Preventative Cardiology and Healthy Aging studies: i*nitial pre-intervention assessment upon enrollment into study*.
41-50	344/249	Preventative Cardiology and Healthy Aging studies: i*nitial pre-intervention assessment upon enrollment into study*.
51-60	414/279	Preventative Cardiology and Healthy Aging studies: i*nitial pre-intervention assessment upon enrollment into study*.
61-70	263/173	Preventative Cardiology and Healthy Aging studies: i*nitial pre-intervention assessment upon enrollment into study*.
71-80	63/33	Preventative Cardiology and Healthy Aging studies: i*nitial pre-intervention assessment upon enrollment into study*.
81-90	29/17	Healthy Aging study: *initial assessment upon enrollment into study*.

Test Administration

The WAVi test begins with an intake, a 4-min audio P300, followed by TMTA and TMTB. Values were extracted using WAVi 9.7.5 software (intended to create the target ranges for WAVi 9.8-10.0 software).

During the TMT, subjects were asked to press the circle containing the correct number or letter on the screen where a connecting line is then drawn for them when the correct choice is made. This is a small deviation from paper-based TMT which asks the subjects to draw the connecting lines between the appropriate items.

The PRT is obtained during the 4-minute 2-tone eyes-closed audio P300 protocol.[22] The P300 test used 200 common tones and 40 randomly interweaved rare tones that were presented through headphones.

Participants were instructed to manually respond by pressing the left button on a standard mouse whenever the rare tone was heard. This test allowed for the recording of the PRT from this manual response alongside P300 voltage and latency times.

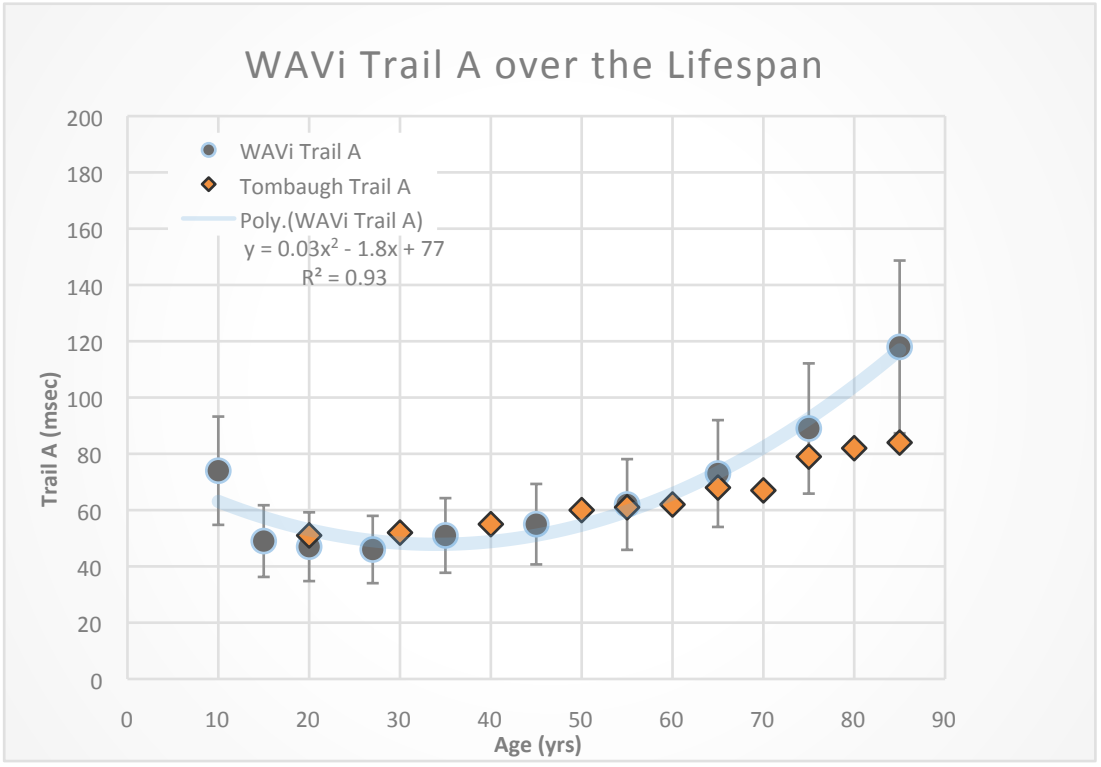
Results: Age Trends

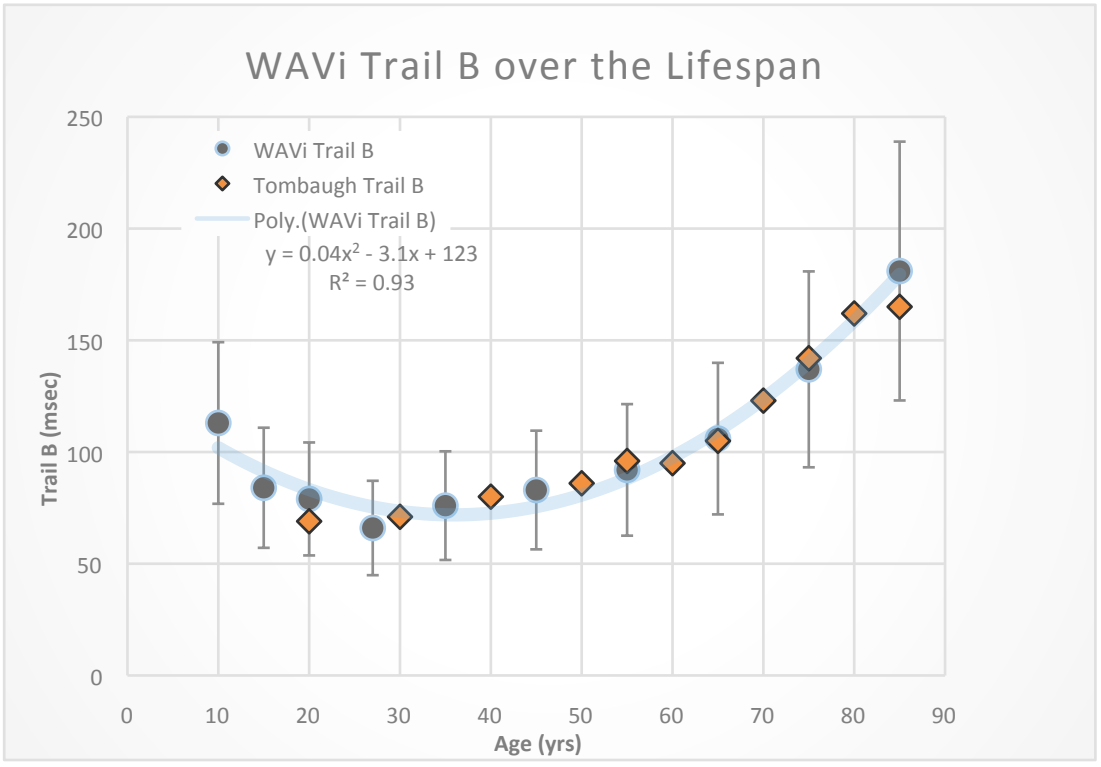
Results for WAVi Trail A and Trail B testing times are shown in Figures 1 and 2 respectively. Here we compare the WAVi in-clinic touchscreen data, with the blue line the best fit, to a previous study of 911 patients over the same age range.[23] Because of the difference in techniques, with the touch screen methods typically slower than the paper methods, the previous TMTA was normalized up by 17 seconds and TMTB normalized by seconds to better guide the eye.

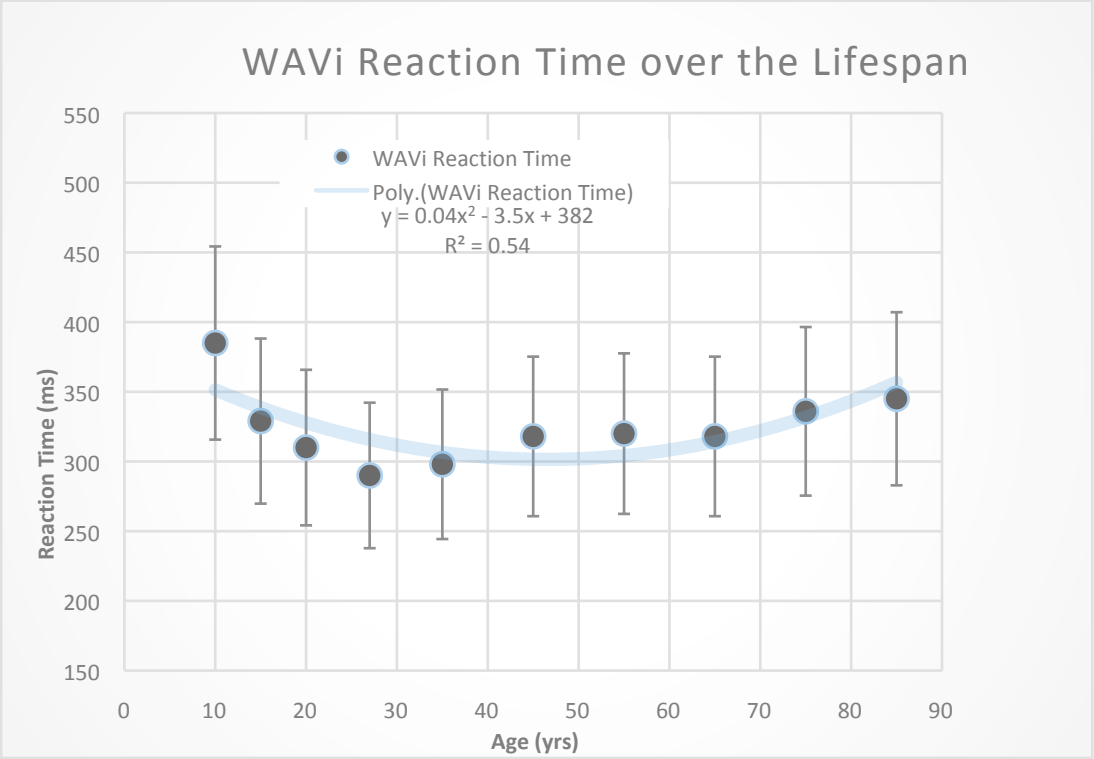
For TMTA time to competion, the WAVi data show the maturation effects that seem to reach maximum performance around 20 years of age that remain flat until the later age. The WAVi trend matches the previous paper studies until the older ages where the age falloff is much steeper. It is not clear if this is actually age-related or an unfamiliarity with computers versus a familiarity with paper for this older group.

Regarding TMTB completion time, however, the match with the previous paper methods is very strong for all ages with a strong age falloff in all cases. This large falloff is accompanied by large person-to-person variation in the older ages. WAVi also sees a maturation prediction that peaks at around 20 years of age as with TMTA.

Finally, the PRT data are shown in Figure 3. Here we see that after maturation peaks at around 20 years of age, the PRT remains fairly flat throughout the lifespan. Large person-to-person variation relative to the variation in age, however, makes the trend line less reliable than in the TMT cases.


WAVi Trail A over the Lifespan
WAVi Trail A
Tombaugh Trail A
Poly.(WAVi Trail A)
y = 0.03x² - 1.8x + 77
R² = 0.93
Trail A (msec)
Age (yrs)
0
20
40
60
80
100
120
140
160
180
200
0
10
20
30
40
50
60
70
80
90


WAVi Trail B over the Lifespan
WAVi Trail B
Tombaugh Trail B
Poly.(WAVi Trail B)
y = 0.04x² - 3.1x + 123
R² = 0.93
Trail B (msec)
Age (yrs)
0
50
100
150
200
250
0
10
20
30
40
50
60
70
80
90



A note on variance: It is common practice in medicine to quote "normal ranges" (middle 68 percent of people in a Gaussian distribution) in order to provide context for both the clinician and client. Because TMT and PRT are a combination of state *and* trait a Gaussian distribution cannot always be assumed, however. For TMTA, TMTB, and PRT we have found variances of +-25%; +-29%; and +-18% provide useful target ranges that adequately capture the middle ranges.

Conclusion

Test completion times for a tablet-based TMT collected in-vivo produce trends that decline with age in agreement with previous paper-based studies. These data also see a decline for younger ages presumably due to maturation. Physical reaction time, on the other hand, appears to be more flat across the lifespan. A target reference of TMT completion times and PRT is a useful tool to compares trends with end points of high functioning people on both ends of the age range.

[1] Butler M, Retzlaff P, Vanderploeg R. Neuropsychological test usage. Professional Psychology: Research and Practice. 1991;6:510–512.

[2] Rabin LA, Barr WH, Burton LA. Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. Archives of Clinical Neuropsychology. 2005;20:33–65. [

[3] Sellers AH, Nadler JD. A survey of current neuropsychological assessment procedures used for different age groups. Psychotherapy in Private Practice. 1992;11:47–57.

[4] Lezak MD, Howieson DB, Loring DW. Neuropsychological Assessment. 4. Oxford University Press; New York, NY: 2004.

[5] Salthouse TA, Fristoe NM. A process analysis of adult age effects on a computer-administered trail making test. Neuropsychology. 1995;9:518–528.

[6] Strauss E, Sherman EMS, Spreen O. A compendium of neuropsychological tests: Administration, norms, and commentary. 3. Oxford University Press; New York, NY: 2006.

[7] ((Frank Ref))

[8] ((Clayton))

[9] ((Robins-Wahlin et al., 1996)).

[10] Salthouse TA, What cognitive abilities are involved in trail-making performance? Intelligence. 2011 July-August; 39(4): 222–232.

[11] Lee Ashendorf, Angela L. Jefferson, Maureen K. O'Connor, Christine Chaisson, Robert C. Green, and Robert A. Stern. Trail Making Test Errors in Normal Aging, Mild Cognitive Impairment, and Dementia. Arch Clin Neuropsychol. 2008 March ; 23(2): 129–137.

[12] James T. Eckner, Jeffrey S. Kutcher, James K Richardson Effect of concussion on clinically measured reaction time in nine NCAA Division I collegiate athletes: a preliminary study. PM R . 2011 March ; 3(3): 212–218.

[13] Eckner, J., Kutcher, J., Broglio, S., & Richardson, J. (2014). British J Sports Med , 48(2), 112-118.

[14] Johnson JK, Lui L, Yaffe K. Executive function, more than global cognition, predicts functional decline and mortality in elderly women. Journal of Gerontology: Medical Sciences. 2007;62:1134–1141.

[15] Wetter SR, Delis DC, Houston WS, Jacobson MW, Lansing A, Cobell K, et al. Deficits in inhibition and flexibility are associated with the APOE-E4 allele in nondemented older adults. Journal of Clinical and Experimental Neuropsychology. 2005;27:943–952.

[17] Whitmer AJ, Banich MT. Inhibition versus switching deficits in different forms of rumination. Psychological Science. 2007;18:546–553.

[18] Carlson MC, Fried LP, Xue QL, Tekwe C, Brandt J. Validation of the Hopkins Medication Schedule to identify difficulties in taking medications. Journal of Gerontology: Medical Sciences. 2005;60:217–223.

[19] Tun, PA, Lachman, ME. Age Differences in Reaction time and Attention in a National Telephone Sample of Adults: Education, Sex, and Task Complexity Matter. Dev Psychol. Sep 2008; 44(5): 1421–1429.

[20] Maroney, D. et al. (2019). *In Progress*.

[21] Wilson, J. et al. (2019). *In Progress*

[22] Joffe, D. et al. (2019). *In Progress*

[23] Tom N. Tombaugh. Trail Making Test A and B: Normative data stratified by age and education. Archives of Clinical Neuropsychology 19 (2004) 203–214