Reliability of a Battery of cognitive tests in Young, healthy people

Ramírez-delaCruz M[1], Bravo-Sánchez A[2], Esteban-García P[3], Jiménez F[4], Abián-Vicen J[5]*

Abstract

Cognitive performance is an essential aspect of sports. Still, few studies provide reliable cognitive tests for training and evaluating cognitive aspects related to sports, particularly in a healthy and/or youthful population. This study assessed the intraday and interday reliability of a cognitive test battery related to sports performance. In the research, seventeen healthy volunteers participated. DynavisionTMD2, Bassin Anticipation Timer, Go/No-Go, Eriksen Flanker, and the Trail Making Test (TMT) were analyzed. TMT-A and TMT-B (intraday), as well as the following variables in the interday analysis, exhibited significant differences: "Physical RT - Fastest" (Test 1), "Visual RT - Fastest" and "Physical RT - Fastest" (Test 2), "Visual RT - Fastest" (Test 4) and TMT-A. Thirteen of the thirty-one variables analyzed on the intraday exam had moderate Intraclass Correlation Coefficients (ICC), six were good, and one was exceptional. In the interday analysis, 15 variables with moderate ICC and 5 variables with excellent ICC were identified. Reliable tests include "Reaction Time" (Tests 5 and 6) and "Red and Green Lights" of the DynavisionTMD2, the 15 mph speed of the Bassin Anticipation Timer, and the Go/No-Go, Eriksen Flanker, and TMT tests.

Keywords: sport; psychology; neurocognitive; reaction time; anticipation.

1. Introduction

Lindenberger, Scherer, and Baltes (2001) define cognitive performance as the skills and capacities of our brain to function and utilize the information it receives through the five senses. In recent years, cognitive training (CT) has encountered a significant research boom (Simons et al., 2016). Cognitive training consists of the systematic practice of tasks designed to develop skills such as working memory and attention to transfer them to other tasks and environments. We can distinguish between two types of cognitive training (CT): (I) general domain, which aims to develop essential functions applicable to a variety of tasks, and (II) context-specific, such as perceptual-cognitive skills training using the expert performance approach, which targets cognitive skills in a specific task (such as anticipation in a tennis serve) (Harris, Wilson, & Vine, 2018). Mahncke et al. (2006); Simons et al. (2016) Both types of CT could result in a broad range of benefits for memory, attention, processing speed, fluid intelligence, problem-solving, and learning ability in young and old subjects. Commonly, these cognitive aspects are evaluated using cognitive batteries, which can be defined as a set of tests administered in a grouped fashion and provide the opportunity to assess a wide range of cognitive abilities (Nagahara, Bernot, & Tuszynski, 2010). However, compared to testing in older adults and populations with health problems where the device is intended to improve cognitive function (Harris et al., 2018), devices that assess cognitive performance or are used for CT have undergone minimal direct testing in athletes and other healthy populations. These findings are essential for determining the overall efficacy of CT devices, but generalizing them to athletes is exceedingly difficult.

[1]Performance and Sport Rehabilitation Laboratory (DEPORSALUD), Faculty of Sport Sciences, University of Castilla-La Mancha, Toledo, Spain. María Ramírez-delaCruz, MSc, Performance and Sport Rehabilitation Laboratory. Avda. Carlos III s/n. 45071 Toledo, Spain. Email: maria.ramirez@uclm.es. Orcid: https://orcid.org/0000-0002-7898-1254

[2]Faculty of Health Sciences, Universidad Francisco de Vitoria, Ctra. Pozuelo-Majadahonda km 1,800, 28223 Pozuelo de Alarcón, Spain. Alfredo Bravo-Sánchez, PT, PhD, Faculty of Health Sciences, Universidad Francisco de Vitoria, Ctra. Pozuelo-Majadahonda km 1,800, 28223 Pozuelo de Alarcón, Spain. Email: alfredo.bravo@ufv.es. Orcid: https://orcid.org/0000-0002-0476-066X

[3]Performance and Sport Rehabilitation Laboratory (DEPORSALUD), Faculty of Sport Sciences, University of Castilla-La Mancha, Toledo, Spain. Paula Esteban-García, PhD, Performance and Sport Rehabilitation Laboratory. Avda. Carlos III s/n. 45071 Toledo, Spain. Email: paula.estebangarc@gmail.com. Orcid: https://orcid.org/0000-0002-7690-9734

[4]Performance and Sport Rehabilitation Laboratory (DEPORSALUD), Faculty of Sport Sciences, University of Castilla-La Mancha, Toledo, Spain. Fernando Jiménez, MD, PhD, Performance and Sport Rehabilitation Laboratory. Avda. Carlos III s/n. 45071 Toledo, Spain. Email: josefernando.jimenez@uclm.es. Orcid: https://orcid.org/0000-0002-5407-668X

[5]Performance and Sport Rehabilitation Laboratory (DEPORSALUD), Faculty of Sport Sciences, University of Castilla-La Mancha, Toledo, Spain. Javier Abián-Vicen, PT, PhD, Performance and Sport Rehabilitation Laboratory. Avda. Carlos III s/n. 45071 Toledo, Email: Spain. javier.abian@uclm.es. Orcid: https://orcid.org/0000-0001-9635-3289

***Corresponding author:** Address all correspondence to Dr Abián-Vicén at: javier.abian@uclm.es

Sports research has evaluated cognitive performance using visual, timing, signal detection, recall, and complex decision tasks. This type of study aimed to assess the contribution of various perceptual and cognitive factors in predicting the performance of expert athletes (Helsen & Starkes, 1999). Cognitive skills such as selective attention (Abernethy, 1989), divided attention (Memmert, 2009), and working memory (Furley & Memmert, 2010), among others, are required for optimal performance in any discipline. Recent research, such as that conducted by Ducrocq et al. (2017), indicates that training in these cognitive functions can be transferred to sports and enhance athletic performance. For example, this study demonstrated that a training program that included a customized attentional task that targeted the inhibition function of working memory enhanced volley striking in recreational tennis players.

Most research on athletes has been restricted to evaluating cognitive performance in a single sport and, consequently, has focused on a limited number of variables. In addition, traditional methods have measured cognitive function using procedures administered by trained technicians in clinical and laboratory contexts, typically on a single occasion (Sliwinski et al., 2018). The artificial nature of these testing environments and sources of intrapersonal variability may have a negative impact on the validity and reliability of cognitive assessment tests and instruments (Allard et al., 2014; Timmers et al., 2014).

This issue and the fact that few studies evaluate cognitive variables of sport in general and the absence of these evaluations in a healthy and/or youthful population makes our study necessary for conducting a thorough cognitive evaluation. Furthermore, since reliability and validity are analyzed, it is possible to determine which instruments or tests are most appropriate in each situation. Consequently, this methodological study evaluated the validity of cognitive difficulties related to athletic performance in young, healthy individuals.

2. Methods

2.1 Participants

Participating in the survey were seventeen healthy and physically active volunteers (14 men and 3 women; age: 21.4 ±1.6 years; weight: 74.17 ±11.89 kg; and height: 174.4 ±8.2 cm). All of them were students of the Degree in Physical Activity and Sport Sciences and signed the Consent of Participation voluntarily after reading the study's Information Letter. The inclusion criteria were legal age (over 18) and the absence of any known active disease or pathology. Participants with severe injuries and/or colorblindness were disqualified from the study. The local ethics committee approved this research, and all procedures adhered to the Helsinki Declaration.

2.2 Study Design and Procedure

The intraday and interday reliability of a battery of cognitive tests related to sports performance was determined using repeated measures. Each participant visited the Laboratory four times (Familiarization, Series 1, Series 2, and Series 3) throughout three sessions (Figure 1). Following previous research (Bernecké, Pukėnas, & Brazaitis, 2016; Schmidt, Germano, & Milani, 2015), the breaks between assessments on the same day were one hour (intraday reliability) and 48 hours (interday reliability) between measurement days.

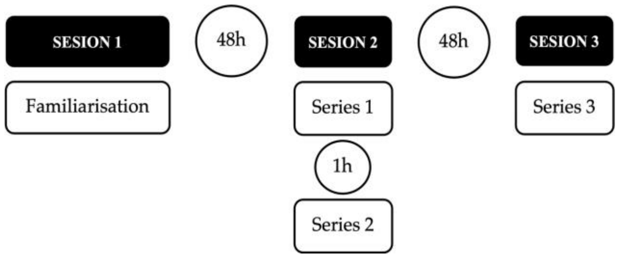

Figure 1. Study design.

The initial session aimed to familiarize participants with the study instruments and examinations. Each cognitive test was practiced by the participants with explanations and corrections. Stewart, Turner, and Miller (2014) Participants were required to attend each subsequent day under identical conditions of rest and nutrition. During the second and third sessions, the participants completed five cognitive tests in the same order described below.

Dynavision™ D2 Visuomotor Device

Reigal et al. (2019) cite the Dynavision™ D2 as one of the most popular instruments for training and evaluating reaction time (RT). It comprises a 1.21 1.21-meter board with 64 target buttons arranged in five concentric circles surrounding a central screen. This display can additionally be illuminated to invigorate the participant. The accompanying software measures RT in seconds (with an accuracy of 1/100 of a second). Before beginning, the participant should be positioned with the central screen at eye level and examined to ensure they are comfortable and can press all buttons. The "Reaction Time" and "Red and Green Lights" tests were selected for this study because they are two of the most frequently utilized tests in studies that evaluated RT (Bigsby et al., 2014; Hunzinger et al., 2020; Wells et al., 2014; Wells & Johnson, 2022). These two DynavisionTM D2 experiments are described in detail below:

a) “Reaction Time”:

There are six assessments in various areas of the board. The participant must press the "home button" and, when a new red light appears on the board, quickly release it to press the new red light. Each exam is completed three times with the dominant hand only. The fastest performance in each of the six tests ("Fastest") is analyzed for each variable. Variables obtained from the test include:

- Visual reaction time ("Visual RT"): The time elapsed between the appearance of the new red light and the actuation of the home button.
- Motor reaction time ("Motor RT"): The time between the discharge of the home button and the pressing of the new red light.
- Physical reaction time ("Physical RT"): The time elapsed between the appearance of the new red light and its activation.

b) “Red and Green Lights”:

Red and green lights should illuminate the controls for 30 seconds. Only the red lights must be pressed before they vanish in 0.75 seconds, while the green lights must be disregarded/not pressed. It is performed with both hands, and only one attempt is allowed. Variables obtained from the test include:

- Reaction Time (RT): The time required to activate the red lights. We have analyzed the quickest time ("Fastest") and the average time ("Average") within this variable.
- Red lights ("Red"): The number (N) of effective hits ("Hits") and the percentage of hits ("Percentage") are examined.

Bassin Anticipation Timer

This tool evaluates visual acuity, eye-hand coordination, and anticipation. It comprises a 1.54-meter horizontal beam with 33 LEDs (Tallis et al., 2013). There were three iterations of each of the following conditions or speeds: 5, 10, and 15 miles per hour (mph), 8, 16, and 24 kilometers per hour. This protocol was selected based on positive findings from previous studies (Brady, 1996; Kuan et al., 2018; Wrisberg, Hardy, & Beitel, 1982). The response time and whether it was early or late were recorded for each trial. The delay signal (visual warning system) was set randomly with the timer to reduce the likelihood of the participant timing the trial internally. The participant was instructed to stand in front of the bar with the push button in their dominant hand and press it as near as possible to the target (light 33). The best of three attempts for each condition/speed was selected for data analysis.

Go/No-Go Test

The Go/No-Go test evaluates sustained attention and response control and the capacity to inhibit an inappropriate response. This test was administered on a computer using the online software "Cognitive fun" (2008), which has been utilized in recent years by several studies to assess cognitive function (Thomas et al., 2019; van Campen et al., 2020). Votruba and Langenecker (2013) found moderate to strong correlations between Go/No-Go measures and measures of psychomotor speed, visuomotor coordination, interference control, inhibition, and set-shifting. Times of reaction were measured in milliseconds. For each trial, an image or stimulus is displayed on the screen. If the image is a green circular shape, it corresponds to a "go" stimulus. If an engraved green circle appears, this indicates a "no-go" stimulus. Participants must press the space bar on the keyboard with their dominant hand as quickly as possible in response to a "go" stimulus and refrain from pressing it in response to a "no-go" stimulus. We obtained the following variables for data analysis:

1. Fastest reaction time (“*Fastest*”) for a correct answer.
2. The average reaction time (“*Average*”) during the test.

Eriksen Flanker Test

This examination was also administered using the online application "Cognitive fun" (2008). Thomas et al. (2019) measure attentional capacity, specifically focused attention and the ability to inhibit engagement from distractions. The participant must strike the “←” or “→” key on the computer keyboard with their dominant hand as quickly as possible, depending on the direction the target arrow is pointing. This target arrow will be located in the center of the screen and surrounded by other arrows whose orientation is either congruent (same direction) or incongruent (opposite direction). All indicators are horizontally positioned and point to the left or right. The software itself logs response times in milliseconds. These were the variables obtained from this test:

1. Congruent Response Time (CRT): Time is taken to press the correct arrow when all arrows are in the same direction.
2. Incongruent Response Time (IRT): Time is taken to press the correct arrow when the arrows are going in the opposite direction to the target arrow.

Trail Making Test (TMT)

This test measures visual scanning, working memory, and cognitive flexibility. According to Smith et al. (2008), TMT results have sufficient validity for assessing these factors. The participant must have a pen, a stopwatch, and a printed page with the TMT instructions. The TMT has two components:

This examination measures memory and processing efficiency. It involves connecting consecutively and as rapidly as possible 25 numbered circles from 1 to 25.

b) TMT-B: measures cognitive adaptability and executive function. In this case, the circles contain the numerals 1 through 13 and the letters A through L. Therefore, the participant must connect the circles sequentially, alternating numbers and letters.
The obtained variable will be the time in seconds required to conclude the examination. The time includes correcting errors, so the evaluator must be vigilant during the examination and alert the student of the error without halting the stopwatch.

2.3 Statistical analysis

IBM SPSS Statistic 24.0 (SPSS, Chicago, Illinois, United States) was used for statistical analysis. All information was expressed as the mean, standard deviation (SD). Normality was determined using the Shapiro-Wilk test, and variables with p-values greater than 0.05 were deemed to have a normal distribution. All variables examined had normal distributions. In addition, a paired samples t-test was conducted to determine the difference in mean between intraday and interday values. Using Intraclass Correlation Coefficients (ICC), intraday (Series 1 versus Series 2) and interday (Series 1 versus Series 3) reliability were assessed (Hopkins, 2000). Koo and Li (2016) established the following reliability thresholds for ICC values: poor (>0.50), moderate (0.50-0.75), good (0.75-0.90), and outstanding (>0.90). In addition, Bland-Altman plots were used to determine the interday and intraday reliability of the most representative variables of each instrument or cognitive test. The following equations were used to calculate the standard error of measurement (SEM) and the minimum detectable change (MDC): SEM = SDdiff √/2, while MDC = SEM x 1.96 x 2 (Weir, 2005). SDdiff is the standard deviation of the differences between them. Weir (2005) Only differences between two measurements that surpass the MDC represent an actual (error-free) change in a subject's scores. The level of significance was p 0.05.

3. Results

Paired samples t-test analyses showed no significant differences between intraday and interday sessions for most of the test variables analyzed ($p > 0.05$). Significant intraday differences were only found in the variables TMT-A ($p = 0.001$) and TMT-B ($p = 0.014$). In the interday comparison, significant differences were found in the TMT-A ($p = 0.001$) and in the following variables of the "Reaction Time" test: "Physical RT - Fastest" ($p < 0.001$) of Test 1; in the "Visual RT - Fastest" ($p = 0.036$) and "Physical RT - Fastest" ($p < 0.001$) of Test 2; and in the "Visual RT - Fastest" ($p = 0.004$) of Test 4 (Tables 1 and 2).
Overall, the intraday and interday reliability values were moderate for the cognitive tests performed. Of the 31 variables studied in the intraday analysis, 11 have a poor ICC (<0.50), 13 have a moderate ICC (0.50-0.75), 6 variables are considered good (0.75-0.90), and 1 variable has an excellent ICC (>0.90). In the results of the interday analysis, 11 variables have a poor ICC, 15 have a moderate ICC, and 5 have a good ICC out of a total of 31 variables studied (Tables 1 and 2).

Table 1

Results of the Dynavision™ D2 Instrument (Mean ± SD)

Test		Variables		Series 1	Series 2	Series 3	P_{1-2}	P_{1-3}	ICC_{1-2} (95%CI)	ICC_{1-3} (95%CI)	MDC_{1-2}	MDC_{1-3}
Dynavision™ D2 Visuomotor Device	*"Reaction Time"* Test 1	*"Visual RT"*	*"Fastest"* (s)	0.29±0.04	0.28±0.03	0.30±0.05	0.073	0.525	0.63 (-0.01; 0.02)	0.70 (-0.01;0.02)	0.05	0.06
		"Motor RT"	*"Fastest"* (s)	0.16±0.04	0.15±0.05	0.14±0.04	0.613	0.345	0.07 (-0.02;0.04)	0.40 (-0.04;0.01)	0.13	0.10
		"Physical RT"	*"Fastest"* (s)	0.48±0.06	0.46±0.06	0.46±0.05	0.253	<0.001*	0.41 (-0.01;0.05)	0.57 (0.27;0.33)	0.12	0.10
	"Reaction Time" Test 2	*"Visual RT"*	*"Fastest"* (s)	0.28±0.02	0.29±0.03	0.29±0.02	0.096	0.036*	0.59 (-0.02;0.01)	0.46 (-0.02;-0.01)	0.04	0.04
		"Motor RT"	*"Fastest"* (s)	0.13±0.05	0.16±0.05	0.16±0.05	0.134	0.783	0.19 -0.06;0.01)	0.38 (-0.04;0.03)	0.13	0.11

Test	Variables		Series 1	Series 2	Series 3	P_{1-2}	P_{1-3}	ICC_{1-2} (95%CI)	ICC_{1-3} (95%CI)	MDC_{1-2}	MDC_{1-3}
	“Physical RT	*“Fastest”* (s)	0.46±0.06	0.47±0.05	0.47±0.05	0.327	<0.001*	0.53 (-0.04;0.01)	0.45 (0.28;0.33)	0.11	0.10
“Reaction Time” Test 3	*“Visual RT”*	*“Fastest”* (s)	0.33±0.04	0.31±0.04	0.32±0.03	0.228	0.642	0.70 (-0.01;0.04)	0.32 (-0.02;0.03)	0.06	0.08
	“Motor RT”	*“Fastest”* (s)	0.21±0.03	0.20±0.05	0.19±0.04	0.318	0.231	0.45 (-0.01;0.03)	0.47 (-0.01;0.04)	0.08	0.09
	“Physical RT”	*“Fastest”* (s)	0.54±0.05	0.53±0.06	0.54±0.07	0.585	0.948	0.35 (-0.2;0.04)	0.77 (-0.02;0.02)	0.13	0.08
“Reaction Time” Test 4	*“Visual RT”*	*“Fastest”* (s)	0.33±0.04	0.34±0.06	0.35±0.03	0.206	0.004*	0.63 (-0.03;0.01)	0.74 (-0.03;-0.01)	0.08	0.05
	“Motor RT”	*“Fastest”* (s)	0.18±0.05	0.18±0.06	0.19±0.10	0.862	0.449	0.38 (-0.03;0.03)	0.47 (-0.04;0.02)	0.12	0.11
	“Physical RT”	*“Fastest”* (s)	0.54±0.06	0.57±0.11	0.58±0.12	0.157	0.155	0.43 (-0.09;0.01)	0.47 (-0.05;0.01)	0.18	0.13
“Reaction Time” Test 5	*“Visual RT”*	*“Fastest”* (s)	0.26±0.04	0.26±0.03	0.26±0.03	0.832	0.542	0.76 (-0.01;0.01)	0.68 (-0.01;0.01)	0.05	0.05
	“Motor RT”	*“Fastest”* (s)	0.14±0.03	0.14±0.04	0.14±0.04	0.381	0.912	0.62 (-0.01;0.02)	0.81 (-0.01;0.01)	0.06	0.05
	“Physical RT”	*“Fastest”* (s)	0.42±0.06	0.42±0.06	0.41±0.06	1.000	0.747	0.80 (-0.02;0.02)	0.86 (-0.01;0.01)	0.07	0.06
“Reaction Time” Test 6	*“Visual RT”*	*“Fastest”* (s)	0.27±0.03	0.26±0.03	0.26±0.03	0.136	0.332	0.63 (-0.01;0.02)	0.60 (-0.01;0.02)	0.06	0.05
	“Motor RT”	*“Fastest”* (s)	0.13±0.02	0.14±0.03	0.13±0.05	0.263	0.760	0.83 (-0.01;0.01)	0.23 (-0.02;0.02)	0.03	0.09
	“Physical RT”	*“Fastest”* (s)	0.40±0.04	0.42±0.05	0.41±0.06	0.156	0.705	0.71 (-0.03;0.01)	0.59 (-0.03;0.02)	0.07	0.10
“Red and Green Lights”	*“Reaction Time”*	*“Fastest”* (s)	0.43±0.05	0.43±0.05	0.43±0.03	0.754	1.000	0.38 (-0.02;0.03)	0.36 (-0.02;0.02)	0.11	0.09
		“Average” (s)	0.56±0.03	0.56±0.03	0.56±0.03	1.000	0.725	0.76 (-0.01;0.01)	0.71 (-0.01;0.01)	0.05	0.05
	“Red”	*“Hits”* (N)	21.47±6.37	21±5.79	20.64±7.97	0.936	0.849	0.50 (-3.14;3.39)	0.66 (-3.15;3.76)	11.84	11.02
		“Percentage” (%)	68.76±12.26	61.92±18.70	68.01±13.52	0.232	0.792	0.54 (-3.61;13.61)	0.59 (-5.27;6.80)	29.48	22.67

* p<0.05 t-test for paired samples.

Table 2

Results of Bassin Anticipation Timer, Go/No-Go, Eriksen Flanker Test, and TMT (Mean ± SD).

Test	Variables	Series 1	Series 2	Series 3	P_{1-2}	P_{1-3}	ICC_{1-2} (95%CI)	ICC_{1-3} (95%CI)	MDC_{1-2}	MDC_{1-3}
Bassin Anticipation Timer	5 mph (s)	0.029±0.023	0.021±0.014	0.021±0.020	0.128	0.303	0.26 (-0.01;0.02)	0.03 (-0.01;0.02)	0.05	0.06
	10 mph (s)	0.022±0.018	0.021±0.024	0.017±0.013	0.820	0.164	0.67 (-0.01;0.01)	0.66 (-0.01;0.01)	0.03	0.03
	15 mph (s)	0.021±0.030	0.025±0.031	0.017±0.032	0.670	0.321	0.94 (-0.01;0.01)	0.89 (-0.01;0.01)	0.02	0.03
Go/No-Go	*"Fastest"* (ms)	318.76±28.52	325.93±42.61	335.64±34.63	0.506	0.350	0.48 (-25.81;13.31)	0.52 (-11.01;29.26)	71.22	73.53
	"Average" (ms)	388.27±41.86	385.76±33.96	391.98±35.67	0.784	0.519	0.62 (-15.53;20.21)	0.55 (-12.15;23.08)	64.74	64.06
Eriksen Flanker Test	CRT (ms)	441.97±45.89	421.72±40.57	456.01±76.15	0.085	0.261	0.63 (-2.71;37.29)	0.69 (-11.53;39.61)	74.50	96.63
	IRT (ms)	493.14±66.46	473.24±49.27	485.13±46.82	0.131	0.803	0.85 (-4.29;30.08)	0.70 (-19.73;24.97)	63.18	71.82
TMT	TMT-A (s)	28.94±9.90	22.42±7.30	16.72±5.79	0.001*	0.001*	0.76 (3.18;10.20)	0.76 (2.66;8.17)	12.53	9.63
	TMT-B (s)	53.28±26.85	39.27±10.49	51.14±16.96	0.014*	0.231	0.41 (3.60;27.57)	0.67 (-17.05;4.51)	45.97	36.04

* $p<0.05$ t-test for paired samples.

The Bland-Altman plots (Figures 2 and 3) do not show a high systematic bias, as most data points are distributed close to the mean and within confidence limits.

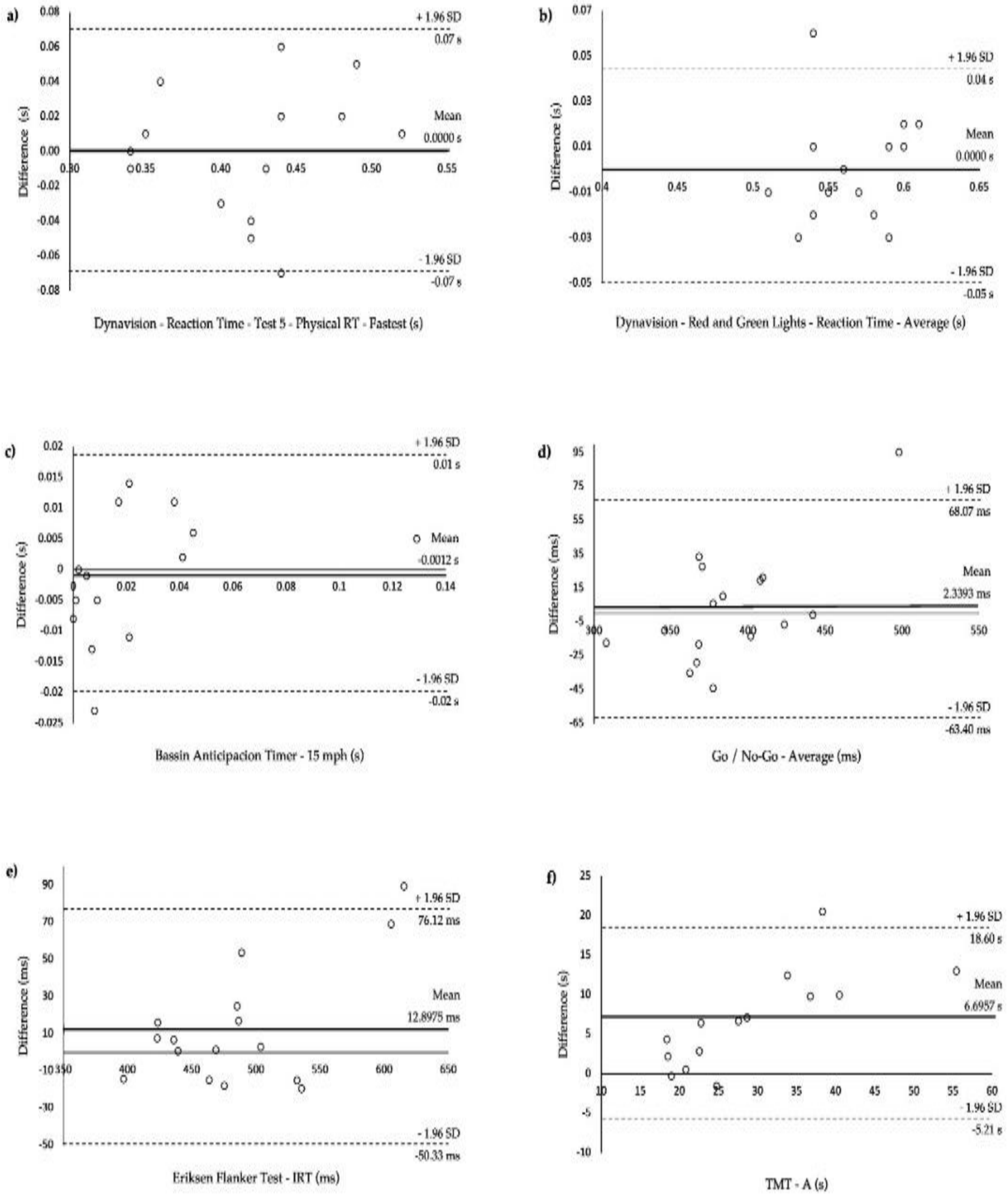

***Figure* 2.** Bland-Altman plots for the intraday reliability of the most representative variables of each cognitive test: (a) Intraday reliability of the variable "Physical RT – Fastest" measured with the DynavisionTM D2 instrument of the Reaction Time (Test 5), (b) Intraday reliability of the variable "Reaction Time - Average" measured with the DynavisionTM D2 instrument of the Red and Green Lights, (c) Intraday reliability of the variable "15 mph" measured with the Bassin Anticipation Timer instrument, (d) Intraday reliability of the variable "Average" measured with Go/No-Go test, (e) Intraday reliability of the variable "IRT" measured with Eriksen Flanker Test, and (f) Intraday reliability of the variable "TMT-A" measured with TMT.

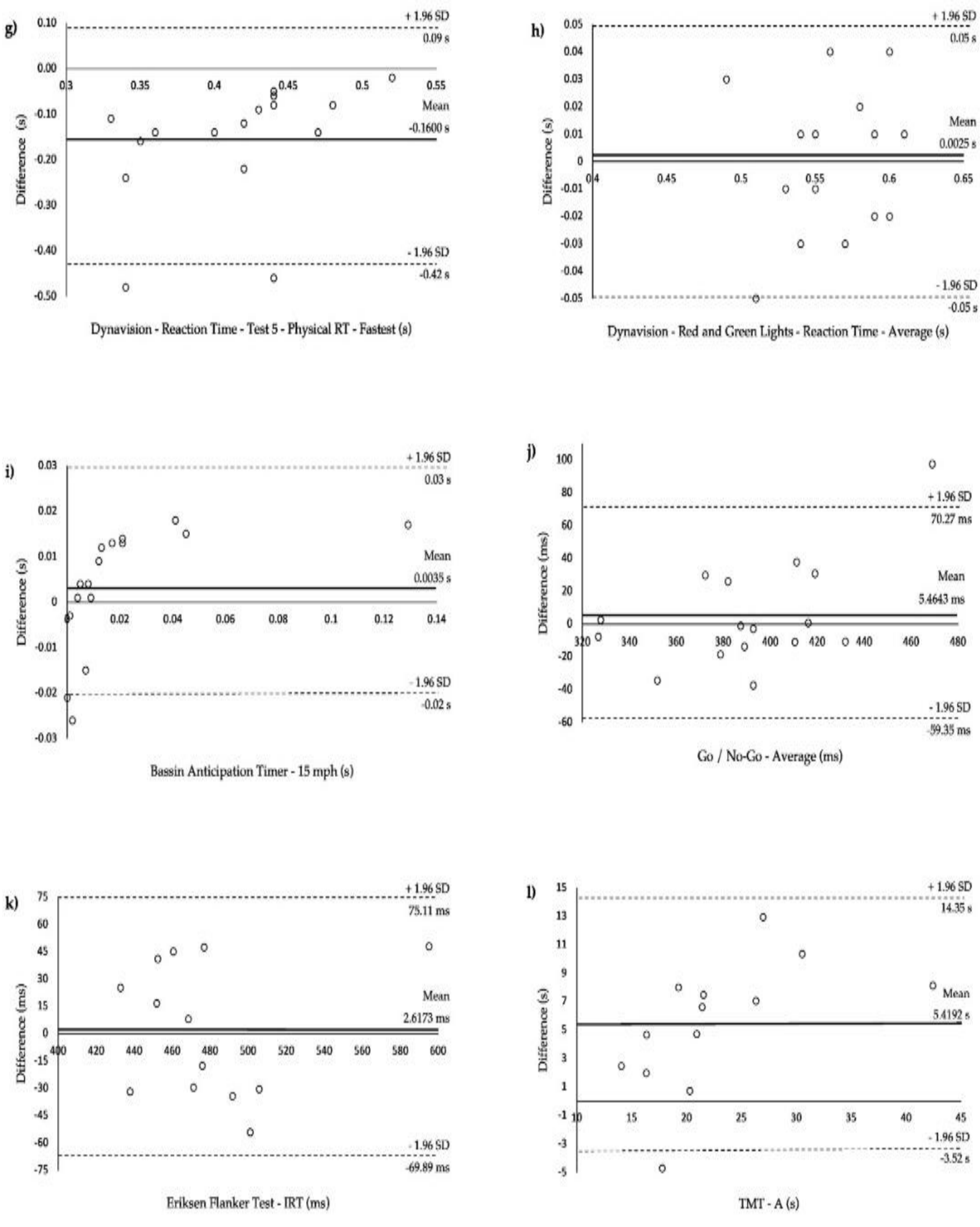

Figure 3. Bland-Altman plots for the interday reliability of the most representative variables of each cognitive test: (a) Interday reliability of the variable "Physical RT – Fastest" measured with the DynavisionTM D2 instrument of the Reaction Time (Test 5), (b) Interday reliability of the variable "Reaction Time - Average" measured with the DynavisionTM D2 instrument of the Red and Green Lights, (c) Interday reliability of the variable "15 mph" measured with the Bassin Anticipation Timer instrument, (d) Interday reliability of the variable "Average" measured with Go/No-Go test, (e) Interday reliability of the variable "IRT" measured with Eriksen Flanker Test, and (f) Interday reliability of the variable "TMT-A" measured with TMT.

4. Discussion

This study demonstrates the viability of a battery of cognitive tests to evaluate and train various cognitive aspects of sports performance. Intraday and interday reliability analyses were conducted, yielding satisfactory reproducibility in the DynavisionTM D2 for the "Average," "Hits," and "Percentage" variables in the "Red and Green Lights" test. The DynavisionTM D2 "Reaction Time" test yielded contradictory results across its six tests, indicating that Tests 5 and 6 are more reliable than the other tests. The Bassin Anticipation Timer instrument demonstrated a distinct tendency to be more reproducible as the condition's speed increased. The computer-based assessments (Go/No-Go and Eriksen Flanker Test) were deemed moderate reliability, with the IRT variable standing out as particularly significant. The TMT-A test was found to be highly reliable, whereas the TMT-B test was found to be less reproducible.

The DynavisionTM D2 Visuomotor Device offers an assortment of options and tests. Numerous studies have demonstrated that this instrument is a valid method for evaluating the reaction times of juvenile students (Bigsby et al., 2014; Hunzinger et al., 2020; Wells et al., 2014; Wells & Johnson, 2022). However, the "Reaction Time" test results are seen as contradictory and highly variable. Therefore, it is suggested that Tests 5 and 6 of the "Reaction Time" test evaluate all variables except for "Motor TR" in Test 6's intraday. Only Test 3 is recommended for assessing the "Physical TR - Fastest" interday. The most significant finding was the low reliability of the "Motor TR - Fastest" test, for which all results of Tests 1, 2, 3, and 4 were deemed inadequate. Good performance on Tests 5 and 6 may be attributable to the test's relative simplicity compared to Tests 1 through 4. In addition, some participants did not perform well on the test due to errors such as not observing the light when covering it with their arm, not comprehending where the lights could emerge, and glancing elsewhere on the board. These errors, particularly in the initial runs, may have contributed to the lack of dependability in Tests 1-4. We believe that familiarization could have been a crucial factor. In addition, we believe that improved outcomes could have been achieved with greater familiarization.

Good reliability results were discovered for the "Red and Green Lights" test, except for the "Reaction Time - Fastest" variable, whose results were deemed unreliable in intraday and interday analyses. Therefore, we concur with the findings of Wells et al. (2014), and we advise using this test to evaluate RT ("Reaction Time - Average"), as this variable had the highest reliability values. Similarly, assessing the "Hits" and "Percentage" variables is suggested, which yielded intermediate results. Therefore, we concur with previous research indicating that the DynavisionTM D2 has no "practice effect" after 3 trials and that a baseline score can be obtained after only one trial of the "Red and Green Lights" test, whereas 2-3 familiarisation trials are required for the "Reaction Time" test (Klavora, Gaskovski, & Forsyth, 1995; Vesia et al., 2008; Wells et al., 2014).

As with the previous instrument, the Bassin Anticipation Timer measures anticipation time under various conditions or rates. Three repetitions of three speeds (5, 10, and 15 mph) were conducted in our investigation. Higher velocities were associated with higher ICC values in both intraday and interday analyses. These ICC values were considered low for the 5-mph speed, moderate for the 10-mph speed, and acceptable and excellent for the 15-mph condition. Our findings concur with those of Coker (2004), Williams, Katene, and Fleming (2002), and Williams (2000), who found that higher speeds resulted in more precise measurements. Other studies, however, found the opposite (Coker, 2005; Williams, 1985). Therefore, we believe the assessment protocol should be a fundamental aspect of this examination. Some authors emphasize the need to substantiate the "learning effect" when analyzing test-retest reliability, as differences between measurement instants may be attributable to exposure to the same instrument over a brief period or to familiarity with the test procedures (Crocetta et al., 2019). With the slower speed, they anticipated too much, whereas, with the faster speed, they had "rehearsed" by accruing 6 repetitions (3 at 5 mph and 3 at 10 mph) before beginning the 15-mph condition.

The "Average" variable had the highest reliability values for the Go/No-Go test, yielding moderate ICC values for the intraday and interday analyses. The "Fastest" variable was rated as inadequate for the intraday test and satisfactory for the interday test. According to studies by Williams and Kaufmann (2012) and Votruba and Langenecker (2013), the Go/No-Go test can attain a high level of reliability. The Eriken Flanker Test, on the other hand, was measured using the same software as the initial test and yielded significantly better results than the remaining tests. It has moderate to good ICC values for all of its variables, so it can be recommended as a reliable test for measuring CRT, particularly IRT. Our results are consistent with other research demonstrating the test's high reliability (Hooper et al., 2022; Wöstmann et al., 2013). Following the previous study, we concur that response times for incongruent versus congruent stimuli were slower in the Eriksen Flanker task (Konishi et al., 2017; Ludyga et al., 2019; O'Leary et al., 2011).

It should be noted, however, that numerous studies have used both tasks (Go/No-Go and Eriksen Flanker) for their measurements, but not all of them in the same manner, as different software, instruments, and protocols are employed. In addition, the website we used, "Cognitive fun" (2008), has been used in recent years by several studies to assess cognitive function (Thomas et al., 2019; van Campen et al., 2020), but we were unable to locate any evidence demonstrating its reproducibility. Therefore, based on our findings, we can recommend both assessments; however, it would be beneficial for future research to use a larger sample size and to establish correlations between measurements of other versions of the same tasks.

Finally, significant differences were detected between intraday and interday measurements for the TMT. This may be because the participants mastered the test, improving their performance with each trial ("learning effect"). The TMT is a reliable test, with the TMT-A variant being more reproducible than the TMT-B variant, as indicated by the reliability results. However, the literature contains some inconsistencies. For some authors, it is a reliable and reproducible test in all versions (Makizako et al., 2013; Wagner et al., 2011), whereas, contrary to our findings, TMT-B has more stable reliability than Part A (Dikmen et al., 1999; Levine et al., 2004). Accordingly, our results confirm the pattern in psychometrics, in which reliability tends to increase as the number of trials increases (Wöstmann et al., 2013). In addition, participants found the TMT-B test to be more challenging than the TMT-A test, as confirmed by Gaudino, Geisler, and Squires (1995) in their analysis of the test properties. This increased complexity of the TMT-B, combined with the small number of B tests discovered, may have resulted in the TMT-B exhibiting a daily variation in complexity. Perhaps using the same B-test for each assessment would have been preferable, but we wished to avoid the previously discussed repetition learning.

5. Conclusion

We can affirm that the battery of cognitive tests presented in our study is reliable for training various sport-related cognitive aspects. We recommend measuring RT with the "Reaction Time" (Tests 5 and 6) and "Red and Green Lights" tests of the DynavisionTM D2 instrument. From the Bassin Anticipation Timer device, measuring the anticipation time at 15 mph is recommended, though it is suggested to modify the measurement protocol. It is claimed that the Go/No-Go and Eriksen Flanker tests are reproducible with the software in use. Finally, the TMT test is considered reliable, although the TMT-A is recommended due to the limited availability of TMT-B versions.

This study provides some practical implications and substantial information on tests that could be used to assess or train cognitive aspects of sports in general. These tests have been categorized, and a reliable battery of cognitive tests comprising various cognitive aspects crucial for athletic performance has been developed. In addition, it has been determined that the employed instruments and protocols are suitable for measuring and training these cognitive variables. And due to the study sample, a small number of reference values for a less-studied population, such as young and healthy individuals, are provided. Future research must therefore examine this cognitive battery with a larger sample size and provide reference values that can be contrasted with other populations (e.g., older individuals, individuals with intellectual disabilities...). This cognitive battery is also recommended for the training and improvement of cognitive performance in a variety of sports, as well as the study of cognitive performance under various conditions, such as exposure to different heights, temperatures, ingestion of ergogenic aids... etc., to obtain more pertinent information for the improvement of sports performance.

References

Abernethy, B. (1989). Expert-novice differences in perception: How expert does the expert have to be? *Canadian Journal of Sport Sciences, 14*(1), 27–30. https://pubmed.ncbi.nlm.nih.gov/2924219

Allard, M., Husky, M., Catheline, G., Pelletier, A., Dilharreguy, B., Amieva, H., Pérès, K., Foubert-Samier, A., Dartigues, J.-F., & Swendsen, J. (2014). Mobile Technologies in the Early Detection of Cognitive Decline. *PLOS ONE, 9*(12), e112197. https://doi.org/10.1371/journal.pone.0112197

Bernecké, V., Pukėnas, K., & Brazaitis, M. (2016). Sex differences in reliability of tests to assess cognitive function. *Baltic journal of sport and health sciences, 2*(101), 17-25. http://dx.doi.org/10.33607/bjshs.v2i101.57

Bigsby, K., Mangine, R. E., Clark, J. F., Rauch, J. T., Bixenmann, B., Susaret, A. W., Hasselfeld, K. A., & Colosimo, A. J. (2014). Effects of postural control manipulation on visuomotor training performance: comparative data in healthy athletes. *International Journal of Sports Physical Therapy, 9*(4), 436-446. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4127506

Brady, F. (1996). Anticipation of Coincidence, Gender, and Sports Classification. *Perceptual and Motor Skills, 82*(1), 227-239. https://doi.org/10.2466/pms.1996.82.1.227

Cognitive fun. (2008). *cognitivefun.net*. https://new.cognitivefun.net/

Coker, C. A. (2004). Bilateral Symmetry in Coincident Timing: A Preliminary Investigation. *Perceptual and Motor Skills, 98*(1), 359-365. https://doi.org/10.2466/pms.98.1.359-365

Coker, C. A. (2005). Direction of an Approaching Stimulus on Coincident Timing Performance of a Ballistic Striking Task. *Perceptual and Motor Skills, 100*(3), 859-860. https://doi.org/10.2466/pms.100.3.859-860

Crocetta, T. B., Guarnieri, R., Massetti, T., Silva, T. D. d., de Almeida Barbosa, R. T., Ferreira de Lima Antão, J. Y., Abreu, L. C. d., & Monteiro, C. B. d. M. (2019). Concurrent Validity and Reliability of Alternative Computer Game for the Coincidence-Anticipation Timing Task. *Measurement in Physical Education and Exercise Science, 23*(2), 169-185. https://doi.org/10.1080/1091367X.2018.1560297

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test–retest reliability and practice effects of Expanded Halstead–Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society, 5*(4), 346-356. https://doi.org/10.1017/S1355617799544056

Ducrocq, E., Wilson, M., Smith, T. J., & Derakshan, N. (2017). Adaptive working memory training reduces the negative impact of anxiety on competitive motor performance. *Journal of Sport and Exercise Psychology, 39*(6), 412-422. https://doi.org/10.1123/jsep.2017-0217

Furley, P., & Memmert, D. (2010). Differences in Spatial Working Memory as a Function of Team Sports Expertise: The Corsi Block-Tapping Task in Sport Psychological Assessment. *Perceptual and Motor Skills, 110*(3), 801-808. https://doi.org/10.2466/pms.110.3.801-808

Gaudino, E. A., Geisler, M. W., & Squires, N. K. (1995). Construct validity in the trail making test: What makes part B harder? *Journal of Clinical and Experimental Neuropsychology, 17*(4), 529-535. https://doi.org/10.1080/01688639508405143

Harris, D. J., Wilson, M. R., & Vine, S. J. (2018). A Systematic Review of Commercial Cognitive Training Devices: Implications for Use in Sport. *Frontiers in Psychology, 9*, 709. https://doi.org/10.3389/fpsyg.2018.00709

Helsen, W. F., & Starkes, J. L. (1999). A multidimensional approach to skilled perception and performance in sport. *Applied Cognitive Psychology, 13*(1), 1-27. https://doi.org/10.1002/(SICI)1099-0720(199902)13:1<1::AID-ACP540>3.0.CO;2-T

Hooper, B., Faria, L. O., Fortes, L. d. S., Wanner, S. P., & Albuquerque, M. R. (2022). Development and reliability of a test for assessing executive functions during exercise. *Applied Neuropsychology: Adult, 29*(4), 750-760. https://doi.org/10.1080/23279095.2020.1807984

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine, 30*(1), 1-15. https://doi.org/10.2165/00007256-200030010-00001

Hunzinger, K. J., Sanders, E. W., Deal, H. E., Langdon, J. L., Evans, K. M., Clouse, B. A., Munkasy, B. A., & Buckley, T. A. (2020). The use of a visual motor test to identify lingering deficits in concussed collegiate athletes. *Journal of Clinical and Translational Research, 5*(4), 178-185. http://dx.doi.org/10.18053/jctres.05.202004.004

Klavora, P., Gaskovski, P., & Forsyth, R. D. (1995). Test-Retest Reliability of Three Dynavision Tasks. *Perceptual and Motor Skills, 80*(2), 607-610. https://doi.org/10.2466/pms.1995.80.2.607

Konishi, K., Kimura, T., Yuhaku, A., Kurihara, T., Fujimoto, M., Hamaoka, T., & Sanada, K. (2017). Mouth rinsing with a carbohydrate solution attenuates exercise-induced decline in executive function. *Journal of the International Society of Sports Nutrition, 14*(1), 45. https://doi.org/10.1186/s12970-017-0200-0

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine, 15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kuan, Y. M., Zuhairi, N. A., Manan, F. A., Knight, V. F., & Omar, R. (2018). Visual reaction time and visual anticipation time between athletes and non-athletes. *Malaysian Journal of Public Health Medicine, 1*, 135-141. https://www.researchgate.net/publication/324441797

Levine, A. J., Miller, E. N., Becker, J. T., Selnes, O. A., & Cohen, B. A. (2004). Normative Data for Determining Significance of Test–Retest Differences on Eight Common Neuropsychological Instruments. *The Clinical Neuropsychologist, 18*(3), 373-384. https://doi.org/10.1080/1385404049052420

Lindenberger, U., Scherer, H., & Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: Not due to sensory acuity reductions operating during cognitive assessment. *Psychology and Aging, 16*(2), 196-205. https://doi.org/10.1037/0882-7974.16.2.196

Ludyga, S., Pühse, U., Lucchi, S., Marti, J., & Gerber, M. (2019). Immediate and sustained effects of intermittent exercise on inhibitory control and task-related heart rate variability in adolescents. *Journal of Science and Medicine in Sport, 22*(1), 96-100. https://doi.org/10.1016/j.jsams.2018.05.027

Mahncke, H. W., Connor, B. B., Appelman, J., Ahsanuddin, O. N., Hardy, J. L., Wood, R. A., Joyce, N. M., Boniske, T., Atkins, S. M., & Merzenich, M. M. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: A randomized, controlled study. *Proceedings of the National Academy of Sciences, 103*(33), 12523-12528. https://doi.org/10.1073/pnas.0605194103

Makizako, H., Shimada, H., Park, H., Doi, T., Yoshida, D., Uemura, K., Tsutsumimoto, K., & Suzuki, T. (2013). Evaluation of multidimensional neurocognitive function using a tablet personal computer: Test–retest reliability and validity in community-dwelling older adults. *Geriatrics & Gerontology International, 13*(4), 860-866. https://doi.org/10.1111/ggi.12014

Memmert, D. (2009). Pay attention! A review of visual attentional expertise in sport. *International Review of Sport and Exercise Psychology, 2*(2), 119-138. https://doi.org/10.1080/17509840802641372

Nagahara, A. H., Bernot, T., & Tuszynski, M. H. (2010). Age-related cognitive deficits in rhesus monkeys mirror human deficits on an automated test battery. *Neurobiology of Aging, 31*(6), 1020-1031. https://doi.org/10.1016/j.neurobiolaging.2008.07.007

O'Leary, K. C., Pontifex, M. B., Scudder, M. R., Brown, M. L., & Hillman, C. H. (2011). The effects of single bouts of aerobic exercise, exergaming, and videogame play on cognitive control. *Clinical Neurophysiology, 122*(8), 1518-1525. https://doi.org/10.1016/j.clinph.2011.01.049

Reigal, R. E., Barrero, S., Martín, I., Morales-Sánchez, V., Juárez-Ruiz de Mier, R., & Hernández-Mendo, A. (2019). Relationships Between Reaction Time, Selective Attention, Physical Activity, and Physical Fitness in Children. *Frontiers in Psychology, 10*, 2278. https://doi.org/10.3389/fpsyg.2019.02278

Schmidt, D., Germano, A. M. C., & Milani, T. L. (2015). Aspects of Dynamic Balance Responses: Inter- and Intra-Day Reliability. *PLOS ONE, 10*(9), e0136551. https://doi.org/10.1371/journal.pone.0136551

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "Brain-Training" Programs Work? *Psychological Science in the Public Interest, 17*(3), 103-186. https://doi.org/10.1177/1529100616661983

Sliwinski, M. J., Mogle, J. A., Hyun, J., Munoz, E., Smyth, J. M., & Lipton, R. B. (2018). Reliability and Validity of Ambulatory Cognitive Assessments. *Assessment, 25*(1), 14-30. https://doi.org/10.1177/1073191116643164

Smith, S. R., Servesco, A. M., Edwards, J. W., Rahban, R., Barazani, S., Nowinski, L. A., Little, J. A., Blazer, A. L., & Green, J. G. (2008). Exploring the Validity of the Comprehensive Trail Making Test. *The Clinical Neuropsychologist, 22*(3), 507-518. https://doi.org/10.1080/13854040701399269

Stewart, P. F., Turner, A. N., & Miller, S. C. (2014). Reliability, factorial validity, and interrelationships of five commonly used change of direction speed tests. *Scandinavian Journal of Medicine & Science in Sports, 24*(3), 500-506. https://doi.org/10.1111/sms.12019

Tallis, J., Duncan, M. J., Wright, S. L., Eyre, E. L. J., Bryant, E., Langdon, D., & James, R. S. (2013). Assessment of the ergogenic effect of caffeine supplementation on mood, anticipation timing, and muscular strength in older adults. *Physiological Reports, 1*(3), e00072. https://doi.org/10.1002/phy2.72

Thomas, C. J., Rothschild, J., Earnest, C. P., & Blaisdell, A. (2019). The Effects of Energy Drink Consumption on Cognitive and Physical Performance in Elite League of Legends Players. *Sports, 7*(9), 196. https://doi.org/10.3390/sports7090196

Timmers, C., Maeghs, A., Vestjens, M., Bonnemayer, C., Hamers, H., & Blokland, A. (2014). Ambulant Cognitive Assessment Using a Smartphone. *Applied Neuropsychology: Adult, 21*(2), 136-142. https://doi.org/10.1080/09084282.2013.778261

van Campen, C. M. C., Rowe, P. C., Verheugt, F. W. A., & Visser, F. C. (2020). Cognitive Function Declines Following Orthostatic Stress in Adults With Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *Frontiers in Neuroscience, 14*, 688. https://doi.org/10.3389/fnins.2020.00688

Vesia, M., Esposito, J., Prime, S. L., & Klavora, P. (2008). Correlations of Selected Psychomotor and Visuomotor Tests with Initial Dynavision Performance. *Perceptual and Motor Skills, 107*(1), 14-20. https://doi.org/10.2466/pms.107.1.14-20

Votruba, K. L., & Langenecker, S. A. (2013). Factor structure, construct validity, and age- and education-based normative data for the Parametric Go/No-Go Test. *Journal of Clinical and Experimental Neuropsychology, 35*(2), 132-146. https://doi.org/10.1080/13803395.2012.758239

Wagner, S., Helmreich, I., Dahmen, N., Lieb, K., & Tadić, A. (2011). Reliability of Three Alternate Forms of the Trail Making Tests A and B. *Archives of Clinical Neuropsychology, 26*(4), 314-321. https://doi.org/10.1093/arclin/acr024

Weir, J. P. (2005). Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. *The Journal of Strength & Conditioning Research, 19*(1), 231-240. https://doi.org/10.1519/15184.1

Wells, A. J., Hoffman, J. R., Beyer, K. S., Jajtner, A. R., Gonzalez, A. M., Townsend, J. R., Mangine, G. T., Robinson, E. H. t., McCormack, W. P., Fragala, M. S., & Stout, J. R. (2014). Reliability of the dynavision™ d2 for assessing reaction time performance. *Journal of Sports Science and Medicine, 13*(1), 145-150. http://www.jssm.org/researchjssm-13-145.xml.xml

Wells, A. J., & Johnson, B.-a. D. I. (2022). Test–Retest Reliability, Training, and Detraining Effects Associated With the Dynavision D2™ Mode A Visuomotor Reaction Time Test. *Journal of Sport Rehabilitation, 31*(2), 253-261. https://doi.org/10.1123/jsr.2020-0550

Williams, B. J., & Kaufmann, L. M. (2012). Reliability of the Go/No Go Association Task. *Journal of Experimental Social Psychology, 48*(4), 879-891. https://doi.org/10.1016/j.jesp.2012.03.001

Williams, K. (1985). Age Differences on a Coincident Anticipation Task. *Journal of Motor Behavior, 17*(4), 389-410. https://doi.org/10.1080/00222895.1985.10735357

Williams, L. R. T. (2000). Coincidence Timing of a Soccer Pass: Effects of Stimulus Velocity and Movement Distance. *Perceptual and Motor Skills, 91*(1), 39-52. https://doi.org/10.2466/pms.2000.91.1.39

Williams, L. R. T., Katene, W. H., & Fleming, K. (2002). Coincidence Timing of a Tennis Stroke: Effects of Age, Skill Level, Gender, Stimulus Velocity, and Attention Demand. *Research Quarterly for Exercise and Sport, 73*(1), 28-37. https://doi.org/10.1080/02701367.2002.10608989

Wöstmann, N. M., Aichert, D. S., Costa, A., Rubia, K., Möller, H.-J., & Ettinger, U. (2013). Reliability and plasticity of response inhibition and interference control. *Brain and Cognition, 81*(1), 82-94. https://doi.org/10.1016/j.bandc.2012.09.010

Wrisberg, C. A., Hardy, C. J., & Beitel, P. A. (1982). Stimulus Velocity and Movement Distance as Determiners of Movement Velocity and Coincident Timing Accuracy. *Human Factors, 24*(5), 599-608. https://doi.org/10.1177/001872088202400510