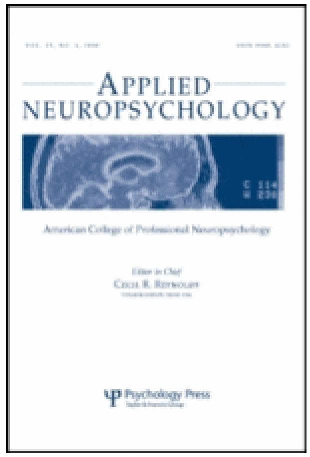
“Encephalapp Stroop”: Validity and reliability of a smartphone app to measure cognitive performance in physically active subjects

Published Version:
Luiz José Frota Solon-Júnior, ., Luiz Vieira Da Silva Neto, ., Alves Araujo De Lima Junior, D.R., Costa, Y.P., José Klinger Da Silva Oliveira, ., Lenamar Fiorese &, et al. (2026). “Encephalapp Stroop”: Validity and reliability of a smartphone app to measure cognitive performance in physically active subjects. APPLIED NEUROPSYCHOLOGY. ADULT, 33(1), 219-224 [10.1080/23279095.2024.2343024].

Availability:
This version is available at: https://hdl.handle.net/11585/1017166 since: 2025-06-05

Published:
DOI: http://doi.org/10.1080/23279095.2024.2343024

(Article begins on next page)



"Encephalapp Stroop": validity and reliability of a smartphone app to measure cognitive performance in physically active subjects.

Journal:	*Applied Neuropsychology: Adult*
Manuscript ID	HAPN-2024-0028.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Solon, Luis; Paraiba Federal University Vieira da Silva Neto, Luiz; Universidade Estadual Vale do Acarau Lima-Júnior, Dalton; University of Bologna Costa, Yago; Universidade Federal da Paraíba Klinger da Silva Oliveira, José; Universidade Estadual Vale do Acarau Fiorese, Lenamar; State University of Maringá Fortes, Leonardo; Universidade Federal da Paraíba
Keywords:	brain, neuroscience, app, validity, cognitive effort

Title: "Encephalapp Stroop": validity and reliability of a smartphone app to measure cognitive performance in physically active subjects.

Running Head: app for measuring cognitive performance

Authors: Luiz José Frota Solon-Júnior[1], Luiz Vieira da Silva Neto[2], Dalton de Lima-Junior[3], Yago P. Costa [1], José Klinger da Silva Oliveira[2], Lenamar Fiorese[4], and Leonardo de Sousa Fortes[1].

Affiliations:

[1] Federal University of Paraíba. Department of Physical Education, João Pessoa, Paraíba, Brazil.

[2] Associate Physical Education at University of Vale do Acaraú, Sobral, Ceará, Brazil.

[3] University of Bologna. Department for Life Quality Studies, Rimini, Italy.

[4] Associate Program of Physical Education, State University of Maringá, Maringá, Brazil.

Corresponding author:

Dalton de Lima-Junior

Phone: +39 344 403 8098

E-mail: limajunior.dalton@gmail.com

Title: "Encephalapp Stroop": validity and reliability **of a smartphone app to measure cognitive performance in physically active subjects.**

Abstract

This study aimed to analyze the validity and reliability of a smartphone app (i.e., Encephalapp Stroop) to measure cognitive performance. Ten physically active individuals (mean age: 25.2 ± 4.3 years) participated in three laboratory visits. In a randomized sequence, subjects completed cognitive assessments using the app and a computer separately. The response times from 100 measurements for congruent and incongruent stimuli were compared between the two devices (app and computer) using the intraclass correlation coefficient, Pearson product-moment correlation coefficient (r), Cronbach's alpha (α), and Bland-Altman plots. There was excellent agreement between the computerized Stroop and the app version for response times for congruent (ICC = 0.806, P < 0.001; Bland-Altman bias = 41.3 ± 56.4 ms, P < 0.001) and incongruent stimuli (ICC = 0.755, P < 0.001; Bland-Altman bias = 76.4 ± 75.1 ms, P < 0.001). In comparison to the computerized version, the app demonstrated a moderate correlation for response times of congruent (r = 0.678, P < 0.001) and incongruent (r = 0.623, P < 0.001) stimuli. The results of the present study indicate that the response times of congruent and incongruent stimuli can be easily, accurately, and reliably evaluated using a mobile app.

Keywords: brain, neuroscience, app, validity, cognitive effort.

Introduction

Cognitive tests are frequent scientific practice and research tools are employed across various outcomes, including health (e. g., cognitive assessment in the elderly) (Spaner et al., 2019), disease (e. g., diagnosis of Covert Hepatic Encephalopathy) (Bajaj et al., 2015), and sport (e. g., assessments of the cognitive ability of players.) (Trecroci et al., 2021). Different computerized behavioral tests have been used to assess executive function performance (Diamond, 2013). Notable examples include the Psychomotor Vigilance Task to assess reaction time, the Flanker task to assess cognitive flexibility (Van Cutsem et al., 2020), and the Stroop Color-Word Test to measure inhibitory control and selective attention (Fortes et al., 2019).

Originally developed by American psychologist John Riddly Stroop (Stroop, 1935), the Stroop Task (ST) is one of the oldest and most widely utilized tests in psychology to examine cognitive functioning and response inhibition in different populations, including children, adults, and clinical patients (Schiehser & Bondi, 2010). In the standard form of the task, a participant is asked to (1) respond to both congruent (when the color of a word aligns with the spelling of the word) and (2) incongruent (when the color of a word conflicts with the spelling of the word) stimuli (Schiehser & Bondi, 2010). For incongruent stimuli, response inhibition is required because two processes are in competition (spelling and word color). In this test, the main results utilized as indicators of cognitive performance have been response time and accuracy (Fortes et al., 2021). Additionally, the Stroop effect, measured by the time taken to respond to incongruent stimuli compared to congruent stimuli, is utilized.

Initially, the stroop test was administered with paper (i. e. a name of one color printed in the ink of another color—a word stimulus and a color stimulus) (Stroop, 1935), however, despite its low cost and ease of use, this way of applying the stroop test has limitations, mainly associated with human error (e.g., possible incorrect application and delay in response time) (Henik, 1996; Linnman et al., 2006). For this reason, to overcome the limitations mentioned above, the administration of the ST has evolved from a card presentation format to a computerized version (Dalgleish, 1995; MacLeod, 1991). These computerized versions feature wide-ranging variations, including individual stimulus versions, numeric, spatial, picture versions, auditory, and emotional versions (MacLeod, 1991).

Computerized versions of individual stimuli have been widely utilized in recent years to explore neural correlates in brain imaging studies. These functional imaging surveys usually preclude oral responses from the volunteers and, for this reason, the stroop test with manual responses, such as button presses, has been utilized (Penner et al., 2012). On the other hand, although these computerized versions still have advantages (e.g., consistency of administration and ease of application and interpretation of data), alternatives should be created and tested. In recent years, the use of mobile devices has been ubiquitous in the population. Neuropsychology researchers have already begun experimenting with smartphones as behavioral data collection tools due to their advantages as continuous and refined data collection in naturalistic environments (Harari et al., 2016).

In 2015, researchers developed a mobile app based on the Stroop Test, called 'Encephalapp Stroop' (ES) (Bajaj et al., 2015). The ES has been utilized to diagnose patients with occult hepatic encephalopathy (Bajaj et al., 2015) and predict poor sleep quality in cirrhotic patients (Luo et al., 2020). However, although this app has been utilized to measure attention, psychomotor speed, and cognitive flexibility in healthy older adults (Spaner et al., 2019), no study to date has verified the validity and reliability of the ES when compared to a computerized version (stroop test with manual response), a critical step to ensure data dependability (Hopkins, 2000).

Thus, the current study aimed to assess the validity and reliability of the ES app, a alternative that may offer a cost-effective and convenient method for evaluating cognitive skills in healthy and physically active individuals. This approach can potentially expedite data collection for cognitive assessments across diverse settings (e.g., clinical, sport, and work), enhancing efficiency, and the time needed for data tabulation and analysis.

Materials and methods

Participants

Ten university students healthy and physically active (eight males and two females, mean age 25.2 ± 4.3 years, height 1.72 ± 0.06 m, body mass 73.3 ± 17.1 kg, BMI 24.4 ± 3.9 Kg/m^2) volunteered to participate in this study. The inclusion criteria adopted in the present study were: a) participants aged between 18 and 35 years; b) physically active [i.e., engaged in a minimum of 150 minutes of moderate-intensity physical activity or at least 75 to 150 minutes of vigorous-intensity activity per week (McKay et al., 2022)]; c) non-smokers and non-consumers of alcohol; d) no history of cognitive, neurological, or psychological disorders; e) no self-reported sleep or vision disturbances; f) not taking any medications or supplements during the study.

All participants were physically active (4.4 ± 1.5 training sessions per week with session duration of 68 ± 7 minutes) and engaged in cognitive activities such as studying and working (7.7 ± 2.1 hours per day). They participated in various sports and other activities, including resistance training, beach volleyball, Brazilian Jiu-Jitsu, running, basketball, futsal, and yoga. We classify them as physically active based on the guidelines and recommendations proposed by McKay et al., (2022). We chose physically active subjects because evidence supports the effect of exercise on improving cognition throughout life, so that lack of exercise is considered a risk for the incidence of various neurological disorders (Fernandes et al., 2017). Additionally, this group represents a significant portion of the global population [approximately 35%–42% (McKay et al., 2022)].

Eligible participants voluntarily signed the informed consent, which included information about possible risks and study procedures. This study was approved by the Human Research Ethics Committee of the Vale do Acaraú State University (protocol: 5.321.850).

Procedures

Participants visited the lab thrice, with two to seven days intervals between each visit. During the initial visit, they provided informed consent, underwent anthropometric measurements, and were introduced to cognitive tests, including computerized and app-based Stroop tests. Additionally, subjects received a checklist during this first visit with instructions that included abstaining from alcohol and stimulant consumption (e.g., caffeine) the night before and on the day of the visit, maintaining regular dietary and exercise routines, avoiding social media and cognitive efforts for more than 30 consecutive minutes before data collection, and ensuring 7 to 9 hours of sleep starting between 22:00 and 00:00, with a typical waking time between 06:00 and 08:00.

Following the initial visit, participants were invited to two more visits, one test and another retest. Upon arrival at the laboratory, researchers verified adherence to the checklist recommendations. Then, participants had a 20-minute rest period without engaging in cognitive-demanding tasks (e.g. using smartphones). Subsequently, participants completed the Stroop test five times in a randomized order (www.randomization.com) and alternating between the app (Bajaj et al., 2015) and the computer (Fortes et al., 2021). Each test consisted of 120 trials (60 congruent and 60 incongruent), with a 2-minute rest period in between (see Figure 1a).

All tests were conducted in a quiet room with subdued lighting, the subjects were seated, and only one researcher was present to minimize potential distractions influencing participant responses. Response times for congruent and incongruent stimuli were recorded as the criteria for analysis, with data accuracy below 95% excluded from the analysis. Detailed descriptions of the test protocols are presented in the subsequent sections.

*****Insert Figure 1*****

Cognitive tests

App "Encephalapp Stroop."

The app was developed in 2013 and is very practical (Bajaj et al., 2015). The app's interface includes trials with Stroop "on" (incongruent stimuli) and "off" (congruent stimuli). To administer the test, researchers configured the number of trials according to the app's settings, similar to the computer. For this, in "settings", we insert the number 6 in "test phases" in both "stroop off" and "stroop on" (For more details, see https://apps.apple.com/br/app/encephalapp-stroop-test/id589478721?platform=iphone). Each trial included stimuli that appeared in the text "####" (Stroop "off," that is, congruent stimuli) or words "red, blue, and green" (Stroop "on," that is, incongruent stimuli). In both types of stimuli, the subjects had to touch the screen of an iPhone 8 plus at the location where it matched the color in which the words were written as quickly and accurately as possible. In incongruous trials, the subjects were instructed to touch the button corresponding to the color in which the words were written but not the colors they would signify. For example, if the word 'green' was written in blue, participants should tap the button for the blue color (see Figure 1b). On the other hand, in incongruent trials, the subjects touched the button corresponding to the color of the pound sign (i.e., ####). Finally, data were collected regarding response time (mean time for correct responses, analyzed separately for congruent and incongruent stimuli, i.e., mean of the 60 congruent and 60 incongruent trials). The accuracies (mean percentage ±SD) for congruent and incongruent stimuli were 98.9 ± 1.2 and 98.9 ± 1.4%, respectively.

Computer Stroop Test

The computers utilized for the Stroop test had a 20-inch display (HP Compaq, LA2006x) with a resolution of 1600x900, a 60Hz refresh rate, and an eight-bit intensity. The Stroop test on the computer consisted of 120 trials (Fortes et al., 2021; Gualtieri & Johnson, 2006) to standardize the number of stimuli with the app. Therefore, 60 congruent trials (Stroop "off," with sign #### in different colors) and 60 incongruent trials (Stroop "on," with words in different colors) were presented. In the congruent trials, the subjects answered the color (red, blue, and green) of the sign (####), and in incongruent trials, the subjects answered the color (red, blue, and green) of the words (red, blue, and green) that appeared, inhibiting their spelling. Participants provided motor responses using the computer keyboard: "L" for red, "D" for blue, and "J" for green. It is worth noting that these computer letters were marked with stickers corresponding to the colors. An example stimulus for the color blue is illustrated in Figure 1c. The test was built and applied in the Psychopy3 software version 2021.2.3 (GNU General Public License). The accuracies (mean percentage ±SD) for congruent and incongruent stimuli were 98.2 ± 1.8 and 97.7 ± 2.15 %, respectively.

Statistical analysis

All data were reported as mean ± standard deviation. Paired Student's t-test and the bivariate Pearson product-moment correlation coefficient (r) were applied to assess validity. Additionally, Bland–Altman plots were generated to visualize the agreement between the two instruments (Bland & Altman, 1986). The intraclass correlation coefficient was used to analyze the app's reliability in measuring response times compared to the computer assessment.. Cronbach's alpha (α) was utilized to analyze the app's consistency when measuring response times across five trials for each participant. Finally, an intraclass correlation coefficient (ICC) was used to analyze the app's test-retest reliability between different days. The level of statistical significance was set at $P < 0.05$. All statistical analyses were conducted using IBM SPSS Statistics 22 for Mac (IBM Co., USA).

Results

Validity

The app-based Stroop test and the computer-based version obtained a Pearson product-moment correlation of 100 measures (10 individuals x 5 measures x 2 days) of response times (congruent and incongruent stimuli). Positive and significant correlations were observed between the app and the computer for both congruent ($r = 0.678$, $P < 0.001$) and incongruent stimuli ($r = 0.623$, $P < 0.001$).

Bland-Altman plots were utilized to display the bias and respective 95% limits of agreement for the response time for each of the evaluated stimuli. The mean differences between the instruments for the congruent and incongruent stimuli were 41.3 ± 56.4 (CI: 52.5–30.1) and 76.4 ± 75.1 (CI: 91.3–61.5) milliseconds, respectively. The response times values for stimuli (i.e., congruent and incongruent) in the app (490.2 ± 66.8 e 516.6 ± 74.1, respectively) were significantly lower ($P < 0.001$ e $P < 0.001$, respectively) than those obtained with the computerized version (531.5 ± 73.1 e 595.0 ± 94.0, respectively). However, there was high agreement between the app and the computer in the response time for congruent (ICC = 0.806, IC 95%: 0,712-0,870, $P < 0.001$) and incongruent stimuli (ICC = 0.755, IC 95%: 0.635-0.835, $P < 0.001$).

*****Insert figure 2*****

Internal consistency and test-retest reliability

For the reliability analysis of the five response time measurements obtained from each participant, the app showed good stability (internal consistency) in the test and retest visits for the congruent (Cronbach's α = 0.980 and Cronbach's α = 0.985, respectively) and incongruent (Cronbach's α = 0.987 and Cronbach's α = 0.982, respectively) stimuli. Finally, when analyzing the test-retest reliability of the app, a good agreement was found between test-days for the congruent (ICC = 0.924, IC 95%: 0.862-0.957, $P < 0.001$) and incongruent (ICC = 0.904, IC 95%: 0.811-0.949, $P < 0.001$) stimuli.

Discussion

This study aimed to assess the validity and reliability of a mobile application, a alternative to measure response times to congruent and incongruent stimuli, which are important outcomes of cognitive performance. The app demonstrated validity and reliability in measuring response times in a Stroop test compared to a computerized version. Also, the data presented in Bland–Altman plots (Figure 2) showed that most values of congruent and incongruent response times are close to the mean of the differences between instruments (41.3 ± 56.4 e 76.4 ± 75.1 ms, respectively), thereby representing a high level of agreement. The app's reliability in measuring five response times for each participant showed excellent internal consistency values. Additionally, the test-retest reliability of the app was almost perfect between days. These results suggest that ES can be considered as a valid and reliable alternative to assess cognitive performance, particularly response time, on smartphone devices.

The Stroop test has found diverse applications in clinical practice and various research studies involving neuropsychological assessment (Bajaj et al., 2015; Bezdicek et al., 2015; Emek Savaş et al., 2020). The computerized version is a valid and essential tool for routine clinical screening (Gualtieri & Johnson, 2006). While the app has been validated for diagnosing covert hepatic encephalopathy using the paper-and-pencil Stroop test (Bajaj et al., 2015), this study represents the first direct comparison with a computerized version. Research that aims to verify the validity and accuracy of the response time of mobile applications uses computer versions as a comparison (Coswig et al., 2019). Thus, the present study was the first to compare and show that the mobile application version has validity and reliability compared to the computerized version.

Taking this information into account, the rapid advance of new technologies suggests that the use of smartphones has been greater than that of computers due to their popularity, accessibility, portability, and advanced technology; so, smartphones are poised to become a common platform for accurate health-related measurements (Bort-Roig et al., 2014; Ghaffari et al., 2023; Majumder & Deen, 2019). It is worth noting that the primary limitation of the ES is the relatively small screen size of cell phones compared to computers. This factor may lead volunteers to use peripheral vision, particularly during incongruent trials that demand inhibitory control.

In fact, previous studies have proposed comparing the computerized versions (manual response) with the original (oral response). Their findings showed an influence of the response modality (oral vs manual), pointing to a reduced interference effect (Penner et al., 2012). In this sense, perhaps one of the limitations of the present study was that it did not compare the app version with the original paper version. Although this comparison is important in future studies, we chose to compare two manual versions, which require motor responses.

On the other hand, although two manual versions (computer and app) were compared, it is possible to notice a motor difference in the method of input of responses in the application (touch screen) and in the computer (keyboard buttons), which may also affect response times. These factors could explain the correlation for incongruent stimuli being smaller concerning congruent stimuli. It may be interesting for future studies to investigate this app's use on larger portable devices such as tablets as well as assessing touchscreen computer tests. Another aspect that should be considered in the present study is that the app was tested on healthy individuals. Therefore, we do not know if the correlations could be altered in clinical populations. Thus, future studies could investigate the validity of the app in other populations.

Conclusion

The ability to assess and monitor cognitive performance capacity, particularly through response times, is important in several areas of neuropsychology and research. The results of the current study demonstrate that Stroop test's response times for congruent and incongruent stimuli can be easily, accurately, and reliably evaluated using an app called "Encephalapp Stroop." These findings offer valuable insights for researchers and professionals seeking a valid and cost-effective means to assess the cognitive capacity of their subjects and/or patients.

References

The University of Sydney Manual of PIEL Survey Application. [Mobile app] 2012. [Accessed January 20, 2014]. Available from: http://pielsurvey.org/

Bajaj, J. S., Heuman, D. M., Sterling, R. K., Sanyal, A. J., Siddiqui, M., Matherly, S., Luketic, V., Stravitz, R. T., Fuchs, M., Thacker, L. R., Gilles, H., White, M. B., Unser, A., Hovermale, J., Gavis, E., Noble, N. A., & Wade, J. B. (2015). Validation of EncephalApp, Smartphone-Based Stroop Test, for the Diagnosis of Covert Hepatic Encephalopathy. *Clinical Gastroenterology and Hepatology*, *13*(10), 1828-1835.e1. https://doi.org/10.1016/j.cgh.2014.05.011

Bezdicek, O., Lukavsky, J., Stepankova, H., Nikolai, T., Axelrod, B. N., Michalec, J., Růžička, E., & Kopecek, M. (2015). The Prague Stroop Test: Normative standards in older Czech adults and discriminative validity for mild cognitive impairment in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, *37*(8), 794–807. https://doi.org/10.1080/13803395.2015.1057106

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet (London, England)*, *1*(8476), 307–310.

Bort-Roig, J., Gilson, N. D., Puig-Ribera, A., Contreras, R. S., & Trost, S. G. (2014). Measuring and influencing physical activity with smartphone technology: A systematic review. *Sports Medicine (Auckland, N.Z.)*, *44*(5), 671–686. https://doi.org/10.1007/s40279-014-0142-5

Coswig, V., Sant’ Ana, J., Coelho, M. N., Pereira Moro, A. R., & Diefenthaeler, F. (2019). Development of a Mobile Phone App for Measuring Striking Response Time in Combat Sports: Cross-Sectional Validation Study. *JMIR mHealth and uHealth*, *7*(11), e14641. https://doi.org/10.2196/14641

Dalgleish, T. (1995). Performance on the emotional Stroop task in groups of anxious, expert, and control subjects: A comparison of computer and card presentation formats. *Cognition and Emotion*, *9*(4), 341–362. https://doi.org/10.1080/02699939508408971

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Emek Savaş, D. D., Yerlikaya, D., G Yener, G., & Öktem Tanör, Ö. (2020). Validity, Reliability and Normative Data of the Stroop Test Çapa Version. *Turk Psikiyatri Dergisi = Turkish Journal of Psychiatry*, *31*(1), 9–21. https://doi.org/10.5080/u23549

Fernandes, J., Arida, R. M., & Gomez-Pinilla, F. (2017). Physical exercise as an epigenetic modulator of brain plasticity and cognition. *Neuroscience and Biobehavioral Reviews*, *80*, 443–456. https://doi.org/10.1016/j.neubiorev.2017.06.012

Fortes, L. S., Lima-Júnior, D. D., Gantois, P., Nasicmento-Júnior, J. R. A., & Fonseca, F. S. (2021). Smartphone Use Among High Level Swimmers Is Associated With Mental Fatigue and Slower 100- and 200- but Not 50-Meter Freestyle Racing. *Perceptual and Motor Skills*, *128*(1), 390–408. https://doi.org/10.1177/0031512520952915

Fortes, L. S., Lima-Junior, D., Nascimento-Júnior, J. R. A., Costa, E. C., Matta, M. O., & Ferreira, M. E. C. (2019). Effect of exposure time to smartphone apps on passing decision-making in male soccer athletes. *Psychology of Sport and Exercise*, *44*, 35–41. https://doi.org/10.1016/j.psychsport.2019.05.001

Ghaffari, A., Kildahl Lauritsen, R. E., Christensen, M., Rolighed Thomsen, T., Mahapatra, H., Heck, R., Kold, S., & Rahbek, O. (2023). Exploring the Feasibility and Usability of Smartphones for Monitoring Physical Activity in Orthopedic Patients: Prospective Observational Study. *JMIR mHealth and uHealth*, *11*, e44442. https://doi.org/10.2196/44442

Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *21*(7), 623–643. https://doi.org/10.1016/j.acn.2006.05.007

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, *11*(6), 838–854. https://doi.org/10.1177/1745691616650285

Henik, A. (1996). Paying attention to the Stroop effect? *Journal of the International Neuropsychological Society*, *2*(5), 467–470. https://doi.org/10.1017/S1355617700001557

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

Linnman, C., Carlbring, P., Åhman, Å., Andersson, H., & Andersson, G. (2006). The Stroop effect on the internet. *Computers in Human Behavior*, *22*(3), 448–455. https://doi.org/10.1016/j.chb.2004.09.010

Luo, M., Yu, X.-B., Hu, S.-J., & Bai, F.-H. (2020). EncephalApp Stroop App predicts poor sleep quality in patients with minimal hepatic encephalopathy due to hepatitis B-induced liver cirrhosis. *Saudi Journal of Gastroenterology*, *26*(3), 120. https://doi.org/10.4103/sjg.SJG_558_19

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Majumder, S., & Deen, M. J. (2019). Smartphone Sensors for Health Monitoring and Diagnosis. *Sensors (Basel, Switzerland)*, *19*(9), 2164. https://doi.org/10.3390/s19092164

McKay, A. K. A., Stellingwerff, T., Smith, E. S., Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J., & Burke, L. M. (2022). Defining Training and Performance Caliber: A Participant Classification Framework. *International Journal of Sports Physiology and Performance*, *17*(2), 317–331. https://doi.org/10.1123/ijspp.2021-0451

Penner, I.-K., Kobel, M., Stöcklin, M., Weber, P., Opwis, K., & Calabrese, P. (2012). The Stroop task: Comparison between the original paradigm and computerized versions in children and adults. *The Clinical Neuropsychologist*, *26*(7), 1142–1153. https://doi.org/10.1080/13854046.2012.713513

Schiehser, D. M., & Bondi, M. W. (2010). Stroop Color-Word Test. In *The Corsini Encyclopedia of Psychology* (pp. 1–1). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470479216.corpsy0951

Spaner, C. R., Musteata, S., Kenny, R. A., Gawryluk, J. R., Hofer, S., & Christie, B. R. (2019). *3-Dimensional Multiple Object Tracking Training Can*.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Trecroci, A., Duca, M., Cavaggioni, L., Rossi, A., Scurati, R., Longo, S., Merati, G., Alberti, G., & Formenti, D. (2021). Relationship between Cognitive Functions and Sport-Specific Physical Performance in Youth Volleyball Players. *Brain Sciences*, *11*(2), 227. https://doi.org/10.3390/brainsci11020227

Van Cutsem, J., Roelands, B., Pluym, B., Tassignon, B., Verschueren, J., De Pauw, K., & Meeusen, R. (2020). Can Creatine Combat the Mental Fatigue–associated Decrease in Visuomotor Skills? *Medicine & Science in Sports & Exercise*, *52*(1), 120–130. https://doi.org/10.1249/MSS.0000000000002122

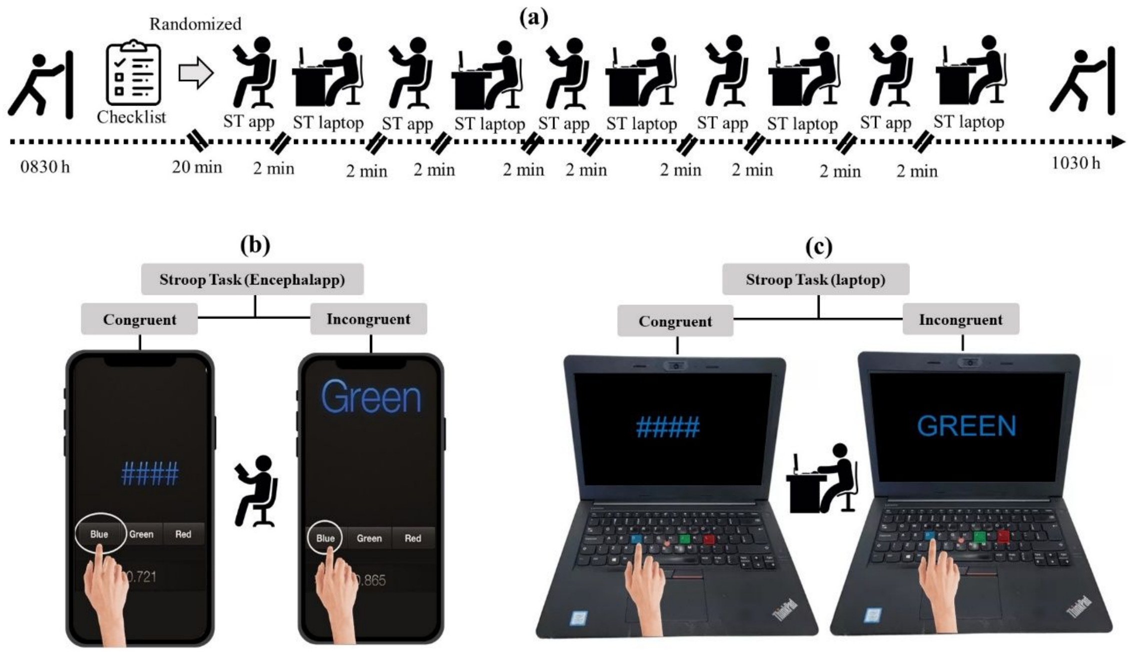


Figura 1. Experimental procedures, app, and computer stroop test. (a) Stages of the test and retest visits consisted of performing five Stroop tests for each equipment. (b) Stroop test version by the "Encephalapp Stroop" app. (c) Stroop test version by computer.

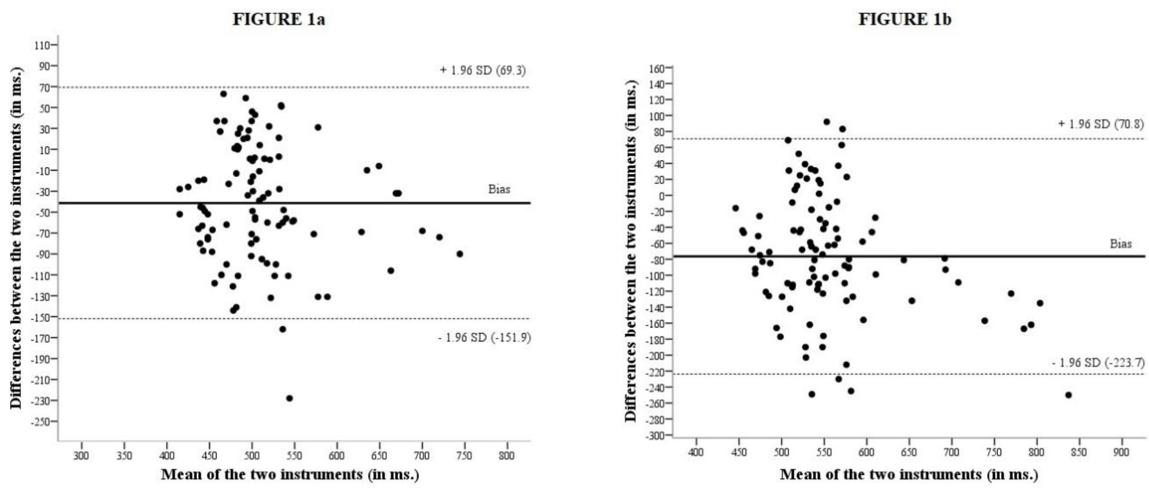


Figure 2. Bland-Altman plots the response times of congruent (Figure 3a) and incongruent (Figure 3b) stimuli for computer and app Stroop.