

RESEARCH ARTICLE

Emphasizing speed or accuracy in an eye-tracking version of the Trail-Making-Test: Towards experimental diagnostics for decomposing executive functions

Lukas Recker^{1*}, Rebecca M. Foerster^{1,2}, Werner X. Schneider¹, Christian H. Poth¹

1 Neuro-cognitive Psychology and Center for Cognitive Interaction Technology (CITEC), Bielefeld University, Bielefeld, Germany, **2** Medical School EWL, Bielefeld University, Bielefeld, Germany

* lukas.recker@uni-bielefeld.de

Abstract

The Trail-Making-Test (TMT) is one of the most widely used neuropsychological tests for assessing executive functions, the brain functions underlying cognitively controlled thought and action. Obtaining a number of test scores at once, the TMT allows to characterize an assortment of executive functions efficiently. Critically, however, as most test scores are derived from test completion times, the scores only provide a summary measure of various cognitive control processes. To address this problem, we extended the TMT in two ways. First, using a computerized eye-tracking version of the TMT, we added specific eye movement measures that deliver a richer set of data with a higher degree of cognitive process specificity. Second, we included an experimental manipulation of a fundamental executive function, namely participants' ability to emphasize speed or accuracy in task performance. Our study of healthy participants showed that eye movement measures differed between TMT conditions that are usually compared to assess the cognitive control process of alternating between task sets for action control. This demonstrates that eye movement measures are indeed sensitive to executive functions implicated in the TMT. Crucially, comparing performance under cognitive control sets of speed vs. accuracy emphasis revealed which test scores primarily varied due to this manipulation (e.g., trial duration, number of fixations), and which were still more sensitive to other differences between individuals (e.g., fixation duration, saccade amplitude). This provided an experimental construct validation of the test scores by distinguishing scores primarily reflecting the executive function of emphasizing speed vs. accuracy and those independent from it. In sum, both the inclusion of eye movement measures and of the experimental manipulation of executive functions in the TMT enabled a more specific interpretation of the TMT in terms of cognitive functions and mechanisms, which offers more precise diagnoses in clinical applications and basic research.

OPEN ACCESS

Citation: Recker L, Foerster RM, Schneider WX, Poth CH (2022) Emphasizing speed or accuracy in an eye-tracking version of the Trail-Making-Test: Towards experimental diagnostics for decomposing executive functions. *PLoS ONE* 17(9): e0274579. <https://doi.org/10.1371/journal.pone.0274579>

Editor: Nicholas V Swindale, University of British Columbia, CANADA

Received: November 30, 2021

Accepted: August 30, 2022

Published: September 12, 2022

Copyright: © 2022 Recker et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data have been made publicly available at the Open Science Framework and can be accessed at: <https://osf.io/E4UR5>.

Funding: This work was funded by the the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, <https://www.dfg.de/>) – project number 418552203. Acquired by R.M.F. The funders had no role in study design, data

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Diagnostic tests are the cornerstones of psychological assessments for assessing cognitive impairments and guiding clinical treatments. The Trail-Making-Test (TMT, [1]) is among the most widely used diagnostic tests in neuropsychology and related psychological assessments. It consists of two parts in which participants manually connect sequences of targets in a pre-determined order. During part A (TMT-A) participants connect numbers 1 to 25 in ascending order. During part B (TMT-B) participants connect numbers 1 to 13 and letters A to L in ascending order, alternating between the two sets (i.e. 1, A, 2, B, 3 . . .). In terms of executive functions, correct performance in TMT-B implies a much stronger involvement compared to test version TMT-A [2, 3]. These brain functions enable cognitively controlled behavior [4, 5] and their impairments prevent neurological and psychiatric patients from coping with everyday life [6]. In clinical standard practice, typically a number of different cognitive domains are assessed using different neuropsychological tests. The TMT often constitutes an important part of such an assessment, for instance to diagnose cognitive control impairments of patients suffering from stroke [7], mild cognitive impairment, and dementia [8]. In addition to such clinical applications, the test is widely-used in basic research to study differences in executive functions between healthy individuals [9, 10]. While the TMT offers a practical way to survey individuals' executive functions, its test scores are based on relatively unspecific compound measures of cognitive performance on a speeded paper-pencil task. Such compound measures can be obtained easily and efficiently and are well suited for cognitive screenings. However, because the measures reflect several different underlying processes collectively, they are hard to interpret in terms of specific cognitive functions and mechanisms. Here, we address this problem by providing a new computer-based version of the TMT that extends the original in two ways: First, it includes eye movement measures that can be interpreted in terms of more specific cognitive functions. Second, it includes an experimental manipulation of participants' executive functions for action control, namely, their task set as emphasizing speed or accuracy of test performance. This cognitive control function is rarely investigated in neuropsychological tests. We use this manipulation of participants' task set to evaluate each of the test scores in light of this executive function. As a result, the test scores can now be interpreted either as measures of this executive function or as measures of other cognitive functions (cognitive control and more basic processes). This validates the test scores as measures of the executive function (i.e. providing converging or diverging construct validity), provides a greater conceptual specificity of the interpretation of the scores, and links the scores with mechanistic theories from the literature of experimental and neuro-cognitive psychology. Thus, taken together, by including eye-movement measures and an experimental manipulation of the cognitive control set speed vs. accuracy, the new version of the TMT extends the test's construct validity and supports a cognitively more specific interpretation for clinical diagnostic decision-making.

Based on accuracy and completion time, the TMT provides a number of different test scores indexing cognitive functions, whereby norms for clinical assessments mostly focus on completion time [3]. Classically, the test is conducted using pencil and paper and participants' completion times and their errors are manually recorded by the experimenter. If the participant makes an error this is pointed out by the experimenter and the sequence is continued from the last correct stimulus. The completion time for test half A is related to cognitive processing speed [3, 11]. The difference and quotient between the completion times for part A and part B are taken as measures of the ability to shift between task-sets [2, 12]. The completion times and errors in the TMT-B are associated with the individual's capabilities for visual-spatial processing, working memory, inhibition and behavioral control, and their general cognitive flexibility [2, 12–14]. Therefore, the TMT-B is often used to screen an individual across

multiple different executive functions [15]. Critically however, the obtained test scores are relatively unspecific and cannot be interpreted as indices of specific cognitive functions or mechanisms. For instance, interpreting the completion time of the TMT-A as a measure of “cognitive processing speed”, is rather descriptive and does not provide insight into possibly affected cognitive functions (e.g. localization of action targets, movement planning, etc.) or cognitive mechanisms (e.g., attention mechanisms performing the localization of action targets, etc.). One way to address this problem is to include eye-tracking measures into the test. Most of these measures are based on saccadic eye movements and intervening fixations. There is a wealth of neurophysiological [16–18] and cognitive research [19–21] on the cognitive functions implicated in saccadic eye movements and on the neuro-cognitive mechanisms fulfilling these functions [22–24]. As a result, currently there exist a number of theories and models enabling a specific interpretation of saccade-related performance measures.

To increase the conceptual specificity and better depict the determinants of performance in the test, we extend the TMT by including a number of different eye movement measures. First, fixation durations enable the closer examination of the processing of visual stimuli. The time participants spend fixating on objects in large parts parallels the time needed to extract the necessary information for subsequent actions and planning [25]. Thus, fixation durations describe in greater detail the ongoing cognitive mechanisms during task performance. Second, the size of executed eye movements represented by saccade amplitudes can provide insights into strategic cognitive processes employed by participants [26]. For instance, participants could either scan their environment using large- or small-scale eye movements. Furthermore, as indicated by previous studies both fixation durations and saccade amplitudes may be determined by stable endogenous factors of individuals and thus could present a useful tool to differentiate between persons [27–29]. Third, the overall number of fixations can indicate applied scanning strategies [30] and can for instance imply the number of times participants shift their attention during the task. Previous studies of the TMT which included eye-tracking indicated differences in the number of fixations across test halves [31–33] and stability of saccade amplitude and fixation duration [33].

The available measures provided by eye tracking further extend to the connection of eye and hand movements in the task. Since the task requires the participants to consecutively connect targets via lines, measures that examine and describe more closely the role and relationship of eye movements during sequences of manual actions can be determined. These measures include characteristics of fixation types [20, 34] or the interplay of the eyes and the hands [35, 36]. In particular, based on the current position in the sequence fixations can be categorized as either falling on the current target, i.e. guiding fixations, or past and future targets, i.e. searching fixations. Guiding fixations are assumed to play a central role in coordinating eye and hand movements. Usually, prior to a hand movement to a target object (e.g., reaching for a glass of water), the eyes land on the target, allowing more efficient arm movements to the foveated location [20, 37]. In our eye tracking based TMT version, a fixation on the current clicking target object (e.g., on number 5) should contribute to more precise clicking on this object. Moreover, searching fixations are required when the target object for hand movements has not been found and foveated yet [37]. For the TMT, prior to clicking on the current target (e.g. after a preceding click on number 4, the current search target should be number 5), searching fixations at potential target locations are made and the search process stops once the current target is found (e.g., number 5). Furthermore, the temporal coupling of eyes and hands is illustrated in the eye-hand span, the time between the fixation on an action target and the subsequent action. This measure is often applied to study differences between novices and experts, e.g. in reading music or other sensory-motor tasks [36, 38, 39], and thus describes expertise in and difficulty of a task more closely. Lastly, the length of the scanpath can describe

how efficient participants are in covering the search area and provide an additional indicator of the effectiveness in gathering information. Taken together, these measures grant additional insights into basic neuro-cognitive functions [20, 37, 40, 41] especially into visual selection processes and thus present a valuable source of complementary information when understanding performance in the TMT.

So, including eye-tracking in the TMT should improve the validity and conceptual specificity of the test by providing additional measures interpretable in terms of their respective cognitive functions. This has proven helpful not only for the TMT but a wide range of neuropsychological tests before (e.g. Corsi blocks, [42]; spatial working memory, [43]). Still, relating test scores to differences between the test halves leaves a lot of room for interpretation since both differ greatly in their cognitive demands and in turn their underlying cognitive functions. However, this problem can be addressed by combining the TMT with an experimental manipulation of participants' task set whose influence on executive and other cognitive functions more generally has been thoroughly established in the neuro-cognitive literature. One type of fundamental task set that regulates performance on any given cognitive task is the one trading-off the speed against the accuracy of task-related actions [44, 45]. When participants employ a task set emphasizing speed, actions become faster at the expense of a lower accuracy and more errors, whereas when they employ a task set emphasizing accuracy, actions become slower at the benefit of a higher accuracy and fewer errors (for a review, see [45]). This speed-accuracy trade-off seems to reflect one of the few general laws in psychology, as it is evident in a wide variety of cognitive tasks (e.g., decision-making, [46], visual search [26, 47], memory retrieval [44]), and even extends to other species [48]. Including this manipulation to the TMT targets a set of executive functions that are fundamental to human intelligent behavior, as they are implicated in the dynamic adaptation of behavior to current task demands [4, 49]. Specifically, these functions consist in dynamically adopting specific cognitive task sets, representations of task demands which govern the cognitive processes underlying task performance [50]. Comparing TMT performance under these experimentally manipulated conditions allows to assess how each test score is affected by this manipulation, and the variation of test scores due to this effect can then be compared to the interindividual variation within the test scores. Test scores that primarily reflect the speed-accuracy task set vary more strongly due to the manipulation, whereas test scores that do not primarily reflect this construct vary more strongly due to interindividual differences. Therefore, the focus of including the speed-accuracy manipulation in the test lies not in specifying the exact effects of focusing on either one task set but in enabling connections of test scores to the underlying cognitive control function. In other words, using this approach, one can use a theory-based manipulation of a cognitive construct to test whether it is reflected or not reflected in a given test score of the TMT. In this way, the approach extends the test scores' validity by offering an experimentally-based construct validity, which allows a more specific interpretation of the test score in terms of the cognitive functions and mechanisms given by the neuro-cognitive theories the manipulation is based on. To the best of our knowledge, this study is the first that aims to use speed-accuracy instructions in this way.

In the present study, we aimed to increase conceptual specificity with which the test scores of the TMT can be interpreted in terms of cognitive functions and mechanisms. To this end, we provide a computer-based version of the TMT that extends the original with eye movement measures allowing a more specific examination of reflected cognitive processes and with an experimental manipulation of participants' executive functions for action control, namely, their task set as emphasizing speed or accuracy of test performance. We test the new version of the TMT in an experiment comprising the TMT-A and the TMT-B.

First, the additionally included eye tracking measures were investigated with regard to differences between TMT-A and TMT-B. If a test score differed between test half A and B, then the underlying cognitive function could provide additional information about the performance related to known differences in demands of the test halves. However, since differences between TMT-A and TMT-B could still reflect a number of different cognitive functions in either test half we aimed to furthermore link test scores to processes involved in action control. If test scores were affected by the speed and accuracy instructions, then these scores should reflect the state of the current task set. This effect could then provide the link to active processes in action control when emphasizing either speed or accuracy.

In addition, for each test score we could compare the variance stemming from interindividual differences with the variance induced by the experimental manipulation of participants' speed vs. accuracy task sets. If a test score primarily reflected the speed vs. accuracy task set, then its variance should be dominated by the task set manipulation and not by interindividual differences. In contrast, if a test score primarily reflected a cognitive construct different from the speed vs. accuracy task set, then the variance of this score should be dominated by interindividual differences and not by the task manipulation.

Methods

Participants

Sixty-two healthy participants (43 female, 19 male) from Bielefeld University were paid to take part in the experiment, two of which had to be excluded due to missing data. They were between 18 and 28 years old ($M = 21.71$, $SD = 2.24$). All of them were right-handed and reported normal or corrected-to-normal vision. Before the experimental procedure, participants gave written informed consent. The experiment was conducted in accordance with the ethical guidelines of the German Psychological Association (Deutsche Gesellschaft für Psychologie) and was approved by the ethics committee at Bielefeld University.

Apparatus and stimuli

The experiment was conducted in a dimly lit room, where participants sat down in front of a 19 inch monitor (ViewSonic G90fb) with a resolution of 1024 x 768 pixels (corresponding to physical dimensions of 36 cm x 27 cm (width x height) and a refresh rate of 100 Hz (pre-heated for at least 5 min, [51]). Their head was stabilized on a chin rest 71 cm from the screen. Movements of their right eyes were recorded at 1000 Hz, using the desktop-mounted eye-tracker EyeLink 1000 (SR Research, Ontario, Canada). During the experimental trials, they navigated the cursor with a computer mouse (Logitech RX250), whose positions were sampled at 100 Hz (time-locked to screen refresh). Stimulus presentation was controlled by the Experiment Builder software (SR Research, Ontario, Canada).

Stimuli consisted of black letters and/or numbers surrounded by a black, unfilled circle. In version A of the stimulus set, numbers 1 to 25 constituted the target stimuli. In version B of the stimulus set, numbers 1 to 13 and letters A to L constituted the target stimuli. In both versions, they were presented in Arial with the font size set to 16 pt (approximate height 0.59° va). The surrounding circle measured 1.35° va diameter, its line width was set to 5. The mouse cursor was represented by a small black, filled circle measuring 0.42° va diameter. Stimuli were presented on a uniform grey background (30 cd/m^2 , measured with a MAVOLUX-digital luminance meter). Stimulus positions were the same for all participants within TMT-A and -B respectively. That is, each participant completed the same spatial arrangement in TMT-A and TMT-B trials, while the order of the speed or accuracy instructions was counterbalanced across arrangements within the blocks. The positions consisted in points scattered around the

center points of a randomly chosen field from a 5-by-5 grid spanning the entire screen (i.e. with a field size of 5.72 x 4.32 (width x height) ° va). However, they were not specifically matched across test halves. Sounds were provided as feedback for the participants' answers. A high-pitched sound (53 dB(A), measured with a Mengshen MS-M80A) played for 70 ms indicated a hit, a low-pitched sound (46 dB(A)) played for 100 ms indicated a miss.

Procedure

The experiment was divided into two blocks (A and B), each consisting of one training and two experimental trials. Each block started with a written instruction on the screen after which the eye tracker was calibrated using a 9-point calibration and validation procedure. In block A (i.e. TMT-A) participants' task was to click the target stimuli, numbers 1 to 25, in ascending order (cf. Fig 1 for a depiction of the trial procedure). In block B (i.e. TMT-B) participants' task was to click the target stimuli, numbers 1 to 13 and letters A to L, in ascending order, alternating between the two sequences (e.g. 1, A, 2, B, 3, . . . , 13). Different from the paper-pencil version of the test, targets were not connected via lines. This also meant that targets next to each other in the sequence did not need to be in adjacent circles in the stimulus display and intermediate targets could be moved across by the participant. Participants first completed the training trial with a reduced number of stimuli (8 targets). The experimental trials then were each preceded by a short, written instruction, introducing the speed or accuracy condition of the following run. In the speed condition, participants were asked to complete the sequence as fast as possible. In the accuracy condition, participants were asked to hit the targets as central as possible. Either way, responses were counted as correct if they fell within the circle surrounding the number/ letter. A correct response was followed by a high-pitched sound, an incorrect response was followed by a low-pitched sound. Participants could only move on to

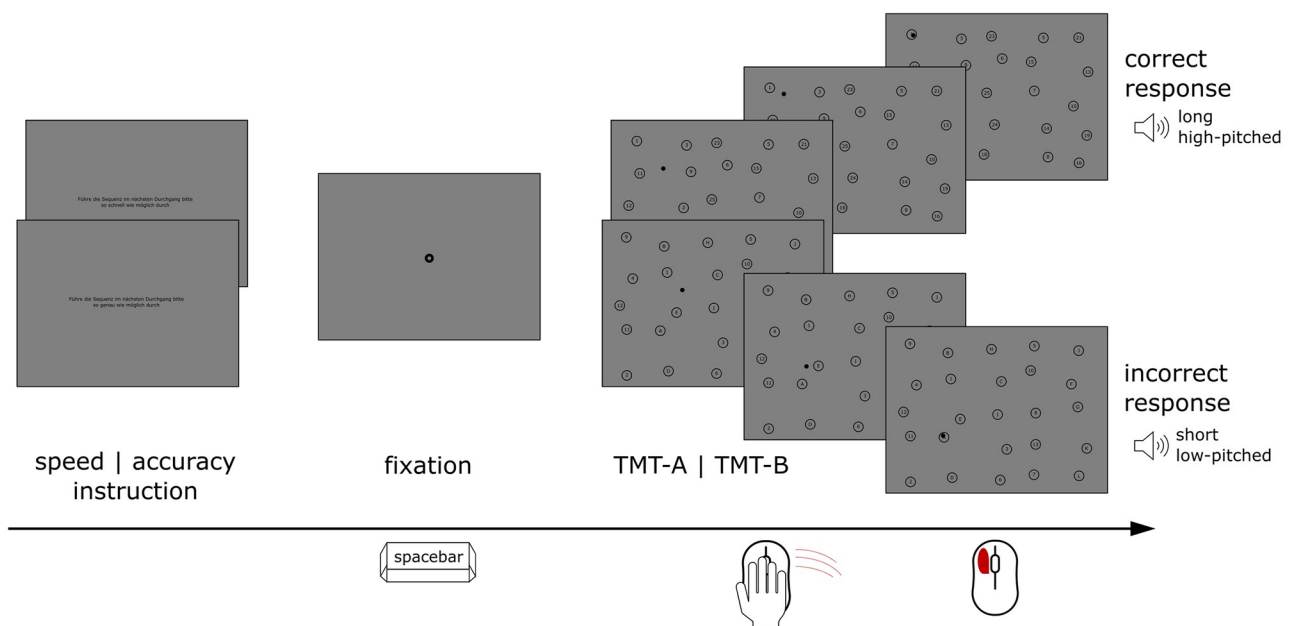


Fig 1. Trial procedure. Time course of a trial in this version of the TMT. Participants were first instructed to hit targets in the following trial either as fast (speed) or as central (accuracy) as possible. After a fixation display, they started the TMT sequence by pressing the space bar. Participants clicked the sequence of numbers (TMT-A) or numbers and letters (TMT-B) in ascending order while their eye movements were recorded. Correct answers were indicated by a high-pitched sound, wrong answers were indicated by a low-pitched sound. The black dot indicates the current position of the mouse.

<https://doi.org/10.1371/journal.pone.0274579.g001>

the next target in sequence if they successfully hit the previous one. A centrally presented black ring, serving as a fixation stimulus preceded each trial (0.45° va outer circle, 0.11° va inner circle). Participants could start each trial by pressing the space bar. Altogether, the experiment comprising the two blocks and three trials each was completed within 10 minutes. The order of the blocks and the trial instructions within the blocks was counterbalanced across participants.

Dependent variables

To detect fixations, saccades and blinks, the algorithm of the EyeLink 1000 was used. Saccades and fixations were detected online using a velocity threshold of $30^\circ \text{ va } \times \text{ s}^{-1}$ and an acceleration threshold of $8000^\circ \text{ va } \times \text{ s}^{-2}$. Fixations were merged if they were less than 1° va apart. The analysis includes the following basic dependent variables: the trial duration provides the measure for the completion time in each trial, it was computed as the time between participants signal to start (hitting the space bar) and the successful completion of the last target (mouse click on target); the fixation durations, saccade amplitudes, and number of fixations as provided by the online detection algorithm of the EyeLink 1000. Furthermore, the following dependent variables were added: the number of searching fixations, i.e. fixations visiting future or past targets., guiding fixations, i.e. fixations visiting the present target [20, 40] scanpath length, and eye-hand span, i.e. the time between a fixation on the target and a consequent action. To define a fixation as either searching or guiding, the position of the fixation must fall within a circle with 3.25° va diameter around the target center. To calculate the eye-hand span the timestamp of the first guiding fixation is taken and compared to the timestamp of the matching click. Resulting spans indicate the time by which the eyes lead the hands [36].

Results

Statistical analyses

The data and fully documented analysis code have been made publicly available at the Open Science Framework and can be accessed at: <https://osf.io/E4UR5/>. Data analysis was done in R (Version 4.1.3 [52]). We divided the data analysis into two steps: First, we examined the effects of the speed-accuracy instructions on the outcome measures provided in this TMT. We therefore used linear mixed-effect models with subjects as random factors employing the lme4 package (Version 1.1–28 [53]). Fixed effects could comprise the within subject factors test type (TMT-A vs. TMT-B) and instruction (accuracy vs. speed), as well as their interaction term. Factors are dummy-coded for the analyses. Effects of test type are always in reference to factor level TMT-A. Effects of instruction are always in reference to factor level accuracy. Estimates of p-values were provided by the lmerTest-package (Version 3.1–3 [54]) using a significance-criterion of $p < .05$, however because estimation of p-values in mixed models is not without issues, we focus more on the confidence intervals of the estimates. The data of two participants were excluded from the analysis due to missing values. Additionally, we report the best fitting models which means for some dependent variables up to two additional data sets were excluded. However, this did not alter results.

Second, we developed a procedure to determine the degree to which dependent variables are dominated by the effects of the task set manipulation emphasizing speed or accuracy in TMT-A and TMT-B or interindividual variability respectively. Therefore, we compared the influence of the speed-accuracy manipulation with the interindividual variability, i.e. the fixed and the random effects, in the data. The comparison of fixed and random effects was achieved by calculating the quotient of the Bayes Factor (BF_{10} , BayesFactor-package; [55]) of a model comprising only the fixed effects structure divided by the Bayes Factor of a model comprising

Table 1. Results of the linear mixed-effect models.

DV	Fixed Parts				Random Parts		
	Betas				N _{group}	Observations	SD _{Intercept}
	(Intercept)	test type	instruction	test type × instruction			
trial duration	72.55 [65.58–79.52]	22.97 *** [15.37–30.58]	-28.18 *** [-35.58 –-20.57]	14.20 * [3.42–24.98]	58	231	17.22
fixation duration	160.80 [155.45–166.15]	3.40 ** [1.14–5.65]	2.22 [-0.03–4.48]	-1.90 [-5.08–1.29]	58	232	19.85
saccade amplitude	4.62 [4.46–4.77]	-0.18 ** [-0.31 –-0.06]	-0.21 ** [-0.33 –-0.08]	0.45 *** [0.27–0.63]	59	236	0.50
n fixations ^a	5.54 [5.45–5.62]	0.25 *** [0.17–0.33]	-0.33 *** [-0.41 –-0.26]	0.30 *** [0.19–0.40]	59	236	0.26
n guiding fixations	53.73 [50.20–57.26]	-1.27 [-4.42–1.88]	-2.58 [-5.73–0.57]	1.83 [-2.62–6.28]	59	236	10.74
n searching fixations ^a	4.31 [4.20–4.42]	0.34 *** [0.23–0.45]	-0.43 *** [-0.54–-0.32]	0.26 ** [0.11–0.41]	59	236	0.30
eye-hand span ^a	7.48 [7.39–7.57]	0.05 [-0.04–0.14]	-0.49 *** [-0.58–-0.40]	0.27 *** [0.15–0.40]	57	228	0.23
scanpath length ^a	11.17 [11.09–11.27]	0.31 *** [0.23–0.39]	-0.39 *** [-0.49 –-0.31]	0.35 *** [0.24–0.47]	59	236	0.27

DV = dependent variable;

^a = marked variables were log-transformed;

n = number of . . . ; SD = standard deviation; Numbers in squared brackets represent 95% confidence intervals; p-value estimates are indicated by

* < .05,

** < .01,

*** < .001 (Intercepts all had $p < .001$)

<https://doi.org/10.1371/journal.pone.0274579.t001>

only the random effects. The result indicated whether the fixed effects-only or the random effects-only model was more likely to find the given data compared to an intercept-only model. Additionally, we calculated marginal R-squared and conditional R-squared values for the model [56]. These values represent estimates describing the variance that is explained by the model's fixed effects (marginal) and the combination of fixed and random effects (conditional).

We repeated the described procedure for each dependent variable. The results of the analyses are described in the following. The first analysis step is depicted in Table 1 and Fig 2. Additionally, we provide a table including the descriptive statistics illustrated in Fig 2 in the supplementary material (see S1 Table). The second analysis step is depicted in Fig 3. A commented version of the analysis code further describing the newly applied method is included in the OSF repository.

Dependent variables

Trial durations. Like in all versions of the TMT the trial duration was overall longer in TMT-B (test type: $b = 22.974$, $SE = 3.879$, $t = 5.921$, $p < .001$). Participants were faster under speed-instructions ($b = -28.177$, $SE = 3.879$, $t = -7.263$, $p < .001$). This effect was less pronounced in TMT-B ($M_{diff} = 14.09$ s, $SD = 35.26$ s) than it was in TMT-A ($M_{diff} = 28.18$ s, $SD = 21.82$ s; $b = 14.200$, $SE = 5.501$, $t = 2.581$, $p < .05$).

The additional analyses indicated a prevailing effect of the experimental speed-accuracy manipulation over interindividual variability in the TMT-A but not the TMT-B. The Bayes Factor comparison strongly favored the fixed effects-only model ($BF_{10} = 4.07 \times 10^{11}$) in the

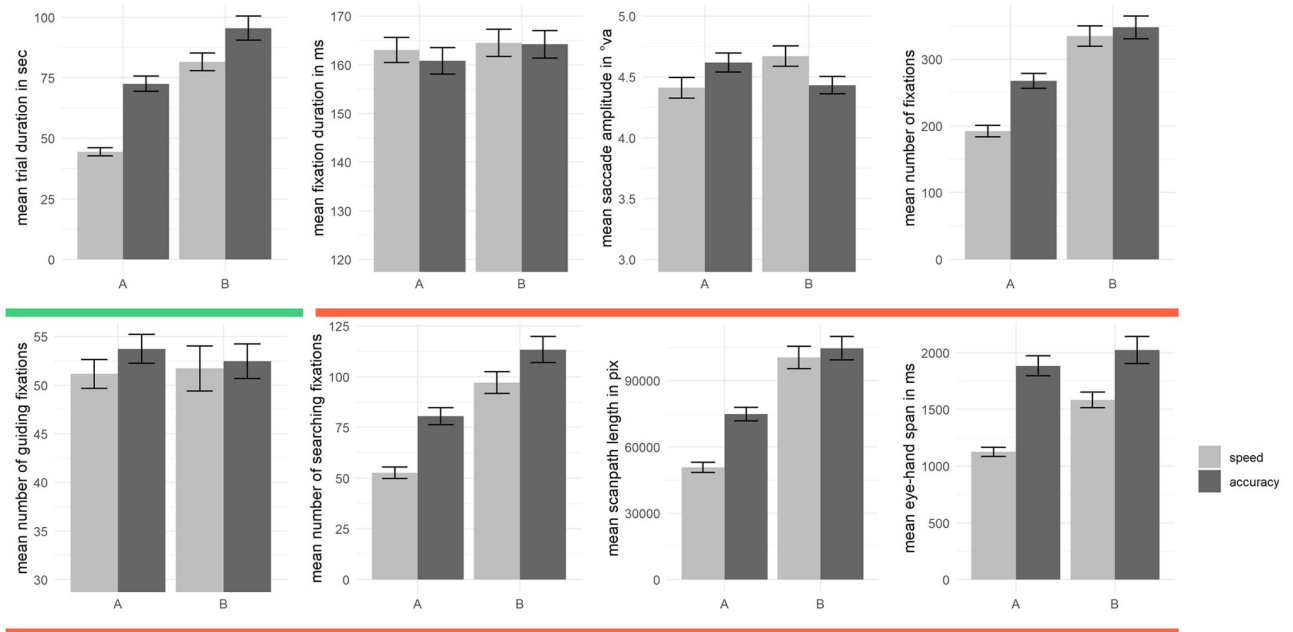


Fig 2. Test scores in TMT-A/-B under speed/accuracy instructions. Bar plots for the dependent variables across all conditions (test halves A and B; speed and accuracy instruction). Error bars represent the standard error of the mean. The red line underneath the plots indicates variables representing eye movement control measures. The green underneath the plots indicates variables representing classic TMT measures, i.e. trial duration.

<https://doi.org/10.1371/journal.pone.0274579.g002>

TMT-A. In the TMT-B the random effects-only model was favored ($BF_{10} = 0.07$). The R-squared values supported these results (TMT-A: $R_m^2 = .355$ [.240, .463]; $R_c^2 = .578$ [.408, .720]; TMT-B: $R_m^2 = .043$ [.005, .120]; $R_c^2 = .465$ [.270, .652]) indicating predominant influences of the speed-accuracy manipulation in the TMT-A but not the TMT-B.

Fixation duration. Fixation durations were slightly longer in TMT-B ($M = 164.36$ ms, $SD = 21.31$ ms) than in TMT-A ($M = 161.91$ ms, $SD = 20.11$ ms), although the effect was small ($b = 2.448$, $SE = 0.817$, $t = 2.996$, $p < .01$). Also, this was not influenced by speed or accuracy instructions. Generally, the instructions had no effects on fixation durations (all $bs < 2.224$, $ts < 1.935$, $ps > .05$).

The second part of our analysis for the fixation durations of participants revealed a strong influence of interindividual variability in either half of the test. The comparison of the Bayes Factors underlined the presenting picture, suggesting strong evidence for the random effects-only model in the TMT-A, $BF_{10} = 1.75 \times 10^{-18}$, and the TMT-B, $BF_{10} = 6.99 \times 10^{-27}$. Accordingly, the marginal R-squared of the final model was close to 0 in both test halves (TMT-A: $R_m^2 = .003$ [.000, .015]; TMT-B: $R_m^2 = .000$ [.000, .003]), while the conditional R-squared was much greater (TMT-A: $R_c^2 = .878$ [.810, .929]; TMT-B: $R_c^2 = .937$ [.896, .962]).

Saccade amplitude. Results for participants' saccade amplitude suggested influences from all experimental factors. However, all effects were of small magnitude as evident by the confidence intervals close to 0 (see Table 1).

This pattern further established in the second part of the analysis. Bayes Factor comparisons, too, favored the random effects-only model in the TMT-A, $BF_{10} = 9.27 \times 10^{-11}$ and the TMT-B, $BF_{10} = 6.00 \times 10^{-10}$. Taken together with the calculations for the R-squared estimates, this suggested a predominant influence of interindividual variability (TMT-A: $R_m^2 = .027$ [.008, .061]; $R_c^2 = .806$ [.699, .881]; TMT-B: $R_m^2 = .040$ [.014, .079]; $R_c^2 = .819$ [.726, .890]).

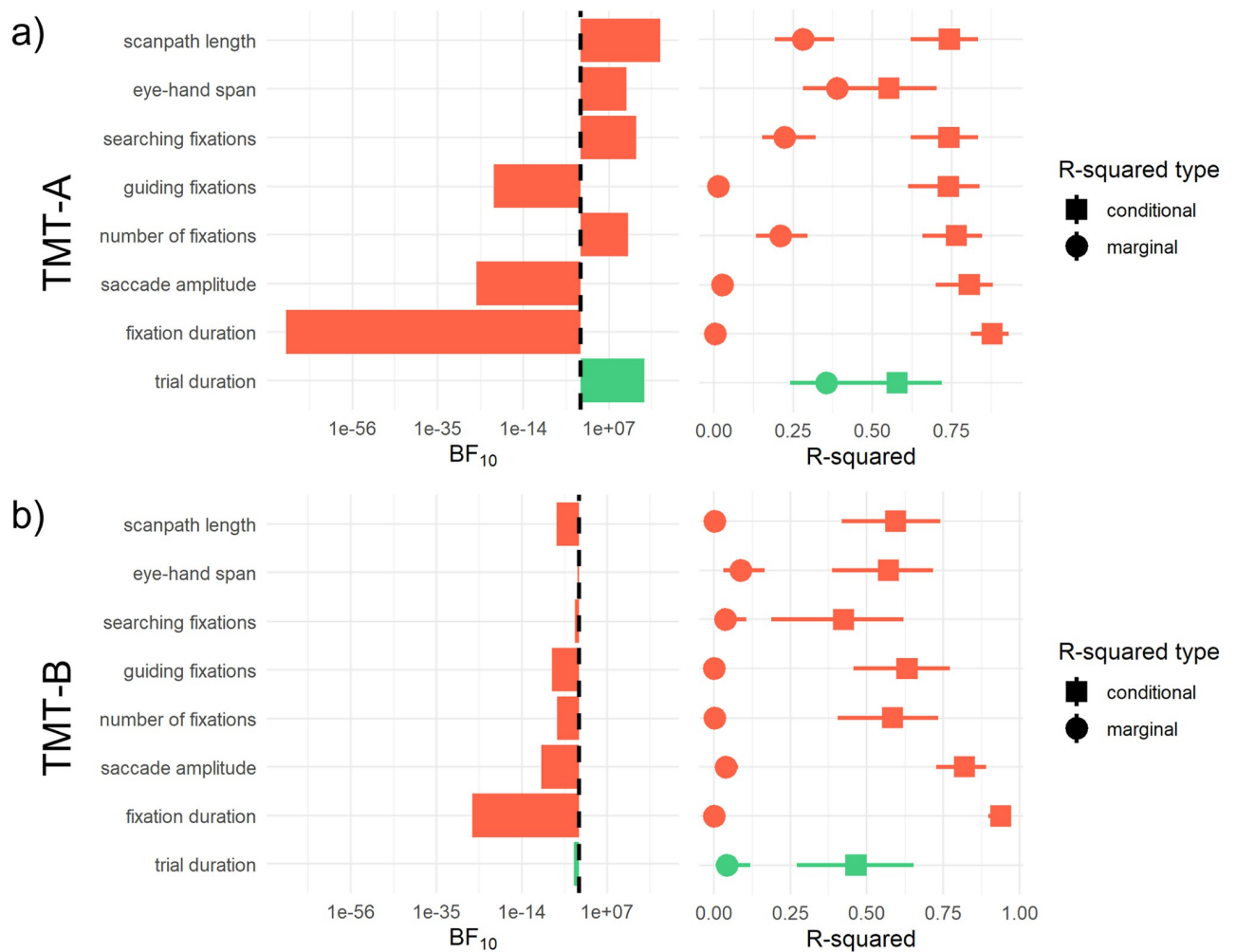


Fig 3. Variables dominated by the individual variability vs. the speed-accuracy manipulation. Results of the analyses for R-squared values and the Bayes Factor comparison of fixed versus random effects in the model for each examined dependent variable in TMT-A (Fig 3a) and TMT-B (Fig 3b). Error bars in the R-squared plot represent bootstrapped 95% C.I. All variables related to eye movement control are plotted in red. All variables related to classic TMT measures, i.e. trial duration, are plotted in green.

<https://doi.org/10.1371/journal.pone.0274579.g003>

Number of fixations. The model for participants' number of fixation was fitted to log-transformed values, since they indicated a better fit than the original values. Both the test type ($b = 0.250$, $SE = 0.039$, $t = 6.460$, $p < .001$), the instruction ($b = -0.331$, $SE = 0.039$, $t = -8.557$, $p < .001$) and their interaction ($b = 0.297$, $SE = 0.055$, $t = 5.430$, $p < .001$) influenced the number of fixations. Since participants also spent more time on the respective task, these results were in line with the results of the trial duration, i.e. they performed more fixations in TMT-B, less during speed instructions, while the latter effect was less pronounced in TMT-B than in TMT-A.

Analyzing the number of fixations applying the second analysis method comparing the Bayes Factors for fixed effects-only and random effects-only models, the results indicated a better fit of the fixed effects-only model in the TMT-A, $BF_{10} = 7190$. Similar to the trial durations, this was not the case in the TMT-B, $BF_{10} = 4.92 \times 10^{-6}$. R-squared values supported these results, revealing that the fixed effects accounted for .211 [.133, .296] proportion of the variance, fixed and random effects for .765 [.658, .846] in the TMT-A. In the TMT-B fixed

effects accounted for .002 [.000, .027] proportion of the variance, fixed and random effect for .584 [.404, .733]. Thus, the number of fixations was predominantly determined by the speed-accuracy instructions in the TMT-A but not the TMT-B.

Number of guiding fixations. Results for the number of guiding fixations mirrored the results for the saccade amplitudes. Here, all confidence intervals for the estimates of the factors included 0. Therefore, we found no differences in participants' number of guiding fixations depending on test type or instruction.

The second part of our analysis revealed that the comparison of the Bayes Factors supported the evidence for the importance of interindividual variability compared to the task influences in both TMT-A, $BF_{10} = 1.24 \times 10^{-9}$, and TMT-B, $BF_{10} = 2.69 \times 10^{-7}$. Estimates of the R-squared values were .013 [.001, .043] for the marginal and .740 [.613, .838] for the conditional R-squared in the TMT-A and .001 [.000, .019] for the marginal and .631 [.456, .770] for the conditional R-squared in the TMT-B respectively. Similarly to fixation durations and saccade amplitudes, the number of guiding fixations were predominantly determined by interindividual variability in both test halves.

Number of searching fixations. The model for the number of searching fixations indicated influences for both factors (test type: $b = 0.340$, $SE = 0.055$, $t = 6.152$, $p < .001$; instruction: $b = -0.432$, $SE = 0.055$, $t = -7.813$, $p < .001$) and their interaction ($b = 0.258$, $SE = 0.078$, $t = 3.305$, $p < .01$), similar to the general number of fixations. Here too, we log-transformed the dependent measure to assure the better model fit. Taking together the results of the overall fixations and the separate fixation types, the pattern indicates that the change in overall number of fixations stems from influences on the number of searching fixations.

The Bayes Factors calculated for the comparison of fixed and random effects influence in the second part of our analysis, came out in favor of fixed effects impact, $BF_{10} = 4.73 \times 10^4$ in the TMT-A and in favor of the random effects impact, $BF_{10} = 0.12$ in the TMT-B. We obtained R-squared values of $R_m^2 = .222$ [.151, .322] and $R_c^2 = .742$ [.622, .834] in the TMT-A and $R_m^2 = .038$ [.004, .105] and $R_c^2 = .424$ [.188, .619] in the TMT-B. Akin to the general number of fixations, the results present a more divided picture, but stress the effects of the experimental task manipulation on the number of searching fixations in the TMT-A.

Eye-hand span. Results of the model for the eye-hand span indicated that the instruction (instruction: $b = -0.493$, $SE = 0.049$, $t = -10.088$, $p < .001$) and the interaction between test type and instruction ($b = 0.273$, $SE = 0.057$, $t = 4.802$, $p < .001$) had an impact on the temporal distance between participants fixation on a target and the following click. For the instruction the eye-hand span was shorter under speed- than under accuracy instructions ($M_{\text{speed}} = 1.35$ s, $SD_{\text{speed}} = 0.48$ s; $M_{\text{accuracy}} = 1.95$ s, $SD_{\text{accuracy}} = 0.79$ s). Furthermore, this effect was smaller for eye-hand spans in TMT-B. Descriptively, the eye-hand span was shorter in TMT-A ($M = 1.50$ s, $SD = 0.64$ s) than in TMT-B ($M = 1.80$ s, $SD = 0.76$ s). However, the confidence interval of the estimate included 0.

The Bayes Factor comparisons for the eye-hand span suggested an advantage of the fixed effects-only model over the random effects-only model in the TMT-A, $BF_{10} = 1.13 \times 10^{13}$. But again, turned around in the TMT-B, $BF_{10} = 0.55$. The resulting R-squared values for the model were $R_m^2 = .389$ [.282, .511] and $R_c^2 = .552$ [.401, .703] in the TMT-A, and $R_m^2 = .086$ [.030, .165] and $R_c^2 = .570$ [.386, .716] in the TMT-B. These results suggest a predominant influence of the experimental task manipulation on the eye-hand span in the TMT-A only.

Scanpath length. Finally, for the scanpath length we found comparable results to the trial duration and number of fixations. The model fitted to the log-transformed values included both factors test type, $b = 0.310$, $SE = 0.043$, $t = 7.156$, $p < .001$ and instruction, $b = -0.395$, $SE = 0.040$, $t = -9.750$, $p < .001$, as well as their interaction, $b = 0.354$, $SE = 0.057$, $t = 6.816$, $p < .001$. Participants scanpath length fitted the pattern that more time and more fixations were

spent in TMT-B, while speed-instructions shortened the scanpath length, especially in TMT-A.

The second part of our analysis further established the reoccurring picture. Again, the comparison of the Bayes Factors suggested a greater influence of experimental task manipulation on the length of the scanpath than interindividual variability, $BF_{10} = 2.95 \times 10^7$ but only in TMT-A. The Bayes Factor in the TMT-B suggested a greater influence of individual variability on the length of the scanpath than of the experimental task manipulation, $BF_{10} = 3.24 \times 10^{-6}$. The resulting R-squared values were $R_m^2 = .280$ [.192, .379] and $R_c^2 = .743$ [.621, .833] in the TMT-A, and $R_m^2 = .003$ [.000, .030] and $R_c^2 = .593$ [.418, .739] in the TMT-B.

Discussion

Diagnostics tests provide the tool for neuropsychological assessment to identify cognitive impairments and profile the strengths and weaknesses of individuals. As one of the most popular tests in neuropsychological diagnostics, the TMT is frequently applied to survey participants' processing speed and executive functions. Despite the test's widespread use, the provided test scores portray rather unspecific compound measures of cognitive functions since they are based on global test completion times. Here, we provided a new version of the TMT aiming to increase the test scores' conceptual specificity and their interpretation in terms of cognitive processes and mechanisms. To this end, we included additional eye tracking measures to the test and added an experimental manipulation of participants' task set. Contrasting the eye-tracking-based test scores from the different test halves revealed which of these measures were sensitive to the compound measure of task set switching of the classic TMT. As the cognitive processes captured by these measures are better understood in terms of neurophysiology and cognition [16, 19, 57], they can aid interpretation by revealing specific component processes of the broader compound measure of the classic TMT. Furthermore, for each test score we compared the interindividual variance with the variance introduced by the experimental manipulation of emphasizing speed or accuracy. This comparison allowed us to dissociate test scores from or associate test scores with the experimental manipulation and evaluate them in relation to the cognitive constructs linked with the speed vs. accuracy task set. Taken together, this demonstrated how expanding the TMT with experimentally-based resources can improve the understanding of examined test scores and validate them by providing links to cognitive processes and mechanisms grounded in cognitive theory and experimental basic research.

Eye-tracking measures in the TMT

The new version of the TMT included two distinct additions: added test scores provided by eye-tracking and a manipulation of participants' task set, emphasizing speed or accuracy. The aim of the inclusion of eye-tracking measures in the TMT was to increase the specificity of provided test scores in terms of enabled interpretations considering cognitive functions associated with the execution of eye movements. *Fixation durations* and *saccade amplitudes* were relatively stable across test halves and speed-accuracy instructions. Neither the numerous differences between TMT-A and -B in affected cognitive functions [2], nor the additional manipulation of participants' task set had substantial impact on the duration of their fixations or the magnitude of their eye movements. Due to the different spatial arrangement of targets in the test halves, we have to be cautious when interpreting measures that are closely related to spatial components of the task like the saccadic amplitude. However, since the stability of fixation durations and saccade amplitudes parallels previous work examining these measures across different tasks [27, 28, 58] and the TMT [33] we are confident that these measures might reflect

stable interindividual differences also in the context of this TMT. The *number of fixations*, increased from TMT-A to TMT-B and interacted with the speed-accuracy instructions. Previous investigations of eye movements in the TMT which found differences in the number of fixations across test halves [31–33] provided consistent results. Performance in the TMT-B is thought to reflect many underlying cognitive processes including behavioral control and inhibition [12, 13]. An increased number of fixations aids interpreting differences between test halves by demonstrating concrete effects of the additional demands on participants behavior in terms of their eye movement control. The *number of searching fixations* and the *scanpath length* both increased from TMT-A to TMT-B and also interacted with the speed-accuracy instructions. Especially the scanpath length is heavily influenced by characteristics of the search display like the location of stimuli. However, the magnitude of differences in the scanpath lengths between TMT-A and TMT-B together with the paralleling results from trial durations and the general number of fixations increasing from part A to B suggest effects beyond influences of spatially different arranged stimuli. Different cognitive demands in the test halves are therefore also reflected in characteristics of the scanpath and the number of searching fixations. The *eye-hand span*, i.e. the temporal distance between participants' fixation on a target and the following click, did not differ between TMT-A and TMT-B but was affected by the speed-accuracy instructions. Up until now, no measures specific to action sequence tasks, like the number of different fixation types or the temporal relation between eye and hand movements, were included in a study investigating the TMT. Studies examining the effects of affordances [41] and learning of action sequences [40], suggest that these kinds of measures might be sensitive to task-related differences. Our results for these eye movement measures in part support this assumption. We found that the number of searching fixations, the eye-hand span, and the scanpath length were affected by the current task set. The *number of guiding fixations* in contrast was relatively stable across both test halves and the speed-accuracy instructions. Also considering the results on the overall number of fixations which can be attributed to differences between searching but not guiding fixations suggests an important role of search related processes in the TMT and enhances the understanding of underlying cognitive processes in that way. Taken together, the included eye-tracking measures can increase the specificity of interpretations of test scores in the TMT in two ways. Test scores varying with the specific demands of the task contain information about the established executive functions underlying the TMT. Furthermore, test scores that are stable across the task can help narrow down which aspects of the executive functions are influenced and which are not. It should be noted that the spatial arrangement and thus the length of the overall trail participants had to complete was randomized and thus not separately controlled for between TMT-A and TMT-B. Therefore, test scores particularly sensitive to influences from the spatial stimulus arrangement (e.g. saccade amplitude and scanpath length) should be interpreted with caution. Nevertheless, our findings from the comparison of these measures between the TMT-A and TMT-B should not have been driven by differences in the randomized spatial arrangements, because a recent follow-up study provided converging results under conditions where the spatial arrangement of stimuli had been controlled for (see the pre-registration of the project as well as its results at: <https://osf.io/fwhjb/>)

Speed vs. accuracy task set influences in the TMT

Next to the addition of eye-tracking measures, this version of the TMT also included a manipulation of participants' task set, namely their emphasis on either speed or accuracy during task performance. We used this manipulation to compare the variance from the experimental manipulation with the variance due to interindividual differences and examined each test

score considering the predominant influence of either experimental manipulation or interindividual variability in each test half. Test scores that primarily reflect the speed-accuracy task set vary more strongly due to the manipulation, whereas test scores that do not primarily reflect this construct vary more strongly due to interindividual variability. In this way, we link test scores with, or we demarcate them from the cognitive construct of the speed-accuracy task set and test whether it is reflected or not reflected in a given test score of the TMT. In the TMT-A, the trial duration, the number of fixations, the number of searching fixations, the eye-hand span, and the scanpath length were dominated by the experimental task manipulation of emphasizing speed or accuracy. As such, the test scores provided by these measures reflect influences of participants' task set, i.e. their adaptation to current task demands in terms of emphasizing speed or accuracy. Since the trade-off between speed and accuracy of responses affects presumably every task in a large variety of everyday situations [26, 59], it has a high ecological validity. The trade-off also successfully provided a tool to study and explain possible differences for instance in aging [60–62] or the progression of Mild Cognitive Impairment [63]. Therefore, relating the set of test scores to this construct can offer valuable insights into participants' task performance while practicing the TMT because they provide valid measures of the applied experimental manipulation. On the other hand, we found that fixation durations, saccade amplitudes and the number of guiding fixations were predominantly determined by the interindividual variability in the TMT-A. Furthermore, in the TMT-B all test scores were predominantly determined by the interindividual variability. That is, all of these test scores were not dominated by the current emphasis on speed or accuracy of the participant but rather the variability between participants. In the TMT-A these measures thus represent test scores which are relatively stable across adaptation of participants' task set of emphasizing speed or accuracy. This confirms previous results on fixation durations and saccade amplitudes [27, 28, 58] and provides divergent validity for their stability in the TMT because of their relative independence from the experimental manipulation. In case of the TMT-B, the relative dominance of interindividual variability speaks to the complexity of the involved functions in the determinants of task. As the trade-off between speed and accuracy is a phenomenon impacting a great number of tasks and processes it presumably depicts a strong effect. Still, variability between individuals in the TMT-B is relatively stronger than this effect. That is not to say, that there are no influences of emphasizing speed or accuracy in the TMT-B, as our previous analyses illustrated, but rather that test scores in the TMT-B are more strongly determined by interindividual variability not caused by the speed accuracy manipulation.

Enabling a cognitively specific interpretation of the TMT

The new version of the TMT improves the construct validity of the test in two ways. First, the inclusion of additional eye-tracking measures provides test scores directly relatable to specific cognitive functions. Since there is a body of work investigating eye movements with respect to their functions while performing for example manual actions [20] or visual search [64] or influences of cognitive control [21], they can establish connections to such constructs. For instance, the number of fixations and their respective function as searching or guiding, demonstrates that while searching behavior increases in the TMT-B, the guidance of specific actions is constant across test halves. Second, combining the test with an experimental manipulation targeting a specific and established cognitive construct enables the assessment of each test scores with respect to this manipulation. Here, instructing participants to either emphasize their speed or their accuracy during performance allowed us to compare test scores relative to this task set manipulation and evaluate them based on the predominant influence. By comparing the variation in the test scores that is due to the cognitive construct captured by the

experimental manipulation (i.e. here the task set for cognitively adjusting speed and accuracy of task performance) with the variation due to interindividual differences we create a measure supporting the test score's construct validity. That is, test scores that are dominated by the task set manipulation receive a converging construct validation showing that the test score is a valid measure of the construct. At the same time, test scores that are dominated by interindividual variability even despite the experimental manipulation receive a diverging construct validation that demonstrates their relative independence from the construct. This evaluation helps identifying the set of scores best suited for the diagnostic needs. Test scores varying due to the experimental manipulation represent valid measures of the included construct, while test scores varying due to individual differences can be demarcated from the included construct. Taken together using an experimental manipulation in a neuropsychological test as a benchmark for the test scores enables interpretations in light of the experimental construct. Interpretations can be linked to current models and theories of cognitive function and thus improve the understanding of sampled test scores.

Practical implications

Clinical decisions rely on a clear and differentiated assessment of the neuropsychological functionality of individuals. Based on this assessment conclusions are drawn and treatments are chosen. Extending previous studies on eye-tracking studies in the TMT [31–33], our results reveal how extending neuropsychological tests with eye-tracking supports neuropsychological assessments of cognition. First, eye-tracking measures are easier to interpret in terms of neuro-cognitive functions, because compared with behavioral measures, their neurophysiological underpinnings are better understood [16, 18]. Second, eye movements are often intact even when neurological syndromes impair patients' manual behavior, so that eye-tracking also enables to cognitively assess patients that would otherwise be untestable due to these impairments [32, 65]. Besides the advantages of eye-tracking for neuropsychological assessments, our findings also show how including experimental manipulations in diagnostic tests supports validation of test scores with respect to the cognitive construct the test aims to assess and increases the conceptual specificity of the test's scores. That is, one can evaluate the test scores against experimental manipulations that have been established to target a highly specific and theory-grounded cognitive construct. In this way, the high conceptual specificity of the manipulations is transferred to the test scores in question. Here, test scores could be identified as primarily reflecting the executive function underlying the adaptation of performance to speed vs. accuracy demands. For instance, in the TMT-A fixation durations, saccade amplitudes and the number of guiding fixations were dominated by interindividual variability. Thus, their underlying cognitive processes should not be primarily dependent on the task set. This supports and clarifies the interpretation of the scores. Taken together, the here applied approach offers a more mechanistic interpretation of the TMT test scores which helps the diagnostic assessment of executive functions and cognitive impairments, thus supporting clinical decision-making. [32]

Conclusions

In conclusion, we provide a new version of the TMT comprising additional eye-tracking measures and an experimental manipulation of participants' task set on emphasizing speed vs. accuracy. This approach allows for a more specific interpretation of test scores in the TMT, going beyond the usually provided compound measures of cognitive functions in the classic test. By including an experimental manipulation into the test, we were able to differentiate between test scores that predominantly reflect the speed-accuracy manipulation vs. the inter-

individual variability derived from other cognitive processes. In this way, test scores can be linked to and interpreted considering cognitive theories the manipulation is based on. Using established experimental manipulations in this way, presents a unique approach in including cognitive constructs into tests. Experimental and differential/ correlational approaches in psychology have long been called to move together and benefit from another [66, 67], but only recently this call began to attract answers [68, 69]. The present approach attempts to join this path and suggests a widely applicable way to actively incorporate cognitive theory in neuropsychological settings.

Supporting information

S1 Table. Descriptive results of dependent variables per experimental condition.

(DOCX)

Author Contributions

Conceptualization: Lukas Recker, Christian H. Poth.

Formal analysis: Lukas Recker.

Funding acquisition: Rebecca M. Foerster.

Software: Lukas Recker.

Supervision: Christian H. Poth.

Visualization: Lukas Recker.

Writing – original draft: Lukas Recker, Christian H. Poth.

Writing – review & editing: Rebecca M. Foerster, Werner X. Schneider, Christian H. Poth.

References

1. Reitan RM. Validity of the Trail Making Test as an Indicator of Organic Brain Damage. *Percept Mot Skills*. 1958; 8: 271–276. <https://doi.org/10.2466/pms.1958.8.3.271>
2. Sánchez-Cubillo I, Periáñez JA, Adrover-Roig D, Rodríguez-Sánchez JM, Ríos-Lago M, Tirapu J, et al. Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuospatial abilities. *Journal of the International Neuropsychological Society*. Cambridge University Press; 2009. pp. 438–450. <https://doi.org/10.1017/S1355617709090626> PMID: 19402930
3. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. *Nat Protoc*. 2006; 1: 2277–2281. <https://doi.org/10.1038/nprot.2006.390> PMID: 17406468
4. Diamond A. Executive functions. *Annual Review of Psychology*. Annual Reviews Inc.; 2013. pp. 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750> PMID: 23020641
5. Egner T, Wiley J. *The Wiley Handbook of Cognitive Control*. The Wiley Handbook of Cognitive Control. John Wiley & Sons, Ltd; 2017. <https://doi.org/10.1002/9781118920497>
6. Grafman J, Litvan I. Importance of deficits in executive functions. *Lancet*. Elsevier Limited; 1999. pp. 1921–1923. [https://doi.org/10.1016/S0140-6736\(99\)90438-5](https://doi.org/10.1016/S0140-6736(99)90438-5) PMID: 10622291
7. Kopp B, Rössler N, Tabeling S, Stürenburg HJ, De Haan B, Karnath HO, et al. Errors on the trail making test are associated with right hemispheric frontal lobe damage in stroke patients. *Behav Neurol*. 2015; 2015. <https://doi.org/10.1155/2015/309235> PMID: 26074673
8. Ashendorf L, Jefferson AL, O'Connor MK, Chaisson C, Green RC, Stern RA. Trail Making Test errors in normal aging, mild cognitive impairment, and dementia. *Arch Clin Neuropsychol*. 2008; 23: 129–137. <https://doi.org/10.1016/j.acn.2007.11.005> PMID: 18178372
9. Crowe SF. The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the trail making test. *J Clin Psychol*. 1998; 54: 585–591. PMID: 9696108

10. Salthouse TA. What cognitive abilities are involved in trail-making performance? *Intelligence*. 2011; 39: 222–232. <https://doi.org/10.1016/j.intell.2011.03.001> PMID: 21789028
11. Ríos Lago M, Periañez JA, Muñoz-Céspedes JM. Attentional control and slowness of information processing after severe traumatic brain injury. *Brain Inj*. 2004; 18: 257–272. <https://doi.org/10.1080/02699050310001617442> PMID: 14726285
12. Arbuthnott K, Frank J. Trail Making Test, Part B as a measure of executive control: Validation using a set-switching paradigm. *J Clin Exp Neuropsychol*. 2000; 22: 518–528. [https://doi.org/10.1076/1380-3395\(200008\)22:4;1-0;FT518](https://doi.org/10.1076/1380-3395(200008)22:4;1-0;FT518) PMID: 10923061
13. Kortte KB, Horner MD, Windham WK. The Trail Making Test, Part B: Cognitive Flexibility or Ability to Maintain Set? *Appl Neuropsychol*. 2002; 9: 106–109. https://doi.org/10.1207/S15324826AN0902_5 PMID: 12214820
14. MacPherson SE, Cox SR, Dickie DA, Karama S, Starr JM, Evans AC, et al. Processing speed and the relationship between Trail Making Test-B performance, cortical thinning and white matter microstructure in older adults. *Cortex*. 2017; 95: 92–103. <https://doi.org/10.1016/j.cortex.2017.07.021> PMID: 28865241
15. Lezak MD, Howieson DB, Loring DW, Hannay HJ, Fisher JS. *Neuropsychological assessment*, 4th ed. —PsycNET. Oxford Univ. Press; 2004. https://psycnet.apa.org/record/2004-16637-000?utm_medium=email&utm_source=transaction
16. Schall JD, Thompson KG. NEURAL SELECTION AND CONTROL OF VISUALLY GUIDED EYE MOVEMENTS. 2003; 22: 241–259. <http://dx.doi.org/10.1146/ANNUREV.NEURO.22.1.241> PMID: 10202539
17. Bisley JW, Bisley J. The neural basis of visual attention. *J Physiol*. 2011; 589: 49–57. PMID: 20807786
18. Krauzlis RJ, Goffart L, Hafed ZM. Neuronal control of fixation and fixational eye movements. *Philos Trans R Soc B Biol Sci*. 2017; 372. <https://doi.org/10.1098/rstb.2016.0205> PMID: 28242738
19. Schütz AC, Braun DI, Gegenfurtner KR. Eye movements and perception: A selective review. *J Vis*. 2011; 11: 9–9. <https://doi.org/10.1167/11.5.9> PMID: 21917784
20. Land M, Tatler B. *Looking and Acting: Vision and eye movements in natural behaviour. Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford University Press; 2009. <https://doi.org/10.1093/acprof:oso/9780198570943.001.0001>
21. Hutton SB. Cognitive control of saccadic eye movements. *Brain Cogn*. 2008; 68: 327–340. <https://doi.org/10.1016/j.bandc.2008.08.021> PMID: 19028265
22. Wischniewski M, Belardinelli A, Schneider WX, Steil JJ. Where to Look Next? Combining Static and Dynamic Proto-objects in a TVA-based Model of Visual Attention. *Cognit Comput*. 2010; 2: 326–343.
23. Tatler BW, Brockmole JR, Carpenter RHS. Latest: A model of saccadic decisions in space and time. *Psychol Rev*. 2017; 124: 267–300. <https://doi.org/10.1037/rev0000054> PMID: 28358564
24. Itti L, Koch C. Computational modelling of visual attention. *Nat Rev Neurosci* 2001 23. 2001; 2: 194–203. <https://doi.org/10.1038/35058500> PMID: 11256080
25. Hayhoe MM, Shrivastava A, Mruczek R, Pelz JB. Visual memory and motor planning in a natural task. *J Vis*. 2003; 3: 49–63. <https://doi.org/10.1167/3.1.6> PMID: 12678625
26. McCarley JS. Effects of speed—Accuracy instructions on oculomotor scanning and target recognition in a simulated baggage X-ray screening task. *Ergonomics*. 2009; 52: 325–333. <https://doi.org/10.1080/00140130802376059> PMID: 19296326
27. Andrews TJ, Coppola DM. Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Res*. 1999; 39: 2947–2953. [https://doi.org/10.1016/s0042-6989\(99\)00019-x](https://doi.org/10.1016/s0042-6989(99)00019-x) PMID: 10492820
28. Rayner K, Li X, Williams CC, Cave KR, Well AD. Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Res*. 2007; 47: 2714–2726. <https://doi.org/10.1016/j.visres.2007.05.007> PMID: 17614113
29. Henderson JM, Luke SG. Stable individual differences in saccadic eye movements during reading, pseudoreading, scene viewing, and scene search. *J Exp Psychol Hum Percept Perform*. 2014; 40: 1390–1400. <https://doi.org/10.1037/a0036330> PMID: 24730735
30. Becic E, Kramer AF, Boot WR. Age-related differences in visual search in dynamic displays. *Psychol Aging*. 2007; 22: 67–74. <https://doi.org/10.1037/0882-7974.22.1.67> PMID: 17385984
31. Wölwer W, Gaebel W. Impaired Trail-Making-Test-B performance in patients with acute schizophrenia is related to inefficient sequencing of planning and acting. *J Psychiatr Res*. 2002; 36: 407–416. [https://doi.org/10.1016/s0022-3956\(02\)00050-x](https://doi.org/10.1016/s0022-3956(02)00050-x) PMID: 12393310
32. Hicks SL, Sharma R, Khan AN, Berna CM, Waldecker A, Talbot K, et al. An eye-tracking version of the trail-making test. *PLoS One*. 2013; 8. <https://doi.org/10.1371/journal.pone.0084061> PMID: 24367626

33. Linari I, Juantorena G, Ibañez A, Petroni A, Kamienkowski JE. Unveiling Trail Making Test: Visual and manual trajectories indexing multiple executive processes. arXiv:210915255 [q-bio]. 2021 [cited 8 Oct 2021]. <http://arxiv.org/abs/2109.15255>
34. Epelboim J, Steinman RM, Kowler E, Edwards M, Pizlo Z, Erkelens CJ, et al. The function of visual search and memory in sequential looking tasks. *Vision Res.* 1995; 35: 3401–3422. [https://doi.org/10.1016/0042-6989\(95\)00080-x](https://doi.org/10.1016/0042-6989(95)00080-x) PMID: 8560808
35. Land MF, Hayhoe M. In what ways do eye movements contribute to everyday activities? *Vision Research*. Pergamon; 2001. pp. 3559–3565. [https://doi.org/10.1016/S0042-6989\(01\)00102-X](https://doi.org/10.1016/S0042-6989(01)00102-X) PMID: 11718795
36. Foerster RM, Carbone E, Koesling H, Schneider WX. Saccadic eye movements in a high-speed bimanual stacking task: Changes of attentional control during learning and automatization. *J Vis.* 2011; 11: 1–16.
37. Foerster RM, Schneider WX. Expectation violations in sensorimotor sequences: Shifting from LTM-based attentional selection to visual search. *Ann N Y Acad Sci.* 2015; 1339: 45–59. <https://doi.org/10.1111/nyas.12729> PMID: 25708482
38. Perra J, Poulin-Charronnat B, Baccino T, Drai-Zerbib V. Review on Eye-Hand Span in Sight-Reading of Music. *J Eye Mov Res.* 2021; 14. PMID: 34840670
39. Nivala M, Cichy A, Gruber H. How prior experience, cognitive skills and practice are related with eye-hand span and performance in video gaming. *J Eye Mov Res.* 2018; 11. <https://doi.org/10.16910/jemr.11.3.1> PMID: 33828697
40. Foerster RM, Schneider WX. Anticipatory eye movements in sensorimotor actions: on the role of guiding fixations during learning. *Cogn Process.* 2015; 16: 227–231. <https://doi.org/10.1007/s10339-015-0701-1> PMID: 26239751
41. Belardinelli A, Herbot O, Butz MV. Goal-oriented gaze strategies afforded by object interaction. *Vision Res.* 2015; 106: 47–57. <https://doi.org/10.1016/j.visres.2014.11.003> PMID: 25451241
42. Hodgson TL, Ezard G, Hermens F. Eye Movements in Neuropsychological Tasks. *Current Topics in Behavioral Neurosciences*. Springer; 2019. pp. 393–418. https://doi.org/10.1007/7854_2019_98 PMID: 31446573
43. Hodgson TL, Hermens F, Pennington K, Pickering JS, Ezard G, Clarke R, et al. Eye movements in the “Morris Maze” spatial working memory task reveal deficits in strategic planning. *J Cogn Neurosci.* 2018; 31: 497–509. https://doi.org/10.1162/jocn_a_01362 PMID: 30513043
44. Wickelgren WA. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol (Amst).* 1977; 41: 67–85. [https://doi.org/10.1016/0001-6918\(77\)90012-9](https://doi.org/10.1016/0001-6918(77)90012-9)
45. Heitz RP. The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Front Neurosci.* 2014; 8: 1–19. <https://doi.org/10.3389/fnins.2014.00150> PMID: 24966810
46. Rae B, Heathcote A, Donkin C, Averell L, Brown S. The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *J Exp Psychol Learn Mem Cogn.* 2014; 40: 1226–1243. <https://doi.org/10.1037/a0036801> PMID: 24797438
47. Carrasco M, McElree B. Covert attention accelerates the rate of visual information processing. *Proc Natl Acad Sci U S A.* 2001; 98: 5363–5367. <https://doi.org/10.1073/pnas.081074098> PMID: 11309485
48. Chittka L, Dyer AG, Bock F, Dornhaus A. Bees trade off foraging speed for accuracy. *Nature.* 2003; 424: 388. <https://doi.org/10.1038/424388a> PMID: 12879057
49. Cohen JD. *Cognitive Control*. Wiley Handb Cogn Control. 2017; 1–28. <https://doi.org/10.1002/9781118920497.CH1>
50. Monsell S. Task Set Regulation. *The Wiley Handbook of Cognitive Control*. John Wiley & Sons, Ltd; 2017. pp. 29–49. <https://doi.org/10.1002/9781118920497.ch2>
51. Poth CH, Horstmann G. Assessing the monitor warm-up time required before a psychological experiment can begin. 2017; 13. <https://doi.org/10.20982/tqmp.13.3.p166>
52. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019. <https://www.r-project.org/>
53. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *J Stat Softw.* 2015; 67. <https://doi.org/10.18637/jss.v067.i01>
54. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw.* 2017; 82: 1–26. <https://doi.org/10.18637/JSS.V082.I13>
55. Morey RD, Rouder JN. BayesFactor: Computation of Bayes Factors for Common Designs. 2018. <https://cran.r-project.org/package=BayesFactor>
56. Nakagawa S, Johnson PCD, Schielzeth H. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J R Soc Interface.* 2017; 14: 20170213. <https://doi.org/10.1098/rsif.2017.0213> PMID: 28904005

57. Tas AC, Moore CM, Hollingworth A. An object-mediated updating account of insensitivity to transsaccadic change. *J Vis.* 2012; 12: 1–13. <https://doi.org/10.1167/12.11.18> PMID: 23092946
58. Castelhana MS, Henderson JM. Stable individual differences across images in human saccadic eye movements. *Can J Exp Psychol.* 2008; 62: 1–14. <https://doi.org/10.1037/1196-1961.62.1.1> PMID: 18473624
59. Beilock SL, Bertenthal BI, McCoy AM, Carr TH. Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychon Bull Rev.* 2004; 11: 373–379. <https://doi.org/10.3758/bf03196585> PMID: 15260208
60. Brébion G. Language processing, slowing, and speed/accuracy trade-off in the elderly. *Exp Aging Res.* 2001; 27: 137–150. <https://doi.org/10.1080/036107301750073999> PMID: 11330210
61. Salthouse TA. Adult age and the speed-accuracy trade-off. *Ergonomics.* 1979; 22: 811–821. <https://doi.org/10.1080/00140137908924659> PMID: 488072
62. Ghisletta P, Joly-Burra E, Aichele S, Lindenberger U, Schmiedek F. Age Differences in Day-To-Day Speed-Accuracy Tradeoffs: Results from the COGITO Study. *Multivariate Behav Res.* 2018; 53: 842–852. <https://doi.org/10.1080/00273171.2018.1463194> PMID: 29683724
63. Lassen-Greene CL, Steward K, Okonkwo O, Porter E, Crowe M, Vance DE, et al. Mild Cognitive Impairment and Changes in Everyday Function Over Time: The Importance of Evaluating Both Speed and Accuracy. *J Geriatr Psychiatry Neurol.* 2017; 30: 220–227. <https://doi.org/10.1177/0891988717711807> PMID: 28639877
64. Eckstein MP. Visual search: A retrospective. *J Vis.* 2011; 11: 14–14. <https://doi.org/10.1167/11.5.14> PMID: 22209816
65. Poletti B, Carelli L, Solca F, Lafronza A, Pedroli E, Faini A, et al. An eye-tracker controlled cognitive battery: overcoming verbal-motor limitations in ALS. *J Neurol.* 2017; 264: 1136–1145. <https://doi.org/10.1007/s00415-017-8506-z> PMID: 28503706
66. Cronbach LJ. The two disciplines of scientific psychology. *Am Psychol.* 1957; 12: 671–684. <https://doi.org/10.1037/h0043943>
67. Wilmer JB. How to use individual differences to isolate functional organization, biology, and utility of visual functions; With illustrative proposals for stereopsis. *Spat Vis.* 2008; 21: 561–579. <https://doi.org/10.1163/156856808786451408> PMID: 19017483
68. Kliegl R, Wei P, Dambacher M, Yan M, Zhou X. Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Front Psychol.* 2011; 1: 1–12. <https://doi.org/10.3389/fpsyg.2010.00238> PMID: 21833292
69. Goodhew SC, Edwards M. Translating experimental paradigms into individual-differences research: Contributions, challenges, and practical recommendations. *Conscious Cogn.* 2019; 69: 14–25. <https://doi.org/10.1016/j.concog.2019.01.008> PMID: 30685513