Aus dem Institut für Psychiatrische Phänomik und Genomik (IPPG)
Institut der Ludwig-Maximilians-Universität München
Direktor: Prof. Dr. med. Thomas G. Schulze

Executive Functions Across the Life Span: Methodological and Genomic Aspects

Kumulative Habilitationsschrift zum Erwerb der Lehrbefähigung für das Fachgebiet Psychiatrische Phänomik und Genomik

vorgelegt von Dr. rer. nat. Urs Heilbronner

2022

Dieser Arbeit liegen folgende Publikationen zugrunde:

Heilbronner, U. & Pollmann, S. (2010). Is there a structural limit to 'branch' recursively between more than two tasks? *Psychological Research*, *74*(3), 327–336. https://doi.org/10.1007/s00426-009-0249-8

Heilbronner, U. & Münte, T. F. (2013). Rapid event-related near-infrared spectroscopy detects age-related qualitative changes in the neural correlates of response inhibition. *Neuroimage*, *65*, 408–415. https://doi.org/10.1016/j.neuroimage.2012.09.066

Heilbronner, U., Samara, M., Leucht, S., Falkai, P. & Schulze, T. G. (2016). The longitudinal course of schizophrenia across the lifespan: Clinical, cognitive, and neurobiological aspects. *Harvard Review of Psychiatry*, *24*(2), 118–128. https://doi.org/10.1097/HRP.0000000000000092

Heilbronner, U., Papiol, S., Budde, M., Andlauer, T. F. M., Strohmaier, J., Streit, F., Frank, J., Degenhardt, F., Heilmann-Heimbach, S., Witt, S. H., Forstner, A. J., Loerbroks, A., Amelang, M., Stürmer, T., Müller-Myhsok, B., Nöthen, M. M., Rietschel, M. & Schulze, T. G. (2021). "The Heidelberg Five" personality dimensions: Genome-wide associations, polygenic risk for neuroticism, and psychopathology 20 years after assessment. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *186*(2), 77–89. https://doi.org/10.1002/ajmg.b.32837

Wendel, B., Papiol, S., Andlauer, T. F. M., Zimmermann, J., Wiltfang, J., Spitzer, C., Senner, F., Schulte, E. C., Schmauß, M., Schaupp, S. K., Repple, J., Reininghaus, E., Reimer, J., Reich-Erkelenz, D., Opel, N., Nenadić, I., Meinert, S., Konrad, C., Klöhn-Saghatolislam, F., … Heilbronner, U. (2021). A genome-wide association study of the longitudinal course of executive functions. *Translational Psychiatry*, *11*(1), 386. https://doi.org/10.1038/s41398-021-01510-8

Abstract

The term Executive Functions (EFs) denotes higher cognitive functions that are responsible for controlling and coordinating mental processes, serving the attainment of behavioral goals. The construct, albeit lacking a concise definition, includes an array of meta-cognitive processes such as inhibition, set-shifting, working memory, and verbal fluency. It has been shown that EFs can be divided into subfunctions, having both common and separable behavioral underpinnings that share a polygenic basis, and are distinct from but related to general intelligence. Neurotypical EF critically depends on the integrity of the prefrontal cortex, and its interconnected areas. Importantly, EFs are impaired in many mental disorders, e.g. in schizophrenia. The present thesis summarizes five studies that focus on different areas of the EF construct. Some studies have a methodological emphasis, and most of them are relevant to the longitudinal course of EFs across the lifespan. Behavioral (Heilbronner & Pollmann, 2010), neuronal (Heilbronner & Münte, 2013), clinical (Heilbronner et al., 2016), and genomic (Heilbronner, Papiol, et al., 2021, Wendel et al., 2021) aspects are addressed. Results show EFs to dynamically change across the life span, both in health and disease, with a pronounced genetic influence. Future research on EF could benefit from the integration of cross-sectional and longitudinal data, and from the use of latent phenotypes.

Contents

1 Introduction

This cumulative thesis compiles five studies that shed light on different aspects of the EF construct. While the first study (Heilbronner & Pollmann, 2010) tests the behavioral predictions of a computational model of human prefrontal cortex (PFC) function, methodological aspects permeate the remaining studies. The second study (Heilbronner & Münte, 2013) employs a novel methodological approach in neuroimaging, and the third study (Heilbronner et al., 2016) characterizes the longitudinal course of executive deficits in schizophrenia (SZ). The fourth study (Heilbronner, Papiol, et al., 2021) researches the genetic background of a latent EF personality trait, and observes its effects on psychopathology some 20 years later. Finally, in the fifth study, Wendel et al. (2021) research the genomic underpinnings of the short-term course of EFs, identifying and replicating a genomic variant associated with different trajectories of EF perfomance over time.

After a brief introduction to the construct of EFs, each study is summarized in Sections 2 to 6.

1.1 The Executive Functions Construct in Health and Disease

The psychological construct of EFs is intimately linked to the human PFC. The famous case of the railroad worker Phineas Gage who, in 1848, suffered a grave injury of the left frontal lobe by piercing an iron rod through his skull (Ratiu & Talos, 2004), leaving him miraculously alive, is one of the most famous case studies in medical history. The consequences of the extensive damage to Gage's brain became apparent in his behavior, which was characterized as "hyperactive" and "disinhibited" (Goldstein et al., 2014), resembling the behavior of monkeys with circumscribed lesions in their prefrontal cortices (Pribram, 1973). Today, the clinical condition that is often the consequence of damage to the PFC is well-known as the so-called dysexecutive syndrome, and may encompass a variey of symptoms, including a lack of impulse control (e.g., Thornton, 2017). Goldstein and Naglieri (2014, p. 9) brilliantly summarize the impairments that characterize the syndrome:

> Ironically, individuals experiencing executive function problems, the result of either atypical development or trauma, often retain

> their memory and capacity to master academic skills but they struggle how to efficiently use what they know. They are inconsistent, unpredictable, and often poorly self-governed. They are inefficient in their ability to make plans, keep track of time, evaluate their behavior, and socialize appropriately. Typically they struggle in many critical aspects of life.

As diverse as these symptoms are definitions of the exact nature of EFs in neurotypical individuals. Goldstein et al. (2014) review more than 30 different definitions, and conclude that EFs are "an umbrella term used for a diversity of hypothesized cognitive processes, including planning, working memory (WM), attention, inhibition, self-monitoring, self-regulation, and initiation carried out by prefrontal areas of the frontal lobes". Historically, this high-level construct has received widespread attention from the scientific community since the early 1970s (Pribram, 1973, Royall et al., 2002, Posner & Snyder, 2004). Of note, although the PFC and EFs are closely related, there are individuals with prefrontal lesions that perform normally on EF tests (Shallice & Burgess, 1991), and, conversely, individuals with lesions outside the PFC that are severely impaired on tests of EF (e.g., Reitan & Wolfson, 1994). Consistently, functional neuroimaging experiments show that there are also regions outside of the PFC that are associated with cognitive control (see Figure 1, and also below).

1.2 Common and Specific Components of Executive Functions

There are several influential theories of EFs (e.g., Baddeley & Hitch, 1994, West, 1996, Miller & Cohen, 2001, Braver & Barch, 2002). To provide a coherent framework of the studies presented in this thesis, the seminal work of Miyake et al. (2000) is briefly presented. In the aforementioned publication, the authors used confirmatory factor analysis to identify three latent components of EFs:

1. Inhibiton of prepotent response tendencies
2. Updating of WM contents
3. Shifting of cognitive task-sets (stimulus-response contingencies)

In a later extension of this Unity and Diversity model (Friedman et al., 2008, Friedman & Miyake, 2017), the latent Inhibiton factor was reconceptualized as Common EF factor, thought to be responsible for interindividual "differences in the ability to maintain and manage goals", since it explained virtually the entire genetic overlap with a higher order factor common to all EFs (Friedman & Miyake, 2017). This extended model also better captures that all executive tasks load onto a latent Common EF factor, but specific task-sets with a strong suppression of pre-potent responses do not load on any additional executive factors (Friedman & Miyake, 2017).

Importantly, all of these latent EF factors have the following properties (Friedman et al., 2008, Friedman & Robbins, 2021):

- Are robustly correlated but separable
- Show high heritability and polygenicity
- Activate both common and specific neuronal areas

Importantly, only the Updating component of EFs is highly correlated with measures of fluid and crystallized intelligence (Friedman et al., 2006). This fits well with the clinical observation that some frontally injured individuals perform normally on intelligence tests. On a final note, verbal fluency, requiring executive control to retrieve contents from memory, is not represented in the Unity and Diversity model, but may constitute an additional executive factor (Fisk & Sharp, 2004). Indeed, there may be more EF factors, the theory does not claim to be comprehensive (Friedman & Miyake, 2017).

1.3 Genetic Underpinnings of Executive Functions

There is ample evidence that all psychological traits are heritable to some degree (Turkheimer, 2000), so it is not surprising that also EFs show substantial heritability (see Li & Roberts, 2018, for review). Indeed, twin studies on a latent Common EF factor in adolescents showed individual differences to be mediated almost entirely by genetics (heritability estimates of 77-100%, Coolidge et al., 2000, Friedman et al., 2008, Engelhardt et al., 2015). However, in adolescencent children, twin studies show much smaller heritability estimates on the level of individual EF tests. Regarding cognitive inhibition

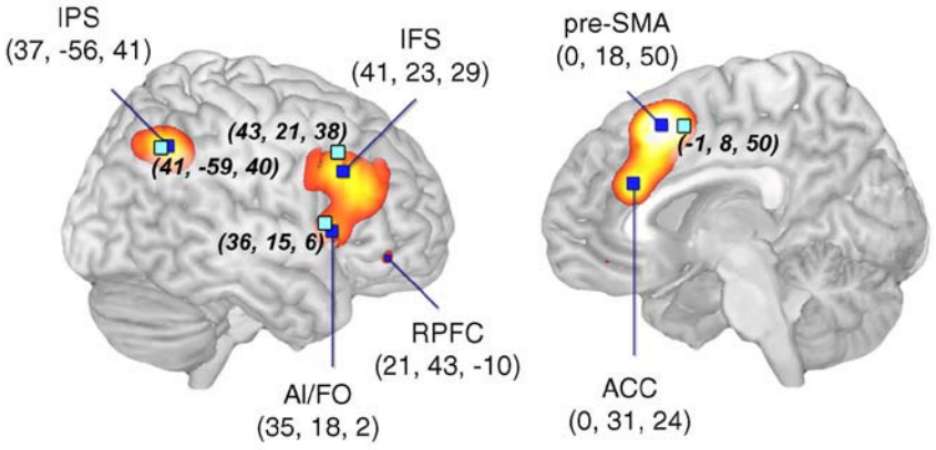
(Go/NoGo) tasks, heritabilities have been shown to vary widely between a lower bound of 10% to an upper bound of 54% (Groot et al., 2004, Kuntsi et al., 2006). Most molecular genetic studies so far (Seshadri et al., 2007, Need et al., 2009, Cirulli et al., 2010, Luciano et al., 2011, Malone et al., 2014, LeBlanc et al., 2012, Ibrahim-Verbaas et al., 2016) did not identify specific genetic variants associated with EF performance. A notable execption is the recent large-scale UK Biobank study of Hatoum et al. (2019) with over 427,000 participants, available as a preprint at the time of writing. This study used factor analytic methods to derive a common EFs score from several EF tests, and identified 112 genome-wide significant independent genomic loci, associated with synaptic transmission. In particular, Hatoum et al. (2019) argue that their results suggest the latent Common EF factor to be "a genetically distinct cognitive construct that is particularly relevant to understanding the genetic variance in psychiatric disorders" (c.f. Friedman & Robbins, 2021).

1.4 Neuroimaging Correlates of Executive Functions

During the past decade, neuronal correlates of EFs have been subject to many neuroimaging studies, conducted primarily by functional magnetic resonance imaging (fMRI). In these studies, regarding the Common EF component mentioned above, a "multiple-demand" system of the primate frontal and parietal cortices has been identified. This system is active across multiple tasks (Figure 1) and concerned with cognitive control, namely the "structure and requirements of complex, multi-component behaviour" (Duncan, 2010). There are also several neuronal areas activated in response to specific EFs, reviewed by Friedman and Miyake (2017).

In the following, each study of this cumulative thesis is briefly summarized, indicating its importance for the research area. The published articles are also included, if copyright restrictions permit.

Figure 1: The "multiple demand" system of the primate brain. Reused from Duncan (2010) with permission. Abbreviations: AI/FO-anterior insula and adjacent frontal operculum, IFS-Inferior frontal sulcus, IPS-intraparietal sulcus, pre-SMA/ACC-pre-supplementary motor area and adjacent dorsal anterior cingulate, RPFC-rostrolateral prefrontal cortex.



2 Cognitive Branching Between More than Two Tasks (Heilbronner & Pollmann, 2010)

In 2007, Koechlin and Hyafil presented a new computational model of the human PFC. Briefly, this model proposes an anterior-to-posterior frontal hierarchy, in which the frontopolar cortex sits on top and coordinates different task-sets stored in lateral PFC. The model makes predictions concerning situations of so-called cognitive branching, the successful reuptake of a task-set after interruption by another task. Specifically, it is predicted that "only a single task-set can be maintained in a pending state at any one time" (Koechlin & Hyafil, 2007, p. 598), i.e., that individuals face severe problems when cognitive branching occurs between more than two tasks. This prediction goes against our everyday experience that most people can coordinate more than two tasks without problems. Rather, we hypothesized that WM load would be the critical factor in such situations. We therefore designed two experiments in which WM load and number of branching steps were varied independently of each other. Both of these experiments used simple mental arithmetic as stimulus material. Briefly, participants were required to solve up to four simple arithmetic tasks in a row, and remember both the result and the operator (addition, subtraction, multiplication, or division) of the calculation. Subsequently, participants had to use the information stored in WM in another calculation. In the first experiment, these subsequent calculations were to be performed in reverse order than the initial arithmetic tasks, in the second experiment, the order of the subsequent calculations was pseudo-randomized. The experimental manipulation by which we varied WM load consisted of some calculations having always the same result, whereas others always had the same operator. Results of these experiments clearly show that, in line with our hypothesis, humans are able to perform cognitive branching between more than two tasks, as long as WM load is kept low. Thus, it it appears that WM storage capacity determines the number of successful cognitive branching steps humans can successfully perform, and not an inbuilt structural limit.

Psychological Research (2010) 74:327–336
DOI 10.1007/s00426-009-0249-8

ORIGINAL ARTICLE

Is there a structural limit to 'branch' recursively between more than two tasks?

Urs Heilbronner · Stefan Pollmann

Received: 11 March 2009 / Accepted: 22 June 2009 / Published online: 8 July 2009

Abstract The term 'branching' refers to processes needed for successful reuptake of a task after interruption by another task. Based on a model of human prefrontal cognitive architecture, it has been postulated that people cannot branch recursively between more than two tasks due to a capacity limit built into the cognitive architecture (Koechlin and Hyafil in Science 318:594–598, 2007). As an alternative to a structural limit for recursive branching between more than two tasks we put forward the hypothesis that working memory capacity is the limiting factor in recursive branching. We tested this hypothesis by independently varying working memory load and number of recursive branching steps. Successful branching between up to four tasks was observed, as long as working memory load was kept low. Our data, thus, do not support the proposition of a structural limit to recursive branching beyond two tasks. Instead, they suggest that working memory capacity limit is the most important factor that limits the capacity for branching. We further observed that the requirement to retain task sets and task contents additively contributed to the difficulty of recursive branching. In a broader context, our data thus support working memory models that conceptualize working memory and executive functions not as separate modules, but as tightly interactive processes.

U. Heilbronner · S. Pollmann
Department of Experimental Psychology,
University of Magdeburg, Magdeburg, Germany

S. Pollmann (✉)
Institut für Psychologie II, Otto-von-Guericke Universität,
Postfach 4120, 39016 Magdeburg, Germany
e-mail: stefan.pollmann@ovgu.de

Introduction

Recently, Koechlin and Hyafil (2007) and Koechlin and Summerfield (2007) have presented a computational model of prefrontal cortex function, which postulates a hierarchical cascade of executive processes linked to distinct prefrontal structures. The model is based on functional neuroimaging data that suggest a frontal posterior to anterior hierarchy of control processes from sensory control in premotor cortex, via contextual and episodic control processes in posterior and anterior lateral prefrontal cortex to branching control in frontopolar cortex (Koechlin, Basso, Pietrini, Panzer, & Grafman 1999; Koechlin, Ody, & Kouneiher 2003). The model is a valuable attempt to understand the structure of executive processes and its relation to substructures of frontal cortex. It needs, however, further empirical tests to assess in how far it can explain human executive functioning. Here, we want to focus on a central assumption of the model that concerns limits of 'branching' between tasks. The term 'branching' refers to the reuptake of a task which was interrupted by another task, such as continuing a conversation after one has answered a telephone call.

Introducing the 'branching' paradigm, Koechlin et al. (1999) presented a pseudorandom sequence of letters. Subjects had to indicate whether the current letter followed the previously presented letter in the word 'tablet' (e.g. l–e would be correct, l–a incorrect). This task had to be carried out independently for uppercase and lowercase letters. Sequences of upper and lowercase letters were presented in unpredictable succession. Specifically in the branching condition, following a case change subjects had to compare the current letter with the last letter presented in the same case, e.g. in the sequence L-t-e-a-B the 'B' needed to be compared to the 'L' to check whether the

sequence L-B occurs in 'tablet'. This requirement to remember the last uppercase item for reuptake of the uppercase sequence when the lowercase sequence ended (and vice versa for the lowercase letters) distinguished branching from a dual task control condition in which the same letter sequence assessment was done separately for lower and uppercase sequences without remembering the last item from the previous same—case sequence. The task was further complicated by the requirement to decide whether every first letter indicating a lowercase change was the letter t. Therefore, not only the last item before a case change needed to be remembered, but also the different task sets following upper and lowercase changes. Koechlin et al. reported that the activation of frontopolar cortex selectively reflected branching demands, but not dual tasking.

In their model, Koechlin and Hyafil (2007) postulate that human beings are unable to branch between more than two tasks: "… the model especially predicts that the FPC (frontopolar cortex; explanation added by us) is unable to recursively perform cognitive branching—resuming a primary and secondary pending task after completion of a third task—because interferences supervene between the two pending tasks" (p. 597). According to the model, this endogenous branching limit is a consequence of the architecture of human cognition: "Computer simulations show that this neuronal system forces LPC (lateral prefrontal cortex, explanation added by us) and FPC neurons to potentially select and maintain only the two most rewarding task sets. The other task sets are discarded" (p. 596). However, that we should be principally unable to branch between more than two tasks is a strong claim which, to our knowledge, is not yet backed by empirical evidence. Because a structural branching limit would have strong implications for human cognition, we found this issue worthy of further investigation.

One alternative explanation, which could make branching between more than two tasks impossible, is the high demand on working memory that the branching paradigm affords. It may be that limitations in branching between several tasks are sufficiently explained by limits of working memory capacity, without the need to invoke a limit of tasks which can be held available for reuptake.

The capacity of working memory has been estimated to lie between about four chunks (Henderson 1972; Luck & Vogel 1997) to up to seven chunks (Lisman & Idiart 1995; Miller 1956). The exact number of chunks may vary depending on stimuli and task affordances (Cavanagh 1972; Cowan 2001). If we consider the branching paradigm, the minimum information that needs to be kept available to resume a previous task once a secondary task has been completed, is the nature of the primary task (the task set) and the last item which has been processed in the primary task. If we assume that for each task these two chunks of information (task set and last processed item) need to be kept available for branching and if we further assume that both compete for the same working memory store—verbal working memory in the case of the paradigm used by Koechlin et al. (1999) as well as in the present study—working memory capacity may be filled to its limit by branching between a primary and secondary task, consequently leading to difficulties in branching between more than two tasks.

Thus, the purpose of our experiments was to show that switching between more than two tasks held in working memory is quite possible and that an upper limit of the tasks between which branching can occur is set solely by working memory capacity and not by an endogenous limit rooted in human cognitive architecture, as proposed by Koechlin and Hyafil (2007). The original branching paradigm is not well suited to quantify working memory load. When asked to judge whether the current letter follows the memorized letter in the word 'tablet', subjects may recall the complete letter order (t-a-b-l-e-t), which would easily fill phonological working memory. However, even if they learn to memorize letter pairs from 'tablet' and retrieve them as chunks for the comparison process, working memory load may still be quite high because all possible letter pairs (t-a, a-b, b-l, l-e, e-t) need to be compared to the current letter pair. Thus, in the branching task used by Koechlin et al. working memory load is difficult to estimate, but it appears to be generally quite high.

In order to test the hypothesis that limits of branching are due to working memory overload, we therefore devised a protocol in which storage of task sets on the one hand and storage of task results on the other hand demanded only minimal working memory capacity. We chose mental arithmetic with one-digit numbers (e.g. $3 + 1 = 4$). When a sequence of arithmetic operations needs to be carried out, it appears reasonable that storing a particular operator (e.g. $+$) will consume only one chunk of working memory capacity. The same may be assumed for the result of an operation (e.g. 4). Such a task would then pose the minimal working memory demands possible: one item of working memory capacity for the operation (task set) plus one item for the result of a task, which is needed for later resumption. Even in this case, branching between more than two tasks may be already taxing for memory capacity. As an example of our experimental protocol, consider the following example sequence of tasks:

$3 + 1 =$ followed by $6 - 1 =$ followed by $2 \times 1 =$

Subjects were instructed that after carrying out each of these calculations, they would be required to retrieve the

results and the respective operators of previously presented tasks in backward order (information from last task to be used first). This information was needed to carry out further calculations. One sequence of tests to assess working memory content could look as follows:

A [Last result] [Last operator]2 = followed by
B [Penultimate result] [Penultimate operator]1 = followed by
C [First result] [First operator]1

Those instructions would, in the present example, translate to sequentially solving the following calculations:

A $2 \times 2 =$ followed by
B $5 - 1 =$ followed by
C $4 + 1 =$ in mind.

Of course, one would only succeed in solving these later tests if all information were correctly recalled. The aforementioned tests would thus require that one stores three results (4, 5, 2) and three operators ($+$, $-$, $\times$) in working memory and uses them in the right order to solve the test equations. Given that working memory capacity may already be filled to its limit in this very simple example of branching, it is obvious that limits of branching between more than two tasks may simply be a consequence of limited working memory capacity, without the need to invoke additional structural processing limits. This is even more the case if more complex tasks are used, as in the literature on branching (Koechlin et al. 1999; see above). In order to demonstrate that branching can occur between more than two tasks if working memory is not overloaded, we kept the results of the equations constant (e.g. $a + b = x$, $c-d = x$, $e \times f = x$, $g/h = x$). Because only one result needs to be memorized in this case for all four equations, there may be enough working memory capacity to store even all four different operators ($1 + 4 = 5$ items in working memory). As a further test of the assumption of a specific limit for branching between task sets, we tested the equivalence of operators (task sets) and results as items in working memory, by also testing the reverse case, in which the operator was kept constant and the results varied (e.g. $a + b = w$, $c + d = x$, $e + f = y$, $g + h = z$). Moreover, we varied the number of tasks that had to be solved sequentially from one to four (Experiment 1). One would expect a sharp increase in error rate when more than two tasks are used if Koechlin and Hyafil (2007) are correct in their assumption of a structural limit to recursive branching. In contrast, our hypothesis that branching deficits are due to working memory capacity limits predicted that there are no qualitative differences between operators or results as working memory contents.

Experiment 1

Method and materials

Participants

In Experiment 1, 26 subjects (18 female) participated. All participants had normal or corrected vision and were either university students or graduates. They received course credit or monetary compensation for their participation. One subject was excluded from analysis as her performance in the No Change condition (minimal working memory load; see below) did not exceed 50%. The age of the remaining subjects ranged from 19 to 38 years (mean 24.3 years).

Apparatus and stimuli

Stimuli were presented on a 24-inch flat-screen display connected to a PC. Experiments were carried out in a sound-proof chamber. The only source of illumination was a desk-lamp located behind the flat-screen display. Viewing distance was 110 cm. Experimental software was written in the Python programming language using the software library PyEPL (Geller, Schlefer, Sederberg, Jacobs, & Kahana 2007; http://pyepl.sourceforge.net). Subjects entered their responses via a numerical keypad. Stimuli were presented in white font on a black background (vertical viewing angle was approximately 2°).

Procedure

Subjects first had to solve a number (up to four) of consecutive arithmetic tasks (task part of the experiment; see Fig. 1) and enter the respective results by key press in the number block of a standard keyboard. Each response was followed by feedback, indicating whether the result was correct or not and displaying the correct result. The results were positive natural numbers between 0 and 7. For details see "Appendix". Immediately following the completion of one to four of these tasks, a test part had to be completed (equal number of tests and tasks). The tests were either $x?1 =$ or $x?2 =$ (determined by randomization), where x stood for the result, ? stood for the operator to be retrieved from memory and 1 and 2 were numerals. The result, which always was a single-digit positive natural number, was to be entered by key press. There was no feedback about the correctness of subjects' answers in the test part of the experiment, only a simple 'OK' message was displayed to acknowledge the response.

The experiment consisted of 48 blocks. One block consisted of four trials. In the first trial, one task was

presented (3 + 1 = in Fig. 1a) followed by a leftward pointing arrow (←; reminiscent of the 'back' symbol) after which a test was presented (x?1 = in Fig. 1a). The correct answer to his test would thus be 4 (result of last task) + (operator of last task) 1 = 5. In the second trial, two tasks were presented consecutively (8/2 = followed by 3 + 1 = in Fig. 1b), followed by an arrow, a test (*x*?2 = in Fig. 1b), another arrow and another test (*x*?1 = in Fig. 1b). The first test (*x*?2 =) corresponded with the last task, the second test (*x*?1 =) with the penultimate task. The correct answer to be entered on the first test would thus be: 4 (result of last task) + (operator of last task) 2 = 6. On the second test, the correct answer would be: 4 (result of penultimate task)/(operator of penultimate task) 1 = 4. Trials three and four (not shown in Fig. 1) were constructed analogously, containing three respectively four tasks and corresponding numbers of tests. Owing to the complexity of the instruction, a training session, which consisted of a shortened version (16 blocks) of the experiment, was completed before the experiment. All results presented here refer exclusively to the test parts of the experiments.

In equal numbers of blocks, either both operators and results changed between tasks, only operators or only results changed, or both operators and results were held constant. Thus, we had a factorial design with Operator Change (yes, no), Result Change (yes, no) and Number of Tasks (1–4) as factors. The different conditions that resulted from combination of the factors Operator Change and Result Change were named No Change (neither operators nor results change), Results Only (only results change), Operators Only (only operators change), and Both Change (both operators and results change). Figure 1 is, thus, an example of an Operators Only block in which operators (but not results) change across the different trials. The order of blocks was pseudorandomized so that not more than two identical conditions followed each other. For each task, the respective operator/result combination was randomly drawn from a pool (see "Appendix"; Tables 4, 5, 6, 7). After data acquisition was completed, we also investigated whether certain results of test tasks or sequences of results of test tasks (i.e. the correct answers to *x*?1 = and *x*?2 =) had occurred more often than others. Our analyses show that the number 4 was the most frequent result in approximately 20% of all test tasks. Based on the distribution of results obtained by analyzing relative frequencies of each factor level combination, we parsed the data set for the most likely sequence of results. We did not find evidence of certain sequences occurring particularly often. The sequence that occurred most often was 3–3 in condition No Change when two test tasks had to be answered (12% of sequences of two tasks in this condition). All other investigated sequences had lower percentages of occurrence.

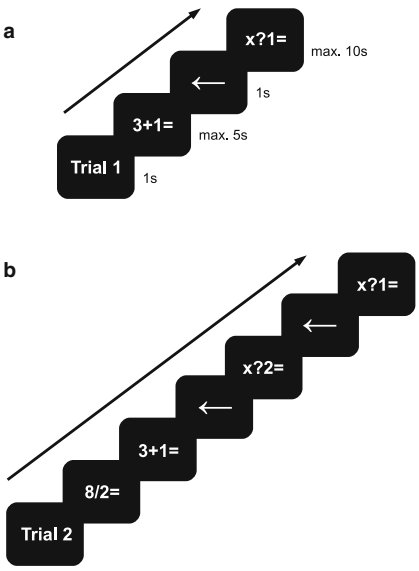


Fig. 1 Schematic illustration of first (**a**) and second (**b**) trial in Experiment 1. The presentation duration is shown next to the first trial. Feedback messages to the answers of subjects are not shown. For further explanations see text

Data analysis

Data from the training sessions were not formally analyzed. Data were aggregated according to the parameter of interest (see below) and analyses of variance (ANOVAs) followed by Tukey's honest significant difference post hoc test were performed. Where indicated, Student's *t* tests were used to compare combinations of conditions. To evade problems concerning the assumption of normal distribution, relative proportions were transformed to logits before analysis (Johnson 1949). For clarity, however, raw proportions are shown in graphs. The statistical software package R (Ihaka & Gentleman 1996, Version 2.8.1) was used for analysis. All data are depicted as mean ± standard error of the mean.

Results

Figure 2 shows error rates across trials. In Table 1, the results of the three-way repeated measures ANOVA (Factors Operator Change × Result Change × Number of Tasks) are reported.

Not surprisingly, all main effects were significant. This showed that both change manipulations as well as the

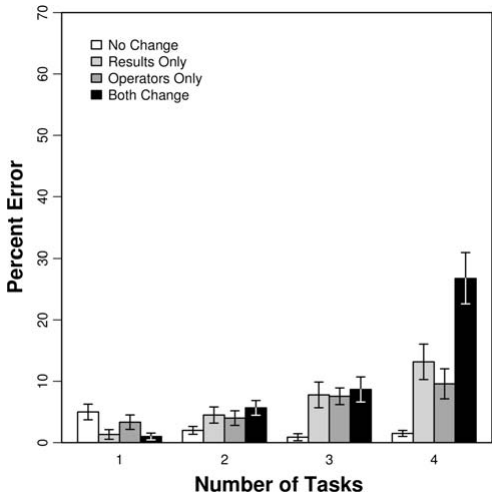


Fig. 2 Error rates in Experiment 1 across levels of the factor number of tasks

Table 1 Three-way ANOVA Operator Change × Result Change × Number of Tasks (Experiment 1; Chg, Change)

Error rate (Logit)	*F*	*df*	*P*	Sign.
Main effects				
Operator Chg	29.55	1, 24	<0.001	Yes
Result Chg	29.65	1, 24	<0.001	Yes
No. of tasks	23.59	3, 72	<0.001	Yes
Interactions				
Operator Chg × Result Chg	0.77	1, 24	0.390	No
Operator Chg × No. of tasks	13.30	3, 24	<0.001	Yes
Result Chg × No. of tasks	28.68	3, 24	<0.001	Yes
Operator Chg × Result Chg × No. of tasks	3.48	3, 72	0.020	Yes

number of tasks increased errors. Performance was comparable in conditions in which only operators or results changed ($P > 0.983$ in all post hoc tests comparing the Results Only and Operator Only conditions across levels of the factor Number of Tasks). The interaction of Operator Change × Result Change was not significant indicating that the number of operators and results to be retained in working memory posed additive burdens on processing. These two features, comparable performance decrements and additivity, suggest that we successfully designed a task in which task sets and task contents to be retained for future task reuptake draw comparably on a common working memory capacity.

All interactions involving the factor Number of Tasks were significant, reflecting the observation that both single change conditions showed increased error rates with increasing number of tasks, in contrast to the No Change condition. Furthermore, as reflected by the three-way interaction, a particularly strong increase was observed when both operator and result changes had to be retained across four tasks. The latter was expected if retention of task sets and contents draw on the same limited working memory capacity. Given that operator and result changes showed equal and additive performance decrements, we can quantify working memory load by the sum of the number of operators and results to be retained.

When the data are grouped by working memory load, it becomes evident that error rate increased from 1.5% at a working memory load of 2 items to 11.4% at a load of 5 items and an increase to 26.8% from 5 to 8 items (see Fig. 3).

We were also interested in which serial position errors occurred when four tasks had to be remembered. Therefore, we examined the position of tasks selectively for the four tasks level of the factor Number of Tasks by a three-factorial repeated measures ANOVA Operator Change × Result Change × Position of Test (Table 2). Please note that Position Four refers to the last test presented (information from the first task was required). The error rates are shown in Fig. 4. Across conditions, a post hoc test revealed significant differences between positions 1 and position 2, position 1 and position 3 as well as between position 1 and position 4 (all $P < 0.001$). The remaining positions did not

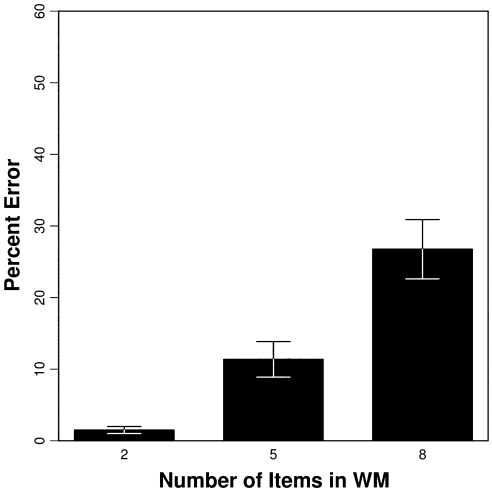


Fig. 3 Error rates in Experiment 1 on four task trials aggregated according to the number of items that have to be kept in working memory (WM) (condition No Change—two items, condition Operators Only and Results Only—five items, condition Both Change—eight items)

Table 2 Three-way ANOVA Operator Change × Result Change × Position of Test on factor level four tasks (Experiment 1; Chg, Change)

Error rate (Logit)	*F*	*df*	*P*	Sign.
Main effects				
Operator Chg	36.78	1, 24	<0.001	Yes
Result Chg	72.35	1, 24	<0.001	Yes
Pos. of Test	20.91	3, 72	<0.001	Yes
Interactions				
Operator Chg × Result Chg	0.019	1, 24	0.893	No
Operator Chg × Pos. of Test	9.76	3, 72	<0.001	Yes
Result Chg × Pos. of Test	11.22	3, 72	<0.001	Yes
Operator Chg × Result Chg × Pos. of Test	2.34	3, 72	0.080	No

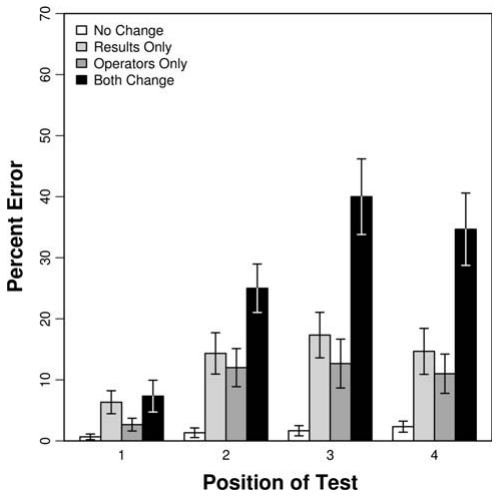


Fig. 4 Error rates in Experiment 1 on four task trials according to the position of test. Note that position 4 refers to the first task presented

differ significantly from each other (all $P > 0.596$). Even if data were aggregated across all three change conditions, there was no indication for less errors in position 4 versus 3 ($t = 1.33$, $df = 24$, $P = 0.195$).

Discussion

We developed a new branching paradigm in which working memory demands can be varied independently of the number of tasks. Subjects carried out up to four simple calculations and had to keep the operators and results of each calculation in working memory. In a test phase, directly after the calculations, they had to recall the results and/or operators for new calculations. In a typical branching task, both the task set (here the operator) and the last content of a task (here the result of the calculation) need to be remembered. Even if both the task set and the content can be very efficiently coded, keeping two tasks in mind for future reuptake while working on a third task may load working memory almost to its limit. Therefore, we introduced a condition in which either the results or the operators of a series of calculations did not change, so that only one result or operator needed to be remembered. In these conditions, branching between up to four tasks was possible, with about 10% of errors. In contrast, if both operators and results had to be remembered for each task, branching between four tasks led to about 25% of errors, while branching between three tasks was still in the range of about 10% of errors. Operators and results loaded working memory to a comparable degree and both kinds of items drew additively on working memory capacity.

The data suggest that branching between our tasks becomes taxing with five chunks to be remembered with a further severe deterioration when eight chunks need to be kept in memory. This measure of working memory capacity, although, may be inflated because subjects may have stored part of the information in longterm memory (Cowan 1999). It should be remembered that we were interested in memory limits of branching rather than an assessment of 'pure' working memory capacity. It is further noteworthy that even in the most difficult condition, error rates were below 30%. Given that the probability of guessing the right response was very low [1/10 alternatives (numbers 0–9)], even a 30% error rate may not be sufficient to say that participants were generally unable to branch recursively between more than two tasks.

Experiment 2

Experiment 1 clearly showed that branching between more than two (actually up to four) tasks is possible. It might be argued, although, that Experiment 1 is not a fair test of a structural branching limit because the sequence of test reuptake was completely predictable, so that subjects may have retained a sequence of tasks instead of individual task sets and contents attached with a task. In order to test this possibility, we carried out Experiment 2. Here, we investigated if branching between more than two tasks would still be possible if tests were presented unpredictably, i.e. task sets and associated contents had to be retrieved in random order.

Methods

Experiment 2 followed Experiment 1 in all methodological details except the following. Experiment 2 was carried out

by 16 subjects (twelve female), who had not taken part in Experiment 1. Again, one female subject was excluded from analysis as her performance in the No Change condition (no working memory load) did not exceed 50%. The age of the remaining participants ranged from 21 to 30 years (mean 23.9). Two blocks of one subject were excluded due to technical problems. We tested only sequences of four tasks in Experiment 2, because Experiment 1 had shown that branching between less than four tasks was unproblematic. The experiment consisted of 64 blocks, the training session of 32 blocks. One block comprised four consecutive tasks, each appearing in a different color. The order of colors in the task part was always the same (yellow, blue, red and green). In contrast, the colors in which test tasks were presented varied in random order. Subject had to solve these tests by mentally inserting the information of the task that had previously been presented in the same color (See Fig. 5 for a schematic description of one block. The stimuli used in Experiment 2 are shown in Table 8 (see "Appendix").

Again, we investigated whether certain results or sequences of results of test tasks had occurred more often and thereby made certain results or sequences of results more likely than others (see "Experiment 1"). As in Experiment 1, the highest proportion of results was the number 4 (approximately 22% of all test tasks). Analyzing the data set for the most likely sequence of results did, again, not support the notion of an accumulation of certain sequences. The sequence that occurred most often was 4-4-4-4 in condition No Change (3.75% of all sequences in this condition).

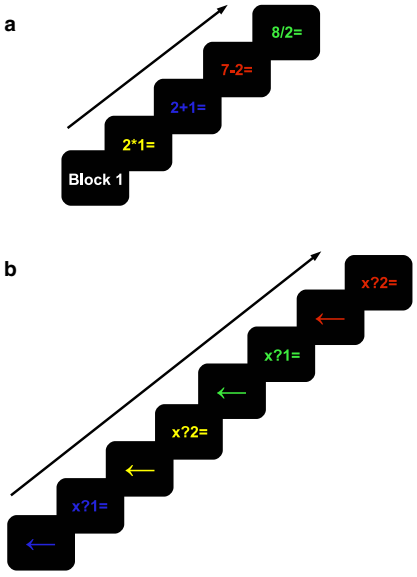


Fig. 5 Schematic illustration of one block in Experiment 2. Task (**a**) and test part (**b**) are shown. Feedback messages to the answers of the subjects are not shown. Presentation times are the same as in Experiment 1

Results

Figure 6 shows the error rates according to the serial position of the previously presented tasks. The results of the three-way repeated measures ANOVA (Factors Operator Change × Result Change × Position of Task) are reported in Table 3. Again, the main effects of operator and result change were significant, indicating that the numbers of operators and results to be retained affected performance. The effect of position was also significant. Figure 6 shows that performance was best for the first and the last task presented suggesting primacy and recency effects. Post hoc tests between levels of the factor Position of Task, aggregated over all change conditions confirmed that there were significantly less errors in position 1 versus 2 ($t = -3.78$, $df = 14$, $P = 0.002$) as well as position 4 versus 3 ($t = 2.75$, $df = 14$, $P = 0.016$). None of the interactions reached significance. Thus, operator and result changes were again additive in their effects on error rate. In addition, the change conditions did not interact with position of tasks. In Fig. 7, data are again displayed according to total memory load. As in Experiment 1, there was a steep increase in errors with increasing working memory load.

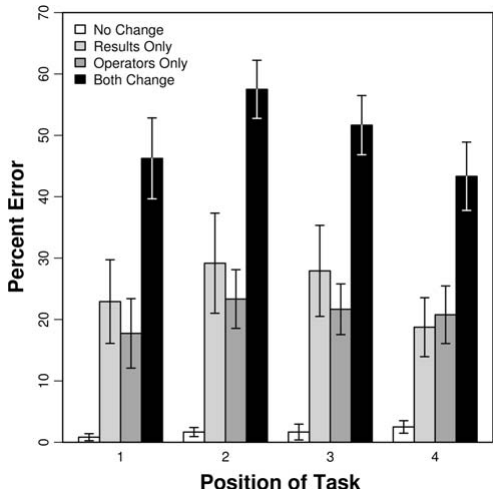


Fig. 6 Error rates in Experiment 2 according to the serial position of tasks

Table 3 Three-way ANOVA Operator Change × Result Change × Position of Task (Experiment 2; Chg, Change; Pos, Position)

Error rate (Logit)	*F*	*df*	*P*	Sign.
Main effects				
Operator Chg	130.96	1, 14	0.001	Yes
Result Chg	73.53	1, 14	0.001	Yes
Pos. of Task	3.85	3, 42	0.016	Yes
Interactions				
Operator Chg × Result Chg	0.32	1, 14	0.580	No
Operator Chg × Pos. of Task	1.28	3, 42	0.292	No
Result Chg × Pos. of Task	1.35	3, 42	0.271	No
Operator Chg × Result Chg × Pos. of Task	0.410	3, 42	0.747	No

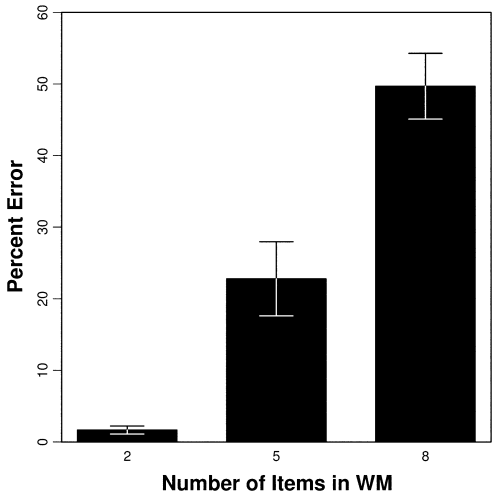


Fig. 7 Error rates in Experiment 2 aggregated according to the number of items that have to be kept in working memory (WM) (condition No Change—two items, condition Operators Only and Results Only—five items, condition Both Change—eight items)

Discussion

Even if the sequence of previous task reuptake was unpredictable, branching was still possible for up to four tasks, clearly in disagreement with the concept of a structural limit of two tasks, as proposed by Koechlin and Hyafil (2007). Error rates increased in comparison with Experiment 1, but successful branching was clearly achieved up to a working memory load of five items. As in Experiment 1, performance was dramatically affected by working memory load, leading to about 50% errors at a working memory load of eight items. The overall higher error rates may indicate that subjects may indeed have been successful in using sequence information for more effective chunking in Experiment 1, which was not possible in Experiment 2. Again, although an error rate of about 20%, observed at an estimated working memory load of 5 chunks, indicates a substantial difficulty level, it does not appear fair to say that subjects were unable to branch between more than two tasks under this load, in contradistinction to the concept of a structural recursive branching limit.

General discussion

In two experiments with a novel branching paradigm, we investigated the limits of branching between different tasks. Branching means that an ongoing task is interrupted for the execution of a second task, but later taken up again at the exact point where it was interrupted. Thus, branching is not only task switching, but affords that both the task set and the parameters of the task at the moment of interruption are stored in memory while the next task is being carried out. In a recent computational model of prefrontal cortex function, Koechlin and Hyafil (2007) postulated, based on as yet unpublished behavioral data, that humans have a structural limit of one task which can be held available in the above manner while a second task is carried out, hence, that a limit of branching between two tasks arises.

The current data clearly do not support such a structural limit. Instead, we showed that a limit in branching between tasks is closely tied to working memory capacity. In the current experiments, remembering task sets and task parameters contributed additively to performance decrements, suggesting that both kinds of information relied on a common (probably verbal) working memory store.

It may be that our data do not reflect 'pure' working memory capacity. Indeed, in Experiment 2, we observed an indication for a primacy effect (reduced errors for the parameters of the first task) which has been taken as an indicator of retrieval from long-term memory (Atkinson & Shiffrin 1971). This, however, does not change the interpretation of our data. In many everyday situations, we may be able to store a limited amount of information in long-term memory and use this capacity in addition to the items stored in working memory. Nevertheless, the still very limited overall capacity sets severe limits to 'branching', as our data show.

In their paper, Koechlin and Hyafil (2007) acknowledge that branching may occur in certain situations between more than two tasks. They suggest that recursive branching may become possible with expertise where spatial or verbal recursive tree structures may guide branching. Expertise may indeed be a crucial point. Working memory capacity is easily overloaded even in branching between tasks that

are highly overlearned, such as simple mental arithmetic, and will even faster reach its limit in less over learned tasks. However, we would turn the argument around in claiming that there is no structural limit to recursive task reuptake (i.e. branching), but that simply working memory capacity determines the limit of recursive branching of which we are capable. This interpretation has several advantages. First, it is more parsimonious in that it relies only on the well-established concept of working memory capacity limits. Second, it does not need to invoke speculative alternative ways which may make recursive branching possible. Finally, it is backed by empirical evidence, which, to our knowledge, is not yet available for the assumption of the structural two-task branching limit postulated in the Koechlin and Hyafil model.

This is not to say that we suggest the Koechlin and Hyafil model is not suitable to explain prefrontal function. In contrast, we regard the model as a valuable attempt to explain how different prefrontal structures support executive functions. It is not clear whether a structural recursive branching limit is a necessary aspect of the computational model of prefrontal function of Koechlin and Hyafil. Further model developments may possibly show that the model can accommodate recursive branching between more than two tasks with low working memory demands without changing its principal architecture.

The additivity of task content and task set demands on working memory in the current experiments speaks against the possibility of a structural task processing limit which is separate from working memory capacity. It needs to be mentioned that in our task, both task set and results are easily verbalized and probably retained in verbal working memory. It remains to be seen whether retaining visuo-spatial task parameters and verbal task sets would also add up to increase the difficulty of branching. The additivity of retained task sets and task parameters, in the present study suggests that task sets and contents are intricately linked in working memory (Cowan 1999; Hazy, Frank, & O'Reilly 2006; Lovett, Reder, &Lebiere 1999). Future imaging experiments may investigate whether lateral frontopolar cortex, which has been shown to be activated by branching tasks, will respond differentially to working memory demands posed by task contents and task sets during branching.

Acknowledgments The authors would like to thank Michael Hanke for his programming support. This work was supported by a grant of the Deutsche Forschungsgemeinschaft (PO548/6-1-2).

Appendix A: stimulus material in Experiment 1

See Tables 4, 5, 6, 7.

Table 4 Tasks in Experiment 1: condition No Change

Operator	Result	Tasks			
+	2	2 + 0=	1 + 1=	0 + 2=	
+	4	4 + 0=	2 + 2=	1 + 3=	3 + 1=
−	2	3 − 1=	4 − 2=	5 − 3=	6 − 4=
−	4	5 − 1=	6 − 2=	7 − 3=	8 − 4=

Table 5 Tasks in Experiment 1: condition Results Only

Operator	Result	Tasks	
+	3	2 + 1=	1 + 2=
+	4	2 + 2=	3 + 1=
+	5	3 + 2=	4 + 1=
+	6	2 + 4=	5 + 1=
−	3	4 − 1=	5 − 2=
−	4	5 − 1=	6 − 2=
−	5	7 − 2=	6 − 1=
−	6	7 − 1=	8 − 2=

Table 6 Tasks in Experiment 1: condition Operators Only

Operator	Result	Tasks	
+	2	1 + 1=	2 + 0=
+	4	2 + 2=	3 + 1=
−	2	3 − 1=	4 − 2=
−	4	5 − 1=	6 − 2=
×	2	1 × 2=	2 × 1=
×	4	2 × 2=	4 × 1=
/	2	4/2=	2/1=
/	4	8/2=	4/1=

Table 7 Tasks in Experiment 1: condition Both Change

Operator	Result	Tasks
+	3	2 + 1=
+	7	4 + 3=
−	5	7 − 2=
−	6	9 − 3=
×	0	0 × 1=
×	1	1 × 1=
/	2	4/2=
/	4	8/2=

Appendix B: stimulus material in Experiment 2

For each task, the respective operator/result combination was randomly drawn from a pool (see Table 8).

Table 8 Tasks in Experiment 2

Operator	Result	Tasks			
+	2	1 + 1=	2 + 0=	0 + 2=	
	3	2 + 1=	1 + 2=	3 + 0=	0 + 3=
	4	2 + 2=	3 + 1=	1 + 3=	
	5	3 + 2=	2 + 3=	1 + 4=	4 + 1=
	6	2 + 4=	4 + 2=	1 + 5=	5 + 1=
	7	4 + 3=	3 + 4=	5 + 2=	2 + 5=
−	2	3 − 1=	4 − 2=	5 − 3=	6 − 4=
	3	4 − 1=	5 − 2=	7 − 4=	8 − 5=
	4	5 − 1=	6 − 2=	7 − 3=	8 − 4=
	5	6 − 1=	7 − 2=	8 − 3=	9 − 4=
	6	7 − 1=	8 − 2=	9 − 3=	
×	1	1 × 1=			
	2	2 × 1=	1 × 2=		
	4	2 × 2=	4 × 1=		
/	2	4/2=	2/1=	8/4=	
	4	4/1=	8/2=		

Condition No Change

Operators were either + or −, results were either 3, 4, 5 or 6 (four blocks each), so that eight different combinations resulted (two blocks each).

Condition Results Only

The results were either 3, 4, 5 or 6 (four blocks each), operators remained constant. Operators were either + or − (eight blocks each), so that eight different combinations resulted (two blocks each).

Condition Operators Only

Operators were either +, −, × or / (four blocks each), results remained constant. The results were either 2 or 4 (eight blocks each) so that eight different combinations resulted (two blocks each).

Condition Both Change

Both operators (either +, −, × or /) and results (either 1, 2, 3, 4, 5, 6 or 7) were different. All four operators were presented; the respective results were determined by randomization.

References

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 25*, 82–90.

Cavanagh, J. P. (1972). Relation between the immediate memory span and the memory search rate. *Psychological Review, 79*, 525–530.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory (pp. 28–61)*. New York: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–114.

Geller, A. S., Schlefer, I. K., Sederberg, P. B., Jacobs, J., & Kahana, M. J. (2007). PyEPL: a cross platform experiment-programming library. *Behavior Research Methods, 39*, 950–958.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006). Banishing the homunculus: making working memory work. *Neuroscience, 139*, 105–118.

Henderson, L. (1972). Spatial and verbal codes and the capacity of STM. *The Quarterly Journal of Experimental Psychology, 24*(4), 485–495.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics, 5*(3), 299–314.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika, 36*(1/2), 149–176.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature, 399*, 148–151.

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*, 594–598.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302*, 1181–1185.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11*, 229–235.

Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science, 267*, 1512–1515.

Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R-perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 135–182). New York: Cambridge University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279–281.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

3 Rapid Event-Related Near-Infrared Spectroscopy for the Study of Cognitive Aging (Heilbronner & Münte, 2013)

Many studies that research EFs with neuroimaging techniques use fMRI. Albeit offering superior spatial resolution, there are also disadvantages to this imaging modality. For example, neuroimaging techniques that are not confined to the supine position of individuals allow for greater flexibility in experimental design. Also, studies that seek to assess older individuals are faced with the problem that this group often fulfills exclusion criteria for fMRI, such as cardiac pacemakers or metal implants. In such cases, an alternative to fMRI is functional near-infrared spectroscopy (fNIRS), a neuroimaging technique that uses near-infrared light to visualize changes in oxy- (HbO) and deoxyhemoglobin (HbR) on the cortical surface (Jöbsis, 1977, Hillman, 2007). As in fMRI, these changes are interpreted as readout of neuronal activity. Early fNIRS studies of EFs used the so-called block design (e.g., Herrmann et al., 2005), in which an activity is carried out for an extended period of time (e.g., 30 s), to obtain a robust hemodynamic signal. In fMRI research, this experimental design has been superseded by the so-called "fast event-related" design, which (together with a model-based analysis strategy) permits rapid presentation of brief (e.g., 100 ms) stimuli (Burock et al., 1998, Zarahn et al., 1997). Fast event-related designs have several advantages over block designs. These include the ability to post-hoc sort single responses of individuals participants, allowing e.g., to separate successful from unsuccessful inhibitory responses. This is especially helpful when two experimental groups differ in their trial-by-trial response patterns. Although there have been several successful attempts to use fast event-related designs also in fNIRS (Obrig et al., 2000, Schroeter et al., 2002, Schroeter et al., 2003, Boecker et al., 2007, Taga & Asakawa, 2007), it had not been shown yet that age-associated changes can be detected using this method. In this study, we compared young (<30 years) and older (>60 years) individuals using a classic Go/NoGo experiment assessed by fast event-related fNIRS. In the behavioral task, participants were instructed to press a button in response to a frequently occurring stimulus, and suppress their response to another, less frequently occurring, stimulus. Results of this study show that age-associated changes in frontal activation can be detected

using fast event-related fNIRS, taking advantage of the possibility to disentangle different response patterns from neuronal activation. In contrast to young individuals, older individuals showed activation in the dorsolateral PFC instead of the inferior frontal cortex. Moreover, our results are consistent with fMRI activation patterns between age groups (Nielson et al., 2002). Fast event-related fNIRS may thus be useful in future longitudinal studies on the effects of aging.

NeuroImage 65 (2013) 408–415

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Rapid event-related near-infrared spectroscopy detects age-related qualitative changes in the neural correlates of response inhibition

Urs Heilbronner [a,b,*], Thomas F. Münte [c]

[a] DZNE, German Center for Neurodegenerative Diseases, Haus 15, Otto-von-Guericke University Magdeburg, Medical Faculty, Leipziger Straße 44, 39120 Magdeburg, Germany
[b] Department of Psychiatry and Psychotherapy, Section on Psychiatric Genetics, University Medical Center Göttingen, Georg August University, Von-Siebold-Str. 5, 37075 Göttingen, Germany
[c] Department of Neurology, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

ARTICLE INFO

Article history:
Accepted 22 September 2012
Available online 4 October 2012

Keywords:
NIRS
Go/NoGo
Executive
Aging
Phenotyping

ABSTRACT

Near-infrared spectroscopy (NIRS) is a promising neuroimaging tool for the study of human cognition. Here, we show that event-related NIRS is able to detect age-related differences in the neural processing in a simple visual Go/NoGo task using a relatively fast (stimulus onset asynchrony approx. 1.4 s) event-related design together with a model-based analysis approach. Subjects were healthy young (<30 years) and elderly (>60 years) adults. Behaviorally, old adults were slower but more accurate than young adults. The event-related analysis approach of NIRS data allowed us to contrast activation of successfully inhibited NoGo stimuli with that of correctly answered Go stimuli. Both age-groups showed frontal activation differences between these events in oxy- (HbO; increase) and deoxyhemoglobin (HbR; decrease). Between age groups, differences in HbR were found in right dorsolateral frontal (old>young), right temporal/postcentral/precentral and left precentral/inferior frontal (young>old) channels. These differences are in line with age-associated activation changes in inhibition detected with functional magnetic resonance imaging. The present study successfully separated the neural correlates of response inhibition from errors of commission/omission and provides data from multiple simultaneously recorded optodes. Furthermore, these results demonstrate the feasibility of using NIRS to investigate neural processes related to aging and dementia, in particular in patients for which other neuroimaging techniques are contraindicated. In the future, functional phenotyping of successful aging in respect to executive performance may be feasible.

Introduction

Near-infrared spectroscopy (NIRS) is an optical technique that can be used to assess variations in the content of oxy- (HbO) and deoxyhemoglobin (HbR) in superficial layers of brain tissue through the intact skull (Jöbsis, 1977; for review see Hillman, 2007). NIRS measurements are correlated with signals obtained by functional magnetic resonance imaging (fMRI) in which a closely related vascular signal, the BOLD contrast, is measured (for review see Steinbrink et al., 2006). Although inferior to fMRI in spatial resolution, NIRS has several advantages: there are few restrictions in the type of experimental paradigm that can be investigated and subjects are not confined to the supine position which facilitates application of complex stimuli and response characteristics (e.g. Strangman et al., 2002). Also, unlike fMRI, subjects with cardiac pacemakers, metal implants or large tattoos need not to be excluded. The time resolution of NIRS is superior to that of fMRI in that the sampling rate is an order of magnitude higher, although a similar, inherently slow, hemodynamic response is measured with both techniques. In fMRI research, a significant leap has been made through the use of event-related compared to block designs (Josephs et al., 1997). Whereas in the latter a task of interest is performed for a period of 30 s or so, event related designs allow to measure brain responses to individual, short duration stimuli. Event-related designs have several advantages compared to block designs, among which is the ability of post hoc trial-by-trial sorting according to subject performance and a greater overall flexibility in experimental design (Burock et al., 1998; Zarahn et al., 1997). A number of event-related NIRS experiments have shown that the same rationale can also be applied to this imaging technique (Boecker et al., 2007; Obrig et al., 2000; Schroeter et al., 2002; Taga and Asakawa, 2007). Schroeter et al. (2004) have shown that the length of the intertrial interval can be shortened to 2 s without a reduction of the HbR amplitude. Different analysis approaches have been explored with regard to NIRS data. Many researchers analyze NIRS studies by simply measuring concentration changes in HbO and HbR in response to task blocks (e.g. Herrmann et al., 2005) or single events (Boecker et al., 2007). A different way is a model-based general linear model

* Corresponding author at: University Medical Center, Georg August University, Department of Psychiatry and Psychotherapy, Section on Psychiatric Genetics, Von-Siebold-Str. 5, 37075, Göttingen, Germany. Fax: +49 551 39 14272.
E-mail address: urs.heilbronner@med.uni-goettingen.de (U. Heilbronner).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.09.066

(GLM) approach which is widely used in fMRI research. Plichta et al. (2007) have shown that analyzing fast event-related NIRS data with this approach is feasible.

A research area for which the above mentioned advantages of NIRS seem especially important is the field of aging and dementia as many elderly and demented patients have contraindications against examinations with fMRI or show excessive movement artifacts. Also, the versatility of the NIRS technique appears to be a major asset when higher cognitive functions are to be investigated. In the past, cognitive inhibition has been of particular interest to scientists, as a decline on various levels of this domain was observed both in normal aging (reviewed by Hasher and Zacks, 1988 and by West, 1996; see also Nielson et al., 2002; Spieler et al., 1996) and early Alzheimer's disease (reviewed by Amieva et al., 2004). Inhibition is a separable entity within the metacognitive executive function framework (Miyake et al., 2000), which is considered to have general importance for a variety of other downstream cognitive functions such as language, reading, memory, attention and working memory. However, there appear to be different aspects of cognitive inhibition, which, in respect to aging, may show different trajectories (reviewed by McDowd, 1997). A widely used paradigm in cognitive neuroimaging is the Go/NoGo task in which a prepotent motor response has to be actively inhibited. As there are a number of event-related fMRI studies using various Go/NoGo tasks (see Discussion), we chose this paradigm for our feasibility investigation with respect to cognitive inhibition. Moreover, there exists a large body of literature which compares young and old populations using fMRI research, also with respect to executive function (for a meta-analysis see Spreng et al., 2010). Thus, we are interested in whether event-related NIRS with short intertrial intervals and a GLM based analysis approach is similarly able to detect age-associated changes in neural correlates of simple motor response inhibition. To this end, we measure healthy young (<30 years) and elderly (>60 years) subjects performing a simple Go/NoGo task to explore potential changes in brain activation evoked by response inhibition between age-groups using rapid event-related NIRS. As mentioned, our primary aspect of interest is to test the feasibility of this expansion to the rapid event-related nature of the experiment, similar to investigations in fMRI research (Burock et al., 1998; Dale, 1999; Dale and Buckner, 1997). There are several competing theories on functional activation changes during healthy cognitive aging (reviewed by Dennis and Cabeza, 2008) which make different predictions, often tied to specific paradigms. Apart from the general feasibility hypothesis, we do therefore not predict the exact nature of the age effects between groups but are rather interested in whether both within- as well as between-group differences can be observed using this approach. In an attempt to align the obtained results with existing evidence on the cognitive neuroimaging of healthy aging (see Spreng et al., 2010), we therefore conduct post-hoc exploratory analyses, focusing on differential activation patterns in the two age groups.

Material and methods

Participants

Young subjects (<30 years) were recruited via a student social network, and old subjects (>60 years) were recruited by contacting representatives of the senior lecturing program of the University of Magdeburg and through the subject database of the German Center for Neurodegenerative Diseases (Magdeburg). All subjects were informed about the procedures of the experiment in written form. Additionally, subjects had to fill out a short questionnaire about their general health status and life habits. Data from 22 old and 27 young participants were acquired. After data acquisition 14 (6 old; 8 young) participants were excluded from the analysis because of technical problems during data acquisition (4 young), health reasons (assessed by a medical doctor (TFM): 3 old due to depression and/or psychotropic medication; 2 young, one participant with frequent migraine attacks, one participant due to recreational drug use on the evening preceding the experiment), or because they exceeded the a-priori set criterion of 50% failed inhibitions (2 old; 1 young). The remaining 2 subjects (1 old; 1 young) were excluded because the data on their handedness was either not collected (1 young) or did not match that of the other subjects (1 old left-handed subject). All analyzed participants had a high school diploma that entitled them to access higher (college-level) education ("Abitur" or "Fachhochschulreife"). The finally analyzed sample consisted of 16 old (68.4 ± 1.4 (mean ± SEM) years, range 60–76, 5 women) and 19 young (23.1 ± 0.4 years of age, range 20–26, 9 women) subjects. The study was approved by the Ethics Council of the Faculty of Medicine of the Otto-von-Guericke University Magdeburg.

Experimental procedure

A fast event-related design (Burock et al., 1998; Dale, 1999; Dale and Buckner, 1997) was used to exploit the possibility of post-hoc response sorting. In each run, participants saw a stream of + (Go) and × (NoGo) symbols which were presented in pseudorandom order (optimized for efficiency; see below) for 100 ms each. A total of 1083 stimuli appeared during the experiment (20.2% of these were NoGo stimuli) in five runs. Participants had to press a mouse button for frequent Go stimuli and had to inhibit this prepotent response when a NoGo stimulus occurred. Before the actual experiment, subjects completed a short test run to become familiar with the task. They had the possibility to rest between runs and continue the experiment in a self-paced manner. The mean stimulus onset asynchrony (SOA) was approximately 1.4 s and was jittered between 1 and 2.5 s, optimized for both power and efficiency of the event-related design (Dale, 1999).

Near-infrared spectroscopy

We recorded the concentration of oxygenated and deoxygenated hemoglobin at a sampling rate of 10 Hz using a Hitachi ETG-4000 Optical Topography System (Hitachi Medical Systems) which uses a modified Beer–Lambert Law to calculate hemoglobin concentrations. A thin plastic stick with a smooth ending was used to remove hair under the sockets in the optode grid to ensure direct skin contact of the optodes. Thirty-three optodes were placed on the subject's forehead from which 52 channels were recorded (see Fig. 1). The positioning of the optode grid was performed such that the middle optode of the most inferior row on the 3×11 optode grid was located on the point Fpz of the international 10/20 electroencephalography (EEG) system (Jasper, 1958). Also, the distances between the optode grid and both preauricular points were kept equivalent. The built-in 3D digitizer of the Hitachi system was used to record individual channel positions in each subject and to transform these to Montreal Neurological Institute (MNI) space during analysis using the toolbox of NFRI functions (Singh et al., 2005) included in the analysis software NIRS-SPM (Ye et al., 2009; see below). In three subjects, several (not more than three) channels failed to record a meaningful signal. In these cases, the channels were excluded from further analysis. Following transformation of individual channel positions to MNI space, we used the mean MNI coordinates of all subjects to visualize the channel positions on the MNI brain (Fig. 1).

Data analysis

Behavioral data and beta weights yielded by the NIRS analysis (see below) were analyzed with the software package R (version 2.12.1; R Development Core Team, 2010) with the package nlme (Pinheiro et al., 2011). Reaction times for Go stimuli and error rates (ERs) on NoGo trials were compared between age groups using unpaired

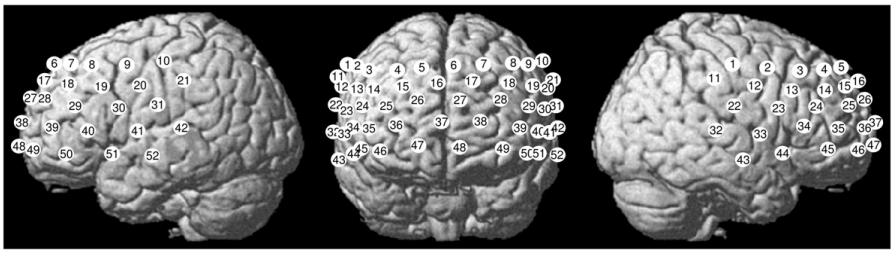


Fig. 1. Approximate location of the NIRS channel positions (averaged over subjects) in MNI space. The optode grid was placed on each subject's forehead and aligned to positions of the standard 10/20 EEG system (Jasper, 1958; see text). Thereafter, the position of each optode was recorded using a 3D digitizer and transformed to MNI space (Singh et al., 2005).

t-tests. Before all t-tests, we tested equality of variances with Bartlett's test and, if significant, subsequently used the Welch approximation to the degrees of freedom. Prior to statistical comparison, ERs were arcsine-transformed to evade problems concerning normal distribution of data. Furthermore, signal detection theory (SDT; Green and Swets, 1966) statistics (d′ and c) were computed for old and young subjects. The function of Pallier (2002) for R was used to compute d′. SDT statistics were compared between groups using unpaired t-tests.

NIRS data were analyzed with NIRS-SPM (Jang et al., 2009; Tak et al., 2010, 2011; Ye et al., 2009). This toolbox for the neuroimaging suite SPM5 (http://www.fil.ion.ucl.ac.uk/spm) analyzes NIRS data using a model-based analysis approach according to the general linear model (GLM) and allows for the creation of activation maps with super-resolution localization (Ye et al., 2009). Models for HbO and HbR were specified, each containing five regressors convolved with the corresponding hemodynamic response function with time derivative: (1) successfully answered Go stimuli (Go hit), (2) missed Go stimuli (Go miss), (3) successfully inhibited NoGo stimuli (NoGo CR [for correct rejections]), (4) failed attempts to inhibit NoGo stimuli (NoGo FA [for false alarms]) and (5) all corresponding responses. NIRS-SPM creates the models for HbO and HbR with opposing polarity so that a significant model fit for HbO indicates increased concentration and for HbR decreased concentration. We used the wavelet minimum description length algorithm to decompose NIRS measurements into global trends (such as pulse or movement artifacts), hemodynamic signals and uncorrelated noise components. Also, the precoloring method (Worsley and Friston, 1995) was used to remove temporal correlations from NIRS data using a low-pass filter with the shape of the hemodynamic response function. The model was estimated and group activation maps were created. When comparing different age groups in brain imaging studies, confounding comparisons of interest with neurovascular, morphological or variance/noise changes have to be avoided (Samanez-Larkin and D'Esposito, 2008). The hemodynamic response in event-related NIRS studies is age-dependent (Schroeter et al., 2003) and aging is associated with prefrontal volume changes (e.g. Raz et al., 1997) which might contribute to the increased differential path length factor (DPF) of aging subjects (Duncan et al., 1996; Strangman et al. (2002) refer to the DPF as "a measure of the path length the scattered light actually traveled through the tissue"). Also, Cui et al. (2011) have found that scalp–brain distance influences NIRS/BOLD correlations. Therefore, Schroeter et al. (2003) suggest comparing effect sizes in order to circumvent this pitfall as these are independent of DPFs. For a comparison between age groups, we used differential betas of the contrast [NoGo CR > Go hit] for HbO and HbR. This statistic should therefore neither be influenced by different DPFs in the experimental groups nor by other altered hemodynamic responses in elderly subjects (e.g. Ances et al., 2009). Differential betas from each channel were compared using a mixed Age Group (old/young) × Channel (1–52) ANOVA with Age Group as a between- and channel as a within-subject factor both for HbO and HbR (Table 1). The p-values were Greenhouse–Geisser corrected to account for possible violations of the sphericity assumption (Baron and Li, 2011). In an exploratory post-hoc analysis, we then compared channels in areas which showed differential activation in the single-group [NoGo CR > Go Hit] activation maps between the two age groups using unpaired t-tests. Anatomical labels of brain areas lying below the investigated channels were determined by using NIRS-SPMs interface to an atlas of cortical structures (Shattuck et al., 2008) and are reported in Table 2. All data are presented as mean ± SEM. Results were considered significant if $p < 0.05$.

Results

Performance data

Error rates are shown in Fig. 2. Analysis of the arcsine-transformed ERs revealed significant differences between the group of old and young adults on NoGo trials ($t(33) = 2.279$, $p = 0.029$). Comparison of the SDT statistic d′ between groups did not yield significant results for sensitivity (young: 2.916 ± 0.122, old: 2.889 ± 0.091, $t(33) = 0.173$, $p = 0.864$). However, the age groups differed when compared in a measure of reaction tendency (c; young: -0.91 ± 0.08, old: -0.62 ± 0.08, $t(33) = -2.697$, $p = 0.011$). Reaction times (RTs) upon presentation of Go stimuli were significantly longer in old adults (young: 365 ± 8 ms, old: 447 ± 11 ms, $t(28.488) = -5.820$, $p < 0.001$). Failures to inhibit motor responses in NoGo trials were associated with shorter response times (young: 306 ± 5 ms, old: 363 ± 10 ms).

Within-group analysis of NIRS data

Fig. 3 shows activation maps for oxy- and deoxygenated hemoglobin in the two age groups. As explicated in the Material and methods section, we constrained our interest to the difference between successfully inhibited NoGo stimuli and answered Go stimuli (contrast [NoGo CR > Go hit]) as the number of trials in the other conditions

Table 1
Results of the mixed Age Group (old/young) × Channel (1–52) ANOVAs of the differential beta values (contrast [NoGo CR > Go hit]; see also text).

	df	F	p	Significant?
HbO				
Age Group	1, 33	2.675	0.112	No
Channel	51, 1677	2.741	0.104	No
Age Group × Channel	51, 1677	2.746	0.103	No
HbR				
Age Group	1, 33	2.400	0.131	No
Channel	51, 1677	3.080	0.031	Yes
Age Group × Channel	51, 1677	2.874	0.044	Yes

Table 2
Between-group post-hoc comparisons (t-tests) of the differential betas (contrast [NoGo CR > Go hit]) of channels in areas that showed differential activation in Fig. 3 (HbR; see also Figs. 4 and 5). The p-values were adjusted for multiple comparisons using FDR correction (p adj., Benjamini and Hochberg, 1995; * indicates significance).

Channel	MNI coordinate (x, y, z)	t	df	p	p adj.	Area	Overlap (%)
2	56.8, 4.5, 46.9	−2.391	33	0.023	0.034*	Right precentral gyrus	87.5
						Right middle frontal gyrus	7.5
						Right postcentral gyrus	5.1
3	46.9, 25.7, 46.6	−2.876	23.7	0.008	0.034*	Right middle frontal gyrus	100
30	−61.9, 10.4, 22.4	2.207	22.1	0.038	0.044*	Left precentral gyrus	73.0
						Left inferior frontal gyrus	27.0
33	66.7, 0.4, 7	2.370	29.9	0.024	0.034*	Right superior temporal gyrus	49.2
						Right postcentral gyrus	34.2
						Right precentral gyrus	16.6
34	60.1, 27.8, 9.5	1.187	33	0.244	0.243	Right inferior frontal gyrus, triangular part	91.1
						Right inferior frontal gyrus, opercular part	8.9

(missed Go stimuli or failed attempts to inhibit) was too low for a meaningful analysis. In young subjects, changes in both HbO and HbR hemoglobin could be observed extensively in lateral frontal areas of both hemispheres. Right hemisphere activations extended to temporal regions. Changes in HbO were larger than changes in HbR. In old subjects, HbO activations also appeared bilaterally in frontal areas and, on the right hemisphere, also close to the central sulcus. Compared to the group of young subjects, HbO activation foci were shifted rostrally (left hemisphere) and dorsally (right hemisphere) to dorsolateral prefrontal cortical areas. As in the younger subjects, changes in HbR were less extensive than changes in HbO. Frontal activation changes in HbR were largely confined to the right hemisphere where decreases in HbR extended from precentral areas to the middle frontal gyrus (Fig. 3 upper right corner).

Between-group analysis of NIRS data

The results of the mixed Age Group (old/young) × Channel (1–52) ANOVAs of the differential beta values (contrast [NoGo CR > Go hit]) are reported in Table 1. In both hemoglobin species, the factor Age Group was not significant. The factor Channel and the interaction between these two factors yielded significant results for HbR but not HbO. Furthermore, channelwise post-hoc t-tests between age groups, focusing on differentially activated areas, revealed specific differences in the differential beta values of HbR. On the right hemisphere, two channels were identified, covering frontal and precentral areas, for which old adults showed a significantly higher differential beta. Vice versa, the differential betas of young subjects were higher in one channel in the temporal area of the right hemisphere, close to post- and precentral gyri, and in one channel in the left precentral/inferior frontal area (see Table 2, Figs. 4 and 5).

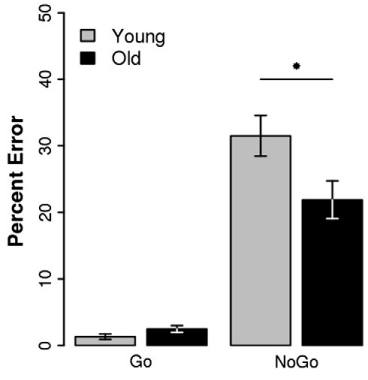


Fig. 2. Error rates in the different experimental conditions (see also text).

Discussion

The purpose of the present study was to investigate whether age-associated changes of neural processing in the Go/NoGo task can be detected with event-related NIRS using a relatively rapid stimulus presentation rate. We used a classic Go/NoGo behavioral paradigm (see Simmonds et al., 2008) in which a prepotent response tendency is established by weighting the task towards Go trials. The advantage of such an approach is that it induces robust signals by naturalistically modeling cognitive inhibition in the context of a stream of ongoing, habitual activity. It might be argued that stimulus frequency and inhibition are confounded in this experimental design, although this dependency is inherent to the paradigm. Although Casey et al. (2001), with a blocked design, have shown frontal brain hemodynamic differences when varying the proportions of NoGo stimuli, activation changes remain difficult to show conclusively as the number of responses is then confounded between the conditions. Importantly, in the latter study, it is well conceivable that a different behavioral task-set is activated depending on the frequency condition. Other studies (e.g. Garavan et al., 1999) try to avoid the frequency difference between Go and NoGo stimuli in the paradigm by requiring participants to respond only to alternating stimuli (e.g. "respond if x and y are alternating"), which has the caveat of introducing an additional working memory load (1-back) to the task. This may be especially relevant in aging research. In the present study, active inhibition was visualized on the baseline of an automated Go task which provides a framework for voluntary inhibition. The stimulus frequency difference between conditions also brings about that different numbers of trials are used to calculate intra-individual (first-level) summary statistics for each condition. One might thus argue that this introduces a sampling bias in the estimation of the true individual statistic which is, in the present study, a differential between two conditions (NoGo CR > Go Hit). Huettel and McCarthy (2001) have investigated the connection between the number of single trials and the resulting hemodynamic response with fMRI. They found that the estimated individual hemodynamic response was stable if 36 or more individual trials were combined. Also, the intra-individual standard deviation at peak remained asymptotically low if more than 45 trials were averaged. When Minati et al. (2011) compared fMRI and NIRS with simultaneous recordings, it was found that the intra-individual variability did not differ between the BOLD signal and HbO or HbR. We have therefore no reason to assume that different numbers of trials in the experimental conditions have biased results of the present study. The aforementioned study by Minati et al. (2011) also shows the inter-subject variability to be increased in NIRS compared to fMRI. Haeussinger et al. (2011) have recently shown that this heightened inter-individual variability appears to be rooted in anatomical differences between subjects such as scalp–brain distance and frontal sinus volume. Another point, which is important to mention in the context of experimental design, is that the Go/NoGo paradigm is analyzed stimulus-locked. In combination with the event-related

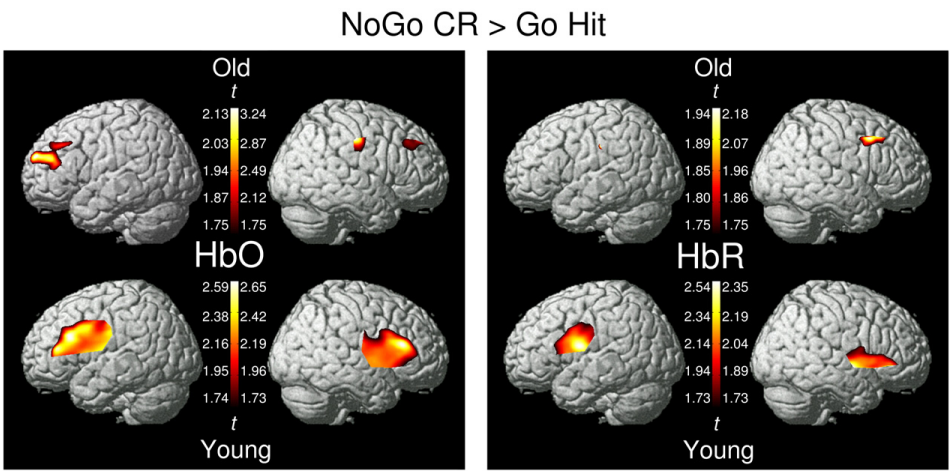


Fig. 3. Response inhibition activation in young and old subjects. All statistical parametric maps were created with a threshold of $p<0.05$ (uncorrected for multiple comparisons). Please note that these supra-threshold t-values indicate increases in HbO and decreases in HbR, respectively.

approach, it is therefore possible to observe neural activity between groups which is independent of differences in behavioral performance or response strategy. A minor disadvantage of our experimental approach is that potential hemodynamic activation of motor and pre-motor areas during inhibition may be masked in the NoGo CR > Go hit contrast. Also, despite the more extensive montage compared to former studies, coverage of the skull is far from complete. Together with the limited penetration depth of the NIRS signal, this does, in the present study, preclude investigation of the complete motor area and the medial frontal cortex.

In both age groups, analysis revealed activation changes in characteristic frontal regions when successfully inhibited NoGo trials with successfully answered Go trials were compared. The results thus demonstrate the feasibility of fast event-related designs when using NIRS to investigate cognitive processing in the elderly. Also, the present results suggest increased activation of temporal areas in young compared to old subjects and an increase of prefrontal activation in old compared to young subjects. Below, we discuss each of these points.

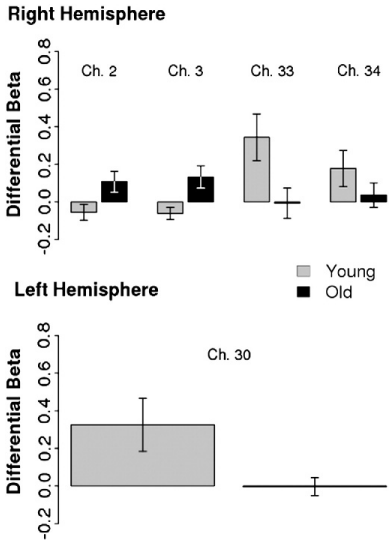


Fig. 4. Differential betas of channels selected for post-hoc comparisons between young and old adults (HbR; [NoGo CR > Go Hit]; for statistical results see Table 2).

Comparison with other NIRS and fMRI studies investigating inhibition

The present study represents a number of significant advances over previous investigations using NIRS in related paradigms, as it (1) used a very rapid presentation rate with an SOA of less than 2 s, and (2) allowed to disentangle brain responses to Go and NoGo trials because of its event-related nature. Herrmann et al. (2005) presented a NIRS study of the Go/NoGo task but used a blocked design contrasting so-called NoGo blocks with Go blocks. Whereas in NoGo blocks half of the stimuli required a Go response and the other half required a NoGo response, all stimuli had to be responded to in Go-blocks. A bilateral prefrontal increase in HbO and a corresponding decrease in HbR were reported for NoGo relative to Go blocks. The block design used in that study may be problematic as in Go blocks participants could be sure that they would never be required to inhibit a response. Such "all-go" conditions will likely lead to a different overall task set than that elicited in mixed Go/NoGo blocks. To study the phasic processes engaged by the ad-hoc need for inhibition of a response, an event-related design with randomized stimulus presentation is preferable. Also, in contrast to the multioptode array employed in the present study, Herrmann et al. (2005) used two 3×3 optode grids placed over the lateral prefrontal cortex on the left and right side and thus could find only changes in these regions. In spite of these differences it is remarkable that the HbO results in the young group are quite similar in the present study and the Herrmann et al. (2005) study. In another NIRS study using an event-related stop-signal-task, Boecker et al. (2007) found similar prefrontal changes when comparing successfully inhibited stop trials with Go trials. This study, however, used only two optodes. Because of the event-related design and the more extensive sensor coverage, the current results lend themselves for comparison with event-related fMRI studies (Braver et al., 2001; Garavan et al., 1999; Kiehl et al., 2000; Konishi et al., 1998, 1999; Menon et al., 2001; Rubia et al., 2001,

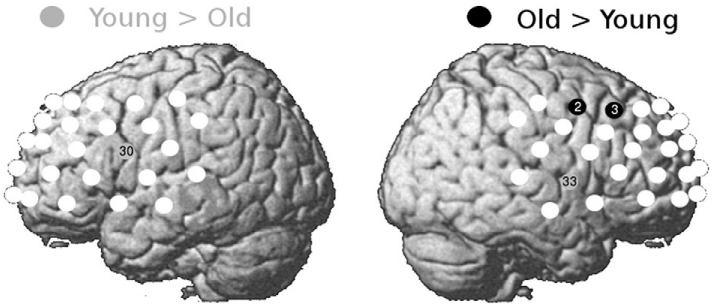


Fig. 5. Approximate locations of channels which differ significantly between age groups (contrast [NoGo CR > Go Hit]; see Table 2) visualized on the MNI brain.

2003). Successful inhibition in NoGo-trials has in many studies been found to result in activation of the right more than left inferior prefrontal cortex, which has been proposed as the major site for inhibition (Aron et al., 2004).

Comparisons between age-groups

The current study revealed differences in the activation pattern between young and old adults. Old adults showed an activation focus for the NoGo CR > Go Hit contrast in the dorsolateral prefrontal cortex (DLPFC) rather than the inferior frontal cortex (IFC). Nielson et al. (2002; see Langenecker and Nielson (2003) for a replication) have also investigated different age groups with a Go/NoGo paradigm using event-related fMRI. When they compared successfully inhibited NoGo stimuli, elderly adults showed more extensive and more bilateral frontal brain activation than young adults, which the authors interpreted to be of compensatory nature. Interestingly, Nielson et al. (2002) also found increased activation for elderly compared to younger subjects in the right middle frontal gyrus and the opposite pattern of results in the right IFC in the response inhibition condition. The current study thus both replicates and extends previous findings on age differences of inhibitory processing (Langenecker and Nielson, 2003; Nielson et al., 2002). In particular the right DLPFC has been clearly associated with inhibition, for example by applying lesions in monkeys (Iversen and Mishkin, 1970; Mishkin, 1964). Also, the right DLPFC as well as the IFC have been found to be activated in both Go/NoGo and stop-signal experiments in humans (de Zubicaray et al., 2000; Garavan et al., 1999) and non-human primates (Sasaki and Gemba, 1986; Sakagami et al., 2001). The role of the DLPFC in inhibition has been stressed on the basis of the meta-analyses of imaging results (Simmonds et al., 2008; Wager et al., 2005). Simmonds et al. (2008) have compared "simple" (fixed NoGo stimuli) and "complex" (NoGo stimuli change depending on context) imaging studies of the Go/NoGo task and found DLPFC activation only in studies using complex tasks which required frequent updating of stimulus–response associations. One might therefore speculate that older adults increasingly recruit brain areas such as the DLPFC even for relatively simple executive tasks, probably to compensate for a decrease in functionality of the IFC. Nielson et al. (2002) also used a complex task according to this terminology. While in the present study, right hemispheric effects between age groups are in line with the former study, increased recruitment of left hemispheric regions in old participants was not observed in the present study. Rather, young subjects showed increased recruitment of the left IFC. In their meta-analysis, Spreng et al. (2010) found that young (compared to aging) subjects show increased activation of ventrolateral prefrontal cortex across several cognitive domains when performance accuracy is equivalent. Insofar as group statistics of the sensitivity index d′ can be interpreted as showing comparable performance, this appears to be a general effect. Spreng et al. (2010) did not take differences in RT into account when classifying performance. To this end, it is well known that performance on executive paradigms such as task switching (e.g. Kramer et al., 1999; Kray and Lindenberger, 2000) and error monitoring (e.g. Band and Kok, 2000; Falkenstein et al., 2001) does decline in normal adult aging and that this demise is paralleled by respective prefrontal neural changes (for review see West, 1996). In most studies of cognitive aging, this performance decrease is measured as increased RT (c.f. Kray et al., 2004), which was also observed in the present study. A general decrease in mental speed is at the core of an influential hypothesis of cognitive aging (Salthouse, 2000). However, it has been shown that age-related slowing is attributable to more than one general factor, emphasizing coordinative complexity as another factor of cognitive aging (Mayr and Kliegl, 1993; Mayr et al., 1996). In the present study, a shift of response strategy towards increased accuracy but prolonged reaction time was observed in the group of old adults. Similar response patterns have been observed in investigations of age-related changes in cognitive experiments (e.g. Salthouse, 1979; Starns and Ratcliff, 2010). Differences in the effects of aging on task performance across studies (ERs and/or RTs) may stem from a variety of reasons such as the experimental paradigm itself or floor/ceiling effects and are difficult to interpret if neither RT nor ER explicitly controlled for. However, as mentioned above, RT differences are not related to brain activation patterns in the current study due to the stimulus-locked analysis and ER differences between age-groups can be ruled out due to the event-related experimental design. When analyzing the ERs within the signal-detection framework, the finding that old adults differ in a measure of reaction tendency from young adults fits with the explanation of a general strategy shift of the aging neurocognitive system. In the present study, we were thus able to show qualitative neural changes accompanied by qualitative behavioral changes. In the past, scientists have remarked on the heterogeneity of older persons and have separated unsuccessful from "successful aging", thereby referring to individuals who are spared from age-associated decline (Rowe and Kahn, 1987). Recently, Düzel et al. (2011) have presented an approach of functionally phenotyping older adults, resulting in a quantitative distinction between successfully and unsuccessfully cognitively aged elders in terms of memory performance. In essence, this attempt uses a comparison between a template brain activation pattern seen in young adults and the individual pattern observed in the older adult being phenotyped. A similar approach may be feasible with respect to successful executive performance using the current or similar paradigms that reflect age-associated alterations.

Conclusions

In the present study, we exploited the advantages of an event-related design in combination with a model-based analysis strategy to

investigate aging effects with NIRS. We show that age-associated alterations in inhibitory function can be detected with fast event-related NIRS and that results are comparable to corresponding fMRI studies. As we are able to separate neural activation and behavior, differences in both measures can be interpreted as neurobehavioral strategy shift between young and old adults. It should be explicitly mentioned that our primary aim was to show general applicability of advances in fMRI research to the method of NIRS. We did therefore not control for type I error accumulation when projecting statistical parametric maps to the MNI brain. Future studies could achieve a higher power by investigating more participants, maybe as part of a clinical routine, to account for the heightened between-subject variability of NIRS compared to fMRI (Minati et al., 2011). Nevertheless, NIRS opens an indispensable window to study the aging brain with experimental designs of contemporary neuroscience.

Acknowledgments

The authors would like to thank Cand.-Psych. Susanne Boge for help in collecting the data and Dr. Michael Scholz for calculation of the optimal trial timing. TFM is supported by the DFG and BMBF.

References

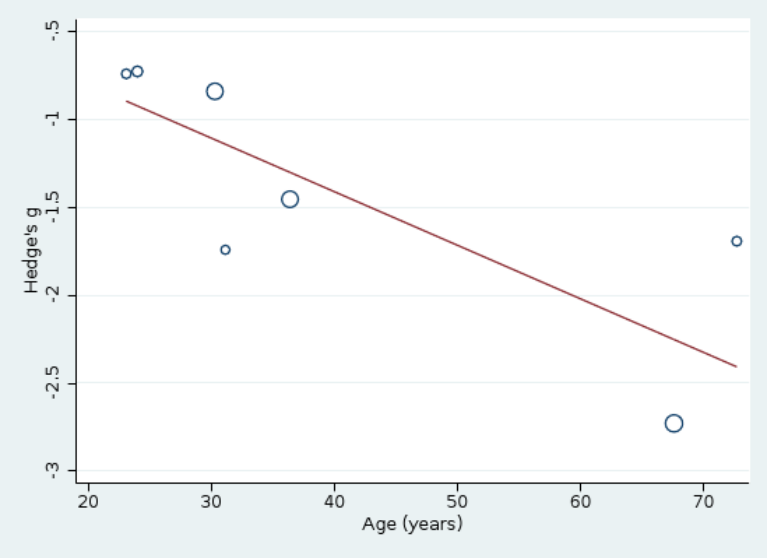
Amieva, H., Phillips, L.H., Della Sala, S., Henry, J.D., 2004. Inhibitory functioning in Alzheimer's disease. Brain 127, 949–964.
Ances, B.M., Liang, C.L., Leontiev, O., Perthen, J.E., Fleisher, A.S., Lansing, A.E., Buxton, R.B., 2009. Effects of aging on cerebral blood flow, oxygen metabolism, and blood oxygenation level dependent responses to visual stimulation. Hum. Brain Mapp. 30, 1120–1132.
Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. 8, 170–177.
Band, G.P., Kok, A., 2000. Age effects on response monitoring in a mental-rotation task. Biol. Psychol. 51, 201–221.
Baron, J., Li, Y., 2011. Notes on the Use of R for Psychology Experiments and Questionnaires. http://www.psych.upenn.edu/~baron/rpsych/rpsych.pdf retrieved May 6, 2011.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist Soc. B 57, 289–300.
Boecker, M., Buecheler, M.M., Schroeter, M.L., Gauggel, S., 2007. Prefrontal brain activation during stop-signal response inhibition: an event-related functional near-infrared spectroscopy study. Behav. Brain Res. 176, 259–266.
Braver, T.S., Barch, D.M., Gray, J.R., Molfese, D.L., Snyder, A., 2001. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. Cereb. Cortex 11, 825–836.
Burock, M.A., Buckner, R.L., Woldorff, M.G., Rosen, B.R., Dale, A.M., 1998. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. Neuroreport 9, 3735–3739.
Casey, B.J., Forman, S.D., Franzen, P., Berkowitz, A., Braver, T.S., Nystrom, L.E., Thomas, K.M., Noll, D.C., 2001. Sensitivity of prefrontal cortex to changes in target probability: a functional MRI study. Hum. Brain Mapp. 13, 26–33.
Cui, X., Bray, S., Bryant, D.M., Glover, G.H., Reiss, A.L., 2011. A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. NeuroImage 54, 2808–2821.
Dale, A.M., 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp. 8, 109–114.
Dale, A.M., Buckner, R.L., 1997. Selective averaging of rapidly presented individual trials using fMRI. Hum. Brain Mapp. 5, 329–340.
de Zubicaray, G.I., Andrew, C., Zelaya, F.O., Williams, S.C.R., Dumanoir, C., 2000. Motor response suppression and the prepotent tendency to respond: a parametric fMRI study. Neuropsychologia 38, 1280–1291.
Dennis, N.A., Cabeza, R., 2008. Neuroimaging of healthy cognitive aging, In: Craik, F.I.M., Salthouse, T.A. (Eds.), The Handbook of Aging and Cognition, 3rd ed. Psychology Press.
Duncan, A., Meek, J.H., Clemence, M., Elwell, C.E., Fallon, P., Tyszczuk, L., Cope, M., Delpy, D.T., 1996. Measurement of cranial optical path length as a function of age using phase resolved near infrared spectroscopy. Pediatr. Res. 39, 889–894.
Düzel, E., Schütze, H., Yonelinas, A.P., Heinze, H.-J., 2011. Functional phenotyping of successful aging in long-term memory: preserved performance in the absence of neural compensation. Hippocampus 21, 803–814.
Falkenstein, M., Hoormann, J., Hohnsbein, J., 2001. Changes of error-related ERPs with age. Exp. Brain Res. 138, 258–262.
Garavan, H., Ross, T.J., Stein, E.A., 1999. Right hemispheric dominance of inhibitory control: an event-related functional MRI study. Proc. Natl. Acad. Sci. U. S. A. 96, 8301–8306.
Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Wiley, New York.
Haeussinger, F.B., Heinzel, S., Hahn, T., Schecklmann, M., Ehlis, A.-C., Fallgatter, A.J., 2011. Simulation of near-infrared light absorption considering individual head and prefrontal cortex anatomy: implications for optical neuroimaging. PLoS One 6, 1–12.
Hasher, L., Zacks, R.T., 1988. Working memory, comprehension, and aging: a review and a new view. In: Bower, G.H. (Ed.), Psychology of Learning and Motivation. Academic Press.
Herrmann, M.J., Plichta, M.M., Ehlis, A.-C., Fallgatter, A.J., 2005. Optical topography during a Go–NoGo task assessed with multi-channel near-infrared spectroscopy. Behav. Brain Res. 160, 135–140.
Hillman, E.M.C., 2007. Optical brain imaging in vivo: techniques and applications from animal to man. J. Biomed. Opt. 12, 051402.
Huettel, S.A., McCarthy, G., 2001. The effects of single trial averaging upon the spatial extend of fMRI activation. Neuroreport 12, 1–6.
Iversen, S., Mishkin, M., 1970. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp. Brain Res. 11, 376–386.
Jang, K.E., Tak, S., Jung, J., Jang, J., Jeong, Y., Ye, J.C., 2009. Wavelet minimum description length detrending for near-infrared spectroscopy. J. Biomed. Opt. 14, 034004.
Jasper, H., 1958. The ten-twenty electrode system of the International Federation. Electroencephalogr. Clin. Neurophysiol. 10, 371–375.
Jöbsis, F.F., 1977. Non-invasive, infra-red monitoring of cerebral O_2 sufficiency, bloodvolume, HbO_2–Hb shifts and bloodflow. Acta Neurol. Scand. 64, 452–453.
Josephs, O., Turner, R., Friston, K., 1997. Event-related fMRI. Hum. Brain Mapp. 5, 243–248.
Kiehl, K.A., Liddle, P.F., Hopfinger, J.B., 2000. Error processing and the rostral anterior cingulate: an event-related fMRI study. Psychophysiology 37, 216–223.
Konishi, S., Nakajima, K., Uchida, I., Sekihara, K., Miyashita, Y., 1998. No-go dominant brain activity in human inferior prefrontal cortex revealed by functional magnetic resonance imaging. Eur. J. Neurosci. 10, 1209–1213.
Konishi, S., Nakajima, K., Uchida, I., Kikyo, H., Kameyama, M., Miyashita, Y., 1999. Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. Brain 122, 981–991.
Kramer, A.F., Hahn, S., Gopher, D., 1999. Task coordination and aging: explorations of executive control processes in the task switching paradigm. Acta Psychol. (Amst) 101, 339–378.
Kray, J., Lindenberger, U., 2000. Adult age differences in task switching. Psychol. Aging 15, 126–147.
Kray, J., Eber, J., Lindenberger, U., 2004. Age differences in executive functioning across the lifespan: the role of verbalization in task preparation. Acta Psychol. (Amst) 115, 143–165.
Langenecker, S.A., Nielson, K.A., 2003. Frontal recruitment during response inhibition in older adults replicated with fMRI. NeuroImage 20, 1384–1392.
Mayr, U., Kliegl, R., 1993. Sequential and coordinative complexity: age-based processing limitations in figural transformations. J. Exp. Psychol. Learn. Mem. Cogn. 19, 1297–1320.
Mayr, U., Kliegl, R., Krampe, R.T., 1996. Sequential and coordinative processing dynamics in figural transformations across the life span. Cognition 59, 61–90.
McDowd, J.M., 1997. Inhibition in attention and aging. J. Gerontol. B Psychol. Sci. Soc. Sci. 52, P265–P273.
Menon, V., Adleman, N.E., White, C.D., Glover, G.H., Reiss, A.L., 2001. Error-related brain activation during a Go/NoGo response inhibition task. Hum. Brain Mapp. 12, 131–143.
Minati, L., Visani, E., Dowell, N.G., Medford, N., Critchley, H.D., 2011. Variability comparison of simultaneous brain near-infrared spectroscopy and functional magnetic resonance imaging during visual stimulation. J. Med. Eng. Technol. 35, 370–376.
Mishkin, M., 1964. Perseveration of central sets after frontal lesions in monkeys. In: Warren, J., Akert, K. (Eds.), The Frontal Granular Cortex and Behavior. McGraw-Hill.
Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., 2000. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. Cogn. Psychol. 41, 49–100.
Nielson, K.A., Langenecker, S.A., Garavan, H., 2002. Differences in the functional neuroanatomy of inhibitory control across the adult lifespan. Psychol. Aging 17, 56–71.
Obrig, H., Wenzel, R., Kohl, M., Horst, S., Wobst, P., Steinbrink, J., Thomas, F., Villringer, A., 2000. Near-infrared spectroscopy: does it function in functional activation studies of the adult brain? Int. J. Psychophysiol. 35, 125–142.
Pallier, C., 2002. Computing Discriminability and Bias with the R Software. http://www.pallier.org/ressources/aprime.pdf retrieved May 6th, 2011.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Development Core Team, 2011. nlme: Linear and Nonlinear Mixed Effects Models. http://cran.r-project.org/web/packages/nlme/index.html.
Plichta, M.M., Heinzel, S., Ehlis, A.-C., Pauli, P., Fallgatter, A.J., 2007. Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: a parametric validation study. NeuroImage 35, 625–634.
R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. http://www.R-project.org/.
Raz, N., Gunning, F.M., Head, D., Dupuis, J.H., McQuain, J., Briggs, S.D., Loken, W.J., Thornton, A.E., Acker, J.D., 1997. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb. Cortex 7, 268–282.
Rowe, J.W., Kahn, R.L., 1987. Human aging: usual and successful. Science 237, 143–149.
Rubia, K., Russell, T., Overmeyer, S., Brammer, M.J., Bullmore, E.T., Sharma, T., Simmons, A., Williams, S.C., Giampietro, V., Andrew, C.M., Taylor, E., 2001. Mapping motor inhibition: conjunctive brain activations across different versions of go/no-go and stop tasks. NeuroImage 13, 250–261.
Rubia, K., Smith, A.B., Brammer, M.J., Taylor, E., 2003. Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. NeuroImage 20, 351–358.
Sakagami, M., Ki, T., Lauwereyns, J., Koizumi, M., Kobayashi, S., Hikosaka, O., 2001. A code for behavioral inhibition on the basis of color, but not motion, in ventrolateral prefrontal cortex of macaque monkey. J. Neurosci. 21, 4801–4808.

Salthouse, T.A., 1979. Adult age and the speed-accuracy trade-off. Ergonomics 22, 811–821.
Salthouse, T.A., 2000. Aging and measures of processing speed. Biol. Psychol. 54, 35–54.
Samanez-Larkin, G.R., D'Esposito, M., 2008. Group comparisons: imaging the aging brain. Soc. Cogn. Affect. Neurosci. 3, 290–297.
Sasaki, K., Gemba, H., 1986. Electrical activity in the prefrontal cortex specific to no-go reaction of conditioned hand movement with colour discrimination in the monkey. Exp. Brain Res. 64, 603–606.
Schroeter, M.L., Zysset, S., Kupka, T., Kruggel, F., von Cramon, D.Y., 2002. Near-infrared spectroscopy can detect brain activity during a color-word matching Stroop task in an event-related design. Hum. Brain Mapp. 17, 61–71.
Schroeter, M.L., Zysset, S., Kruggel, F., von Cramon, D.Y., 2003. Age dependency of the hemodynamic response as measured by functional near-infrared spectroscopy. NeuroImage 19, 555–564.
Schroeter, M.L., Bücheler, M.M., Müller, K., Uludağ, Kâ, Obrig, H., Lohmann, G., Tittgemeyer, M., Villringer, A., von Cramon, D.Y., 2004. Towards a standard analysis for functional near-infrared imaging. NeuroImage 21, 283–290.
Shattuck, D.W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K.L., Poldrack, R.A., Bilder, R.M., Toga, A.W., 2008. Construction of a 3D probabilistic atlas of human cortical structures. NeuroImage 39, 1064–1080.
Simmonds, D.J., Pekar, J.J., Mostofsky, S.H., 2008. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 46, 224–232.
Singh, A.K., Okamoto, M., Dan, H., Jurcak, V., Dan, I., 2005. Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. NeuroImage 27, 842–851.
Spieler, D.H., Balota, D.A., Faust, M.E., 1996. Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. J. Exp. Psychol. Hum. Percept. Perform. 22, 461–479.
Spreng, R.N., Wojtowicz, M., Grady, C.L., 2010. Reliable differences in brain activity between young and old adults: a quantitative meta-analysis across multiple cognitive domains. Neurosci. Biobehav. Rev. 34, 1178–1194.
Starns, J.J., Ratcliff, R., 2010. The effects of aging on the speed-accuracy compromise: boundary optimality in the diffusion model. Psychol. Aging 25, 377–390.
Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., Obrig, H., 2006. Illuminating the BOLD signal: combined fMRI-fNIRS studies. Magn. Reson. Imaging 24, 495–505.
Strangman, G., Boas, D.A., Sutton, J.P., 2002. Non-invasive neuroimaging using near-infrared light. Biol. Psychiatry 52, 679–693.
Taga, G., Asakawa, K., 2007. Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. NeuroImage 36, 1246–1252.
Tak, S., Jang, J., Lee, K., Ye, J.C., 2010. Quantification of CMRO(2) without hypercapnia using simultaneous near-infrared spectroscopy and fMRI measurements. Phys. Med. Biol. 55, 3249–3269.
Tak, S., Yoon, S.J., Jang, J., Yoo, K., Jeong, Y., Ye, J.C., 2011. Quantitative analysis of hemodynamic and metabolic changes in subcortical vascular dementia using simultaneous near-infrared spectroscopy and fMRI measurements. NeuroImage 55, 176–184.
Wager, T.D., Sylvester, C.-Y.C., Lacey, S.C., Nee, D.E., Franklin, M., Jonides, J., 2005. Common and unique components of response inhibition revealed by fMRI. NeuroImage 27, 323–340.
West, R.L., 1996. An application of prefrontal cortex function theory to cognitive aging. Psychol. Bull. 120, 272–292.
Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited—again. NeuroImage 2, 173–181.
Ye, J., Tak, S., Jang, K., Jung, J., Jang, J., 2009. NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. NeuroImage 44, 428–447.
Zarahn, E., Aguirre, G., D'Esposito, M., 1997. A trial-based experimental design for fMRI. NeuroImage 6, 122–138.

4 A Review of the Longitudinal Course of Executive Deficits in Schizophrenia (Heilbronner et al., 2016)

SZ is a heritable psychiatric disorder that is characterized by positive (such as hallucinations) and negative (such as flattened affect) symptoms, as well as cognitive impairments (e.g., Kahn et al., 2015), including impaired EFs. Longitudinal studies in SZ have a long tradition (e.g., Häfner & an der Heiden, 2000), but the available studies are rather heterogeneous in terms of design, diagnostic criteria, available control groups, repeated measurements, and follow-up duration. Also, these studies often emphasize heterogeneity between patient trajectories as defining characteristic (e.g., Modestin et al., 2003). In an attempt to review available longitudinal studies on symptoms, cognition, and neuroimaging, we have employed stringent criteria to select studies. We focussed on a narrow SZ phenotype, excluding schizoaffective individuals, and examined only controlled studies with repeated measures. In summary, EFs can be shown to be differentially affected over time. Whereas cognition in general remained stable over time, some studies researching a follow-up interval of about five years found diagnosis-by-time interactions for verbal fluency, with performance deteriorating in the SZ group. Interestingly, this executive worsening may be related to medication status, as patients in the deteriorating group were more frequently treated with typical antipsychotics. Regarding neuroimaging, despite whole brain volume remaining stable, a lower frontal lobe volume was found over time. Also observed were frontal gray matter reductions, the degree of which covaried with symptom severity. Of note, we carried out a meta-analysis of global cognitive change, which did not yield differences between baseline and follow-up, but detected large heterogenetiy of effect sizes between studies. Intriguingly, when trying to identify variables associated with this heterogeneity, meta-regression (Figure 2) found that, when compared to neurotypical individuals, older individuals with SZ were more severely impaired than younger individuals with SZ. This important finding could be further researched by experimental designs combining cross-sectional and longitudial research, specifically regarding EFs.

Figure 2: Meta-regression of overall (i.e., not specific to EFs) cognitive performance differences between individuals with SZ and neurotypical individuals, with age as an independent moderator. Circles represent studies, with the size of each circle proportional to the precision of the effect size of that study. Reprinted from Heilbronner et al. (2016).



The complete article (Heilbronner et al., 2016) is not included in this thesis due to copyright restrictions. It is, however, available free of charge: https://doi.org/10.1097/HRP.0000000000000092.

5 A Longitudinal Study of a Heritable Executive Function Personality Trait (Heilbronner, Papiol, et al., 2021)

An important result of neurotypical EFs, mediated by the PFC, is the inhibitory regulation of behavior, going beyond mere cognition to personality traits. This is evident when considering individuals with frontal lesions, whose personalities can be permanently changed towards behavioral disinhibition, similar to the case of Phineas Gage mentioned in the Introduction. It is well-known that all complex psychological traits are heritable to some degree (Turkheimer, 2000), and personality traits show substantial heritability (see Turkheimer et al., 2014, for review). Here, we researched the genomic backgrounds of personality dimensions in a population-based study of the city of Heidelberg and surroundings, the *Heidelberger Langzeitstudie zu Risikofaktoren und Diagnose chronischer Erkrankungen* (HeiDE). The initial goals of HeiDE were to research personality traits associated with cardiovascular disease and cancer, and a number of latent personality dimensions (named "The Heidelberg Five") had been identified using factor analysis (Amelang et al., 2004). Interestingly, one of these personality traits had been named Lack of Behavioral Control (LBCN), and is characterized by low anger control, low social desirability, high agression, high irritablity, and high outward expression of anger (Amelang et al., 2004). The extreme end of this personality trait thus phenotypically resembles the behavior of individuals with frontal damage (Szczepanski & Knight, 2014, Tate, 1999). Since a subset of participants of the HeiDE study were genotyped on single-nucleotide polymorphism (SNP) arrays, that interrogate a large fraction of common SNPs in the genome, we performed several genome-wide association studies (GWASs) on latent factor scores of The Heidelberg Five, including LBCN. We also conducted gene-based tests, tissue expression and genet set analyses, and assessed the so-called SNP-based heritability of The Heidelberg Five. The latter is similar to the heritability estimates in twin studies (see e.g., Turkheimer et al., 2014), but based on SNPs, which, for several reasons (Boomsma et al., 2002) is lower than the twin study counterpart. Moreover, since HeiDE is a longitudial study, we characterized associations of factor scores for The Heidelberg Five (assessed at baseline) with symptoms of depression and anxiety, approximately 20 years later. After correction for

multiple testing, results of the GWAS, gene-based tests, tissue expression and gene-set analyses did not reveal significant associations with LBCN. However, a significant SNP-based heritability (29.4%) of LBCN was found. In addition, LBCN scores derived from personality tests conducted in the early 1990s were, some 20 years later, associated with depressive symptoms and lifetime anxitey symptoms. The high SNP-based heritability compared to much larger GWASs of personality traits such as Neuroticism (10-15%, 23andMe Research Team et al., 2018, Docherty et al., 2016, Power & Pluess, 2015) may indicate LBCN to be a clinically valid personality dimension for future genetic research.

Received: 16 December 2020 | Revised: 19 January 2021 | Accepted: 21 January 2021

DOI: 10.1002/ajmg.b.32837

ORIGINAL ARTICLE

American Journal of Medical Genetics Part B: Neuropsychiatric Genetics | WILEY

"The Heidelberg Five" personality dimensions: Genome-wide associations, polygenic risk for neuroticism, and psychopathology 20 years after assessment

Urs Heilbronner[1] | Sergi Papiol[1,2] | Monika Budde[1] | Till F. M. Andlauer[3] | Jana Strohmaier[4] | Fabian Streit[4] | Josef Frank[4] | Franziska Degenhardt[5,6] | Stefanie Heilmann-Heimbach[5] | Stephanie H. Witt[4] | Andreas J. Forstner[5,7,8] | Adrian Loerbroks[9] | Manfred Amelang[10] | Til Stürmer[11] | Bertram Müller-Myhsok[12] | Markus M. Nöthen[5] | Marcella Rietschel[4] | Thomas G. Schulze[1,13]

[1]Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Munich, Germany

[2]Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany

[3]Department of Neurology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany

[4]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[5]Institute of Human Genetics, School of Medicine & University Hospital Bonn, University of Bonn, Bonn, Germany

[6]Department of Child and Adolescent Psychiatry, University Hospital Essen, University of Duisburg-Essen, Essen, Germany

[7]Centre for Human Genetics, University of Marburg, Marburg, Germany

[8]Department of Biomedicine, University of Basel, Basel, Switzerland

[9]Institute of Occupational, Social and Environmental Medicine, Center for Health and Society, Faculty of Medicine, University of Düsseldorf, Düsseldorf, Germany

[10]Department of Psychology, University of Heidelberg, Heidelberg, Germany

[11]Department of Epidemiology, UNC Gillings School of Global Public Health, Chapel Hill, North Carolina

[12]Department of Translational Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany

[13]Department of Psychiatry and Behavioral Sciences, Upstate University Hospital, Syracuse, New York

Correspondence
Urs Heilbronner, Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Nussbaumstr. 7, 80336 Munich, Germany.
Email: urs.heilbronner@med.uni-muenchen.de

Funding information
BONFOR Programme of the University of Bonn; Bundesministerium für Bildung und Forschung, Grant/Award Numbers: 01ZX1614K, 01ZX1614G, 01ZX1614A, 01ZX1614J, 01EE1404H : 01ZZ1804A; Deutsche Forschungsgemeinschaft, Grant/Award Numbers: SCHU 1603/4-1, SCHU 1603/5-1, SCHU 1603/7-1, FA241/16-1 : FOR2107 : RI908/11-2 : WI3429/3-2; Dr. Lisa

Abstract

HeiDE is a longitudinal population-based study that started in the 1990s and, at baseline, assessed an array of health-related personality questionnaires in 5133 individuals. Five latent personality dimensions (The Heidelberg Five) were identified and interpreted as Emotional Lability (ELAB), Lack of Behavioral Control (LBCN), Type A Behavior (TYAB), Locus of Control over Disease (LOCC), and Psychoticism (PSYC). At follow-up, 3268 HeiDE participants (post-QC) were genotyped on single nucleotide polymorphism (SNP) arrays. To further characterize The Heidelberg Five, we analyzed genomic underpinnings, their relations to the genetic basis of the Big Five trait Neuroticism, and longitudinal associations with psychiatric symptoms at follow-up. SNP-based heritability was significant for ELAB (34%) and LBCN (29%). A genome-wide

Oehler Foundation, Kassel; ERA-NET Neuron, Grant/Award Numbers: 01EW1810, 01EW1904; European Union's Horizon 2020 Research and Innovation Programme, Grant/Award Number: RIA 7331; National Institutes of Health, Grant/Award Number: R01 AG056479: R01 HL118255: R01 MD011680; UNC Clinical and Translational Science Award, Grant/Award Number: UL1TR002489

association study for each personality dimension was conducted; only the phenotype PSYC yielded a genome-wide significant finding ($p < 5 \times 10^{-8}$, top SNP rs138223660). Gene-based analyses identified significant findings for ELAB, TYAB, and PSYC. Polygenic risk scores for Neuroticism were only associated with ELAB. Each of The Heidelberg Five was related to depressive symptoms at follow-up. ELAB, LBCN, and PSYC were also associated with lifetime anxiety symptoms. These results highlight the clinical importance of health-related personality traits and identify LBCN as a heritable "executive function" personality trait.

KEYWORDS
control, executive, longitudinal, psychoticism

1 | INTRODUCTION

In its widest sense, personality can be conceptualized as "relatively enduring patterns of thoughts, feelings, and behaviors" (Sanchez-Roige et al., 2018) that constitute hallmarks of individuality. The "Big Five" personality traits (reviewed by Goldberg, 1993) have become the prevailing scientific taxonomy, and an individual's personality can be comprehensively characterized along these latent dimensions. A related but somewhat different scientific approach has been to characterize specific health-related personality dimensions (see Capitanio, 2008; Friedman & Booth-Kewley, 1987), hypothesized to be related to somatic disease such as cardiovascular disease (CVD) and cancer. The longitudinal HeiDE study ("Heidelberger Langzeitstudie zu Risikofaktoren und Diagnose chronischer Erkrankungen") pursues the latter approach (Stürmer et al., 2006). Since the early 1990s, this epidemiological study assesses personality, health, lifestyle, and cognitive variables in a population-based sample of 5133 individuals from the German city of Heidelberg and surroundings. Several follow-up assessments have been conducted, to evaluate the association of psychological factors and disease. Based on an array of questionnaires completed at baseline that assessed depressive symptoms, resilience factors, as well as some broad personality factors (Extraversion, Neuroticism, and Psychoticism), five personality dimensions, named "The Heidelberg Five," were subsequently extracted using exploratory factor analysis (Amelang et al., 2004). These were named Emotional Lability (ELAB, defined by Neuroticism, depression, a tendency to suppress anger, low social support, low optimism, and a low sense of coherence, as well as low Extraversion), Lack of Behavioral Control (LBCN, characterized by low social desirability and low anger control), Type A Behavior (TYAB, defined by high time urgency, exaggerated social control and high Extraversion), Locus of Control over Disease (LOCC, characterized by a high internal locus of control), and Psychoticism (PSYC, defined by high psychoticism). ELAB appears to tap the combination of risk factors for psychopathology, and the absence of resilience factors. LBCN characterizes a low capacity to self-regulate and may thus be regarded as an "executive function" personality trait, primarily defined by a lack of inhibitory control of emotions. Consistently, the dimension LBCN encompasses behaviors associated with executive function during development (e.g., Rohlf et al., 2018), high expression of which resemble some clinical conditions of the prefrontal cortex (e.g., Szczepanski & Knight, 2014; Tate, 1999). The personality dimension TYAB is highly correlated with psychological tests measuring time urgency, exaggerated social control, and Extraversion (for details see Supporting Information). The construct TYAB is "characterized primarily by a chronic incessant struggle to achieve more and more in less and less time" (Smith et al., 1996) and was initially hypothesized to be associated with CVD (Friedman & Rosenman, 1959), but this assumption was later found to have little empirical support (for review see Kuper et al., 2002). In line with this finding, prospective research in the HeiDE study did not identify TYAB as a predictor of the incidence of CVD (Amelang et al., 2004). LOCC is a construct based on Rotter's influential social learning theory (Rotter, 1966) that assesses cognitions of control over health (e.g., Wooldridge et al., 1992). Briefly, social learning theory postulates that individuals differ in their perception of reinforcements and classify these either as being controlled externally, that is, by chance or the specific situation, or internally, by the person's own actions. According to Rotter, generalized expectancies differ between individuals and this constitutes a personality dimension, Locus of Control. The personality dimension determines whether individuals perceive outcomes as rather externally or internally controlled (Weiner et al., 2009). High internal Locus of Control has been shown to be important for a variety of health behaviors including smoking, alcohol consumption, exercise, diet (Steptoe & Wardle, 2001; Strudler Wallston & Wallston, 1978), and medication adherence (Náfrádi et al., 2017). Finally, PSYC is one of the three personality factors in Eysenck's influential theory-based model of personality (Eysenck & Eysenck, 1976) and has been discussed as a core element of maladaptive personality, resembling schizotypy (Chapman et al., 1994; van Kampen, 2009; Wright et al., 2012). Initially, Psychoticism was conceptualized as a continuous dimension, predisposing individuals to psychosis, but a 10-year longitudinal study did not confirm this association (Chapman et al., 1994). The latter study, however, also reported that individuals scoring high on Psychoticism "exceeded controls on ratings of psychotic-like experiences and on symptoms of schizotypal and paranoid personality

disorder." Furthermore, based on a number of analyses, Psychoticism was described as encompassing "impulsivity, lack of socialization and responsibility, aggression, a strong need for independence, and sensation seeking," with clinical extremes (Zuckerman, 1989).

Recently, a subset of HeiDE participants was genotyped on whole-genome arrays, and here, we examine genomic underpinnings of The Heidelberg Five. To establish genetic similarities to and differences from the well-established Big Five trait Neuroticism, we also examined associations of The Heidelberg Five with polygenic risk scores (PRS) for Neuroticism. Also, we research associations of The Heidelberg Five with psychopathological symptoms about 20 years after their initial assessment.

2 | METHODS

Data were analyzed using R (v3.1 or higher; R Core Team, 2014), PLINK 1.9 (GWAS and calculation of PRS; Chang et al., 2015), SHAPEIT/IMPUTE2 (imputation; Delaneau et al., 2012; Howie et al., 2009), MAGMA (v1.07; gene, gene-set, and tissue expression analyses; de Leeuw et al., 2015), and GCTA (v1.92.1beta6, estimation of SNP-based heritabilities and genetic correlations; Yang et al., 2011).

The analyses are covered by an ethics vote of the Medical Faculty of the University of Heidelberg (# 026/2001).

2.1 | Heidelberg Cohort Study of the Elderly

The HeiDE study is a population-based longitudinal cohort study of the inhabitants of Heidelberg (Germany) and was designed to prospectively research the association of personality and somatic diseases. Details on the baseline sample, assessed from 1992 to 1994, can be found in Amelang et al. (2004). The final baseline sample consisted of 5114 individuals (52.2% female) aged between 28 and 74 (99.6% between 40 and 68). Data analyzed in this study are from the baseline assessment (personality phenotypes; see below), the first follow-up (on average 8.5 years later; collection of biomaterials), and from a follow-up conducted in 2013 (psychiatric phenotypes).

2.2 | Personality assessment, principal components factor analysis, and generation of factor scores

At baseline, participants completed an array of personality and health-related questionnaires (see Data S1). We used the original dataset of Amelang et al. (2004) and re-analyzed it using principal components followed by varimax rotation using the R *psych* library, obtaining a virtually identical solution. Regression factor scores were calculated for each latent personality dimension.

2.3 | Genotyping and imputation

DNA from saliva collected with mouthwash samples was extracted on a chemagic platform (PerkinElmer chemagen Technologie GmbH, Germany). DNA collected with Oragene OG500 Kits (DNA Genotek Inc., Canada) was extracted using DNA Genotek's prepIT kit (DNA Genotek Inc., Canada). Samples were genotyped using two different Illumina microarrays (Illumina, San Diego, CA). One subsample ($HeiDE_1$) was genotyped using the Infinium PsychArray-24 BeadChip (n = 2734) and another one using the InfiniumOmniExpressExome-8v1-3_A BeadChip ($HeiDE_2$; n = 1000). The combined dataset (n = 3734 pre-QC) was imputed to the 1000 Genomes phase 3 reference panel. Details on quality control (QC) and imputation can be found in the Supporting Information.

2.4 | Descriptive statistics of the genotyped sample

Of 3320 genotyped HeiDE participants (post-QC), 34 had missing personality phenotypes, and 18 were excluded because the phenotypic sex at baseline was either missing or did not match the sex recorded at follow-up. Thus, 3268 genotyped ($HeiDE_1$: n = 2387, $HeiDE_2$: n = 881) were contained in the final sample. At baseline, these individuals were 52.8 ± 7.0 (mean ± *SD*) years old (range 28–70; 99.6% were between 40 and 68 years old), 52.3% of them were female.

2.5 | Genome-wide association studies

We conducted a genome-wide association study (GWAS) for each personality phenotype. The covariates for each phenotype were the following: age, sex, and the first four multidimensional scaling (MDS) components of the pairwise identity-by-state distance matrix calculated on the nonimputed genotype data.

2.6 | Gene-set and gene property analyses

MAGMA gene-set and gene property tissue-specific expression analysis (GTEx v7, 53 tissue types) were performed as part of the FUMA (Watanabe et al., 2017) pipeline.

2.7 | SNP-based heritabilities and genetic correlation

For each of The Heidelberg Five personality traits, we estimated the aggregate proportion of variance explained by the additive effects of all genetic SNPs/variants and genetic correlations between pairwise combinations of personality traits using GCTA GREML. We estimated the genetic relationships among all HeiDE participants, excluding cryptically related individuals with genetic similarity $\hat{\pi} > 0.025$, and

using the same covariates as in the GWAS analyses. We used the –grm-adj 0 flag and thus assumed that causal loci have a similar distribution of allele frequencies as the genotyped SNPs.

2.8 | Calculation of PRS

We used summary statistics of a large GWAS on Neuroticism (Okbay et al., 2016) by the Social Science Genetic Association Consortium (n = 170911) as training data. PRS were calculated as the sum of the imputation dosage for each risk allele multiplied by the effect size of each genetic variant. SNPs overlapping between the Neuroticism GWAS and the HeiDE sample were clumped with an LD threshold of 0.2 within a 500 kb window. Subsequently, PRS were calculated at 12 different p-value thresholds (from 1×10^{-6} to 1). For each of The Heidelberg Five personality traits, we first evaluated a baseline linear regression model that predicted the factor scores of each individual personality dimension by age, sex, and the first four MDS components. We subsequently regressed residuals of the latter model onto Neuroticism PRS.

2.9 | The Heidelberg Five and psychiatric phenotypes at follow-up

We evaluated whether The Heidelberg Five, assessed at baseline, were associated with current depressive symptoms and lifetime anxiety phenotypes about 20 years later. The HeiDE subsample used in these analyses consisted of n = 2888 individuals, were 71.5 ± 6.6 (mean ± SD, approximated by year of birth) years old (range 53–87), and 47.4% were female (n = 2718 and n = 2660 individuals without missing data were used for analyses of depressive and anxiety symptoms, respectively). Current (past 3 months) depressive symptoms were assessed using the German version of the 15-item CES-D questionnaire (Radloff, 1977; Hautzinger & Bailer, 1993; range of sum scores: 15–60). Using linear regression, we evaluated whether current depressive symptoms were associated with year of birth, sex, and factor scores of each of The Heidelberg Five measured at baseline. Visual inspection of the residuals indicated that these were not normally distributed (data not shown). We therefore log-transformed depression sum scores and subsequent visual inspection of the residuals of this model did not show obvious deviation from normality (see Supporting Information). We also tested, using logistic regression, whether a positive answer to at least one of six yes/no screening questions for lifetime anxiety symptoms (see Supporting Information) at the second follow-up was associated with year of birth, sex, and the factor scores of each of The Heidelberg Five measured at baseline. The R^2 of both models was calculated using the R *rsq* package. For anxiety symptoms, we used a variance-function-based R^2 for generalized linear models (Zhang, 2017).

2.10 | Correction procedures for multiple testing

When analyzing each of The Heidelberg Five personality dimensions separately by GWAS, gene-based, gene-set, and gene property analyses, we used the conservative Bonferroni threshold to correct p-values, to minimize false-positives. In the analyses that compared PRS across different p-value thresholds, and in the analyses in which SNP-based heritabilities were compared across all personality dimensions, we used the more powerful false-discovery rate (FDR; Benjamini & Hochberg, 1995). The latter method was also used when adjusting the p-values of the longitudinal associations of depressive and anxiety symptoms, due to the inherent dependency of both phenotypes.

3 | RESULTS

3.1 | The Heidelberg Five personality dimensions

We extracted the five personality dimensions ELAB, LBCN, TYAB, LOCC, and PSYC (Figure S1). These explained 22%, 14%, 10%, 8%, and 7% of the total variance (cumulative variance explained: 61%). The resulting factor scores had the following ranges: ELAB: −2.78 to 5.07; LBCN: −3.08 to 3.61; TYAB: −3.01 to 4.69; LOCC: −4.12 to 3.66; PSYC: −2.19 to 8.56.

3.2 | Genomic underpinnings of The Heidelberg Five

Tables 1 and 2 detail the results of SNP-based heritability analyses of and genetic correlation analyses between The Heidelberg Five. Nominally significant negative genetic correlations were found between ELAB and LBCN (Table 2) and ELAB and PSYC (Table 3), but these did not remain significant after correction for multiple testing.

3.3 | Emotional Lability

The GWAS of ELAB did not yield a genome-wide significant result (for details see Supporting Information). Gene-based tests identified the gene *Integrin Subunit Beta 5* (*ITGB5*) as significantly associated (z = 4.66, $p = 1.56 \times 10^{-6}$, n = 3268; see Figure 1). Tissue expression and gene-set analyses did not yield significant results (for details see Data S1). The SNP-based heritability was significant (33.9%, Table 1).

TABLE 1 Single nucleotide polymorphism (SNP)-based heritabilities of The Heidelberg Five personality dimensions (n = 2948 for each phenotype)

Phenotype	h^2_{SNP}	*SE*	Nominal *p*	*p*FDR
ELAB	0.339	0.131	0.004	0.019
LBCN	0.294	0.134	0.014	0.034
TYAB	0.063	0.132	0.320	0.320
LOCC	0.093	0.128	0.229	0.320
PSYC	0.079	0.131	0.275	0.320

Abbreviations: ELAB, Emotional Lability; *p*FDR, FDR-corrected p-value; LBCN, Lack of Behavioral Control; LOCC, Locus of Control over Disease; PSYC, Psychoticism; *SE*, standard error; h^2_{SNP}, SNP-based heritability; TYAB, Type A Behavior.

TABLE 2 Bivariate genetic correlations between The Heidelberg Five personality dimensions (1-tailed test, n = 5896 for each phenotype pair)

Phenotype 1	Phenotype 2	r_G	*SE*	Nominal *p*	*p*FDR
ELAB	LBCN	−0.497	0.345	0.0489	0.197
ELAB	LOCC	−0.901	1.015	0.059	0.197
ELAB	PSYC	−1.000	0.942	0.044	0.197
ELAB	TYAB	−0.700	0.982	0.164	0.329
LBCN	LOCC	−0.400	0.731	0.274	0.391
LBCN	PSYC	−0.677	0.847	0.157	0.329
LBCN	TYAB	0.808	1.340	0.246	0.391
TYAB	LOCC	−1.000	2.679	0.5	0.500
TYAB	PSYC	1.000	2.095	0.5	0.500
LOCC	PSYC	−1.000	1.717	0.5	0.500

Note: None of the correlations survived FDR correction.
Abbreviations: ELAB, Emotional Lability; FDR, false-discovery rate; r_G, genetic correlation; LBCN, Lack of Behavioral Control; LOCC, Locus of Control over Disease; PSYC, Psychoticism; *SE*, standard error; TYAB, Type A Behavior.

TABLE 3 Top 10 SNPs from the GWAS of the phenotype PSYC

SNP	CHR	BP	A1	A2	FRQ	INFO	BETA	*SE*	*p*-value
rs138223660	8	130801535	T	C	0.9861	0.8937	−0.6671	0.1087	9.58e-10
rs112196460	8	130842384	C	T	0.9869	0.9353	−0.6669	0.1092	1.139e-09
rs113875761	8	130787390	C	G	0.9859	0.8883	−0.6568	0.1082	1.447e-09
rs117161072	8	131417500	G	A	0.9868	0.9033	−0.6697	0.1109	1.745e-09
rs142975048	8	130743235	A	G	0.9848	0.8279	−0.6466	0.1081	2.472e-09
rs147237681	8	131027478	C	A	0.9863	0.9766	−0.6113	0.1047	5.761e-09
rs139795768	8	131314036	T	A	0.9877	0.8709	−0.6609	0.117	1.748e-08
rs9882438	3	36776413	A	G	0.0225	0.8547	0.4916	0.0879	2.414e-08
rs143320762	8	130749679	C	T	0.9892	0.8892	−0.6883	0.1237	2.827e-08
rs187876956	8	130953808	C	T	0.9873	0.9854	−0.5949	0.1082	4.172e-08

Abbreviations: A1, allele 1; A2, allele 2; BP, position; CHR, chomosome; FRQ, allele 1 frequency; GWAS, genome-wide association study; INFO, R^2 quality metric/information content; PSYC, Psychoticism; *SE*, standard error of effect estimate; SNP, single nucleotide polymorphism.

3.4 | Low Behavioral Control

GWAS, gene-based tests, tissue expression, and gene-set analyses did not show significant results (for details see Supporting Information). We observed, however, a significant SNP-based heritability (29.4%, Table 1). Apart from the nominally significant genetic correlation with ELAB mentioned above, none of the genetic correlations between LBCN and the other personality dimensions were significant.

3.5 | Type A Behavior

There was no genome-wide significant result for TYAB (for details see Supporting Information). The gene-based analysis identified the gene *Coiled-coil Domain Containing 83* (*CCDC83*) as significantly associated (z = 4.63, $p = 1.81 \times 10^{-6}$, n = 3268; see Figure 2) and three SNPs in the *CCDC83* gene (rs56160063, rs35944027, and rs60894727) were among the top 10 SNPs in the GWAS (all $p < 2.07 \times 10^{-6}$; see Supporting Information). Tissue expression and gene-set analyses did not yield significant results, neither did the SNP-based heritability analysis nor analyses of genetic correlations between The Heidelberg Five personality dimensions (for details see Supporting Information).

3.6 | Locus of Control over Disease

For LOCC, neither GWAS, gene-based, gene-set, tissue expression, SNP-based heritability, nor genetic correlation analyses yielded significant results (for details see Supporting Information).

3.7 | Psychoticism

The GWAS of PSYC identified a significantly associated locus on chromosome (top SNP rs138223660, $p = 9.58 \times 10^{-10}$; Figure 3). Genes at this locus include *Gasdermin C* (*GSDMC*), *Family With Sequence*

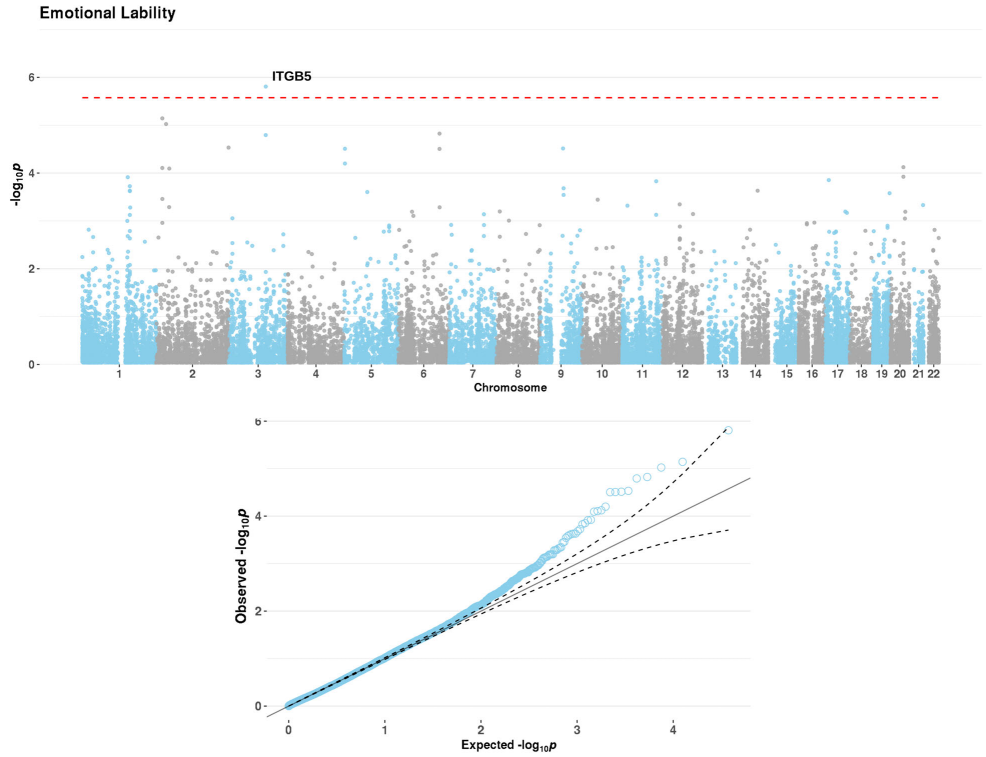


FIGURE 1 Manhattan (top) and Q–Q plots (bottom) of the gene-based test of the phenotype ELAB. Genome-wide significance level (Bonferroni-corrected for 18,776 genes) is indicated by the red dashed line [Color figure can be viewed at wileyonlinelibrary.com]

Similarity 49 Member B (FAM49B), and *ArfGAP With SH3 Domain, Ankyrin Repeat And PH Domain 1 (ASAP1)*. Another SNP on chromosome 3 (rs9882438, $p = 2.41 \times 10^{-8}$), located in an intron of the *Doublecortin Like Kinase 3 (DCLK3)* gene was also significantly associated with PSYC. The 10 SNPs with the lowest *p*-values are shown in Table 3. Gene-based analyses identified the gene *Nuclear Receptor Subfamily 1 Group H Member 4* (*NR1H4*, $z = 4.983$, $p = 3.13 \times 10^{-7}$; Figure 4) on chromosome 12 as associated with PSYC. Also, gene-set analysis identified the gene-set *GO_mf:go_bile_acid_binding* as over-represented in the GWAS results ($p_{Bon} < 0.05$, see Supporting Information). Neither tissue expression nor SNP-based heritability analysis yielded significant results.

3.8 | Associations of The Heidelberg Five with polygenic risk for neuroticism

We assessed the extent to which each of The Heidelberg Five personality dimensions shares a genetic basis with the clinically relevant Big Five personality trait Neuroticism by explaining the residuals of baseline regression models (each containing age, sex, and the first four ancestry principal components) by PRS for Neuroticism. Neuroticism PRS were significantly associated with ELAB (Figure 5), but not with the remaining personality dimensions (see Supporting Information). The direction of the association was positive, and the adjusted R^2s of FDR-significant *p*-value thresholds (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1) were 0.0024, 0.0031, 0.0025, 0.0030, 0.0027, 0.0029, 0.0026, and 0.0027 (see the legend of Figure 5 for *p*-values).

3.9 | Associations of The Heidelberg Five and psychopathology at follow-up

Table 4 lists the result of the regression analyses. All personality dimensions showed significant longitudinal associations with current depressive symptoms about 20 years after assessment. Regarding lifetime anxiety symptoms, ELAB, LBCN, and PSYC, but not LOCC or TYAB, were significantly associated.

4 | DISCUSSION

The aim of this work was to further characterize The Heidelberg Five using information on common genetic variants and long-term follow-

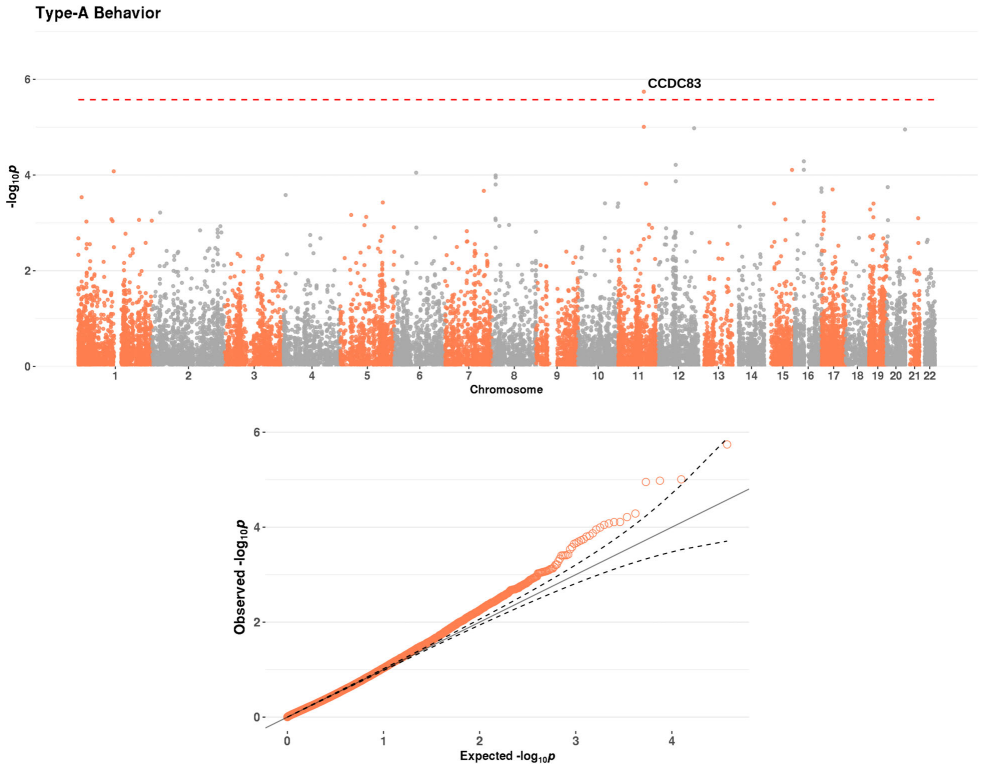


FIGURE 2 Manhattan (top) and Q–Q plots (bottom) of the gene-based test of the phenotype TYAB. Genome-wide significance level (Bonferroni-corrected for 18,776 genes) is indicated by the red dashed line [Color figure can be viewed at wileyonlinelibrary.com]

up data, to gain a more comprehensive understanding of both their biological basis and their putative importance in predicting longitudinal outcomes. Regarding the follow-up analysis, it was surprising that each of The Heidelberg Five (high ELAB, low behavioral control, high TYAB, low internal LOCC, and high PSYC) was associated with more severe depressive symptoms, measured at the 20-year follow-up. These findings alone corroborate the importance of health-related personality traits, providing justification for further research.

Different SNP-based heritabilities across The Heidelberg Five furthermore suggest a varying importance of common genetic variants, albeit this may depend on the population under study (Moore & Shenk, 2017). Specifically, both ELAB and LBCN showed substantial SNP-based heritabilities. Studies that researched SNP-based heritability of Neuroticism, both phenotypically and genetically related to ELAB, report substantially lower heritability estimates around 15% (Docherty et al., 2016; Power & Pluess, 2015). Thus, in an elderly population-based sample, ELAB appears to tap a combination of characteristics that have a relatively strong common genetic basis. Interestingly, high Neuroticism, low Extraversion, and increased age have been also found to be associated with depression scores in a large Norwegian population-based study (Grav et al., 2012), supporting the ELAB construct. Our finding thus underscores the relevance of clinically valid personality dimensions for genetic research. It is conceivable that, in the elderly, ELAB defines a personality dimension having a strong genetic background, being jointly defined by a combination of two Big Five personality factors (Neuroticism and Extraversion).

LBCN also showed a relatively high SNP-based heritability which is supported by previous twin studies that found differences in executive control functions to be almost entirely genetic in origin (Friedman et al., 2008), lending support to the notion of LBCN as a latent "executive function" personality trait. Indeed, the unity/diversity framework for executive functions (Friedman & Miyake, 2017) describes inhibition as a key element. The loadings of low values on Anger Control and Social Desirability scales, and high values on Aggression, Irritability, and Anger Out scales (Amelang et al., 2004) onto the LBCN factor appear to fit well with this interpretation. Furthermore, as described in the case of ELAB, LBCN may define a clinically valid personality dimension for genetic research. Further research is necessary to determine whether an association of LBCN with cognitive tests of

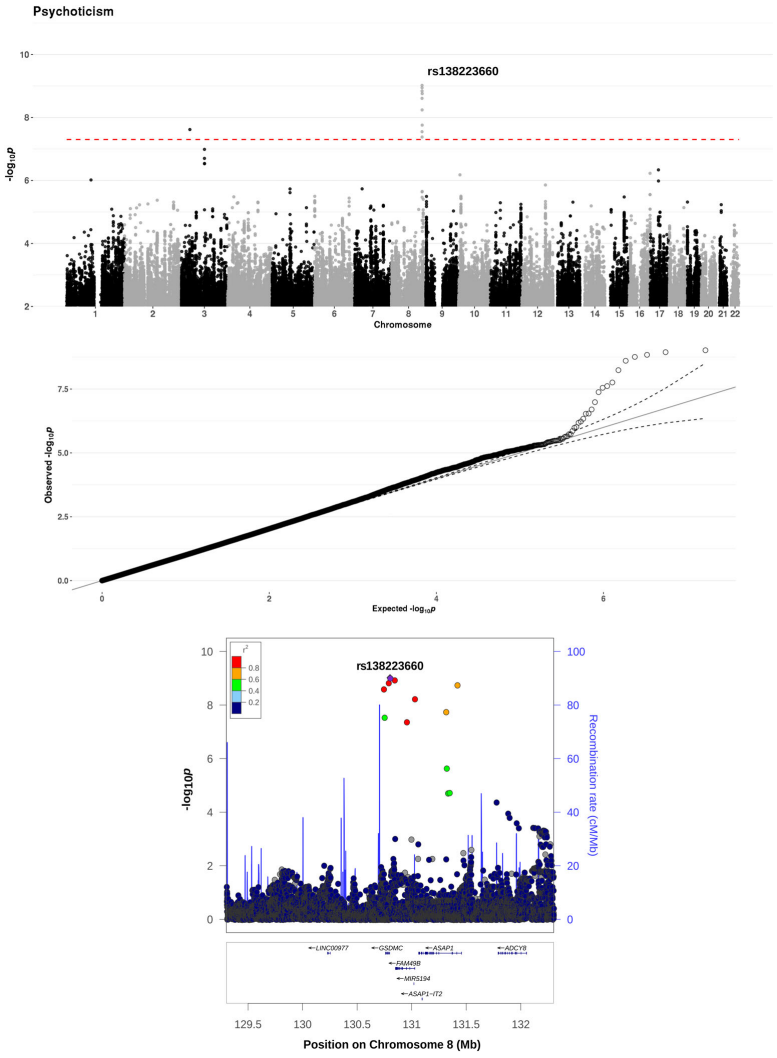


FIGURE 3 Manhattan (top), Q–Q (middle, λ = 1.008), and regional association (bottom) plots of the GWAS of the phenotype PSYC [Color figure can be viewed at wileyonlinelibrary.com]

executive function holds, as has been found for other behavioral constructs (Friedman et al., 2018; Morgan & Lilienfeld, 2000).

In the present study, we did not detect significant h^2_{SNP} for TYAB, LOCC, or PSYC. Similar to Power and Pluess (2015), who found significant SNP-based heritabilities for only two Big Five traits (Neuroticism and Openness), this may be interpreted as emphasizing the putative importance of rare or structural variants for these personality dimensions, as all personality phenotypes are heritable to some degree (Turkheimer et al., 2014). Also, it is possible that unknown environmental covariates exist that explain more phenotypic variance of TYAB, LOCC, and PSYC, and accounting for these would result in larger observed SNP-based heritabilities also for these personality dimensions.

The orthogonality of The Heidelberg Five on the phenotype level is reflected by nonsignificant genetic correlations between them. Conversely, both the phenotypic and genotypic relatedness of Neuroticism and ELAB is reflected in substantial SNP-based heritabilities of both traits (see above) and by the result that Neuroticism PRS explain variation of ELAB. This was not the case for the remaining Heidelberg Five personality dimensions. Gene-based analysis of the ELAB phenotype identified *ITGB5*, encoding a transmembrane protein. The family of integrins, to which *ITGB5* belongs, are membrane proteins that translate intracellular signaling to extracellular interactions. They have been associated with neuropsychiatric disease (Carneiro, 2010) and coordinate both synaptic structure and function (Park & Goda, 2016).

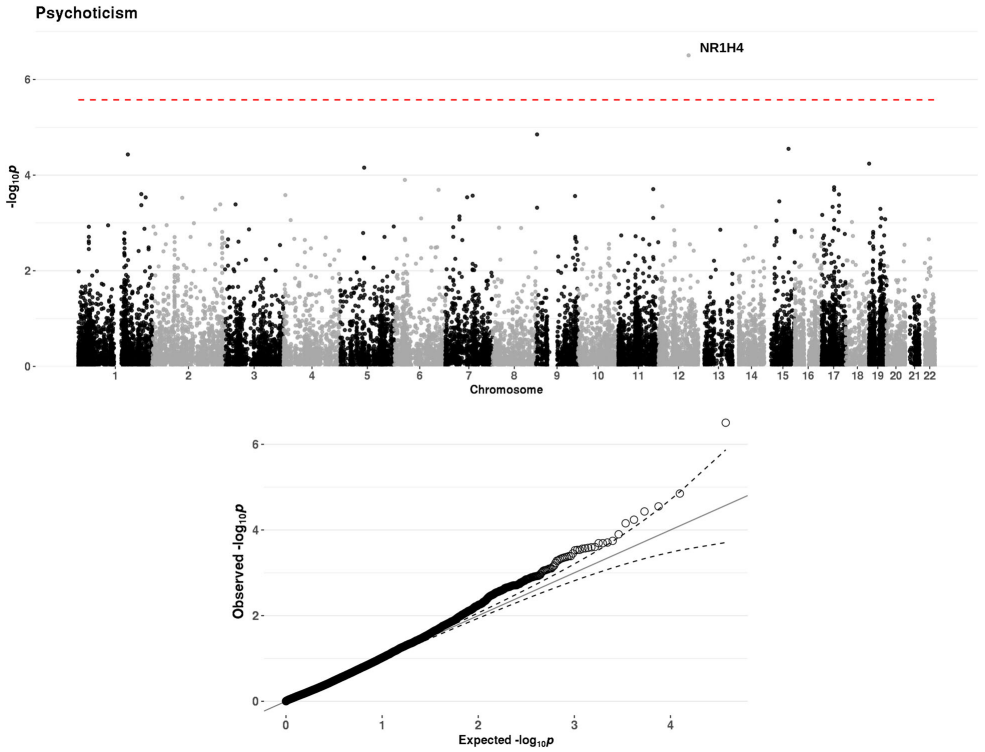


FIGURE 4 Manhattan (top) and Q–Q plots (bottom) of the gene-based test of the phenotype PSYC. Genome-wide significance level (Bonferroni-corrected for 18,776 genes) is indicated by the red dashed line [Color figure can be viewed at wileyonlinelibrary.com]

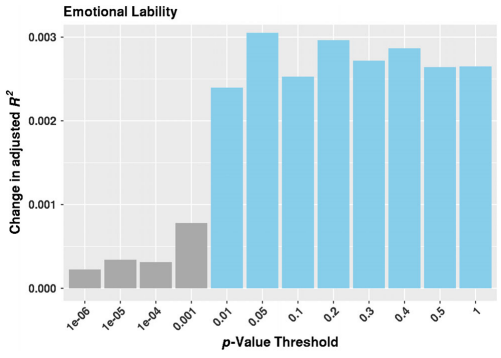


FIGURE 5 Effects (adjusted R^2s) of PRS for neuroticism at different *p*-value thresholds on the residuals of a model regressing the personality dimension ELAB onto a set of baseline variables (see section 2). FDR-corrected *p*-values of the PRS were 0.187, 0.168, 0.168, 0.078, 0.004, 0.004, 0.004, 0.004, 0.004, 0.004, 0.004, and 0.004 [Color figure can be viewed at wileyonlinelibrary.com]

Furthermore, SNPs in the *ITGB5* gene are associated with blood pressure (Giri et al., 2019) and coronary artery disease (Nelson et al., 2017). Interestingly, an association between ELAB and the incidence of CVD was previously identified in longitudinal analyses (Amelang et al., 2004) and thus may suggest a common genetic basis of both. Concerning the longitudinal associations of high ELAB scores with both depressive and anxiety symptoms, observed in the present study, confirm the well-known clinical importance of this Neuroticism-like phenotype (Gale et al., 2016), and are in line with meta-analyses of longitudinal studies of Neuroticism (Hakulinen et al., 2015; Jeronimus et al., 2016).

Regarding TYAB, gene-based analysis pointed to the protein-coding gene *Coiled-coil Domain Containing 83* (*CCDC83*). In European populations, this gene has been linked to urinary tract infection frequency (Tian et al., 2017), but not to behavioral phenotypes.

While no significant SNP-based heritability of PSYC was detected in the present study, GWAS and gene-based analysis revealed significant loci on chromosomes 3, 8, and 12. Of the genes in these loci, a SNP in *FAM49B* showed a suggestive association with post-traumatic

TABLE 4 Regression analyses of current depressive and lifetime anxiety symptoms approximately 20 years after assessment of The Heidelberg Five

Current depressive symptoms					
	Estimate	***SE***	***t*-value**	***p*-value**	p_{FDR}
Intercept	8.5428550	1.1990121	7.125	1.33e-12	5.317157e-12
Year of birth	*−0.0027880*	*0.0006175*	*−4.515*	*6.60e-06*	*1.056070e-05*
Sex (male)	*−0.0234999*	*0.0080501*	*−2.919*	*0.00354*	*4.717726e-03*
ELAB	*0.1315498*	*0.0040296*	*32.646*	*<2e-16*	*<2e-16*
LBCN	*0.0363260*	*0.0041415*	*8.771*	*<2e-16*	*<2e-16*
TYAB	*0.0200983*	*0.0040783*	*4.928*	*8.80e-07*	*1.564546e-06*
LOCC	*−0.0084948*	*0.0040523*	*−2.096*	*0.03615*	*4.131205e-02*
PSYC	*0.0162838*	*0.0039524*	*4.120*	*3.90e-05*	*5.675753e-05*
Lifetime anxiety symptoms					
	Estimate	***SE***	***z*-value**	***p*-value**	p_{FDR}
Intercept	−64.069467	12.953826	−4.946	7.58e-07	1.556365e-06
Year of birth	*0.032959*	*0.006671*	*4.941*	*7.78e-07*	*1.556365e-06*
Sex (male)	*−0.450115*	*0.085660*	*−5.255*	*1.48e-07*	*3.954736e-07*
ELAB	*0.613520*	*0.045513*	*13.480*	*<2e-16*	*<2e-16*
LBCN	*0.295001*	*0.044516*	*6.627*	*3.43e-11*	*1.097851e-10*
TYAB	*0.047679*	0.042997	1.109	0.2675	2.853136e-01
LOCC	0.006232	0.042799	0.146	0.8842	8.842379e-01
PSYC	*0.110470*	*0.043120*	*2.562*	*0.0104*	*1.281183e-02*

Note: The adjusted R^2 of the model (see text) was 31.1% for current depressive symptoms and 11.2% for lifetime anxiety symptoms. FDR-significant regressors are in italics.
Abbreviations: ELAB, Emotional Lability; p_{FDR}, FDR-adjusted *p*-value; LBCN, Lack of Behavioral Control; LOCC, Locus of Control over Disease; PSYC, Psychoticism; *SE*, standard error of the estimate; TYAB, Type A Behavior.

stress disorder (Xie et al., 2013). Furthermore, SNPs in *ASAP1* were suggestively associated with autism spectrum disorder (Grove et al., 2019) and, in individuals with Ashkenazi Jewish ancestry, suggestively associated with schizophrenia (Goes et al., 2015). Finally, the gene-set *GO_mf:go_bile_acid_binding* was overrepresented among the PSYC results. The 10 genes in this gene set include *NR1H4* (also significant in the gene-based analysis), and the *Vitamin-D Receptor*, both of which are ligand-inducible transcription factors. The genes regulated by these transcription factors may thus contribute to the personality dimension PSYC.

5 | CONCLUSIONS

Several findings emerge from the present analyses of The Heidelberg Five. First, each personality dimension is associated with psychiatric phenotypes, measured some 20 years later, which underlines their clinical significance. Second, ELAB is genetically related to Neuroticism. As ELAB explained most of the phenotypic variance in the factor analysis, the behavioral importance of this clinical personality dimension is further underscored. Third, LBCN, a previously unknown latent "executive function" personality dimension has emerged as a heritable trait of clinical importance, warranting further investigation.

Our results need to be interpreted keeping the following limitations in mind: While based on longitudinal data, we used cross-sectional analyses ignoring accrual and mortality. If any of the traits or SNPs are associated with accrual or mortality, this will introduce selection bias. Results of the effects of psychological traits on CVD and cancer, including cause-specific mortality, are reassuring, however, insofar as most had no major impact on these outcomes (Stürmer et al., 2006). Finally, both the sample size, and the lack of a replication sample should be borne in mind when interpreting GWAS results.

ACKNOWLEDGMENTS

The authors are grateful to all investigators and participants of the HeiDE study. Thomas G. Schulze was supported by the Deutsche Forschungsgemeinschaft (DFG) within the framework of the projects www.kfo241.de and www.PsyCourse.de (SCHU 1603/4-1, 5-1, 7-1, FA241/16-1) and by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme with grants awarded to Thomas G. Schulze (01ZX1614K), Marcella Rietschel (01ZX1614G), Markus M. Nöthen (01ZX1614A), and Bertram Müller-Myhsok (01ZX1614J). Thomas G. Schulze received additional support from the BMBF within the framework of the BipoLife network

(01EE1404H) and the Dr. Lisa Oehler Foundation, Kassel (Germany). Marcella Rietschel was additionally supported by the ERA-NET NEURON grants "SynSchiz" (01EW1810) and "EMBED" (01EW1904), and the DFG (FOR2107, RI908/11-2). Franziska Degenhardt received support from the BONFOR Program of the University of Bonn, Germany. Stephanie H. Witt was supported by the DFG (FOR2107, WI3429/3-2). Till F. M. Andlauer was supported by the BMBF through the DIFUTURE consortium of the Medical Informatics Initiative Germany (01ZZ1804A) and the European Union's Horizon 2020 Research and Innovation Programme (MultipleMS, EU RIA 733161). Til Stürmer received investigator-initiated research funding and support as Principal Investigator (R01 AG056479) from the National Institute on Aging (NIA), and as Co-Investigator (R01 HL118255, R01 MD011680), National Institutes of Health (NIH). He also received salary support as Director of Comparative Effectiveness Research (CER), NC TraCS Institute, UNC Clinical and Translational Science Award (UL1TR002489), the Center for Pharmacoepidemiology (current members: GlaxoSmithKline, UCB BioSciences, Takeda, AbbVie, Boehringer Ingelheim), from pharmaceutical companies (Novo Nordisk), and from a generous contribution from Dr. Nancy A. Dreyer to the Department of Epidemiology, University of North Carolina at Chapel Hill. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

Til Stürmer owns stock in Novartis, Roche, and Novo Nordisk, but does not accept personal compensation of any kind from any pharmaceutical company. All other authors declare no Conflict of Interest.

AUTHOR CONTRIBUTIONS

Urs Heilbronner: Conceptualization; formal analysis; visualization; writing – original draft. **Sergi Papiol**: Supervision; writing – review and editing. **Monika Budde**: Conceptualization; writing – review and editing. **Till F. M. Andlauer**: Formal analysis; writing – review and editing. **Jana Strohmaier**: Resources; writing – review and editing. **Fabian Streit**: Resources; writing – review and editing. **Josef Frank**: Resources; writing – review and editing. **Franziska Degenhardt**: Resources; writing – review and editing. **Stefanie Heilmann-Heimbach**: Investigation; resources; writing – review and editing. **Stephanie Witt**: Resources; writing – review and editing. **Andreas J. Forstner**: Writing – review and editing. **Adrian Loerbroks**: Investigation; resources; writing – review and editing. **Manfred Amelang**: Investigation; writing – review and editing. **Til Stürmer**: Investigation; resources; writing – review and editing. **Bertram Müller-Myhsok**: Writing – review and editing. **Markus M. Nöthen**: Resources; writing – review and editing. **Marcella Rietschel**: Conceptualization; resources; project administration; writing – review and editing. **Thomas G. Schulze**: Conceptualization; funding acquisition; project administration; writing – review and editing.

DATA AVAILABILITY STATEMENT

Due to the sensitivity of individual-level genetic data, these data and the corresponding analysis scripts are available from the authors based on reasonable request. Summary statistics of each GWAS are available online (GWAS Catalog, www.ebi.ac.uk/gwas/home), accession numbers GCST90013452 (ELAB), GCST90013453 (LBCN), GCST90013454 (TYAB), GCST90013455 (LOCC), and GCST9 0013456 (PSYC).

ORCID

Urs Heilbronner https://orcid.org/0000-0001-7135-762X
Till F. M. Andlauer https://orcid.org/0000-0002-2917-5889
Fabian Streit https://orcid.org/0000-0003-1080-4339

REFERENCES

Amelang, M., Hasselbach, P., & Stürmer, T. (2004). Personality, cardiovascular disease, and cancer: First results from the Heidelberg cohort study of the elderly. *Zeitschrift für Gesundheitspsychologie*, *12*(3), 102–115.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Capitanio, J. P. (2008). Personality and disease. *Brain, Behavior, and Immunity*, *22*(5), 647–650. https://doi.org/10.1016/j.bbi.2008.02.002

Carneiro, A. M. D. (2010). The emerging role of integrins in neuropsychiatric disorders. *Neuropsychopharmacology*, *35*(1), 338–339. https://doi.org/10.1038/npp.2009.134

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7. https://doi.org/10.1186/s13742-015-0047-8

Chapman, J. P., Chapman, L. J., & Kwapil, T. R. (1994). Does the Eysenck psychoticism scale predict psychosis? A ten year longitudinal study. *Personality and Individual Differences*, *17*(3), 369–375. https://doi.org/10.1016/0191-8869(94)90284-4

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Computational Biology*, *11*(4), e1004219. https://doi.org/10.1371/journal.pcbi.1004219

Delaneau, O., Marchini, J., & Zagury, J.-F. (2012). A linear complexity phasing method for thousands of genomes. *Nature Methods*, *9*(2), 179–181. https://doi.org/10.1038/nmeth.1785

Docherty, A. R., Moscati, A., Peterson, R., Edwards, A. C., Adkins, D. E., Bacanu, S. A., … Kendler, K. S. (2016). SNP-based heritability estimates of the personality dimensions and polygenic prediction of both neuroticism and major depression: Findings from CONVERGE. *Translational Psychiatry*, *6*(10), e926–e926. https://doi.org/10.1038/tp.2016.177

Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. London, England: Hodder and Stoughton.

Friedman, H. S., & Booth-Kewley, S. (1987). Personality, type A behavior, and coronary heart disease: The role of emotional expression. *Journal of Personality and Social Psychology*, *53*(4), 783–792. https://doi.org/10.1037//0022-3514.53.4.783

Friedman, M., & Rosenman, R. H. (1959). Association of specific overt behavior pattern with blood and cardiovascular findings; blood cholesterol level, blood clotting time, incidence of arcus senilis, and clinical coronary artery disease. *Journal of the American Medical Association*, *169*(12), 1286–1296. https://doi.org/10.1001/jama.1959.03000290012005

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology. General*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Friedman, N. P., Rhee, S. H., Ross, J. M., Corley, R. P., & Hewitt, J. K. (2018). Genetic and environmental relations of executive functions to antisocial personality disorder symptoms and psychopathy. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2018.12.007

Gale, C. R., Hagenaars, S. P., Davies, G., Hill, W. D., Liewald, D. C. M., Cullen, B., … Harris, S. E. (2016). Pleiotropy between neuroticism and physical and mental health: Findings from 108038 men and women in UK Biobank. *Translational Psychiatry*, *6*, e791. https://doi.org/10.1038/tp.2016.56

Giri, A., Hellwege, J. N., Keaton, J. M., Park, J., Qiu, C., Warren, H. R., … Edwards, T. L. (2019). Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nature Genetics*, *51*(1), 51–62. https://doi.org/10.1038/s41588-018-0303-9

Goes, F. S., McGrath, J., Avramopoulos, D., Wolyniec, P., Pirooznia, M., Ruczinski, I., … Pulver, A. E. (2015). Genome-wide association study of schizophrenia in Ashkenazi Jews. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *168*(8), 649–659. https://doi.org/10.1002/ajmg.b.32349

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34. https://doi.org/10.1037//0003-066x.48.1.26

Grav, S., Stordal, E., Romild, U. K., & Hellzen, O. (2012). The relationship among neuroticism, extraversion, and depression in the HUNT study: In relation to age and gender. *Issues in Mental Health Nursing*, *33*(11), 777–785. https://doi.org/10.3109/01612840.2012.713082

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., … Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 431–444. https://doi.org/10.1038/s41588-019-0344-8

Hakulinen, C., Elovainio, M., Pulkki-Råback, L., Virtanen, M., Kivimäki, M., & Jokela, M. (2015). Personality and depressive symptoms: individual participant meta-analysis of 10 cohort studies: Personality and depression. *Depression and Anxiety*, *32*(7), 461–470. https://doi.org/10.1002/da.22376

Hautzinger, M., & Bailer, M. (1993). *Allgemeine Depressions Skala: ADS*. Weinheim, Germany: Beltz.

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genetics*, *5*(6), e1000529. https://doi.org/10.1371/journal.pgen.1000529

Jeronimus, B. F., Kotov, R., Riese, H., & Ormel, J. (2016). Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: A meta-analysis on 59 longitudinal/prospective studies with 443 313 participants. *Psychological Medicine*, *46*(14), 2883–2906. https://doi.org/10.1017/S0033291716001653

Kuper, H., Marmot, M., & Hemingway, H. (2002). Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease. *Seminars in Vascular Medicine*, *2*(3), 267–314. https://doi.org/10.1055/s-2002-35401

Moore, D. S., & Shenk, D. (2017). The heritability fallacy. *Wiley Interdisciplinary Reviews Cognitive Science*, *8*(1–2), e1400. https://doi.org/10.1002/wcs.1400

Morgan, A. B., & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, *20*(1), 113–136. https://doi.org/10.1016/S0272-7358(98)00096-8

Náfrádi, L., Nakamoto, K., & Schulz, P. J. (2017). Is patient empowerment the key to promote adherence? A systematic review of the relationship between self-efficacy, health locus of control and medication adherence. *PLoS One*, *12*(10), e0186458. https://doi.org/10.1371/journal.pone.0186458

Nelson, C. P., Goel, A., Butterworth, A. S., Kanoni, S., Webb, T. R., Marouli, E., … Deloukas, P. (2017). Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nature Genetics*, *49*(9), 1385–1391. https://doi.org/10.1038/ng.3913

Okbay, A., Baselmans, B. M. L., De Neve, J.-E., Turley, P., Nivard, M. G., Fontana, M. A., … Cesarini, D. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, *48*(6), 624–633. https://doi.org/10.1038/ng.3552

Park, Y. K., & Goda, Y. (2016). Integrins in synapse regulation. *Nature Reviews. Neuroscience*, *17*(12), 745–756. https://doi.org/10.1038/nrn.2016.138

Power, R. A., & Pluess, M. (2015). Heritability estimates of the Big Five personality traits based on common genetic variants. *Translational Psychiatry*, *5*, e604. https://doi.org/10.1038/tp.2015.96

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Rohlf, H. L., Holl, A. K., Kirsch, F., Krahé, B., & Elsner, B. (2018). Longitudinal links between executive function, anger, and aggression in middle childhood. *Frontiers in Behavioral Neuroscience*, *12*, 27. https://doi.org/10.3389/fnbeh.2018.00027

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, *80*(1), 1–28.

Sanchez-Roige, S., Gray, J. C., MacKillop, J., Chen, C.-H., & Palmer, A. A. (2018). The genetics of human personality. *Genes, Brain, and Behavior*, *17*(3), e12439. https://doi.org/10.1111/gbb.12439

Smith, D. F., Sterndorff, B., Røpcke, G., Gustavsen, E. M., & Hansen, J. K. (1996). Prevalence and severity of anxiety, depression and Type A behaviors in angina pectoris. *Scandinavian Journal of Psychology*, *37*(3), 249–258. https://doi.org/10.1111/j.1467-9450.1996.tb00657.x

Steptoe, A., & Wardle, J. (2001). Locus of control and health behaviour revisited: A multivariate analysis of young adults from 18 countries. *British Journal of Psychology (London, England: 1953)*, *92*(Pt 4), 659–672. https://doi.org/10.1348/000712601162400

Strudler Wallston, B., & Wallston, K. A. (1978). Locus of control and health: A review of the literature. *Health Education Monographs*, *6*(1), 107–117. https://doi.org/10.1177/109019817800600102

Stürmer, T., Hasselbach, P., & Amelang, M. (2006). Personality, lifestyle, and risk of cardiovascular disease and cancer: Follow-up of population based cohort. *BMJ (Clinical Research Ed.)*, *332*(7554), 1359. https://doi.org/10.1136/bmj.38833.479560.80

Szczepanski, S. M., & Knight, R. T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron*, *83*(5), 1002–1018. https://doi.org/10.1016/j.neuron.2014.08.011

Tate, R. L. (1999). Executive dysfunction and characterological changes after traumatic brain injury: Two sides of the same coin? *Cortex*, *35*(1), 39–55. https://doi.org/10.1016/s0010-9452(08)70784-6

Tian, C., Hromatka, B. S., Kiefer, A. K., Eriksson, N., Noble, S. M., Tung, J. Y., & Hinds, D. A. (2017). Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nature Communications*, *8*(1), 599. https://doi.org/10.1038/s41467-017-00257-5

Turkheimer, E., Pettersson, E., & Horn, E. E. (2014). A phenotypic null hypothesis for the genetics of personality. *Annual Review of Psychology*, *65*, 515–540. https://doi.org/10.1146/annurev-psych-113011-143752

van Kampen, D. (2009). Personality and psychopathology: A theory-based revision of Eysenck's PEN model. *Clinical Practice and Epidemiology in Mental Health*, *5*, 9–21. https://doi.org/10.2174/1745017900905010009

Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*(1), 1826. https://doi.org/10.1038/s41467-017-01261-5

Weiner, B., Reisenzein, R., & Pranter, W. (2009). *Motivationspsychologie*. Weinheim, Germany: Beltz, Psychologie-Verl.-Union.

Wooldridge, K. L., Wallston, K. A., Graber, A. L., Brown, A. W., & Davidson, P. (1992). The relationship between health beliefs, adherence, and metabolic control of diabetes. *Diabetes Educator*, *18*(6), 495–500. https://doi.org/10.1177/014572179201800608

Wright, A. G. C., Thomas, K. M., Hopwood, C. J., Markon, K. E., Pincus, A. L., & Krueger, R. F. (2012). The hierarchical structure of DSM-5 pathological personality traits. *Journal of Abnormal Psychology*, *121*(4), 951–957. https://doi.org/10.1037/a0027669

Xie, P., Kranzler, H. R., Yang, C., Zhao, H., Farrer, L. A., & Gelernter, J. (2013). Genome-wide association study identifies new susceptibility loci for posttraumatic stress disorder. *Biological Psychiatry*, *74*(9), 656–663. https://doi.org/10.1016/j.biopsych.2013.04.013

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Zhang, D. (2017). A coefficient of determination for generalized linear models. *American Statistician*, *71*(4), 310–316. https://doi.org/10.1080/00031305.2016.1256839

Zuckerman, M. (1989). Personality in the third dimension: A psychobiological approach. *Personality and Individual Differences*, *10*(4), 391–418. https://doi.org/10.1016/0191-8869(89)90004-4

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Heilbronner U, Papiol S, Budde M, et al. "The Heidelberg Five" personality dimensions: Genome-wide associations, polygenic risk for neuroticism, and psychopathology 20 years after assessment. *Am J Med Genet Part B*. 2021;1–13. https://doi.org/10.1002/ajmg.b.32837

6 A Genome-Wide Association Study of the Longitudinal Course of Executive Functions (Wendel et al., 2021)

Several studies have adressed the genetics and genomics of both common and specific EFs in a developmental context (see Introduction). However, a question that has not received attention to date is whether there are genetic variants underlying differential performance in EFs over relatively short periods of time (months rather than years). To address this question, we used longitudinal data of the PsyCourse Study as discovery sample. This study assessed adult patients from the affective-to-psychotic spectrum and neurotypical individuals at up to four points in time, spaced approximately six months apart (Budde et al., 2019). Data were collected during the past decade, and combined longitudinal deep phenotyping (demographic, cognitive, symptom, and life event variables, see Heilbronner, Adorjan, et al., 2021) with an extensive collection of biomaterial. Two executive phenotypes were used in Wendel et al. (2021): The Trail-Making-Test, Part B (TMT-B; time in seconds needed to complete the test), and the Verbal Digit Span backwards test (VDS-B; total score), corresponding, respectively, to the Set-Shifting and Updating components of the Unity and Diversity EF model presented in the Introduction. Behaviorally, EF peformance improved over time, attributable to re-test effects (c.f. Bartels et al., 2010). Also, we observed the well-documented greater impairment of EFs in psychiatric patients compared to neurotypical individuals. In the discovery sample, we studied the genomics of the course of EFs using linear mixed models, focussing on SNP $\times$ time interactions. For the phenotype TMT-B, we found nine genome-wide significant SNPs on chromosome 5, associated with a pronounced slowing of responses. We addressed replicability of this locus by using TMT-B change-score data of the longitudinal FOR2107 cohort (Kircher et al., 2019). Analysis of these data also identified the top SNP, rs150547358, as significant, and thus replicated our finding. Both in size and direction, the effect estimate for rs150547358 was comparable to that found in data of the PsyCourse Study. Interestingly, this SNP is directly located in the ring finger protein 180 (*RNF180*) gene on chromosome 5q12.3. *RNF180* has been shown to be associated with monoamine levels in the PFC in *RNF180* knockout mice (Kabayama et al., 2013). Another gene

located in the nearby region, *HTR1A* (5-hydroxytryptamine receptor 1A), is an important receptor of serotonin (5-HT) also essential to the regulation of 5-HT concentrations in the frontal lobe (McDevitt & Neumaier, 2011). Furthermore, a polymorphism in the 5-HT system has previously been associated with EFs (Li & Roberts, 2018). Our study demonstrates that genomic factors associated with differential response patterns over time exist. This finding underscores the importance of genomic factors in the course of neurocognitive and psychiatric phenotypes. The more general question of the effects of time on the interplay of genetics and environment has recently been emphasized (Boyce et al., 2020).

ARTICLE OPEN

Check for updates

A genome-wide association study of the longitudinal course of executive functions

Bernadette Wendel[1✉], Sergi Papiol[2,3], Till F. M. Andlauer[4], Jörg Zimmermann[5], Jens Wiltfang[6,7,8], Carsten Spitzer[9], Fanny Senner[2,3], Eva C. Schulte[2,3], Max Schmauß[10], Sabrina K. Schaupp[2], Jonathan Repple[11], Eva Reininghaus[12], Jens Reimer[13,14], Daniela Reich-Erkelenz[2], Nils Opel[11], Igor Nenadić[15,16], Susanne Meinert[11], Carsten Konrad[17], Farahnaz Klöhn-Saghatolislam[2,3], Tilo Kircher[15,16], Janos L. Kalman[2,3,18], Georg Juckel[19], Andreas Jansen[15,16,20], Markus Jäger[21], Maria Heilbronner[2], Martin von Hagen[22], Katrin Gade[6], Christian Figge[23], Andreas J. Fallgatter[24], Detlef E. Dietrich[25,26,27], Udo Dannlowski[11], Ashley L. Comes[2,18], Monika Budde[2], Bernhard T. Baune[28,29,30], Volker Arolt[11], Ion-George Anghelescu[31], Heike Anderson-Schmidt[2,6], Kristina Adorjan[2,3], Peter Falkai[3], Thomas G. Schulze[2,3], Heike Bickeböller[1] and Urs Heilbronner[2]

Executive functions are metacognitive capabilities that control and coordinate mental processes. In the transdiagnostic PsyCourse Study, comprising patients of the affective-to-psychotic spectrum and controls, we investigated the genetic basis of the time course of two core executive subfunctions: set-shifting (Trail Making Test, part B (TMT-B)) and updating (Verbal Digit Span backwards) in 1338 genotyped individuals. Time course was assessed with four measurement points, each 6 months apart. Compared to the initial assessment, executive performance improved across diagnostic groups. We performed a genome-wide association study to identify single nucleotide polymorphisms (SNPs) associated with performance change over time by testing for SNP-by-time interactions using linear mixed models. We identified nine genome-wide significant SNPs for TMT-B in strong linkage disequilibrium with each other on chromosome 5. These were associated with decreased performance on the continuous TMT-B score across time. Variant rs150547358 had the lowest P value $= 7.2 \times 10^{-10}$ with effect estimate beta $= 1.16$ (95% c.i.: 1.11, 1.22). Implementing data of the FOR2107 consortium (1795 individuals), we replicated these findings for the SNP rs150547358 (P value $= 0.015$), analyzing the difference of the two available measurement points two years apart. In the replication study, rs150547358 exhibited a similar effect estimate beta $= 0.85$ (95% c.i.: 0.74, 0.97). Our study demonstrates that longitudinally measured phenotypes have the potential to unmask novel associations, adding time as a dimension to the effects of genomics.

Translational Psychiatry (2021)11:386 ; https://doi.org/10.1038/s41398-021-01510-8

INTRODUCTION

The term "executive functions" (EFs) describes a group of higher-level cognitive abilities [1], including the regulation of thoughts and actions in daily life [1, 2]. As humans age, EFs pass different developmental stages, in which great variability is observed both within and between individuals [3, 4]. EFs naturally decline with

[1]Department of Genetic Epidemiology, University Medical Center Göttingen, Georg-August-University Göttingen, Göttingen 37073, Germany. [2]Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Munich 80336, Germany. [3]Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich 80336, Germany. [4]Department of Neurology, University Hospital rechts der Isar, School of Medicine, Technical University of Munich, Munich 81675, Germany. [5]Psychiatrieverbund Oldenburger Land gGmbH, Karl-Jaspers-Klinik, Bad Zwischenahn 26160, Germany. [6]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen 37075, Germany. [7]German Center for Neurodegenerative Diseases (DZNE), Göttingen 37075, Germany. [8]Neurosciences and Signaling Group, Institute of Biomedicine (iBiMED), Department of Medical Sciences, University of Aveiro, Aveiro 3810-193, Portugal. [9]Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Rostock, Rostock 18147, Germany. [10]Department of Psychiatry and Psychotherapy, Bezirkskrankenhaus Augsburg, Augsburg 86156, Germany. [11]Institute for Translational Psychiatry, University of Münster, Münster 48149, Germany. [12]Department of Psychiatry and Psychotherapeutic Medicine, Research Unit for Bipolar Affective Disorder, Medical University of Graz, Graz 8036, Austria. [13]Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany. [14]Department of Psychiatry, Health North Hospital Group, Bremen 28102, Germany. [15]Department of Psychiatry and Psychotherapy, Philipps-University Marburg, Marburg 35039, Germany. [16]Centre for Mind, Brain and Behaviour, University of Marburg, Marburg 35032, Germany. [17]Department of Psychiatry and Psychotherapy, Agaplesion Diakonieklinikum, Rotenburg 27356, Germany. [18]International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Max Planck Institute of Psychiatry, Munich 80804, Germany. [19]Department of Psychiatry, Ruhr University Bochum, LWL University Hospital, Bochum 44791, Germany. [20]Core-Facility Brainimaging, Faculty of Medicine, University of Marburg, Marburg, Germany. [21]Department of Psychiatry II, Ulm University, Bezirkskrankenhaus Günzburg, Günzburg 89312, Germany. [22]Clinic for Psychiatry and Psychotherapy, Clinical Center Werra-Meißner, Eschwege 37269, Germany. [23]Karl-Jaspers Clinic, European Medical School Oldenburg-Groningen, Oldenburg 26160, Germany. [24]Department of Psychiatry and Psychotherapy, University Tübingen, Tübingen 72076, Germany. [25]AMEOS Clinical Center Hildesheim, Hildesheim 31135, Germany. [26]Center for Systems Neuroscience (ZSN), Hannover 30559, Germany. [27]Department of Psychiatry, Medical School of Hannover, Hannover 30625, Germany. [28]Department of Psychiatry, University of Münster, Münster 48149, Germany. [29]Department of Psychiatry, Melbourne Medical School, The University of Melbourne, Melbourne, Australia. [30]The Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Parkville, VIC, Australia. [31]Department of Psychiatry and Psychotherapy, Mental Health Institute Berlin, Berlin 14050, Germany. ✉email: bernadette.wendel@med.uni-goettingen.de

Received: 2 December 2020 Revised: 4 June 2021 Accepted: 15 June 2021
Published online: 10 July 2021

advanced age [4–6] in a gender-specific manner [7] and diminished EFs are also observed in the longitudinal course of severe mental disorders, such as schizophrenia [8]. In particular, EFs appear to be generally impaired in psychiatric patients suffering from schizophrenia, depression [4], or bipolar disorder [9]. Deficits are also associated, for example with decreased abilities to perform routine tasks [4]. Neurobiologically, EFs are linked intimately to the prefrontal cortex, as exemplified by the famous case of Phineas Gage [10].

There are many definitions of an EF [3], as it represents an umbrella term for multiple cognitive processes [2]. An influential theory of EFs is the "unity and diversity" concept [3, 11] that describes EFs as a "collection of related but separable abilities" [3]. EFs are differentiated into three latent core skills [3, 4, 11]: (i) set-shifting, allowing an individual to approach tasks flexibly and adjust to new conditions [3, 4], (ii) updating (or working memory), with respect to the monitoring, manipulating, and updating of information [4, 11], and (iii) inhibition, enabling an individual to control behavior, emotions, and responses [4, 11]. In general, EFs rank among the "most heritable psychological traits" [3]. On the behavioral genetic level, a highly heritable latent (common) factor affecting all EF aspects accounted for 99% of the variance common to all three skills [3]. Regarding specific EF components, the heritability estimates of set-shifting assessed by the Trail Making Test (TMT) range from 0.34 to 0.65 [12] and the estimates of updating measured by digit span tests range from 0.27 to 0.62 [12] (these results were obtained in twin studies). Recently, several genome-wide association studies (GWASs) on EFs have been undertaken [13–18]; however, genome-wide significance was not attained [2, 12]. Moreover, the genetic basis of variation over time is yet to be elucidated [19].

Here, we performed two longitudinal GWASs for the set-shifting and updating EF abilities assessed by the Trail Making Test, part B (TMT-B) and the Verbal Digit Span backwards (VDS-B), respectively, to identify genetic variation associated with the course of EFs across time. We used a linear mixed model (LMM) to model the dependence structure of the longitudinal PsyCourse Study [20] with four measurements across time. To validate our findings, we also performed a replication study using data from the FOR2107 consortium [21], which assessed two measurements over time.

MATERIALS AND METHODS

Discovery sample: PsyCourse Study

The PsyCourse Study is a multicenter longitudinal study that combines multilevel omics and longitudinal data [20]. We included 1338 genotyped individuals (dataset version 3.0) recruited in different centers in Germany and Austria, comprising patients from the affective-to-psychotic spectrum (377 bipolar I disorder, 100 bipolar II disorder, 420 schizophrenia, 95 schizoaffective disorder, 6 brief psychotic disorder, 9 schizophreniform disorder, and 73 with recurrent depression) and 258 psychiatrically healthy controls. The study protocol was approved by the respective ethics committee for each study center and was carried out following the rules of the Declaration of Helsinki of 1975, revised in 2008 (see ref. [20]). All study participants provided written consent [20]. The patients were diagnosed using parts of the Structured Clinical Interview for DSM (SCID-I) and were classified according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria. The patients were broadly differentiated in patients with predominantly affective symptoms (550 "affective", with recurrent depression, bipolar I and II disorders) and patients with predominantly psychotic symptoms (530, "psychotic", with schizophrenia, schizoaffective, brief psychotic and schizophreniform disorder) [20]. Deep phenotyping was performed during four visits, each ~6 months apart (see ref. [20]), thus corresponding to time t of the longitudinal course.

Set-shifting and updating were assessed with the Trail Making Test, part B (TMT-B) [22] and the Verbal Digit Span backwards (VDS-B) [23], respectively. The TMT-B requires an individual to connect numbers (numbers: 1–26) and letters of the alphabet in ascending alternating order. The test score was the time (in seconds (s)) needed to finish this exercise. As recommended by [24] participants with a time >300 s were set to 300 s. VDS-B measures the updating ability. Here, a trained interviewer verbally presented up to seven pairs of number sequences with increasing length, and the study participant was requested to repeat each sequence in backwards order, receiving a point score for each correctly repeated sequence. The maximum possible score for each sequence pair was 2. The process was terminated when an individual failed to repeat correctly both of the sequences in a pair of given length. The test score was the sum of all correctly repeated sequence pairs (range: 0–14).

Replication sample: FOR2107 consortium

To perform the replication study, we used data from the research consortium FOR2107 [21], a longitudinal cohort with two centers, Marburg and Münster (Germany), in which deep phenotyping was performed twice ~2 years apart [21]. In our analyses, we used a sample comprising 1795 individuals with genotype data available divided into five different diagnostic groups (851 affective: 107 bipolar disorder and 744 depression, 112 psychotic: 68 schizophrenia and 44 schizoaffective disorder, and 832 healthy controls). The participants were classified into the same three broad diagnostic groups (affective, psychotic, and controls) as in the discovery sample. Set-shifting was assessed by the TMT-B. In this cohort, participants with a time >180 s were excluded. For updating, we used the Letter–Number-Sequencing Test (LNST) as a substitute for the VDS-B. Here, a trained interviewer verbally presented an increasing sequence of letters and numbers, which the participant was requested to repeat, starting with the numbers in ascending order and ending with the letters in alphabetical order. The test was terminated when the individual repeated the same sequence incorrectly four times. The sum of the correctly repeated sequences was the test score, with a maximum of 24.

Genotyping and imputation

Discovery sample. The Illumina Infinium PsychArray (Illumina, USA) was used for genotyping purposes [20]. Genotypes were imputed with SHAPEIT2/IMPUTE2 using the 1000 Genomes Project Phase 3 data as a reference panel. Quality control (QC) was performed according to standard procedures, as described previously [25] (details Supplementary List 1) and poorly imputed genetic variants (INFO < 0.8) were excluded [20]. We included ~8.2 million SNPs with minor allele frequency (MAF) ≥ 0.01 in our analysis. Ancestry principal components (PCs) were computed with PLINK v1.9 [26] (http://pngu.mgh.harvard.edu/).

Replication sample. To replicate genome-wide significant SNPs of the discovery sample, we analyzed the genotypes of these nine significant SNPs (SNP_R). We additionally analyzed 187 suggestive SNPs (SNP_{NR}) with a P value $\leq 1 \times 10^{-5}$ in the discovery sample (99 for TMT-B, 88 for VDS-B/LNST) in an exploratory analysis. For the QC in the replication sample, please refer to Supplementary List 2.

Statistical analysis

We performed regression analysis, log-transforming the TMT-B values (lgTMT-B) to fulfill the linear mixed model requirement of normally distributed errors. We present effect estimates with 95% confidence intervals (c.i.s) transformed back to the original scale. Furthermore, we investigated missing data patterns across visits and diagnoses for violation of a missing-at-random (MAR) mechanism [27]. We computed the mean and standard deviation (s.d.) of EFs per visit and diagnostic group, testing for differences in means between diagnostic groups at each visit. For the discovery sample, we fitted LMMs to the longitudinal time course of lgTMT-B and VDS-B, investigating each phenotype first without the SNP terms, and subsequently including them. For each SNP, the fitted model for individual i at visit/time t_{ij} with $j = 1, 2, 3, 4$ was as follows:

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 age_i + \beta_3 gender_i + \beta_4 diagnosis_i + \sum_{k=1}^{5} \beta_{4+k} PC_{ik} + b_{0i} + b_{1i} t_{ij} + c_i center_i + \beta_{10} SNP_i + \beta_{11} SNP_i * t_{ij} + \varepsilon_{ij}$$

The LMM adjusted for age_i, $gender_i$, $diagnosis_i$, PC_{ik}, i.e., age at visit 1, gender, diagnostic group (affective, psychotic, or control), and the top five PCs, for each individual i, the latter to correct for population stratification. We allowed for random intercepts and slopes b_{0i}, b_{1i} of the trajectories and a random center effect.

For the respective SNP under consideration, we integrated the main effect (SNP_i) and the SNP-by-time interaction ($SNP_i * t_{ij}$), where the latter is

Table 1. Characteristics at visit 1 in discovery sample and replication sample by diagnostic group.

Study sample	Phenotypes	Diagnostic groups mean (s.d.) or percentage (%)			Group difference
		Affective	Psychotic	Controls	*P* value
Discovery sample	Age	44.6 (13.4)	41.1 (12.1)	37.1 (15.6)	–
	Females	49.8 %	39.6 %	58.1 %	–
	TMT-B	83.9 (42.6)	92.3 (41.3)	59.4 (25.1)	$<2\times10^{-16}$
	VDS-B	6.2 (2.1)	5.5 (2.0)	7.3 (2.9)	$<2\times10^{-16}$
Replication sample	Age	37.6 (13.4)	38.4 (11.3)	34.1 (12.6)	–
	Females	63.9 %	44.6 %	63.0 %	–
	TMT-B	57.7 (23.9)	73.6 (30.9)	48.8 (18.6)	$<2\times10^{-16}$
	LNST	15.7 (3.3)	13.4 (3.5)	16.8 (3.2)	$<2\times10^{-16}$

The proportion of females (%), means of age (years), TMT-B, and VDS-B/LNST with standard deviation (s.d.).
We tested for differences in means between the diagnostic groups for lgTMT-B and VDS-B. Results are only displayed for visit 1 as results for the other visits proved to be similar.

tested (two-sided) for the influence of the SNP on the longitudinal course (see ref. [28]). The interaction term consisting of SNP × diagnosis × time has not been investigated due to the limited sample size. We assumed an additive genetic model with each considered SNP in dosage format. We set the genome-wide significance level to 5×10^{-8}, yielding replication SNPs (SNP_R), and set the level for suggestive significance to 1×10^{-5} for SNPs to be further explored (SNP_{NR}, not to be replicated). For the replication sample, we separately determined linkage disequilibrium (LD) blocks with $r^2 > 0.8$ for both SNP sets, correcting for multiple testing by dividing 5% by the number of LD blocks for the SNP set [29]. In the end, the SNP_R were contained in a single LD block, so the significance level for replication could be set to 5%. The significance levels for the exploratory analysis of the SNP_{NR} were set to $0.05/24 = 0.0021$ for lgTMT-B and $0.05/12 = 0.0042$ for VDS-B/LNST, respectively.

For the SNP analysis in the replication sample, we analyzed the difference (diff) of lgTMT-B (LNST) between the visits as outcome and SNP, age, gender, diagnosis, and PC's as covariates. We applied the difference model, as the LMM above contained too many parameters for the replication sample with only two measurements (in total: 613 individuals) and incomplete data resulting in low statistical power (data not shown; two-sided test). Here, the SNP effect may be interpreted as the difference between the average change between the genotypes, especially since SNP_R displayed only two genotypes.

We computed LD and haplotypes for Europeans with LDlink [30] and created a regional plot with gene identification using Locus-Zoom [31]. Finally, the average longitudinal course over time per genotype along with 95% c.i. is displayed for the top SNP.

All statistical analyses were performed with R, version 3.5.1 (https://www.r-project.org/). The LMM was fitted with the R package lme4 [32] and *P* values were computed using the Satterthwaite approximation of the lmerTest package [33, 34].

RESULTS

Behavioral characteristics of the EFs

Discovery sample. In comparison with controls, the disease groups were slightly older on average (Table 1). A total of 1272 (1297) individuals had at least one TMT-B (VDS-B) measurement, demonstrating a similar decrease of available data in each diagnostic group (Table 2). Missing value patterns did not hint at any violation of a missing-at-random (MAR) assumption (data not shown). Figure 1 illustrates the mean longitudinal course of TMT-B (left) and VDS-B (right) for each diagnostic group with 95% c.i.s; controls differed significantly from patients (see Fig. 1, c.i.s). Generally, executive performance increased over time, with differences between affective and psychotic patients decreasing over time. An improvement in the respective EF performance is reflected by a decreased TMT-B score for set-shifting and an increased VDS-B score for updating. The individual trajectories were highly variable (Supplementary Fig. 1). The mean difference between diagnostic groups was significant at each visit when adjusting for age and gender (see Table 1). Table 3 displays the time effect estimates in the LMM for each phenotype without SNP stratified by diagnostic group. For lgTMT-B, the time effect within each diagnostic group is highly significant and similar across groups. For VDS-B, the time effects for the two patient groups are similar, very small, and only nominally significant in the psychotic group, but larger and highly significant for controls.

Replication sample. We analyzed 1795 genotyped individuals with at least one TMT-B and LNST measurement (we deleted data for one individual who had a value larger than the maximum score of 24). Phenotypes were measured at both visits for 34.2%. The means of the diagnostic groups at each visit were significantly different (Table 1) during which the controls had again the best EF abilities, followed by affective and then psychotic individuals (Supplementary Fig. 2).

GWAS of the discovery sample

The QQ-plot (Supplementary Fig. 3) demonstrates that the genomic inflation factor was $\lambda = 1.0034$ for lgTMT-B and $\lambda = 0.9999$ for VDS-B, hence not indicating any inflation. As illustrated on the Manhattan plots (lgTMT-B Fig. 2A, VDS-B Fig. 2B) for the SNP-by-time interaction in the LMM, we identified nine genome-wide significant SNPs on chromosome 5 (all imputed) in one LD block ($r^2 > 0.85$) for lgTMT-B, and none for VDS-B. For lgTMT-B, 99 SNPs were suggestive, for VDS-B 88.

For the nine genome-wide significant SNPs of the GWAS, Supplementary Table 1 displays estimates for the effect of the SNP-by-time interaction on lgTMT with 95% c.i. and *P* values. The top SNP rs150547358 (P value $= 7.2\times10^{-10}$) had an effect of 1.16 (95% c.i. 1.11–1.22) seconds per measurement (spm) in the discovery sample on the original TMT-B scale. We present the mean plot for the top SNP in Fig. 2C, where the TMT-B score increases over time for heterozygotes with risk allele "C". Figure 2D displays the regional Manhattan plot with three genes in or near the nine significant SNPs. Four of them, including rs150547358, are located in an intron region of ring finger protein 180 (RNF180) (Supplementary Table 1). Other genes located nearby are regulator of G protein signaling 7 binding protein (RGS7BP) and 5-hydroxytryptamine receptor 1A (HTR1A), but neither contained any of the nine SNPs. For the SNP main effect, also included in the model, we did not observe any genome-wide significant SNPs (Supplementary Fig. 4; $P < 5\times10^{-8}$).

Difference analysis of the replication sample

The analysis of the differences also identified the top SNP, rs150547358, as significant ($P = 0.015$), and thus replicated this GWAS-significant LD block. The effect estimate for the top SNP

Table 2. Available data of TMT-B and VDS-B per visit for the discovery sample.

EF core skill	Diagnostic groups														
	Affective				Psychotic				Controls						
Visit (t)	1	2	3	4	1	2	3	4	1	2	3	4			
TMT-B	506 (92%)	315 (57%)	234 (43%)	182 (33%)	456 (86%)	295 (56%)	252 (46%)	227 (48%)	258 (100%)	225 (82%)	178 (69%)	57 (22%)			
VDS-B	503 (92%)	324 (59%)	234 (43%)	185 (34%)	479 (90%)	320 (60%)	265 (50%)	236 (45%)	257 (99.6%)	225 (87%)	178 (69%)	60 (23%)			

Absolute numbers and percent of group total within the diagnostic group with 550 affective individuals, 530 psychotic individuals, and 258 controls.

was 0.85 (95% c.i. 0.74–0.97) on the original scale and the highest effect size in the scale of the analysis (greatest negative effect). The estimates for the other SNPs were slightly larger when transformed back to the original scale and also positive (see Supplementary Table 1 for the summary).

Exploratory analysis of the GWAS-suggestive SNP_{NR} in the replication sample yielded no significant results after multiple testing corrections for either phenotype (Supplementary Fig. 5).

DISCUSSION

We performed a GWAS on the longitudinal course of EFs and detected nine SNPs within the same LD block associated with change over a relatively short period of time (~1.5 years) in the EF core skill set-shifting. Importantly, we were able to replicate a significant result for this LD block in an independent sample, which was observed in a heterogeneous population including controls and different psychiatric disorders of the affective-to-psychotic spectrum across age groups. Analysis of TMT-B performance of C-allele carriers, in contrast to the AA genotype, revealed a pronounced slowing over time.

Recently, the analysis of longitudinal data has come to the fore in genetic research. Multiple methods have been developed to perform GWAS with longitudinal data [35–40] for binary as well as continuous phenotypes. These analysis methods are mostly applied to analyze long-term developments of the investigated phenotypes [41, 42], as most data comprise multiple measurements over a relatively long period of time. These longitudinal studies often detect group effects [8] based on age or baseline cognitive functions, for example. To date, short-term variability, for example with respect to the longitudinal course of schizophrenia has been found as reviewed [8], but without considering a potential genetic effect. In our longitudinal GWAS, we enter uncharted territory as we study short-term courses of cognitive phenotypes in relation to the genetic background. The discovery sample, the PsyCourse Study, is unique in this sense, as it assesses the phenotypes multiple times in a very heterogeneous sample over a relatively short period of time (18 months). Here, the main interest is the observation of short-term changes specific to a phenotype, such as EF skills, and the use of newly identified characteristics to detect genotype–phenotype associations. The genetic variants found in this study may, if further replicated, be used to improve clinical evaluation of the longitudinal course of EF skills. Knowledge of the genetic status of a patient may, in the future, enhance the interpretation of the course of EF abilities e.g., during psychiatric treatment. Moreover, special training programs could support patients with a known genetic disposition to lack improvement over time. To our knowledge, no other study has performed such analyses to date.

Behavioral results

Prior to our GWAS, we studied the short-term courses of changes in cognitive abilities, focusing on the differences between the diagnostic groups considered. In the discovery sample, we observed an identical pattern for both phenotypes: psychotic individuals demonstrated the lowest EF abilities, followed by those with affective disorders and then the control individuals. This greater EF impairment in psychotic individuals compared to controls is well-documented, as exemplified by [43]. However, regarding the impairment difference between bipolar (affective) and schizophrenic (psychotic) patients, there are various studies [43–48] analyzing these differences. The hypothesis exists that bipolar patients demonstrate less severe impairment in comparison to schizophrenic patients [49]. Some studies [44, 46, 48] lend their support to this hypothesis, though not always statistically significant, whereas others detected similar levels of impairment in symptomatic patients [45, 47]. In our analysis, we observed a statistically significant difference between affective and psychotic

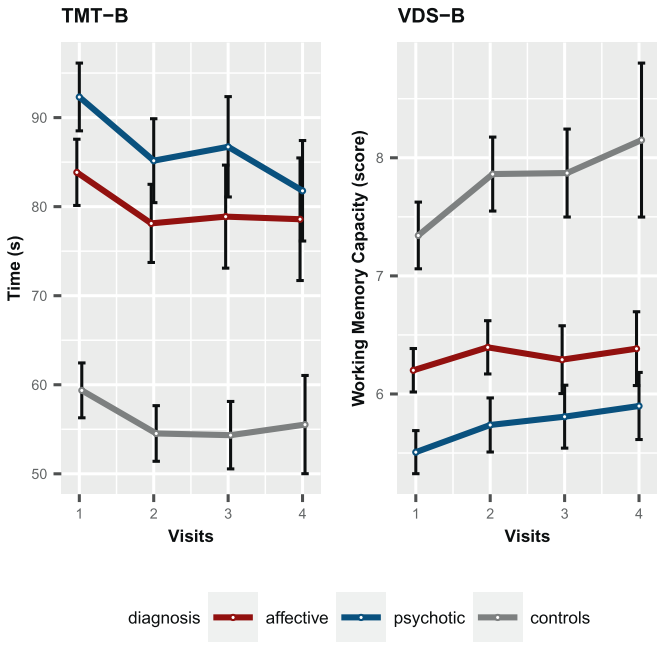


Fig. 1 Longitudinal course of TMT-B score (time in seconds, left) and VDS-B score (working memory capacity, right) for each diagnostic group in the discovery sample. Displayed are means with 95% confidence interval for each visit 1, 2, 3, 4, ~6 months apart.

Table 3. Results of the LMM of the discovery sample to test the time effect on lgTMT-B and VDS-B within each diagnostic group.

EF core skill		TMT-B			VDS-B		
Diagnostic groups	**Time effect (t)**	***β***	**95% c.i.**	***P* value**	***β***	**95% c.i.**	***P* value**
Affective		0.957	0.94, 0.97	9.8×10^{-09}	0.076	0, 0.15	0.053
Psychotic		0.950	0.94, 0.96	$<2 \times 10^{-16}$	0.086	0.02, 0.15	0.011
Controls		0.947	0.93, 0.96	6.1×10^{-11}	0.288	0.17, 0.41	2.7×10^{-06}

The effect estimates β of lgTMT-B are transformed back to their original scale.

individuals at visit 1 but detected a decline in these discrepancies over time. The abilities of these two diagnostic groups converged with patients from the psychotic group displaying an improvement in their skills and patients from the affective group presenting a more constant course. Documentation of the EF convergence is only possible thanks to the longitudinal design of the discovery sample and represents a great advantage of this study design.

Owing to the slightly different age structure of the two study samples, with the discovery sample being minimally older on average at visit 1, we further observed the impact of age reflected by the minimally lower average test score. That is, the discovery sample had lower VDS and greater TMT-B scores than the replication sample. The TMT-B mean scores may also be influenced further by the different cutoff thresholds of 300 s in the discovery sample and 180 s for the replication sample.

Genome-wide association studies

To our knowledge, the LD block comprising the nine SNPs we detected for the set-shifting ability has been not identified in any GWAS before. These SNPs are part of two common haplotypes, that is, 97.7% carry the haplotype consisting of the major alleles and 1.7% have the rare haplotype with only minor alleles in European populations [30]. However, we did not observe different allelic distributions between the three diagnostic groups (Supplementary Table 2). We displayed the longitudinal course for the two genotypes "AC" and "AA" of the top SNP rs150547358, observing a steady increase in the TMT-B score for "AC" and an almost unchanging course for "AA". Consequently, the minor allele C was associated with a decline in the set-shifting ability of ~5 s over a period of 18 months for AC with a large c.i. at the last visit owing to the small number of available heterozygous individuals. This result reflects a relatively high decrease in the ability over this short period. Furthermore, it portrays a highly interesting observation, which is further underpinned when we consider the genetic region of the nine SNPs. Variant rs150547358, the significantly replicated SNP, is one of four associated SNPs directly located in the ring finger protein 180 (RNF180) gene on chromosome 5q12.3. It is an E3 ubiquitin-protein ligase [50], whose product is involved in protein modification. RNF180 is

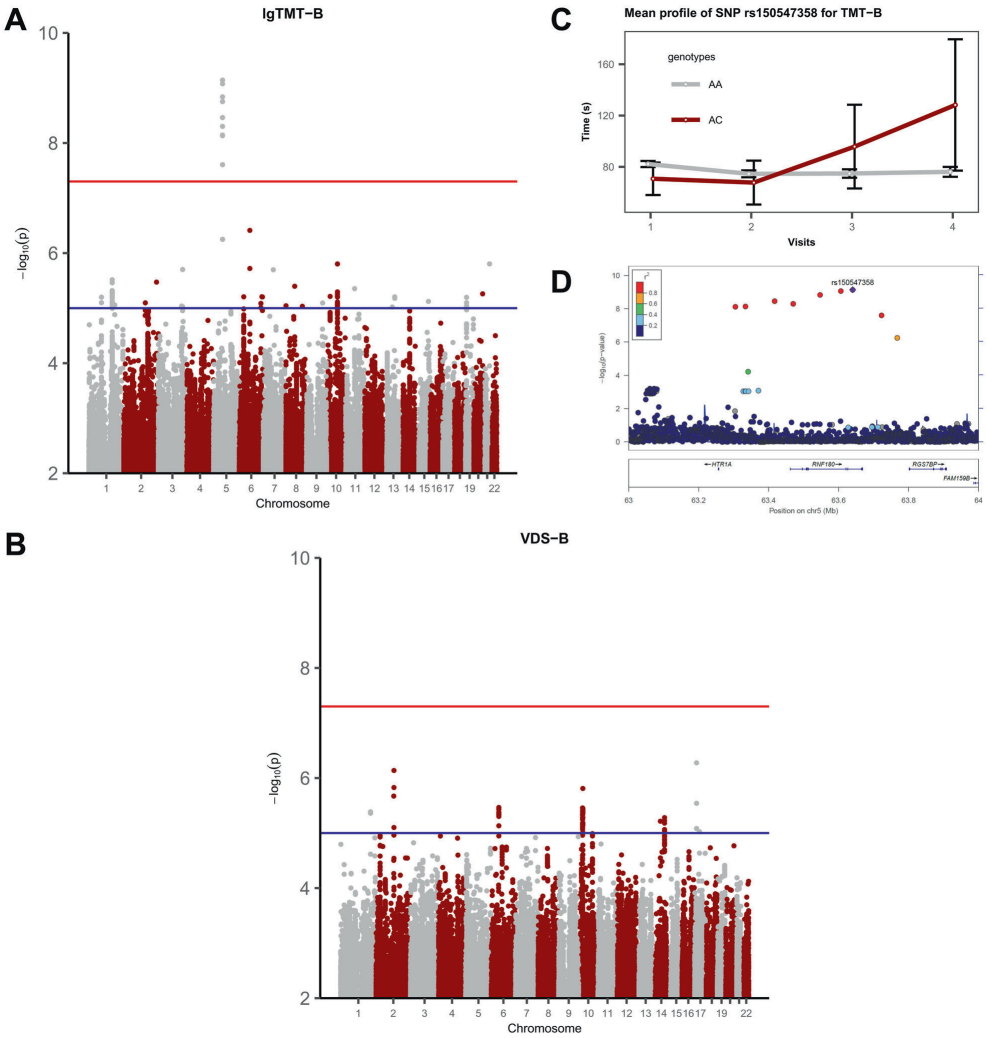


Fig. 2 Results of the genome-wide association studies of the discovery sample. A Manhattan plot of the GWAS of lgTMT-B in the discovery sample. The lines in (**A**) and (**B**) indicate the thresholds for the genome-wide significance of 5×10^{-8} (red) and for suggestive SNPs (blue, $P \leq 1 \times 10^{-5}$). **B** Manhattan plot of the GWAS of VDS-B in the discovery sample. **C** Mean profile of TMT-B by the top SNP rs150547358 genotypes for the discovery sample (1039 AA, 28 AC, 0 CC) with the 95% confidence intervals. **D** GWAS regional Manhattan plot of chromosome 5 for lgTMT-B of the discovery sample. Colors indicate the LD values (r^2) of SNPs with rs150547358 (in purple).

associated with the regulation of monoamine levels in different brain regions, for example, the prefrontal cortex (PFC) in RNF180 knockout mice [51]. The PFC is a critical part of the frontal lobe in the development of EFs [4, 52]. Another gene located in the nearby region, HTR1A (5-hydroxytryptamine receptor 1A), is an important receptor of serotonin (5-HT) also essential to the prefrontal lobe. More importantly, HTR1A is an autoreceptor, located on the cell bodies of serotonin-synthesizing neurons of the brainstem dorsal raphe nucleus, helping to maintain homeostasis in serotonergic function [53]. Furthermore, a genetic polymorphism in the 5-HT system has previously been implicated in EF performance [12].

In an additional exploratory gene-set analysis performed with MAGMA v1.06 as a part of the FUMA pipeline (https://fuma.ctglab.nl/) [54], we did not receive significant (Bonferroni-corrected P values ≤0.05) pathways for either phenotype.

Our results are a first step in the direction of understanding the molecular genetic influences on the longitudinal course of EFs. We were unable to consider the third core ability, inhibition, which also plays an important role for EF, because we could not fulfill a specific assessment requirement resulting from the multicenter and interview-based structure of the discovery sample [20]. Many unknown factors remain, such as the genetic aspects due to the correlation of the different EF abilities, as we only concentrated on

individual EF core skills in two separate analyses. According to the "unity but diversity" concept [11] that also concerns the genetic underpinnings of the EFs, a genetic study of a latent common factor needs to follow. Further, we need to acknowledge the problem of missing data which is a great challenge in longitudinal studies as presented in our samples. Here, selecting the correct analysis method, e.g., linear mixed models are imported but generally, more longitudinal studies with multiple time points and greater sample sizes will be required to unmask further time and genomics interactions [19].

CODE AND DATA AVAILABILITY

R code and data will be available upon reasonable request by the authors. The summary statistics of our analysis will be published in the GWAS Catalog (https://www.ebi.ac.uk/gwas/).

REFERENCES

1. Friedman NP, Miyake A, Altamirano LJ, Corley RP, Young SE, Rhea SA, et al. Stability and change in executive function abilities from late adolescence to early adulthood: a longitudinal twin study. Developmental Psychol. 2016;52:326–40.
2. Barnes JJM, Dean AJ, Nandam LS, O'Connell RG, Bellgrov MA. The molecular genetics of executive function: role of monoamine system genes. Biol Psychiatry. 2011;69:e127–e143.
3. Friedman NP, Miyake A, Young SE, DeFries JC, Corley RP, Hewitt JK. Individual differences in executive functions are almost entirely genetic in origin. J Exp Psychol: General. 2008;137:201–25.
4. Diamond A. Executive functions. Annu Rev Psychol. 2013;64:135–68.
5. Best JR, Miller PH, Jones LL. Executive functions after age 5: changes and correlates. Developmental Rev. 2009;29:180–200.
6. West R. Aging and the neural correlates of executive function. In: Wiebe SA, Karbach JExecutive function. New York: Routledge; 2017. p. 91–105.
7. van Hooren SA, Valentijn AM, Bosma H, Ponds RW, van Boxtel MP, Jolles J. Cognitive functioning in healthy older adults aged 64-81: a cohort study into the effects of age, sex, and education. Aging Neuropsychol Cognition. 2007;24:40–54.
8. Heilbronner U, Samara M, Leucht S, Falkai P, Schulze TG. The longitudinal course of schizophrenia across the lifespan. Harv Rev Psychiatry. 2016;24:118–28.
9. Martínez-Arán A, Vieta E, Colom F, Torrent C, Sánchez-Moreno J, Reinares M, et al. Cognitive impairment in euthymic bipolar patients: implications for clinical and functional outcome. Bipolar Disord. 2004;6:224–32.
10. Ratiu P, Talos IF. The tale of phineas gage, digitally remastered. N Engl J Med. 2004;351:e21.
11. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. Cogn Psychol. 2000;41:49–100.
12. Li JJ, Roberts DK. Genetic influences on executive functions across the life span. In: Wiebe SA, Karbach J, (eds.) Executive function. New York: Routledge; 2017. p. 106–23.
13. Luciano M, Hansell NK, Lahti J, Davies G, Medland SE, Räikkönen K, et al. Whole genome association scan for genetic polymorphisms influencing information processing speed. Biol Psychol. 2011;86:193–202.
14. Seshadri S, DeStefano AL, Au R, Massaro JM, Beiser AS, Kelly-Hayes M, et al. Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham study. BMC Med Genet. 2007;8:1–14.
15. Cirulli ET, Kasperaviciūte D, Attix DK, Need AC, Ge D, Gibson G, et al. Common genetic variation and performance on standardized cognitive tests. Eur J Hum Genet. 2010;18:815–20.
16. Need AC, Attix DK, McEvoy JM, Cirulli ET, Linney KL, Hunt P, et al. A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTB. Hum Mol Genet. 2009;18:4650–61.
17. Malone SM, Vaidyanathan U, Basu S, Miller MB, McGue M, Iacono WG. Heritability and molecular-genetic basis of the P3 event-related brain potential: a genome-wide association study. Psychophysiology. 2014;51:1246–58.
18. LeBlanc M, Kulle B, Sundet K, Agartz I, Melle I, Djurovic S, et al. Genome-wide study identifies PTPRO and WDR72 and FOXQ1-SUMO1P1 interaction associated with neurocognitive function. J Psychiatr Res. 2012;46:271–8.
19. Boyce WT, Sokolowski MB, Robinson GE. Genes and environments, development and time. Proc Natl Acad Sci USA. 2020;117:23235–41.
20. Budde M, Anderson-Schmidt H, Gade K, Reich-Erkelenz D, Adorjan K, Kalman JL, et al. A longitudinal approach to biological psychiatric research: the PsyCourse study. Am J Med Genet Part B: Neuropsychiatr Genet. 2018;180:89–102.
21. Kircher T, Wöhr M, Nenadic I, Schwarting R, Schratt G, Alferink J, et al. Neurobiology of the major psychoses: a translational perspective on brain structure and function—the FOR2107 consortium. Eur Arch Psychiatry Clin Neurosci. 2018;269:949–62.
22. Bowie CR, Harvey PD. Administration and interpretation of the trail making test. Nat Protoc. 2006;1:2277–81.
23. Hilbert S, Nakagawa TT, Puci P, Zech A, Bühner M. The digit span backwards task. Eur J Psychologcal Assess. 2015;31:174–80.
24. Strauss E, Sherman EMS, Spreen O. A compendium of neuropsychological tests—administration, norms, and commentary. New York: Oxford University Press; 2006.
25. Andlauer TF, Buck D, Antony G, Bayas A, Bechmann L, Berthele A, et al. Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. Sci Adv. 2016;2:e1501678.
26. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
27. Molenbergh G, Verbeke, G. Linear mixed models for longitudinal data. Berlin, Heidelberg: Springer; 2000.
28. Sikorska K, Rivadeneira F, Groenen PJF, Hofman A, Uitterlinden AG, Eilers PHC, et al. Fast linear mixed model computations for genome-wide association studies with longitudinal data. Stat. Med. 2012 ;32:165–80.
29. Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. BMC Genomics. 2008;9:516.
30. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31:3555–7.
31. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26:2336–7.
32. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67:1.
33. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. J Stat Softw. 2017;82:1–26.
34. Luke SG. Evaluating significance in linear mixed-effects models in R. Behav Res Methods. 2016;49:1494–502.
35. Sikorska K, Lesaffre E, Groenen PJF, Rivadeneira F, Eilers PHC. Genome-wide analysis of large-scale longitudinal outcomes using penalization GALLOP algorithm. Sci Rep. 2018;8:1–8.
36. Sikorska K, Montazeri NM, Uitterlinden A, Rivadeneira F, Eilers PH, Lesaffre E. GWAS with longitudinal phenotypes: performance of approximate procedures. Eur J Hum Genet. 2015;23:1384–91.
37. Wu W, Wang Z, Xu K, Zhang X, Amei A, Gelernter J, et al. Retrospective association analysis of longitudinal binary traits identifies important loci and pathways in cocaine use. Genetics. 2019;213:1225–36.
38. Rudra P, Broadaway KA, Ware EB, Jhun MA, Bielak LF, Zhao W, et al. Testing cross-phenotype effects of rare variants in longitudinal studies of complex traits. Genet Epidemiol. 2018;41:320–32.
39. Ning C, Wang D, Zhou L, Wei J, Liu Y, Kang H, et al. Efficient multivariate analysis algorithms for longitudinal genome-wide association studies. Bioinformatics. 2019;35:4879–85.
40. Lee Y, Park S, Moon S, Lee J, Elston RC, Lee W, et al. On the analysis of a repeated measure design in genome-wide association analysis. Int J Environ Res Public Health. 2014;11:12283–303.
41. Adkins DE, Clark SL, Copeland WE, Kennedy M, Conway K, Angold A, et al. Genome-wide meta-analysis of longitudinal alcohol consumption across youth and early adulthood. Twin Res Hum Genet. 2015;18:335–47.
42. Tang W, Kowgier M, Loth DW, Soler Artigas M, Joubert BR, Hodge E, et al. Large-scale genome-wide association studies and meta-analyses of longitudinal change in adult lung function. PLoS ONE. 2014;9:e100776.
43. Wobrock T, Ecker UK, Scherk H, Schneider-Axmann T, Falkai P, Gruber O. Cognitive impairment of executive function as a core symptom of schizophrenia. World J Biol Psychiatry. 2009;10:442–51.
44. Szoke A, Meary A, Trandafir A, Bellivier F, Roy I, Schurhoff F, et al. Executive deficits in psychotic and bipolar disorders - Implications for our understanding of schizoaffective disorder. Eur Psychiatry. 2008;23:20–25.
45. Amann B, Gomar JJ, Ortiz-Gil J, McKenna P, Sans-Sansa B, Sarró S, et al. Executive dysfunction and memory impairment in schizoaffective disorder: a comparison with bipolar disorder, schizophrenia and healthy controls. Psychological Med. 2012;42:2127–35.
46. Hill SK, Reilly JL, Keefe RS, Gold JM, Bishop JR, Gershon ES, et al. Neuropsychological Impairments in schizophrenia and psychotic bipolar disorder: findings from the bipolar-schizophrenia network on intermediate phenotypes (B-SNIP) study. Am J Psychiatry. 2013;170:1275–84.

47. Lewandowski KE, Cohen BM, Keshavan MS, Öngür D. Relationship of neurocognitive deficits to diagnosis and symptoms across affective and non-affective psychoses. Schizophrenia Res. 2011;133:212–7.
48. Reichenberg A, Harvey PD, Bowie CR, Mojtabai R, Rabinowitz J, Heaton RK, et al. Neuropsychological function and dysfunction in schizophrenia and psychotic affective disorders. Schizophrenia Bull. 2008;35:1022–9.
49. Lynham AJ, Hubbard L, Tansey KE, Hamshere ML, Legge SE, Owen MJ, et al. Examining cognition across the bipolar/schizophrenia diagnostic spectrum. J Psychiatry Neurosci. 2018;43:245–53.
50. Ogawa M, Mizugishi K, Ishiguro A, Koyabu Y, Imai Y, Takahashi R, et al. Rines/RNF180, a novel RING finger gene-encoded product, is a membrane-bound ubiquitin ligase. Genes Cells. 2008;13:397–409.
51. Kabayama M, Sakoori K, Yamada K, Ornthanalai VG, Ota M, Morimura N, et al. Rines E3 ubiquitin ligase regulates MAO-A levels and emotional responses. J Neurosci. 2013;33:12940–53.
52. Best JR, Miller PH. A developmental perspective on executive function. Child Dev. 2010;81:1641–60.
53. McDevitt RA, Neumaier F. Regulation of dorsal raphe nucleus function by serotonin autoreceptors: a behavioral perspective. J Chem Neuroanat. 2011;41:234–46.
54. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. 2017;8:1–11.

ACKNOWLEDGEMENTS

Thomas G. Schulze and Peter Falkai are supported by the Deutsche Forschungsgemeinschaft (German Research Foundation; DFG) within the framework of the projects http://www.kfo241.de and http://www.PsyCourse.de (SCHU 1603/4-1, 5-1, 7-1; FA241/16-1). Bernadette Wendel and Heike Bickeböller are supported by the DFG (KFO241, BI 576 15-1). Tilo Kircher was also supported by the DFG (FOR2107 KI588/14-1 and FOR2107 KI588/14-2). The genotyping was funded in part by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Program with a grant awarded to Thomas G. Schulze (01ZX1614K). Thomas G. Schulze received additional support from the German Federal Ministry of Education and Research (BMBF) within the framework of the BipoLife network (01EE1404H) and the Dr. Lisa Oehler Foundation (Kassel, Germany). Sergi Papiol was supported by a 2016 NARSAD Young Investigator Grant (25015) from the Brain and Behavior Research Foundation. This work was further funded by the German Research Foundation (DFG, grant FOR2107 DA1151/5-1 and DA1151/5-2 to UD) and the Interdisciplinary Center for Clinical Research (IZKF) of the medical faculty of Münster (grant Dan3/012/17 to UD). Igor Nenadić was supported by the Deutsche Forschungsgemeinschaft (DFG, grants NE2254/1-2, NE2254/2-1, NE2254/3-1, NE2254/4-1). Jens Wiltfang is supported by an Ilídio Pinho professorship, iBiMED (UIDB/04501/2020) at the University of Aveiro, Portugal. Andreas Jansen was supported by the DFG (FOR2107 JA-1890/7-1 and FOR2108 JA-1890/7-2). We acknowledge support by the Open Access Publication Funds of the Göttingen University. The authors would like to thank Frederike Stein for her support in sharing the replication data, and Andrew Entwistle for proofreading the manuscript.

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01510-8.

Correspondence and requests for materials should be addressed to B.W.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

7 Outlook

This thesis compiles several loosely connected studies, undertaken to elucidate the course of the neuronal, phenomic, and genomic basis of EFs over time, with a focus on methodological challenges. Using the Unity and Diversity model of EFs as a framewok, several elements of this model have been focussed on in the present thesis: Set-Shifting and Updating aspects were addressed in Wendel et al. (2021), while the Inhibition aspect was addressed in Heilbronner and Münte (2013). Also, Heilbronner, Papiol, et al. (2021) research an "executive" personality trait. Regarding the latter, it remains to be shown whether the personality dimension LBCN is associated with the genomics and phenomics of cognitive EFs. An important aspect of future work will also be the general and specific genomic and phenomic relationships between EFs and severe mental disorders, research questions that can be addressed in deeply phenotyped individuals, e.g., in participants of the PsyCourse Study (Budde et al., 2019). From a methodological perspective, one general challenge for future genomic research lies in the so-called task impurity (c.f. Friedman & Miyake, 2017) of individual EF tasks. Since EFs involve supervision of lower-level abilities, any EF task includes non-executive processes (Miyake et al., 2000). This obstacle may be overcome by using latent factor scores, derived from multiple EF tests. Since Hatoum et al. (2019) have recently conducted a large scale genomic study of a latent Common EF factor, it should now be feasible to generate polygenic risk scores (aggregated summary measures of the polygenic risk of an individual) in well-phenotyped transdiagnostic psychiatric samples. Eventually, this research could pave the way for individualized clinical applications.

8 References

23andMe Research Team, Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., de Leeuw, C. A., Bryois, J., Savage, J. E., Hammerschlag, A. R., Skene, N. G., Muñoz-Manchado, A. B., White, T., Tiemeier, H., Linnarsson, S., Hjerling-Leffler, J., Polderman, T. J. C., Sullivan, P. F., van der Sluis, S., & Posthuma, D. (2018). Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nature Genetics*, *50*(7), 920–927. https://doi.org/10.1038/s41588-018-0151-7

Amelang, M., Hasselbach, P., & Stürmer, T. (2004). Personality, cardiovascular disease, and cancer: First results from the Heidelberg Cohort Study of the Elderly. *Zeitschrift für Gesundheitspsychologie*, *12*(3), 102–115. https://doi.org/https://doi.org/10.1026/0943-8149.12.3.102

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485–493. https://doi.org/10.1037/0894-4105.8.4.485

Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: A longitudinal study on frequent repetitive cognitive testing. *BMC Neuroscience*, *11*(1), 118. https://doi.org/10.1186/1471-2202-11-118

Boecker, M., Buecheler, M., Schroeter, M., & Gauggel, S. (2007). Prefrontal brain activation during stop-signal response inhibition: An event-related functional near-infrared spectroscopy study. *Behavioural Brain Research*, *176*(2), 259–266. https://doi.org/10.1016/j.bbr.2006.10.009

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, *3*(11), 872–882. https://doi.org/10.1038/nrg932

Boyce, W. T., Sokolowski, M. B., & Robinson, G. E. (2020). Genes and environments, development and time. *Proceedings of the National Academy of Sciences*, *117*(38), 23235–23241. https://doi.org/10.1073/pnas.2016710117

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience & Biobehavioral*

Reviews, *26*(7), 809–817. https://doi.org/10.1016/S0149-7634(02)00067-2

Budde, M., Anderson-Schmidt, H., Gade, K., Reich-Erkelenz, D., Adorjan, K., Kalman, J. L., Senner, F., Papiol, S., Andlauer, T. F. M., Comes, A. L., Schulte, E. C., Klöhn-Saghatolislam, F., Gryaznova, A., Hake, M., Bartholdi, K., Flatau, L., Reitt, M., Quast, S., Stegmaier, S., … Heilbronner, U. (2019). A longitudinal approach to biological psychiatric research: The PsyCourse Study. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *180*(2), 89–102. https://doi.org/10.1002/ajmg.b.32639

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *NeuroReport*, *9*(16), 3735–3739. https://doi.org/10.1097/00001756-199811160-00030

Cirulli, E. T., Kasperavičiūtė, D., Attix, D. K., Need, A. C., Ge, D., Gibson, G., & Goldstein, D. B. (2010). Common genetic variation and performance on standardized cognitive tests. *European Journal of Human Genetics*, *18*(7), 815–820. https://doi.org/10.1038/ejhg.2010.2

Coolidge, F. L., Thede, L. L., & Young, S. E. (2000). Heritability and the comorbidity of attention deficit hyperactivity disorder with behavioral disorders and executive function deficits: A preliminary investigation. *Developmental Neuropsychology*, *17*(3), 273–287. https://doi.org/10.1207/S15326942DN1703_1

Docherty, A. R., Moscati, A., Peterson, R., Edwards, A. C., Adkins, D. E., Bacanu, S. A., Bigdeli, T. B., Webb, B. T., Flint, J., & Kendler, K. S. (2016). SNP-based heritability estimates of the personality dimensions and polygenic prediction of both neuroticism and major depression: Findings from CONVERGE. *Translational Psychiatry*, *6*(10), e926–e926. https://doi.org/10.1038/tp.2016.177

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Engelhardt, L. E., Briley, D. A., Mann, F. D., Harden, K. P., & Tucker-Drob, E. M. (2015). Genes unite executive functions in childhood.

Psychological Science, *26*(8), 1151–1163. https://doi.org/10.1177/0956797615577209

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 874–890. https://doi.org/10.1080/13803390490510680

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology. General*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Friedman, N. P., & Robbins, T. W. (2021). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*. https://doi.org/10.1038/s41386-021-01132-0

Goldstein, S., & Naglieri, J. A. (Eds.). (2014). *Handbook of executive functioning*. New York, NY, Springer New York. https://doi.org/10.1007/978-1-4614-8106-5

Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A history of executive functioning as a theoretical and clinical construct. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 3–12). New York, NY, Springer New York. https://doi.org/10.1007/978-1-4614-8106-5_1

Groot, A. S., de Sonneville, L. M., Stins, J. F., & Boomsma, D. I. (2004). Familial influences on sustained attention and inhibition in preschoolers. *Journal of Child Psychology and Psychiatry*, *45*(2), 306–314. https://doi.org/10.1111/j.1469-7610.2004.00222.x

Häfner, H., & an der Heiden, W. (2000). Methodische Probleme der Verlaufsforschung an der Schizophrenie. *Fortschritte der Neurologie · Psychiatrie*, *68*(5), 193–205. https://doi.org/10.1055/s-2000-12144

Hatoum, A. S., Morrison, C. L., Mitchell, E. C., Lam, M., Benca-Bachman, C. E., Reineberg, A. E., Palmer, R. H. C., Evans, L. M., Keller, M. C., & Friedman, N. P. (2019). Genome-wide association study of over 427,000 individuals establishes executive functioning as a neurocognitive basis of psychiatric disorders influenced by GABAergic processes. *bioRxiv*. https://doi.org/10.1101/674515

Heilbronner, U., Adorjan, K., Anderson-Schmidt, H., Budde, M., Comes, A. L., Gade, K., Heilbronner, M., Kalman, J. L., Kohshour, M. O., Papiol, S., Reich-Erkelenz, D., Schaupp, S. K., Schulte, E. C., Senner, F., Vogl, T., Falkai, P., & Schulze, T. G. (2021). *The PsyCourse codebook, version 5.0*. Universitätsbibliothek der Ludwig-Maximilians-Universität München. https://doi.org/10.5282/UBM/DATA.199

Heilbronner, U., & Münte, T. F. (2013). Rapid event-related near-infrared spectroscopy detects age-related qualitative changes in the neural correlates of response inhibition. *Neuroimage*, *65*, 408–415. https://doi.org/10.1016/j.neuroimage.2012.09.066

Heilbronner, U., Papiol, S., Budde, M., Andlauer, T. F. M., Strohmaier, J., Streit, F., Frank, J., Degenhardt, F., Heilmann-Heimbach, S., Witt, S. H., Forstner, A. J., Loerbroks, A., Amelang, M., Stürmer, T., Müller-Myhsok, B., Nöthen, M. M., Rietschel, M., & Schulze, T. G. (2021). "The Heidelberg Five" personality dimensions: Genome-wide associations, polygenic risk for neuroticism, and psychopathology 20 years after assessment. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *186*(2), 77–89. https://doi.org/10.1002/ajmg.b.32837

Heilbronner, U., & Pollmann, S. (2010). Is there a structural limit to 'branch' recursively between more than two tasks? *Psychological Research*, *74*(3), 327–336. https://doi.org/10.1007/s00426-009-0249-8

Heilbronner, U., Samara, M., Leucht, S., Falkai, P., & Schulze, T. G. (2016). The longitudinal course of schizophrenia across the lifespan: Clinical, cognitive, and neurobiological aspects. *Harvard Review of Psychiatry*, *24*(2), 118–128. https://doi.org/10.1097/HRP.0000000000000092

Herrmann, M. J., Plichta, M. M., Ehlis, A.-C., & Fallgatter, A. J. (2005). Optical topography during a Go–NoGo task assessed with multi-channel near-infrared spectroscopy. *Behavioural Brain Research*, *160*(1), 135–140. https://doi.org/10.1016/j.bbr.2004.11.032

Hillman, E. M. C. (2007). Optical brain imaging in vivo: Techniques and applications from animal to man. *Journal of Biomedical Optics*, *12*(5), 051402. https://doi.org/10.1117/1.2789693

Ibrahim-Verbaas, C. A., Bressler, J., Debette, S., Schuur, M., Smith, A. V., Bis, J. C., Davies, G., Trompet, S., Smith, J. A., Wolf, C., Chibnik, L. B., Liu, Y., Vitart, V., Kirin, M., Petrovic, K., Polasek, O., Zgaga, L., Fawns-Ritchie, C., Hoffmann, P., … Mosley, T. H. (2016). GWAS for executive function and processing speed suggests involvement of the CADM2 gene. *Molecular Psychiatry*, *21*(2), 189–197. https://doi.org/10.1038/mp.2015.37

Jöbsis, F. F. (1977). Non-invasive, infra-red monitoring of cerebral O2 sufficiency, bloodvolume, HbO2-Hb shifts and bloodflow. *Acta Neurologica Scandinavica. Supplementum*, *64*, 452–3.

Kabayama, M., Sakoori, K., Yamada, K., Ornthanalai, V. G., Ota, M., Morimura, N., Katayama, K.-i., Murphy, N. P., & Aruga, J. (2013). Rines E3 ubiquitin ligase regulates MAO-A levels and emotional responses. *Journal of Neuroscience*, *33*(32), 12940–12953. https://doi.org/10.1523/JNEUROSCI.5717-12.2013

Kahn, R. S., Sommer, I. E., Murray, R. M., Meyer-Lindenberg, A., Weinberger, D. R., Cannon, T. D., O'Donovan, M., Correll, C. U., Kane, J. M., van Os, J., & Insel, T. R. (2015). Schizophrenia. *Nature Reviews Disease Primers*, *1*(1), 15067. https://doi.org/10.1038/nrdp.2015.67

Kircher, T., Wöhr, M., Nenadic, I., Schwarting, R., Schratt, G., Alferink, J., Culmsee, C., Garn, H., Hahn, T., Müller-Myhsok, B., Dempfle, A., Hahmann, M., Jansen, A., Pfefferle, P., Renz, H., Rietschel, M., Witt, S. H., Nöthen, M., Krug, A., & Dannlowski, U. (2019). Neurobiology of the major psychoses: A translational perspective on brain structure and function—the FOR2107 consortium. *European Archives of Psychiatry and Clinical Neuroscience*, *269*(8), 949–962. https://doi.org/10.1007/s00406-018-0943-x

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science*, *318*(5850), 594–598. https://doi.org/10.1126/science.1142995

Kuntsi, J., Rogers, H., Swinard, G., Börger, N., van der Meere, J., Rijsdijk, F., & Asherson, P. (2006). Reaction time, inhibition, work-

ing memory and 'delay aversion' performance: Genetic influences and their interpretation. *Psychological Medicine*, *36*(11), 1613–1624. https://doi.org/10.1017/S0033291706008580

LeBlanc, M., Kulle, B., Sundet, K., Agartz, I., Melle, I., Djurovic, S., Frigessi, A., & Andreassen, O. A. (2012). Genome-wide study identifies PTPRO and WDR72 and FOXQ1-SUMO1P1 interaction associated with neurocognitive function. *Journal of Psychiatric Research*, *46*(2), 271–278. https://doi.org/10.1016/j.jpsychires.2011.11.001

Li, J. J., & Roberts, D. K. (2018). Genetic influences on executive functions across the life span. In S. A. Wiebe & J. Karbach (Eds.), *Executive function: Development across the life span*. New York, Routledge, Taylor & Francis Group.

Luciano, M., Hansell, N. K., Lahti, J., Davies, G., Medland, S. E., Räikkönen, K., Tenesa, A., Widen, E., McGhee, K. A., Palotie, A., Liewald, D., Porteous, D. J., Starr, J. M., Montgomery, G. W., Martin, N. G., Eriksson, J. G., Wright, M. J., & Deary, I. J. (2011). Whole genome association scan for genetic polymorphisms influencing information processing speed. *Biological Psychology*, *86*(3), 193–202. https://doi.org/10.1016/j.biopsycho.2010.11.008

Malone, S. M., Vaidyanathan, U., Basu, S., Miller, M. B., Mcgue, M., & Iacono, W. G. (2014). Heritability and molecular-genetic basis of the P3 event-related brain potential: A genome-wide association study: Genome-wide association study of P3 amplitude. *Psychophysiology*, *51*(12), 1246–1258. https://doi.org/10.1111/psyp.12345

McDevitt, R. A., & Neumaier, J. F. (2011). Regulation of dorsal raphe nucleus function by serotonin autoreceptors: A behavioral perspective. *Journal of Chemical Neuroanatomy*, *41*(4), 234–246. https://doi.org/10.1016/j.jchemneu.2011.05.001

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe"tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Modestin, J., Huber, A., Satirli, E., Malti, T., & Hell, D. (2003). Long-term course of schizophrenic illness: Bleuler's study reconsidered. *American Journal of Psychiatry*, *160*(12), 2202–2208. https://doi.org/10.1176/appi.ajp.160.12.2202

Need, A. C., Attix, D. K., McEvoy, J. M., Cirulli, E. T., Linney, K. L., Hunt, P., Ge, D., Heinzen, E. L., Maia, J. M., Shianna, K. V., Weale, M. E., Cherkas, L. F., Clement, G., Spector, T. D., Gibson, G., & Goldstein, D. B. (2009). A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Human Molecular Genetics*, *18*(23), 4650–4661. https://doi.org/10.1093/hmg/ddp413

Nielson, K. A., Langenecker, S. A., & Garavan, H. (2002). Differences in the functional neuroanatomy of inhibitory control across the adult life span. *Psychology and Aging*, *17*(1), 56–71. https://doi.org/10.1037/0882-7974.17.1.56

Obrig, H., Wenzel, R., Kohl, M., Horst, S., Wobst, P., Steinbrink, J., Thomas, F., & Villringer, A. (2000). Near-infrared spectroscopy: Does it function in functional activation studies of the adult brain? *International Journal of Psychophysiology*, *35*(2-3), 125–142. https://doi.org/10.1016/S0167-8760(99)00048-3

Posner, M. I., & Snyder, C. R. R. (2004). Attention and cognitive control. In D. A. Balota & E. J. Marsh (Eds.), *Cognitive psychology: Key readings* (pp. 205–223). New York, NY, US, Psychology Press.

Power, R. A., & Pluess, M. (2015). Heritability estimates of the Big Five personality traits based on common genetic variants. *Translational Psychiatry*, *5*, e604. https://doi.org/10.1038/tp.2015.96

Pribram, K. H. (1973). The primate frontal cortex: Executive of the brain. In *Psychophysiology of the frontal lobes.* (pp. xii, 332–xii, 332). Oxford, England, Academic Press.

Ratiu, P., & Talos, I.-F. (2004). The tale of Phineas Gage, digitally remastered. *New England Journal of Medicine*, *351*(23), e21. https://doi.org/10.1056/NEJMicm031024

Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review*, *4*(3), 161–198. https://doi.org/10.1007/BF01874891

Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., LaFrance, W. C., & Coffey, C. E. (2002). Ex-

ecutive control function: A review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *14*(4). https://doi.org/10.1176/jnp.14.4.377

Schroeter, M. L., Zysset, S., Kruggel, F., & von Cramon, D. (2003). Age dependency of the hemodynamic response as measured by functional near-infrared spectroscopy. *NeuroImage*, *19*(3), 555–564. https://doi.org/10.1016/S1053-8119(03)00155-1

Schroeter, M. L., Zysset, S., Kupka, T., Kruggel, F., & von Cramon, D. Y. (2002). Near-infrared spectroscopy can detect brain activity during a color-word matching Stroop task in an event-related design. *Human Brain Mapping*, *17*(1), 61–71. https://doi.org/10.1002/hbm.10052

Seshadri, S., DeStefano, A. L., Au, R., Massaro, J. M., Beiser, A. S., Kelly-Hayes, M., Kase, C. S., D'Agostino, R. B., DeCarli, C., Atwood, L. D., & Wolf, P. A. (2007). Genetic correlates of brain aging on MRI and cognitive test measures: A genome-wide association and linkage analysis in the Framingham study. *BMC Medical Genetics*, *8*(S1), S15. https://doi.org/10.1186/1471-2350-8-S1-S15

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(2), 727–741. https://doi.org/10.1093/brain/114.2.727

Szczepanski, S. M., & Knight, R. T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron*, *83*(5), 1002–1018. https://doi.org/10.1016/j.neuron.2014.08.011

Taga, G., & Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *NeuroImage*, *36*(4), 1246–1252. https://doi.org/10.1016/j.neuroimage.2007.04.037

Tate, R. L. (1999). Executive dysfunction and characterological changes after traumatic brain injury: Two sides of the same coin? *Cortex*, *35*(1), 39–55. https://doi.org/10.1016/S0010-9452(08)70784-6

Thornton, T. (2017). *Dysexecutive syndrome: Dealing with day-to-day decision making. Information for patients and carers*. Retrieved July 27, 2020, from http://www.dhhs.tas.gov.au/___data/assets/pdf_file/0010/36010/Dysexecutive_Syndrome.pdf

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*(5), 160–164. https://doi.org/10.1111/1467-8721.00084

Turkheimer, E., Pettersson, E., & Horn, E. E. (2014). A phenotypic null hypothesis for the genetics of personality. *Annual Review of Psychology*, *65*(1), 515–540. https://doi.org/10.1146/annurev-psych-113011-143752

Wendel, B., Papiol, S., Andlauer, T. F. M., Zimmermann, J., Wiltfang, J., Spitzer, C., Senner, F., Schulte, E. C., Schmauß, M., Schaupp, S. K., Repple, J., Reininghaus, E., Reimer, J., Reich-Erkelenz, D., Opel, N., Nenadić, I., Meinert, S., Konrad, C., Klöhn-Saghatolislam, F., … Heilbronner, U. (2021). A genome-wide association study of the longitudinal course of executive functions. *Translational Psychiatry*, *11*(1), 386. https://doi.org/10.1038/s41398-021-01510-8

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292. https://doi.org/10.1037/0033-2909.120.2.272

Zarahn, E., Aguirre, G., & D'Esposito, M. (1997). A trial-based experimental design for fMRI. *NeuroImage*, *6*(2), 122–138. https://doi.org/10.1006/nimg.1997.0279

9 Acknowledgements

In my vagabond scientific career, I have been fortunate to work with many gifted individuals. Most of them have shaped my thoughts and actions. Trying to name them all would be a futile endeavour, but I am grateful to each of them. I would like to thank my wife Maria for her continuous support.