


Predicting Working Memory Training Responsiveness in Parkinson's Disease: Both "System Hardware" and Room for Improvement Are Needed

Neurorehabilitation and
Neural Repair
2021, Vol. 35(2) 117–130
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1545968320981956
journals.sagepub.com/home/nnr


Anja Ophey, MSc¹, Sarah Rehberg, MSc¹, Kathrin Giehl, PhD¹ ,
Carsten Eggers, MD^{2,3} , Paul Reker, MD¹, Thilo van Eimeren, MD^{1,4},
and Elke Kalbe, PhD¹

Abstract

Background. Patients with Parkinson's disease (PD) are highly vulnerable to develop cognitive dysfunctions, and the mitigating potential of early cognitive training (CT) is increasingly recognized. Predictors of CT responsiveness, which could help to tailor interventions individually, have rarely been studied in PD. This study aimed to examine individual characteristics of patients with PD associated with responsiveness to targeted working memory training (WMT). **Methods.** Data of 75 patients with PD (age: 63.99 ± 9.74 years, 93% Hoehn & Yahr stage 2) without cognitive dysfunctions from a randomized controlled trial were analyzed using structural equation modeling. Latent change score models with and without covariates were estimated and compared between the WMT group ($n = 37$), who participated in a 5-week adaptive WMT, and a waiting list control group ($n = 38$). **Results.** Latent change score models yielded adequate model fit (χ^2 -test $p > .05$, SRMR $\leq .08$, CFI $\geq .95$). For the near-transfer working memory composite, lower baseline performance, younger age, higher education, and higher fluid intelligence were found to significantly predict higher latent change scores in the WMT group, but not in the control group. For the far-transfer executive function composite, higher self-efficacy expectancy tended to significantly predict larger latent change scores. **Conclusions.** The identified associations between individual characteristics and WMT responsiveness indicate that there has to be room for improvement (e.g., lower baseline performance) and also sufficient "hardware" (e.g., younger age, higher intelligence) to benefit in training-related cognitive plasticity. Our findings are discussed within the compensation versus magnification account. They need to be replicated by methodological high-quality research applying advanced statistical methods with larger samples.

Keywords

Parkinson's disease, working memory, precision medicine, cognitive training, predictors of training responsiveness, structural equation modeling

Introduction

Parkinson's disease (PD) is the second most frequent neurodegenerative disorder of older age, with around 20% of patients showing mild cognitive impairment (PD-MCI) even at time of PD diagnosis¹ and up to 80% developing PD dementia in course of their disease.² Non-pharmacological interventions such as cognitive training (CT) are increasingly discussed to prevent or delay cognitive impairment and the progression to PD dementia.^{3,4}

As working memory (WM) has been ascribed the role of a processing resource for other cognitive abilities⁵ and is a vulnerable function in healthy aging,⁶ WM has already become a main target for CT in healthy older adults. In this population, recent meta-analyses⁷⁻⁹ on WM training (WMT) reveal significant positive training effects in trained tasks

¹University of Cologne, Cologne, Germany

²University Hospital of Marburg, Marburg, Germany

³Universities of Marburg and Gießen, Marburg, Germany

⁴German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

Supplementary material for this article is available on the *Neurorehabilitation & Neural Repair* website along with the online version of this article.

Corresponding Author:

Elke Kalbe, Department of Medical Psychology, Neuropsychology & Gender Studies, Center for Neuropsychological Diagnostic and Intervention (CeNDI), Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Straße 68, Cologne, Nordrhein-Westfalen 50937, Germany.
Email: elke.kalbe@uk-koeln.de

and near-transfer (i.e., untrained WM) tasks. However, the debate on the degree of far-transfer effects (i.e., effects in untrained domains) is unsolved, questioning the practical importance of WMT.⁷⁻⁹ Notably, WM and executive functions belong to the most vulnerable cognitive domains in PD.¹⁰ Furthermore, targeted WMT in healthy individuals has been associated with changes in frontoparietal and subcortical neural networks and particularly dopaminergic functions^{11,12}—neural networks and neurotransmitter systems on the one hand directly associated with WM,¹³ on the other hand, especially vulnerable in PD.¹⁴ Therefore, first randomized controlled trials (RCTs) on targeted WMT in PD have been conducted recently.^{15,16}

However, evidence on the effectiveness of WMT in PD is limited to direct training and partial near-transfer effects¹⁵⁻¹⁷ and the degree of far-transfer effects is unclear. Furthermore, even within near-transfer effects, responsiveness to training is highly heterogeneous, with some patients improving stronger than others, some patients maintaining, and others even decreasing their performance.¹⁶ Individual differences in cognitive plasticity might be able to explain this variance.^{18,19} Identifying such predictors of WMT responsiveness could promote the movement away from a one-treatment-suits-all-approach to the development of a “personalized” or “stratified” precision medicine in the context of the broad spectrum of CT interventions.^{20,21}

Importantly, prognostic research might be especially reasonable in the context of large heterogeneity among treatment effects: This heterogeneity could lead to very small insignificant and even null effects reported in single trials as well as meta-analyses, as effects in opposite directions between participants might eliminate overall reported effects. Prognostic research then bears the potential to identify predictors of treatment responsiveness to answer the question, if there are systematic relationships between individual characteristics and treatment response. One could argue that conducting prognostic analyses on null effects might be dealing with pure noise. However, the adherence to high methodological standards,^{20,22} for example, by comparing prognostic factors between a control group and the group receiving the treatment of interest, allows prognostic research to detangle unsystematic fluctuations and predictors of systematic retest effects, such as practice effects or regression to the mean, from predictors of treatment response.

Previous studies in healthy aging have, for example, investigated baseline performance, intelligence, age, and education as possible predictors of WMT responsiveness.²³ The results were summarized in a recent systematic review,²³ indicating large heterogeneity and several methodological shortcomings across studies. Even though only an elaborated individual participant data meta-analysis will be able to overcome the current research gaps regarding both general effectiveness and mechanisms of WMT responsiveness in healthy aging, the available evidence so far indicated that individuals with lower baseline performance, less education,

younger age, and higher intelligence seemed to benefit most from WMT. Furthermore, motivational processes seem to constitute an important variable regarding CT success in healthy older adults.²⁴

Fellman and colleagues¹⁵ investigated predictors of direct WMT effects in PD. Trained task improvement over the course of WMT was positively predicted by higher baseline general cognitive abilities, higher education, and shorter disease duration. Others, however, found lower baseline performance^{25,26} or longer disease duration^{27,28} to be positive predictors for CT responsiveness in PD. Notably, the applied statistical methods were highly heterogeneous ranging from simple correlational approaches and group comparisons to advanced methods such as growth curve analyses.

By analyzing data of an RCT on the effectiveness of WMT in PD,¹⁶ the aim of this study was to investigate individual characteristics associated with WMT responsiveness in PD, including both posttest and 3-month follow-up data, with an advanced statistical approach, that is, structural equation modelling (SEM). Since previous research in this field is inconclusive and data are rare, we pursue a rather exploratory approach, and no directional hypotheses were formulated. We investigated predictors of immediate versus 3-month follow-up WMT responsiveness separately for near-transfer and far-transfer outcomes, as underlying mechanisms might be heterogeneous and their analysis will further help to develop individualized interventions strengthening cognition in the long term.

Methods

Study Design

The original study was designed as a single-blind RCT to evaluate the effects of a 5-week computerized WMT compared with a waiting list control group (CG). Clinical and neuropsychological assessment took place at baseline, the week after the 5-week training/waiting period (posttest, 5.67 ± 0.58 weeks after baseline), and at 3-month follow-up (14.03 ± 0.86 weeks after posttest). Patients were recruited via the University Hospital of Cologne, Germany, as well as regional neurologists and PD support groups. During all assessments, patients were on their regular medication. Outcome assessors were blinded for group allocation, patients were not.

Summarizing the results of the 2 prior publications on the neuropsychological and clinical outcomes in this trial, statistically significant positive training effects for the WMT group compared with the CG were reported in the WM domain only.^{16,17} Reliable far-transfer effects in cognitive and clinical variables were not found for either point of time. However, variability of training effects was large across participants with some patients improving over time and some patients remaining stable or even showing declined performance over time.¹⁶ The WMT effect on neural correlates investigated with a newly developed WM paradigm and functional MRI was small, but decreased activation in areas relevant for working

memory could be observed, with activation changes correlating with behavioral change.²⁹ Further details are reported in Ophey and colleagues¹⁶ and Giehl and colleagues.^{17,29}

Standard Protocol Approvals, Registrations, and Patient Consents

The study protocol was approved by the local ethics committee of the Medical Faculty of the University of Cologne (vote-no.16-043), conducted in compliance with the Helsinki Declaration of 1975, as revised in 2008, and registered with the German Clinical Trials Register (drks.de, DRKS00009379). Before the baseline assessment, all participants gave written informed consent for participation. The reporting of the original RCT¹⁶ followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines.

Participants

Inclusion criteria for study participation were age between 45 and 85 years, diagnosis of idiopathic PD according to UK brain bank criteria,³⁰ and normal or corrected-to-normal vision and hearing. Exclusion criteria were cognitive dysfunctions according to the level-II diagnostic criteria for PD-MCI³¹ or PD dementia,³² severe depressive symptoms operationalized by the Geriatric Depression Scale ($GDS \geq 11$),³³ deep brain stimulation and other reported psychiatric, neurological, and life-threatening diseases. Included patients were randomized to either the WMT group ($n = 37$) or CG ($n = 39$). 75 of these 76 patients completed the posttest, and 72 patients completed the 3-month follow-up. For details on inclusion and exclusion criteria, see Ophey and colleagues,¹⁶ and for details on study flow, refer to the flow chart in Figure 1.

For this post hoc analysis, we included all patients for which at least one follow-up assessment was carried out (i.e., posttest and/or 3-month follow-up). Therefore, our final sample consisted of $n = 37$ patients in the WMT group, $n = 38$ patients in the CG. On average, our patients (47% female) were 63.99 ± 9.74 years old, and diagnosed with PD 6.36 ± 5.39 years ago, with the wide majority of patients showing mild to moderate motor impairment (93% Hoehn & Yahr stage 2, Unified Parkinson's Disease Rating Scale-III, UPDRS-III, 16.81 ± 8.28). Further details on patient characteristics can be obtained from Table 1.

Intervention

The computerized WMT consisted of 5 varying WM tasks per training session selected out of a total of 9 different WM tasks based on the online CT program NeuroNation (Synaptikon GmbH, Berlin, Germany). The WMT was accessed online and administered at home for 30 minutes a day, 5 days a week, over 5 weeks and adapted to user performance across training sessions. More details are described in

Ophey and colleagues.¹⁶ No training was applied between posttest and 3-month follow-up. The CG was a waiting list CG and was granted access to the WMT after their last follow-up.

Statistical Analysis

All statistical analyses were conducted in R (<https://www.r-project.org>). SEM was carried out using the R package *lavaan*. The R code can be obtained from Supplementary Material 1. We fitted latent difference score models (LDSM), where the difference between the measured outcome at posttest and baseline (or 3-month follow-up and baseline), that is, the change through WMT, is represented as a latent variable (Δ_{posttest} , $\Delta_{\text{follow-up}}$). The measured outcome at posttest or 3-month follow-up in turn is the sum of the measured outcome at baseline and the corresponding latent difference score.

To specify this in terms of SEM,* as represented in Figure 2a for the basic LDSM, both, the path coefficients from baseline to posttest/3-months follow-up and from the latent difference score to the measured outcome at the respective point of time, are set equal to 1. By setting the intercepts and residuals for measured outcomes at posttest and 3-months follow-up equal to 0, the estimated intercept of $\Delta_{\text{posttest}}/\Delta_{\text{follow-up}}$ (α_i) represents the average observed difference score at the respective point of time, and the estimated variance ($\psi_{i,i}$) of $\Delta_{\text{posttest}}/\Delta_{\text{follow-up}}$ represents the variability of the change through training. The covariance ($\psi_{r-1,i}$) between baseline and $\Delta_{\text{posttest}}/\Delta_{\text{follow-up}}$ represents the correlation between the measured outcome at baseline and the change through training.

To identify characteristics associated with individual variation in difference scores beyond individual baseline performance, we extended the basic LDSM by covariates predicting the variance of latent change scores ($\psi_{i,i}$ by $\beta_{i,i}$), Figure 2b. Note that several global terms for individual differences predicting training responsiveness are used interchangeably in literature,^{20,34} for example, “predictors”, “prognostic factors”, “covariates”, and “moderators”. Here, we call them “predictors” on a content-related level, and “covariates” on a methodological level, referring to their properties in the applied LDSM. For further details on outcome measures and covariates, see the corresponding sections below.

Overall model fit was evaluated using the χ^2 -test. Further approximative fit indices as the standardized root mean square residual (SRMR), and comparative fit index (CFI) are reported. Values of $SRMR \leq .08$ and $CFI \geq .95$ were considered to reflect adequate model fit.^{35,36} No model fit indices are reported for the basic LDSM, as the model is just identified (degrees of freedom, $df = 0$) and no fit indices can be assessed.

*Notation follows the LISREL notation system, the most widely applied notation in structural equation modeling literature.

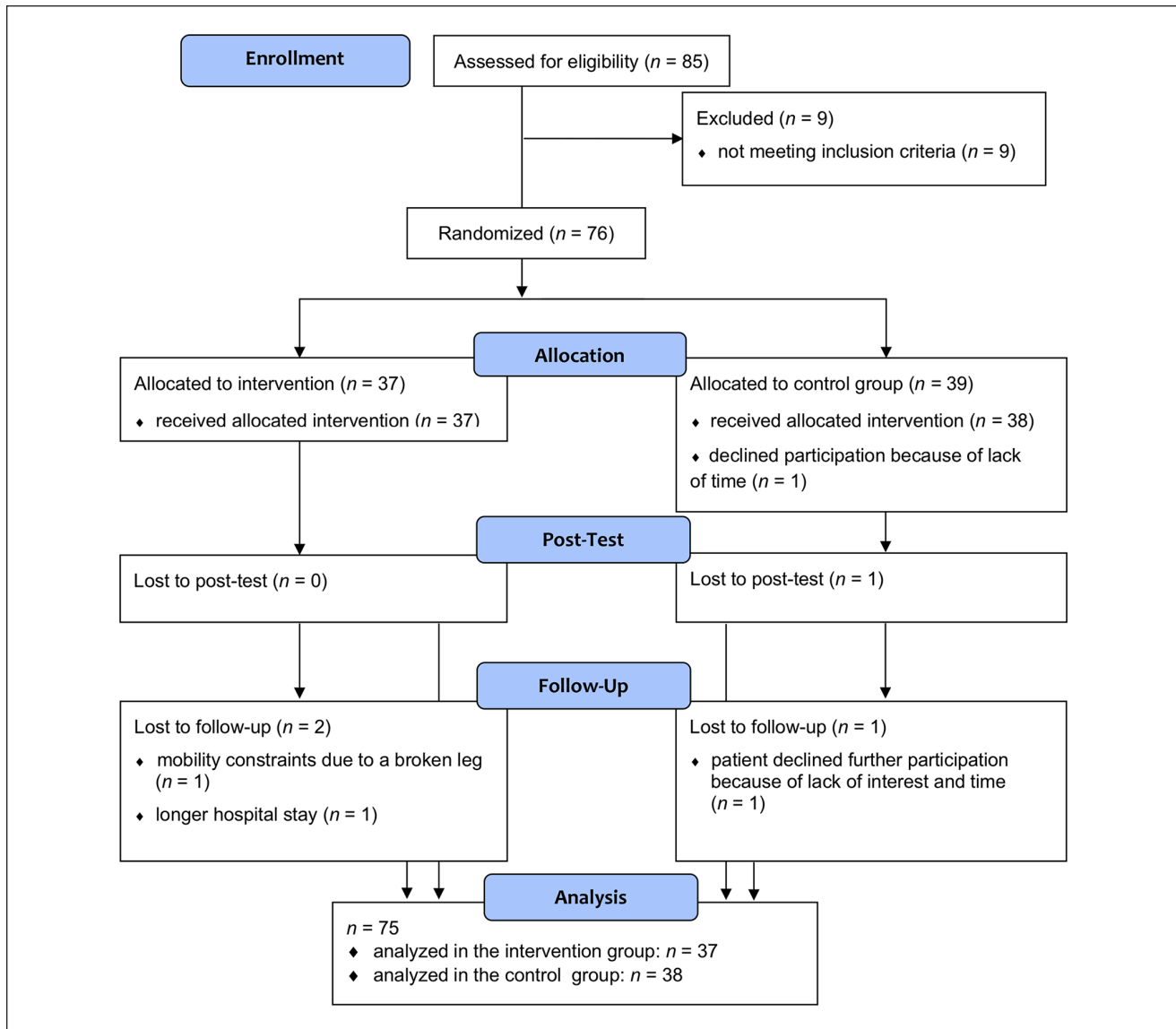


Figure 1. Flowchart: Enrollment, treatment allocation, posttest, and 3-month follow-up of study participants.

Full information maximum likelihood estimation (FIML) was used to deal with missing data patterns. As suggested for SEM with small samples,³⁷ Bollen-Stine bootstrapping with $n = 1000$ bootstrap replicates was performed for both model parameter estimates and fit indices using the R package *semTools*. Significance level for model parameters was set to $\alpha = .05$. Following the explorative nature of our analyses on individual characteristics predicting WMT responsiveness, we interpreted $\alpha \leq .10$ as a trend toward statistical significance for path coefficients of covariates to Δ_{posttest} and $\Delta_{\text{follow-up}}$ ($\beta_{i,j}$). Ninety-five percent CI are reported for model parameters and fit indices.

For the extended LDSM, we compared model fit between a model where the variances ($\psi_{i,i}$), covariances ($\psi_{i-1,i}$), and path coefficients between covariates and the variance of Δ_{posttest} and $\Delta_{\text{follow-up}}$ ($\beta_{i,i}$) were assumed to be equal across the

WMT group and CG versus a model where those parameters were left free to vary using a χ^2 -difference test ($\Delta\chi^2$). These group comparisons enable us to draw inferences on predictors specific for training responsiveness compared to retest effects. Regarding the influence of baseline performance on $\Delta_{\text{posttest}}/\Delta_{\text{follow-up}}$, we additionally compared the standardized covariances (i.e., correlations, $\psi_{\alpha\alpha}$) between baseline performance and $\Delta_{\text{posttest}}/\Delta_{\text{follow-up}}$ between the WMT group and CG using a z -test for independent sample correlations. The fitted models are separately reported for the WMT group and the CG.

Outcomes

We analyzed data of two cognitive domains: WM as near-transfer effect measure and executive functions as a

Table 1. Patient Characteristics.^a

Characteristic	Working memory training group (n = 37)	Control group (n = 38)
Age in years	64.09 (8.56)	63.88 (8.39)
Sex, n (%)		
Female	18 (48.6)	17 (44.7)
Male	19 (51.4)	21 (55.3)
Education in years	15.00 (11.00-22.00)	15.50 (10.00-23.00)
Global cognition MoCA score	27 (24-30)	28 (25-30)
Fluid intelligence LPS-4 T-score	53.84 (10.60)	54.32 (8.32)
Working memory composite T-score		
Baseline	49.08 (5.60)	51.82 (6.16)
Posttest	50.59 (4.60)	52.97 (5.88)
Follow-up	51.10 (5.76)	52.85 (4.77)
Executive functions composite T-score		
Baseline	52.75 (5.76)	54.28 (4.77)
Posttest	53.74 (4.45)	55.17 (4.78)
Follow-up	54.09 (5.35)	54.50 (4.52)
Disease duration in years	5.20 (0.47-22.05)	4.62 (0.35-27.04)
H&Y, n (%)		
Stage 2	36 (94.7)	35 (94.6)
Stage 3	2 (5.3)	2 (5.4)
LEDD in mg	600.00 (0 - 1785)	590.00 (100 - 2120)
Self-efficacy expectancy SWE T-score	49.11 (9.57)	53.65 (11.24)

Abbreviations: H&Y, Hoehn & Yahr scale; LEDD, levodopa equivalent daily dose; LPS-4, Leistungsprüfsystem substest 4—reasoning; MoCA, Montreal Cognitive Assessment; SWE, Skala zur Allgemeinen Selbstwirksamkeitserwartung.

^aData are mean (standard deviation) or median (range by minimum to maximum) unless indicated otherwise. Variables were previously inspected visually by qq-plots and statistically by Shapiro-Wilk tests for normal distribution.

far-transfer effect measure, assessed by several single test measures each, including both paper-pencil assessments and computer tests and parallel test forms, if available. *T*-standardized ($M = 50$, $SD = 10$) test scores were calculated using published normative data of age-, sex-, and education-matched healthy controls. Domain composite scores were computed as the average of the corresponding equally weighted single test *T*-scores, which strengthened the interpretability of scores on the cognitive domain level and reduced overall type-I-error.

WM was assessed by the digit span forward and backward task (Wechsler Memory Scale revised)³⁸ the n-back verbal and n-back nonverbal tasks and the CORSI block tapping tasks forward and backward (Wiener Testsystem).³⁹ Tests for executive functions included semantic and phonemic verbal fluency (CERAD-Plus),⁴⁰ Trail Making Test B/A (TMTB divided by TMTA, CERAD-Plus),⁴⁰ and the interference measure of the Stroop Color-Word-Interference test.⁴¹ For details on further cognitive and clinical outcomes in the study, see Ophey and colleagues.¹⁶

Covariates

Age in years, education in years, fluid intelligence as measured with the LPS-4 (Leistungsprüfsystem substest 4:

reasoning, standardized *T*-score),⁴² PD motor impairment as measured with the UPDRS-III,⁴³ and self-efficacy expectancy as measured with the SWE (Skala zur Allgemeinen Selbstwirksamkeitserwartung, standardized *T*-score)⁴⁴ were included in the extended LDSM.[†] Covariates were assessed at baseline assessment with both patients and outcome assessors blinded for group allocation and knowledge of candidate predictors. Covariates were grand-mean-centered to ensure interpretability of the difference score intercepts (α_j). Covariances between all combinations of covariates were allowed in the model.

[†]With this set of covariates, we derogated from the trial registration, which listed age, education, PD motor impairment, self-efficacy expectancy, and training motivation as potential predictors. As already described above, we aimed at fitting the models in both the WMT group and CG to ensure validity of our results concerning predictors of training responsiveness compared to retest effects. Training motivation was only assessed in the WMT group, and, therefore, had to be excluded from our set of covariates. Furthermore, by including fluid intelligence as one potential predictor, we followed the literature on WMT responsiveness in healthy aging, where fluid intelligence, next to age and baseline performance (which our models account for by including the covariance, $\Psi_{i-1,p}$, between baseline and $\Delta_{\text{posttest}} / \Delta_{\text{follow-up}}$), is one of the most frequently investigated predictors.

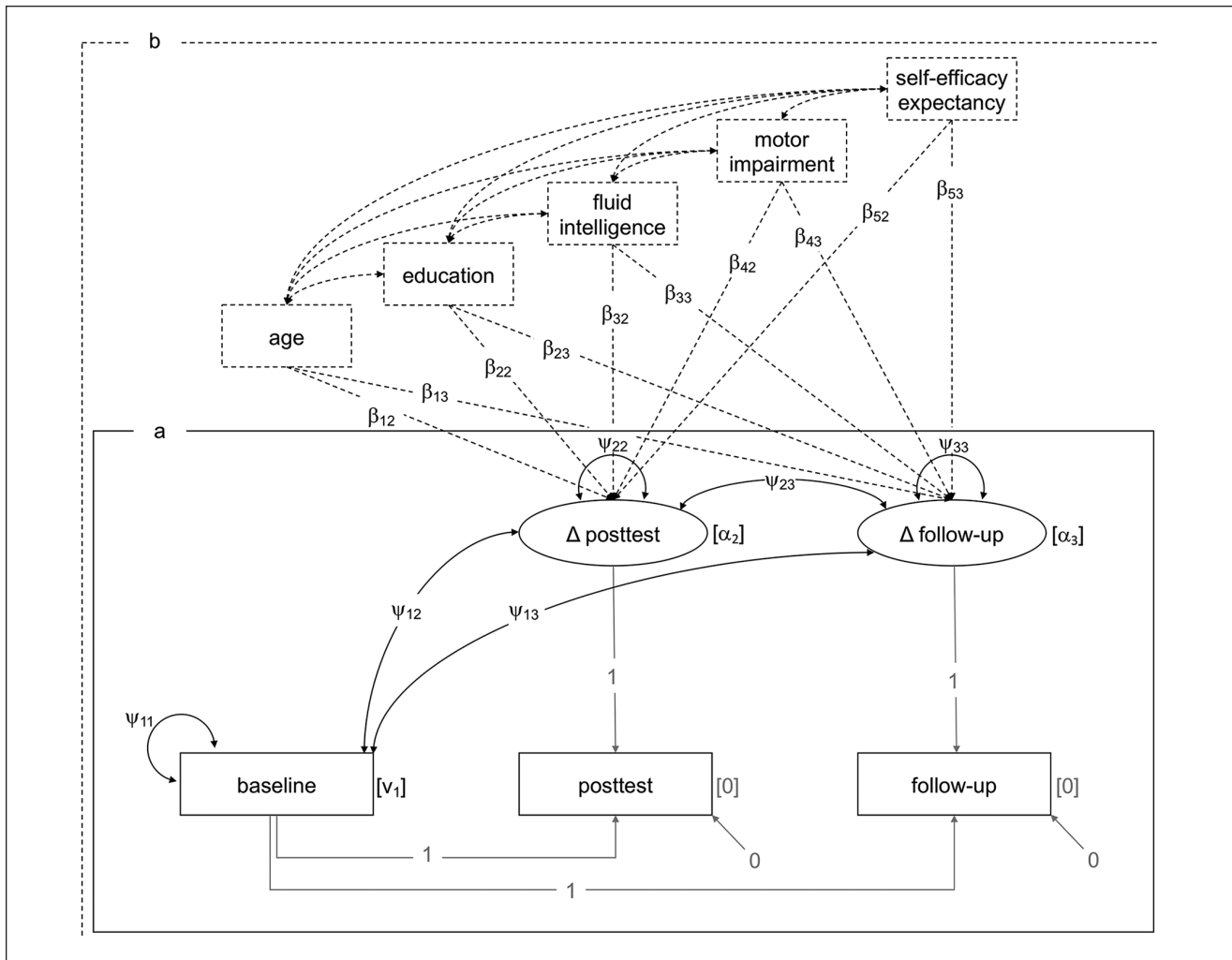


Figure 2. (a) Basic latent difference score model, (b) extended latent difference score model including covariates. Rectangles represent observed variables, and ellipses represent latent variables. Fixed parameters and path coefficients appear in gray. Free (i.e., to be estimated) parameters and path coefficients appear in black. Covariates and corresponding paths introduced for the extended latent difference score model appear with dashed lines. Covariances between covariates introduced for the extended models are estimated for control reasons, however, are not of main interest for our research question. α_i = intercept of latent variable, β_{ji} = path coefficient for covariates, Δ = difference, v_i = intercept of observed variable, $\psi_{i,i}$ = variance, $\psi_{i-1,i}$ = covariance.

Results

Near-Transfer: Working Memory

Details on fitted models for WM without equality constraints between the WMT group and CG can be obtained from Table 2. The basic LDSM was just identified ($df = 0$), therefore neither model fit nor a test for measurement invariance between the WMT and CG was assessed. The estimated mean of Δ posttest was significant in both the WMT group and the CG. However, the estimated mean of Δ follow-up was only significant for the WMT group, but not for the CG. Variance estimates indicate significant individual variation in both Δ posttest and Δ follow-up for both the WMT group and CG.

In the WMT group, the covariances of baseline performance and both Δ posttest ($\psi_{12,WMT} = -10.51, p = .005$) and Δ follow-up ($\psi_{13,WMT} = -13.73, p = .010$) were significantly negative, suggesting that individuals with lower baseline performance showed both larger immediate and follow-up gains through training than individuals with higher baseline performance. In the CG, only baseline performance and Δ follow-up ($\psi_{13,CG} = -7.85, p = .044$) were significantly negatively correlated. A comparison of standardized covariances (i. e., correlations) of the basic LDSM between WMT group and CG indicated descriptively stronger negative relationships between baseline performance and Δ posttest, as well as Δ follow-up, in the WMT group compared with the CG; however, these differences did not reach statistical significance.

Table 2. Latent Difference Score Models of Near-Transfer Working Memory.^a

	Working memory training group (n = 37)				Control group (n = 38)			
	Estimate	(95% CI)	SE	p	Estimate	(95% CI)	SE	p
<i>Basic model</i>								
<i>Intercepts</i>								
Baseline	48.96	(47.14, 50.78)	0.90	<.001***	51.42	(49.44, 53.4)	0.99	<.001***
ΔPosttest	1.62	(0.49, 2.76)	0.56	.006**	1.56	(0.12, 3.01)	0.72	.035*
ΔFollow-up	1.98	(0.26, 3.69)	0.84	.025*	1.16	(-0.05, 2.37)	0.60	.059 ⁺
<i>Variances</i>								
Baseline	29.74	(15.69, 43.8)	6.93	<.001***	36.47	(19.62, 53.32)	8.43	<.001***
ΔPosttest	11.38	(5.96, 16.81)	2.68	<.001***	19.08	(10.14, 28.02)	4.47	<.001***
ΔFollow-up	25.62	(13.27, 37.97)	6.09	<.001***	13.18	(6.92, 19.44)	3.13	<.001***
<i>Covariances</i>								
Baseline~ΔPosttest	-10.51	(-17.62, -3.41)	3.50	.005**	-8.60	(-17.75, 0.54)	4.57	.064 ⁺
Baseline~ΔFollow-up	-13.73	(-24.01, -3.46)	5.07	.010**	-7.85	(-15.49, -0.2)	3.82	.044*
ΔPosttest~ΔFollow-up	8.57	(2.06, 15.07)	3.21	.011*	10.21	(3.9, 16.52)	3.16	.002*
<i>Extended model with covariates</i>								
<i>Intercepts</i>								
Baseline	48.95	(47.12, 50.78)	0.90	<.001***	51.50	(49.54, 53.45)	0.98	<.001***
ΔPosttest	1.79	(0.6, 2.99)	0.59	.004**	1.76	(0.25, 3.27)	0.75	.023*
ΔFollow-up	2.26	(0.49, 4.03)	0.87	.014*	1.06	(-0.12, 2.24)	0.59	.076 ⁺
<i>Variances</i>								
Baseline	30.28	(15.98, 44.59)	7.05	<.001***	35.67	(19.15, 52.19)	8.26	<.001***
ΔPosttest	11.69	(5.92, 17.46)	2.85	<.001***	17.72	(9.21, 26.23)	4.26	<.001***
ΔFollow-up	23.79	(12, 35.58)	5.81	<.001***	10.47	(5.38, 15.56)	2.55	<.001***
<i>Covariances</i>								
Baseline~ΔPosttest	-15.35	(-23.61, -7.09)	4.07	.001**	-10.95	(-20.41, -1.5)	4.73	.024*
Baseline~ΔFollow-up	-16.62	(-27.54, -5.7)	5.38	.004**	-7.31	(-14.46, -0.15)	3.58	.046*
ΔPosttest~ΔFollow-up	10.63	(3.79, 17.47)	3.37	.003**	9.21	(3.58, 14.84)	2.82	.002**
<i>Regressions</i>								
<i>ΔPosttest R²</i>								
~ Age	-0.10	(-0.19, -0.01)	0.04	.024*	-0.05	(-0.23, 0.13)	0.09	.587
~ Education	0.05	(-0.21, 0.31)	0.13	.707	0.20	(-0.36, 0.75)	0.28	.476
~ Fluid intelligence	0.18	(0.11, 0.26)	0.04	<.001***	-0.01	(-0.19, 0.18)	0.09	.940
~ Motor impairment	0.04	(-0.07, 0.15)	0.05	.452	-0.04	(-0.24, 0.16)	0.10	.659
~ Self-efficacy	0.02	(-0.05, 0.1)	0.04	.521	-0.07	(-0.2, 0.06)	0.07	.288
<i>ΔFollow-up R²</i>								
~ Age	-0.12	(-0.29, 0.05)	0.08	.163	-0.07	(-0.22, 0.07)	0.07	.302
~ Education	0.44	(-0.06, 0.93)	0.25	.083 ⁺	0.12	(-0.32, 0.56)	0.22	.589
~ Fluid intelligence	0.07	(-0.07, 0.21)	0.07	.325	-0.05	(-0.2, 0.1)	0.07	.521
~ Motor impairment	0.12	(-0.09, 0.33)	0.10	.254	0.02	(-0.14, 0.18)	0.08	.778
~ Self-efficacy	0.03	(-0.12, 0.19)	0.08	.675	0.04	(-0.07, 0.14)	0.05	.505
<i>Fit indices</i>								
χ ² -test	χ ² (5) = 6.01	(0.92, 16.57)		.305	χ ² (5) = 5.32	(0.84, 14.20)		.378
CFI	0.97	(0.85, 1)			0.98	(0.89, 1)		
SRMR	0.08	(0.03, 0.14)			0.08	(0.03, 0.15)		

Abbreviations; CFI, comparative fit index; SRMR, standardized root mean square residual; Δ, latent difference score; ~, regressed on; ~~, covaries with.

^aBollen-Stine bootstrap parameter estimates with n = 1000 bootstrap samples are reported.

⁺p < .100; *p < .050; **p < .010; ***p < .001.

(ψ_{12} , $z = -1.25$, $p = .106$; ψ_{13} , $z = -0.76$, $p = .225$). The covariance of Δ posttest and Δ follow-up was significantly positive for both the WMT group and CG, indicating that Δ posttest and Δ follow-up were positively correlated.

The extended LDSM further included the covariates age, education, fluid intelligence, PD motor impairment and self-efficacy. The model fit difference between the extended LDSM with, $\chi^2(25) = 57.30$, $p \leq .001$, SRMR = 0.12

[95% CI 0.07, 0.18], CFI = 0.80 [95% CI 0.61, 0.95], and without, $\chi^2(10) = 11.31, p = .305$, SRMR = 0.08 [95% CI 0.04, 0.12], CFI = 0.98 [95% CI 0.92, 1.00], equality constraints between the WMT group and CG was significant, $\Delta\chi^2(15) = 45.99, p \leq .001$, revealing a benefit for model fit by allowing model parameters to differ between the WMT group and CG. The extended LDSM of the WMT group revealed comparable parameter estimates for intercepts, variances, and covariances as the basic LDSM. Even when controlling for the above-mentioned covariates, estimates for $\Delta\text{posttest}$, $\Delta\text{follow-up}$, variances, and covariances remain significant. Age was found to be a significant negative predictor of $\Delta\text{posttest}$ ($\beta_{12} = -0.10, p = .024$) and fluid intelligence was found to be a significant positive predictor of $\Delta\text{posttest}$ ($\beta_{32} = 0.18, p \leq .001$), indicating that younger individuals with higher levels of fluid intelligence show largest immediate training gains. For $\Delta\text{follow-up}$, only for education a trend toward significance as a positive predictor was reported ($\beta_{23} = 0.44, p = .083$), that is, the more educational years, the higher the training gains. No other covariate was found to significantly predict variance in $\Delta\text{posttest}$ and $\Delta\text{follow-up}$ within the WMT group.

The extended LDSM of the CG revealed comparable parameter estimates for intercepts, variances, and covariances as the basic LDSM as well. However, when controlling for the aforementioned covariates, the covariance between baseline and $\Delta\text{posttest}$ reached statistical significance ($\psi_{12,CG} = -10.95, p = .024$). A comparison of standardized covariances of the extended LDSM between WMT group and CG indicated a significantly stronger negative relationship between baseline performance and $\Delta\text{posttest}$ in the WMT group compared with the CG ($\psi_{12}, z = -2.631, p = .004$), as well as a trend in the same direction between baseline performance and $\Delta\text{follow-up}$ ($\psi_{13}, z = -1.251, p = .105$). None of the covariates was found to be a significant predictor of variance in $\Delta\text{posttest}$ and $\Delta\text{follow-up}$ within the CG.

Far-Transfer: Executive Functions

Details on fitted models for executive functions without equality constraints between the WMT group and CG can be obtained from Table 3. The basic LDSM was just identified ($df = 0$). The estimated means of $\Delta\text{posttest}$ and $\Delta\text{follow-up}$ were not statistically significant, neither in the WMT group nor in the CG. However, for both points of time, they were descriptively larger in the WMT group than in the CG. Variance estimates indicate significant individual variation in both $\Delta\text{posttest}$ and $\Delta\text{follow-up}$ for both the WMT group and CG. The covariance of baseline performance and both $\Delta\text{posttest}$, as well as $\Delta\text{follow-up}$, for both the WMT group ($\psi_{12,WMT} = -18.36, p = .001$; $\psi_{13,WMT} = -14.27, p = .004$) and the CG ($\psi_{12,CG} = -16.17, p = .002$; $\psi_{13,CG} = -8.64, p = .009$) were significantly negative;

however, it did not significantly differ between the WMT group and CG ($\psi_{12}, z = -0.98, p = .164$; $\psi_{13}, z = -0.021, p = .492$). The covariance of $\Delta\text{posttest}$ and $\Delta\text{follow-up}$ was significantly positive for both the WMT group and CG.

The model fit difference between the extended LDSM with, $\chi^2(25) = 56.00, p \leq .001$, SRMR = 0.11 [95% CI 0.08, 0.16], CFI = 0.74 [95% CI 0.52, 0.94], and without, $\chi^2(10) = 12.63, p = .245$, SRMR = 0.07 [95% CI 0.04, 0.12], CFI = 0.97 [95% CI 0.85, 1.00], equality constraints between the WMT group and CG was significant, $\Delta\chi^2(15) = 43.37, p \leq .001$, revealing a benefit for model fit by allowing model parameters to differ between the WMT group and CG. The extended LDSM of the WMT group revealed comparable parameter estimates for intercepts, variances, and covariances as the basic LDSM. Regarding the influence of covariates on variance at $\Delta\text{posttest}$ and $\Delta\text{follow-up}$, self-efficacy expectancy only turned out to be a significant positive predictor of $\Delta\text{posttest}$ ($\beta_{52} = 0.14, p = .019$) in the WMT group, indicating that individuals with higher levels of self-efficacy expectancy show larger immediate training gains. No other covariate was found to be a significant predictor of variance in $\Delta\text{posttest}$ and $\Delta\text{follow-up}$ within the WMT group. The extended LDSM of the CG revealed comparable parameter estimates for intercepts, variances, and covariances as the basic LDSM as well. Age was found to be a significantly positive predictor of the variance in $\Delta\text{posttest}$ ($\beta_{13} = 0.20, p = .022$) for the CG. Furthermore, a trend of fluid intelligence being a positive predictor ($\beta_{33} = 0.14, p = .093$) and self-efficacy expectancy being a negative predictor ($\beta_{53} = -0.11, p = .064$) of the variance in $\Delta\text{posttest}$ was found for the CG.

Discussion

The aim of this study was to investigate predictors of WMT responsiveness in patients with PD by analyzing data of an RCT on the effectiveness of a computerized 5-week targeted WMT in PD, including both posttest and 3-months follow-up assessments¹⁶ with an advanced SEM approach. Summarizing, the extended LDSM including covariates to predict individual variance in latent change scores revealed lower baseline performance, younger age, and higher fluid intelligence to significantly predict higher change scores in near-transfer WM outcomes directly after WMT (posttest), as well as lower baseline performance and higher education to predict larger change scores at 3-month follow-up. For far-transfer executive function outcomes, higher self-efficacy expectancy showed a trend to significantly predict larger positive WMT responsiveness at posttest.

The predictive power of individual differences regarding training responsiveness is often discussed within the compensation versus magnification framework.^{45,46} Following the magnification account, individuals with more cognitive resources “to

Table 3. Latent Difference Score Models of Far-Transfer Executive Functions.

	Working memory training group (n = 37)				Control group (n = 38)			
	Estimate	(95% CI)	SE	p	Estimate	(95% CI)	SE	p
<i>Basic model</i>								
<i>Intercepts</i>								
Baseline	52.76	(50.87, 54.65)	0.93	<.001***	54.24	(52.72, 55.77)	0.76	<.001***
ΔPosttest	0.97	(-0.61, 2.56)	0.78	.222	0.93	(-0.67, 2.52)	0.80	.248
ΔFollow-up	1.27	(-0.17, 2.71)	0.71	.083 ⁺	0.18	(-1.08, 1.44)	0.63	.773
<i>Variances</i>								
Baseline	31.58	(16.48, 46.67)	7.44	<.001***	21.47	(11.42, 31.52)	5.02	<.001***
ΔPosttest	22.17	(11.57, 32.77)	5.23	<.001***	23.98	(12.49, 35.48)	5.75	<.001***
ΔFollow-up	17.81	(9.09, 26.53)	4.30	<.001***	14.30	(7.32, 21.28)	3.49	<.001***
<i>Covariances</i>								
Baseline~ΔPosttest	-17.48	(-28.17, -6.8)	5.27	.002**	-11.42	(-20.04, -2.81)	4.31	.011*
Baseline~ΔFollow-up	-11.38	(-20.4, -2.35)	4.45	.015*	-8.40	(-14.96, -1.85)	3.28	.013*
ΔPosttest~ΔFollow-up	13.72	(5.47, 21.97)	4.07	.002**	11.79	(4.38, 19.19)	3.70	.003**
<i>Extended model with covariates</i>								
<i>Intercepts</i>								
Baseline	52.73	(50.86, 54.61)	0.92	<.001***	54.24	(52.72, 55.75)	0.76	<.001***
ΔPosttest	1.37	(-0.21, 2.95)	0.78	.088 ⁺	1.17	(-0.54, 2.87)	0.85	.175
ΔFollow-up	1.33	(-0.17, 2.83)	0.74	.081 ⁺	0.26	(-0.96, 1.48)	0.61	.669
<i>Variances</i>								
Baseline	31.14	(16.25, 46.03)	7.34	<.001***	21.40	(11.42, 31.52)	5.07	<.001***
ΔPosttest	19.88	(9.88, 29.87)	4.93	<.001***	24.94	(12.71, 37.16)	6.12	<.001***
ΔFollow-up	16.75	(8.31, 25.19)	4.16	<.001***	11.43	(5.64, 17.21)	2.89	<.001***
<i>Covariances</i>								
Baseline~ΔPosttest	-18.36	(-29.1, -7.63)	5.29	.001**	-16.17	(-25.89, -6.44)	4.87	.002**
Baseline~ΔFollow-up	-14.27	(-23.8, -4.75)	4.70	.004**	-8.64	(-15, -2.28)	3.18	.009**
ΔPosttest~ΔFollow-up	13.75	(5.64, 21.86)	4.00	.001**	11.67	(4.42, 18.91)	3.62	.002**
<i>Regressions</i>								
<i>ΔPosttest R²</i>								
~ Age	0.01	(-0.12, 0.14)	0.06	.840	0.20	(0.03, 0.37)	0.09	.022*
~ Education	-0.12	(-0.51, 0.27)	0.19	.525	-0.07	(-0.6, 0.45)	0.26	.782
~ Fluid intelligence	0.03	(-0.08, 0.14)	0.06	.582	0.14	(-0.02, 0.31)	0.08	.093 ⁺
~ Motor impairment	-0.00	(-0.17, 0.16)	0.08	.982	-0.07	(-0.26, 0.12)	0.10	.464
~ Self-efficacy	0.14	(0.02, 0.27)	0.06	.019*	-0.11	(-0.24, 0.01)	0.06	.064 ⁺
<i>ΔFollow-up R²</i>								
~ Age	0.09	(-0.06, 0.23)	0.07	.226	0.06	(-0.08, 0.2)	0.07	.362
~ Education	-0.10	(-0.52, 0.32)	0.21	.619	-0.30	(-0.72, 0.12)	0.21	.158
~ Fluid intelligence	0.06	(-0.07, 0.18)	0.06	.364	0.11	(-0.03, 0.24)	0.07	.119
~ Motor impairment	-0.07	(-0.25, 0.1)	0.09	.404	0.12	(-0.04, 0.28)	0.08	.152
~ Self-efficacy	0.06	(-0.07, 0.18)	0.06	.389	0.01	(-0.08, 0.11)	0.05	.783
<i>Fit indices</i>								
χ ² -test	χ ² (5) = 7.28	(1.00, 21.16)		.201	χ ² (5) = 5.23	(0.89, 14.21)		.388
CFI	0.95	(0.76, 1)			0.974	(0.84, 1)		
SRMR	0.08	(0.03, 0.15)			0.066	(0.03, 0.12)		

Abbreviations: CFI, comparative fit index; SRMR, standardized root mean square residual; Δ, latent difference score; ~, regressed on; ~~, covaries with.
^aBollen-Stine bootstrap parameter estimates with n = 1000 bootstrap samples are reported.

⁺p < .100; *p < .050; **p < .010; ***p < .001.

acquire, implement, and sharpen effortful cognitive strategies^{24,45} would show largest training benefits. On the contrary, the compensation hypothesis argues that individuals with lower

performance prior to training have more room for improvement than individuals already performing at or near optimal levels, and therefore would show largest training benefits⁴⁵.

Baseline Performance, the Compensation Hypothesis, and Methodological Considerations

Some studies found positive associations between baseline performance and change through training,^{15,47} whereas others found evidence for negative associations.^{25,26} As broadly discussed in Smolén and colleagues,⁴⁸ one fallacy linking those negative associations between baseline performance and change through training directly to the compensation account is that other sources for this association are neglected and not accounted for in analyses. One such important artefact is regression to the mean, a statistical artefact that causes negative correlations between baseline performance and gain by noisy repeated measurements, where extreme values at the first point of time tend to be closer to the mean at the second point of time without reflecting real change.⁴⁸ Furthermore, analyses to identify predictors of training responsiveness are frequently conducted with data of the experimental group only.⁴⁸ Therefore, they do not control for effects in the control group,²⁰ which impedes disentangling predictors of training responsiveness from predictors of retest and practice effects.⁴⁹

The analytical approach of fitting LDSM that allowed to explore the (statistical) properties of change without actually calculating change scores, and the comparison of models fitted in the WMT group versus models fitted in the CG, enabled us to draw inferences on predictors specific for training responsiveness compared with regression to the mean and practice effects.⁴⁸ In our LDSM, the negative covariances between baseline performance and latent difference scores indicated that individuals with lower baseline performance in the respective outcome measure tend to show larger change scores than individuals with higher baseline performance. For near-transfer WM, not only descriptive but also statistically significant differences between standardized covariances in the WMT group and CG were present, revealing significantly stronger negative relationships between baseline performance and change in the WMT group compared with the CG. We hypothesize this difference to be attributable to an increased compensation effect through training in the WMT group. Therefore, our data indicate that for both post-test and long-term near-transfer responsiveness of WMT in PD, “room for improvement” regarding participants’ WM abilities at baseline seems necessary.

In the CG, the negative association is in line with the compensation account as well, however, this compensation does not reflect change through training, but rather compensation through practice effects and regression to the mean. Similarly, the negative relationship between baseline performance and change in the LDSM for executive functions might rather represent compensation through practice effects and regression to the mean, as the covariances were significant for both groups and differences between the two groups were present on a descriptive level only.

Other Predictors and the Magnification Hypothesis

Notably, our data regarding associations of other predictors than baseline performance rather point to evidence for the magnification account: latent change score estimates were higher for individuals of the WMT group with higher intelligence, younger age, more educational years, and higher self-efficacy expectancy. Higher intelligence might constitute the required “hardware” to utilize the possibilities given by WMT to extend the cognitive repertoire and to engage in true plasticity. Corroborating this hypothesis, higher intelligence was found to be a positive predictor for change through training in the first study investigating predictors of WMT responsiveness for patients with PD,¹⁵ as well as studies with healthy older adults.^{50,51}

Age might constitute a proxy for the general potential to engage in neural and cognitive plasticity.^{52,53} Due to age-related reductions in processing resources,⁶ the ability to engage in plastic changes after WMT might be reduced with increasing age. Our results are in line with several findings from WMT in healthy older adults, which also identified younger age as a positive predictor for training responsiveness.^{50,51,54} In contrast to our results, Fellman and colleagues¹⁵ did not find an association between age and training responsiveness. However, they identified shorter PD disease duration as a positive predictor for change through training. As the prevalence of PD increases with age⁵⁵ and neuropathological changes progress with disease duration,⁵⁶ disease duration might be another (and possibly even the more conclusive) measure as a proxy for the potential of neural and cognitive plasticity in PD. However, conflicting results identifying longer PD duration to be predictive of higher training gains exist.^{27,28} This discrepancy regarding the direction of predictive effects and corresponding mechanisms might be dependent on the type of CT and the applied analytical approach.²³

As in Fellman and colleagues,¹⁵ more educational years were predictive for more WMT related positive change in our cohort of patients with PD. Education might be a proxy for cognitive reserve,⁵⁷ which in terms of the magnification hypothesis enables individuals to engage in processes of plasticity.^{45,46} However, contradicting findings in the CT literature exist in healthy older adults, which discuss that CT might be able to diminish the cognitive reserve disadvantage of less-educated older adults.^{23,58} Furthermore, the discussion can be taken to the next level, considering a higher cognitive reserve to be commonly associated with less cognitive deficits given the same brain pathology.⁵⁹ We did not find a significant predictive value of PD motor impairment, which can be considered a proxy to the amount of PD-related brain pathology.⁶⁰ Therefore, our analyses controlled for a proxy of PD brain pathology, which is why education might have emerged as a proxy of cognitive reserve rather than brain reserve.

A trend of a higher self-efficacy expectancy being a positive predictor for training responsiveness was found for far-transfer executive functions. Self-efficacy refers to an individual's belief in his/her ability to influence and affect certain situations by his/her actions,⁶¹ and is, therefore, strongly linked to motivational processes and general training success. Especially the role of self-efficacy expectancy for transfer effects was discussed⁶²: Higher self-efficacy expectancy would lead to higher motivation to learn and to transfer. Self-efficacy expectancy has been identified as a predictor for CT responsiveness before,²⁴ strengthening the hypothesis that it might be beneficial for CT outcomes to ensure self-efficacy expectancy and motivation at the beginning of the training period.

Timeline of Predictors of Training Responsiveness

For near-transfer WM, evidence for both the compensation and magnification account was found at posttest and 3-month follow-up. For far-transfer executive functions, evidence for the magnification account was present at posttest only. Therefore, the underlying mechanisms of training responsiveness might at least partly be dependent on timing aspects.⁶³ With more time to elaborate and consolidate the WMT content beyond the training period, individual differences might become less important in predicting WMT responsiveness.⁶⁴

Strengths and Limitations

As a major strength, by applying SEM as a highly flexible statistical approach, we were able to circumvent several statistical fallacies clinical trial data often include such as violations of multivariate normality assumptions, nonlinear change trajectories, and missing data patterns, therefore applying high methodological standards to answer the question "who benefits most" from WMT.^{20,34} Furthermore, by fitting LDSM for both the WMT group and CG, we were able to detangle WMT effects from retest effects such as practice effects and regression to the mean. The discussed associations were present in the WMT group only, which further promotes the specificity of predictors for training responsiveness versus retest effects.

By applying a bootstrapping technique to estimate model parameters and fit indices of our LDSM, we made this modeling technique suitable even for small sample sizes.³⁷ However, our sample size still was fairly small to apply SEM, which constitutes a limitation concerning the reliability and validity of our results. We acknowledge that the limited statistical power increases the risk of missing important effects and of reporting false alarms. Therefore, results should be treated cautiously and will have to survive sound methodological replication. So far, only few studies applied

SEM techniques in the CT framework.^{18,45} More methodological high-quality research applying advanced statistical methods beyond simple correlational approaches is urgently needed, especially with larger samples. Hence, also individual participant data meta-analytic approaches might be one suitable format in this context.

Analyses on internal consistency (Supplementary Material 2) of the used cognitive domain composite scores revealed acceptable internal consistency for the working memory composite. For the executive function composite score, however, internal consistency was low. This might limit the interpretability of our findings, as the use of composite scores might obscure fine-grained and differential effects of the WMT on single-task performance. However, the use of composite scores reduces type I error and ensures interpretability on the cognitive domain level, which might be of more clinical importance than the single-test level. Alternatively, possible extensions of the present LDSM framework would be the simultaneous modeling of 2 or more domains of interest, allowing the investigation of cross-domain couplings, and the use of multiple indicators for latent factors instead of using equally weighted (observed) composite scores.⁶⁵ Especially the latter would reduce the influence of measurement errors and, therefore, improve validity and reliability of the analyses and should be implemented where possible. In the context of the present manuscript, post hoc attempts of implementing a multiple indicator LDSM approach failed to converge.

Another limitation is that a substantial proportion of variance in latent difference scores at posttest and 3-month follow-up remained unexplained ($0.17 \leq R^2 \leq 0.28$). Future SEM approaches might, next to demographic, neuropsychological, and clinical individual characteristics, consider individual training trajectories to predict transfer to near- and far-transfer effect measures. Moreover, the clinical significance of our results remains a matter of debate, which is not only important in the context of prognostic research for WMT responsiveness, but also regarding general effectiveness of WMT and other CT approaches.⁶⁶ Future research on effects and mechanisms of WMT and other CT approaches should more consistently include patient-centered outcomes such as daily life functioning or health-related quality of life assessments.

Further strengths of this study include the adherence to high methodological standards in general study design following the CONSORT recommendations. Additionally, the reporting of the prognostic analyses includes the relevant items of the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) Checklist,²² allowing future research to more reliably synthesize findings from single prognostic trials. Another strength is that patients and outcome assessors were blinded for candidate predictors, which might reduce a possible incorporation bias. However, only outcome assessors

were blinded for group allocation across the whole trial. Finally, by including posttest and 3-months follow-up data, a more comprehensive and encompassing view on predictors of WMT responsiveness was derived.

Conclusion

Considering the high vulnerability of patients with PD to develop cognitive dysfunctions¹ and the large heterogeneity in CT outcomes, there is an increasing urge for personalized interventions.²¹ This study contributes to the examination of individual characteristics of patients with PD associated with responsiveness to WMT, a promising nonpharmacological intervention option against cognitive dysfunction in PD. WMT might be especially beneficial for patients with PD of younger age, higher education, higher intelligence, and greater self-efficacy expectancy. However, there has to be room for improvement, as lower baseline performance is associated with a greater positive WMT responsiveness as well. Individual characteristics predicting immediate versus long-term WMT responsiveness need further research to achieve a better understanding of mechanisms leading to enduring WMT effects. In general, the present findings should be interpreted cautiously, as they need sound methodological replication with larger sample sizes or individual participant data meta-analytic approaches.

Acknowledgments

We thank all patients for participation in the study and all staff members who contributed to data collection, entry, and monitoring. Furthermore, we thank the German Parkinson Association (Deutsche Parkinson Vereinigung Bundesverband e.V.) for establishing contact to several regional Parkinson support groups and support in recruitment. We would also like to thank Schuhfried GmbH for providing free access to the computerized subtests of the Wiener Testsystem used in this study and NeuroNation (Synaptikon GmbH) for providing the working memory training tasks, administrative support, and cognitive training licenses for our participants after study completion.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: AO, KG, and SR declare that there is no conflict of interest. CE has received grants from the German Research Foundation (KFO219, TP 10), the Medical Faculty of the Philipps University Marburg, Germany, the German Ministry of Education and Research; honoraria from AbbVie, Wiesbaden, Germany; UCB, Monheim, Germany; Daiichi Sankyo, Munich; Medtronic, Meerbusch, Germany; Bayer Vital, Leverkusen, Germany; Bial, Mörfelden-Walldorf, Germany. PR received a travel grant from AbbVie. TvE reports having received consulting and speaker honoraria as well as research support from Siemens Healthcare, AVID Radiopharmaceuticals, Lilly, Shire Germany, Piramal (now Life Molecular Imaging) and GE Healthcare. EK has

received grants from the German Ministry of Education and Research, Parkinson Fonds Deutschland gGmbH, the German Parkinson Society; honoraria from: Oticon GmbH, Hamburg, Germany; Lilly Pharma GmbH, Bad Homburg, Germany; Bernafon AG, Bern, Switzerland; Desitin GmbH, Hamburg, Germany.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by department budget only and did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ORCID iDs

Kathrin Giehl  <https://orcid.org/0000-0002-0092-5164>

Carsten Eggers  <https://orcid.org/0000-0001-7564-6701>

References

1. Aarsland D, Brønnick K, Fladby T. Mild cognitive impairment in Parkinson's disease. *Curr Neurol Neurosci Rep.* 2011;11:371-378.
2. Hely MA, Reid WG, Adena MA, Halliday GM, Morris JG. The Sydney multicenter study of Parkinson's disease: the inevitability of dementia at 20 years. *Mov Disord.* 2008;23:837-844.
3. Leung IH, Walton CC, Hallock H, Lewis SJ, Valenzuela M, Lampit A. Cognitive training in Parkinson disease: a systematic review and meta-analysis. *Neurology.* 2015;85:1843-1851.
4. Nousia A, Martzoukou M, Tsouris Z, et al. The beneficial effects of computer-based cognitive training in Parkinson's disease: a systematic review. *Arch Clin Neuropsychol.* 2020; 35:434-437.
5. Chai WJ, Hamid AIA, Abdullah JM. Working memory from the psychological and neurosciences perspectives: a review. *Front Psychol.* 2018;9:401.
6. Paraskevoudi N, Balci F, Vatakis A. "Walking" through the sensory, cognitive, and temporal degradations of healthy aging. *Ann N Y Acad Sci.* 2018;1426:72-92.
7. Teixeira-Santos AC, Moreira CS, Magalhães R, et al. Reviewing working memory training gains in healthy older adults: a meta-analytic review of transfer for cognitive outcomes. *Neurosci Biobehav Rev.* 2019;103:163-177.
8. Soveri A, Antfolk J, Karlsson L, Salo B, Laine M. Working memory training revisited: a multi-level meta-analysis of n-back training studies. *Psychon Bull Rev.* 2017;24:1077-1096.
9. Melby-Lervåg M, Redick TS, Hulme C. Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": evidence from a meta-analytic review. *Perspect Psychol Sci.* 2016;11:512-534.
10. Kalbe E, Rehberg SP, Heber I, et al. Subtypes of mild cognitive impairment in patients with Parkinson's disease: evidence from the LANDSCAPE study. *J Neurol Neurosurg Psychiatry.* 2016;87:1099-1105.
11. Bäckman L, Nyberg L. Dopamine and training-related working-memory improvement. *Neurosci Biobehav Rev.* 2013;37(9 pt B):2209-2219.

12. Jordan AD, Cooke KA, Moored KD, et al. Neural correlates of working memory training: evidence for plasticity in older adults. *Neuroimage*. 2020;217:116887.
13. Eriksson J, Vogel EK, Lansner A, Bergström F, Nyberg L. Neurocognitive architecture of working memory. *Neuron*. 2015;88:33-46.
14. Trujillo JP, Gerrits NJHM, Veltman DJ, Berendse HW, van der Werf YD, van den Heuvel OA. Reduced neural connectivity but increased task-related activity during working memory in de novo Parkinson patients. *Hum Brain Mapp*. 2015;36:1554-1566.
15. Fellman D, Salmi J, Ritakallio L, Ellfolk U, Rinne JO, Laine M. Training working memory updating in Parkinson's disease: a randomised controlled trial. *Neuropsychol Rehabil*. 2020;30:673-708.
16. Ophey A, Giehl K, Rehberg S, et al. Effects of working memory training in patients with Parkinson's disease without cognitive impairment: a randomized controlled trial. *Parkinsonism Relat Disord*. 2020;72:13-22.
17. Giehl K, Ophey A, Reker P, et al. Effects of home-based working memory training on visuo-spatial working memory in Parkinson's disease: a randomized controlled trial. *J Cent Nerv Syst Dis*. 2020;12:1179573519899469.
18. Bürki CN, Ludwig C, Chicherio C, de Ribaupierre A. Individual differences in cognitive plasticity: an investigation of training curves in younger and older adults. *Psychol Res*. 2014;78:821-835.
19. Noack H, Lövdén M, Schmiedek F, Lindenberger U. Cognitive plasticity in adulthood and old age: gauging the generality of cognitive intervention effects. *Restor Neurol Neurosci*. 2009;27:435-453.
20. Hingorani AD, van der Windt DA, Riley RD, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ*. 2013;346:e5793.
21. Kalbe E, Aarsland D, Folkerts AK. Cognitive interventions in Parkinson's disease: where we want to go within 20 years. *J Parkinsons Dis*. 2018;8(suppl 1):S107-S113.
22. Moons KG, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11:e1001744.
23. Ophey A, Roheger M, Folkerts A, Skoetz N, Kalbe E. A systematic review on predictors of working memory training responsiveness in healthy older adults: methodological challenges and future directions. *Front Aging Neurosci*. 2020;12:575804.
24. West RL, Bagwell DK, Dark-Freudeman A. Self-efficacy and memory aging: the impact of a memory intervention based on self-efficacy. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*. 2008;15:302-329.
25. Paris AP, Saleta HG, de la Cruz Crespo Maraver M, et al. Blind randomized controlled study of the efficacy of cognitive training in Parkinson's disease. *Mov Disord*. 2011;26:1251-1258.
26. Zimmermann R, Gschwandtner U, Benz N, et al. Cognitive training in Parkinson disease: cognition-specific vs nonspecific computer training. *Neurology*. 2014;82:1219-1226.
27. Edwards JD, Hauser RA, O'Connor ML, Valdés EG, Zesiewicz TA, Uc EY. Randomized trial of cognitive speed of processing training in Parkinson disease. *Neurology*. 2013;81:1284-1290.
28. Nguyen HM, Aravindakshan A, Ross JM, Disbrow EA. Time course of cognitive training in Parkinson disease. *NeuroRehabilitation*. 2020;46:311-320.
29. Giehl K, Ophey A, Hammes J, et al. Working memory training increases neural efficiency in Parkinson's disease: a randomized controlled trial. *Brain Commun*. 2020;2:fcaa115.
30. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry*. 1992;55:181-184.
31. Litvan I, Goldman JG, Tröster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Mov Disord*. 2012;27:349-356.
32. Emre M. Dementia associated with Parkinson's disease. *Lancet Neurol*. 2003;2:229-237.
33. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1983;17:37-49.
34. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338:b375.
35. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1-55.
36. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res*. 2003;8:23-74.
37. Krebsbach CM. *Bootstrapping With Small Samples in Structural Equation Modeling: Goodness of Fit and Confidence Intervals* [master's thesis]. University of Rhode Island; 2014.
38. Wechsler D. *WMS-R: Wechsler Memory Scale-Revised: Manual*. Psychological Corporation; 1984.
39. Schuhfried G. *Wiener Testsystem, Vienna Reaction Unit, Basic Program. Development and Production of Scientific Equipment*. Molding, Austria; 1992.
40. Aebi C. *Validierung der neuropsychologischen Testbatterie CERAD-NP: eine Multi-Center Studie* [thesis]. Basel, Switzerland: University of Basel; 2002.
41. Bäumler G, Stroop J. *Farbe-Wort-Interferenztest nach JR Stroop (FWIT)*. Hogrefe; 1985.
42. Horn W. *Leistungsprüfsystem Manual*. Hogrefe; 1983.
43. Fahn S. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Goldstein M, Calne DB, eds. *Recent Developments in Parkinson's Disease*. Vol 2. Macmillan Healthcare Information; 1987:153-163.
44. Jerusalem M, Schwarzer R. *Skala zur allgemeinen Selbstwirksamkeitserwartung. Skalen zur Erfassung von Lehrer-und Schülermerkmalen Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Freie Universität Berlin; 1999.
45. Lövdén M, Brehmer Y, Li SC, Lindenberger U. Training-induced compensation versus magnification of individual

- differences in memory performance. *Front Hum Neurosci.* 2012;6:141.
46. Lövdén M, Bäckman L, Lindenberger U, Schaefer S, Schmiedek F. A theoretical framework for the study of adult cognitive plasticity. *Psychol Bull.* 2010;136:659-676.
 47. Matysiak O, Kroemeke A, Brzezicka A. Working memory capacity as a predictor of cognitive training efficacy in the elderly population. *Front Aging Neurosci.* 2019;11:126.
 48. Smoleń T, Jastrzebski J, Estrada E, Chuderski A. Most evidence for the compensation account of cognitive training is unreliable. *Mem Cognit.* 2018;46:1315-1330.
 49. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *Clin Neuropsychol.* 2012;26:543-570.
 50. Borella E, Carbone E, Pastore M, De Beni R, Carretti B. Working memory training for healthy older adults: the role of individual characteristics in explaining short-and long-term gains. *Front Hum Neurosci.* 2017;11:99.
 51. Zinke K, Zeintl M, Rose NS, Putzmann J, Pydde A, Kliegel M. Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Dev Psychol.* 2014;50:304-315.
 52. Li SC. Neuromodulation and developmental contextual influences on neural and cognitive plasticity across the lifespan. *Neurosci Biobehav Rev.* 2013;37(9 pt B):2201-2208.
 53. Greenwood PM, Parasuraman R. Neuronal and cognitive plasticity: a neurocognitive framework for ameliorating cognitive aging. *Front Aging Neurosci.* 2010;2:150.
 54. Borella E, Carretti B, Zanoni G, Zavagnin M, De Beni R. Working memory training in old age: an examination of transfer and maintenance effects. *Arch Clin Neuropsychol.* 2013;28:331-347.
 55. de Rijk MC, Launer L, Berger K, et al. Prevalence of Parkinson's disease in Europe: a collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology.* 2000;54(11 suppl 5):S21-S23.
 56. Hu MT, White SJ, Herlihy AH, Chaudhuri KR, Hajnal JV, Brooks DJ. A comparison of 18F-dopa PET and inversion recovery MRI in the diagnosis of Parkinson's disease. *Neurology.* 2001;56:1195-1200.
 57. Stern Y. Cognitive reserve. *Neuropsychologia.* 2009;47:2015-2028.
 58. Mondini S, Madella I, Zangrossi A, et al. Cognitive reserve in dementia: implications for cognitive training. *Front Aging Neurosci.* 2016;8:84.
 59. Hoening MC, Bischof GN, Onur ÖA, et al. Level of education mitigates the impact of tau pathology on neuronal function. *Eur J Nucl Med Mol Imaging.* 2019;46:1787-1795.
 60. Braak H, Del Tredici K, Rüb U, De Vos RAI, Steur ENHJ, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging.* 2003;24:197-211.
 61. Bandura A. Self-efficacy mechanism in human agency. *Am Psychol.* 1982;37:122.
 62. Chiaburu DS, Lindsay DR. Can do or will do? The importance of self-efficacy and instrumentality for training transfer. *Hum Resource Dev Int.* 2008;11:199-206.
 63. Roheger M, Kalbe E, Corbett A, Brooker H, Ballard C. Lower cognitive baseline scores predict cognitive training success after 6 months in healthy older adults: results of an online RCT. *Int J Geriatr Psychiatry.* 2020;35:1000-1008.
 64. Penner IK, Vogt A, Stöcklin M, Gschwind L, Opwis K, Calabrese P. Computerised working memory training in healthy adults: a comparison of two different training schedules. *Neuropsychol Rehabil.* 2012;22:716-733.
 65. Kievit RA, Brandmaier AM, Ziegler G, et al. Developmental cognitive neuroscience using latent change score models: a tutorial and applications. *Dev Cogn Neurosci.* 2018;33:99-117.
 66. Gavelin HM, Lampit A, Hallock H, Sabates J, Bahar-Fuchs A. Cognition-oriented treatments for older adults: a systematic overview of systematic reviews. *Neuropsychol Rev.* 2020;30:167-193.