

Predictors of training efficacy during n-back training

Andrea Vranić^{1*}, Marina Martinčević¹ and Vedran Prpić²

¹Department of Psychology, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia

²Ina d.d., Human resources, Croatia

Abstract: Numerous studies have investigated the efficacy of various cognitive trainings, with working memory being the most often trained cognitive aspect. In this regard, executive aspects of working memory have received the most attention, with updating training being vastly explored. In this study, we aimed to examine the differential contribution of some individual characteristics to the efficacy of updating training using a well-established n-back training paradigm. More specifically, we examined the contribution of fluid reasoning (gf), and personality (neuroticism, conscientiousness) to training efficacy. Participants ($N = 47$) took part in a 15-session, dual n-back training, spread over 4 weeks. They were pretested for fluid reasoning (CFT-3), personality (IPIP-100), and performed the initial testing on the OSPAN task. OSPAN was measured in three additional measurement points (after 5th, 10th, 15th session). The data was analyzed within the multilevel modeling approach. Initial hypotheses were partly confirmed, in that: 1) training was efficient in terms of OSPAN score, which grew linearly over time and the trajectory was similar between participants, 2) although the growth was similar for all participants, differences were found in intercepts, and 3) these differences could be partly explained by differences in fluid reasoning, but not with personality traits of conscientiousness and neuroticism.

Keywords: working memory training, working memory capacity, dual n-back, conscientiousness, neuroticism

Napovedniki učinkovitosti treninga v različnih fazah treninga n-nazaj

Andrea Vranić^{1*}, Marina Martinčević¹ in Vedran Prpić²

¹Department of Psychology, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia

²Ina d.d., Human resources, Croatia

Povzetek: Številne raziskave so preučevale učinkovitost različnih kognitivnih treningov, pri čemer je bil delovni spomin najpogosteje trenirana kognitivna funkcija. Največ pozornosti so bili deležni izvršilni vidiki delovnega spomina in pri tem je bilo še posebej temeljito raziskano posodabljanje informacij. V pričujoči raziskavi smo z uporabo dobro uveljavljene paradigme, tj. paradigme n-nazaj, želeli preučiti, kakšen je prispevek fluidne inteligentnosti oz. sklepanja (gf) in osebnosti (nevroticizma in vestnosti) k učinkovitosti treninga posodabljanja informacij. Udeleženci ($N = 47$) so v 15 seansah skozi 4 tedne izvajali trening z nalogo dvojnega n-nazaj. Pred treningom smo s testom CFT-3 izmerili njihovo sposobnost fluidnega sklepanja, s testom IPIP-100 njihovo osebnost, reševali pa so tudi nalogo obsega operativnega spomina (t. i. nalogo OSPAN). OSPAN smo izmerili še v treh dodatnih časovnih točkah (po 5., 10. in 15. seansi treninga). Podatke smo analizirali s postopki večnivojskega modeliranja. Prvotna hipoteza je bila delno potrjena: (1) trening je bil učinkovit, saj je dosežek na nalogi OSPAN s časom linearno naraščal, naraščanje pa je bilo podobno pri različnih udeležencih; (2) čeprav je bila funkcija naraščanja dosežka pri različnih udeležencih podobna, smo našli medosebne razlike v presečiščih; (3) te razlike smo lahko delno pojasnili z medosebnimi razlikami v sposobnosti fluidnega sklepanja, ne pa tudi z razlikami v osebnostnih potezah vestnosti in nevroticizma.

Ključne besede: trening delovnega spomina, obseg delovnega spomina, dvojni n-nazaj, vestnost, nevroticizem

*Naslov/Address: Andrea Vranic, Department of Psychology, Faculty of Humanities and Social Sciences, University of Zagreb, Ivana Lučića 3, 10000 Zagreb, Croatia, e-mail: avranic@ffzg.hr

Working memory (WM) is a flexible multi-component mental workspace conceptualized via complementary theories for over five decades. Converging assumption of these theories define WM as a system of multiple capacity-limited domain-specific cognitive mechanisms, regulated by an executive unit. What was once viewed as a single central executive component involved in the integration and storage of multimodal information and simultaneous supervision of control processes, is now conceptualized as a range of executive functions (EF), such as attentional control or task-switching (Logie, 2011).

EF are a set of top-down cognitive processes indispensable for all aspects of cognitive functioning, spanning from basic (cognitive control, inhibition, flexibility) to higher order processes, such as planning, organization or reasoning (Diamond, 2013). Probably the most prominent theoretical and empirical framework in this field is the tripartite model, which distinguishes three rudimentary EF: mental set shifting, information updating and monitoring, and inhibition of prepotent responses (Miyake et al., 2000). *Updating* requires monitoring and coding of incoming information relevant to the task, and appropriately replacing old, no longer relevant information held in WM with newer, more relevant information (Morris & Jones, 1990; Miyake et al., 2000). *Shifting* refers to flexibility in altering back and forth between multiple tasks, operations, or mental sets (Altmann & Gray, 2008). *Inhibition* is ability to deliberately inhibit a dominant, automatic, or prepotent response when necessary (Miyake et al., 2000). This three EF play an important role in memory and learning. Even though all three EF are considered of major importance to the functioning of WM, some authors emphasize updating to be central to WM (Ecker et al., 2010). This notion is further corroborated by the findings showing significant correlations of complex span tasks (WM tasks) and updating (Schmiedek et al., 2009; 2014).

Updating training

Given the importance of WM in everyday life, a large number of studies has addressed the question of whether WM can be enhanced through cognitive exercise. Cognitive training is based on the mounting evidence of brain plasticity, not only at an early age, but also much later in life (Pauwels et al., 2018). Intensive learning of new or non-routinized skills, typical for cognitive training, is thus an experiential setting established to promote brain plasticity. Indeed, imaging studies show altered brain activity, as well as structural grey and white matter changes, related to improvement and refinement in performance on various (cognitive) tasks, in young and old alike (Zatorre et al., 2012). Overall, it seems that WM training results in quantitative shifts in brain activity (Buschkuhl et al., 2012).

Conventional WM training consists of regular performance on a complex WM task. A common feature of such a training is the adaptive task tailored to challenge participant's level of ability. The most frequently used updating training task is the *n*-back, which requires participants to indicate whether the currently presented stimulus matches the one presented *n* trials previously. *N*-back was first

introduced by Kirchner (1958) in an investigation of age differences in short-term memory, and although studies have criticized it as a measure of WM (Kane et al., 2007), its face validity – short-term maintenance and constant *updating* of information required to perform the task – has helped it become the signature paradigm of WM updating training. A recent meta-analysis showed that apart from benefiting young adults' updating ability, *n*-back training can benefit other cognitive functions, such as fluid reasoning, and the effects can last up to 18 months (Au et al., 2015; but for older adults see Lampit et al., 2014).

In this study we have used OSPAN as a WM span task and a proxy to training efficacy. What is the relation of *n*-back to OSPAN scores? *N*-back has face validity as a WM task since it requires continuous maintaining, updating and processing of information (Gajewski et al., 2018). Dual *n*-back requires a skill to manage two *n*-back tasks simultaneously, which is a common requirement of tasks tapping *Gf*. Studies which have employed OSPAN as a measure of WM training efficacy suggest a positive interrelation of OSPAN and dual *n*-back training efficacy. For example, Matysiak and colleagues (2019) have found that among a number of variables only baseline OSPAN performance was found to be a significant predictor of the *n*-back result at the first day of training and a moderator of the whole training course.

Individual predictors of training benefits

Among potential moderators of transfer effects studies agree on age, with younger participants having larger improvements compared to older ones (e.g. Melby-Lervåg & Hulme, 2013). However, a moderating role of training dose is disputed. While no effect of dosage is found when this variable is dichotomized based on a median-split (Melby-Lervåg & Hulme, 2013; Soveri et al., 2017), authors which have used continuous moderation find that higher dose training yields larger mean effects (Schwaighofer et al., 2015). In sum, training dose varies substantially between studies and studying its effects should not be easily abandoned.

An after-analysis of studies investigating the efficacy of updating training revealed differential changes among trained participants; positively challenged participants have demonstrated large gains, while others have shown no improvement with some showing signs of regressing (Jaeggi et al., 2011). It might be that updating trainings produce such mixed results for they do not single out the factors which might catalyze its efficacy (Wiemers et al., 2019). Identifying factors which contribute to training benefits has become an important new research avenue. Our study follows this line of research in dealing with the impact of individual differences on training outcomes (e.g. Colquitt et al., 2000; Studer-Luethi et al., 2012a, 2012b).

Suggesting an interaction of aptitude and treatment, a straightforward candidate for moderation of pretest-posttest scores on various cognitive tasks are general individual differences in cognitive ability. Indeed, evidence suggests that individual differences in training-related gains can be partly explained by initial cognitive resources; magnification hypothesis suggests that individuals with good

default abilities will profit more leading to magnification of performance (Lövden et al., 2012). This so-called “rich get richer” effect speaks of initial differences in WM and *Gf* as the predictor of training-related gains and transfer effects (Redick et al. 2015; von Bastian & Oberauer, 2014). On the other hand, compensatory account assumes that people with lower initial abilities can benefit more from cognitive training because they have more room for improvement, i.e. training can aid more to individuals with lower cognitive abilities (Titz & Karbach, 2014).

Other than cognition-related factors, specific personality characteristics have been identified as a potential contributor to training success. Studies consistently show that anxiety, emotional instability, and depression are negative predictors of training success, and these personality characteristics usually relate to neuroticism (Naquin & Holton, 2002; Studer-Luethi et al., 2012a). Individuals with high neuroticism often have the least benefit from trainings. The unsubstantiated explanation offered by studies investigating the relation of neuroticism and training outcomes is in line with the attentional control theory which postulates that anxiety (e.g., worry) exerts additional cognitive load and leaves insufficient resources for further general processing (Eysenck et al., 2007; Grimley et al., 2008). Furthermore, a comparison of after-training performance between high and low demanding training conditions (single vs. dual *n*-back) shows lower gains for participants with higher neuroticism in more complex dual *n*-back condition (e.g. Studer-Luethi et al., 2012a). In simpler tasks, higher initial arousal level sustains vigilance and attention in individuals with high neuroticism, while individuals with lower arousal level seemed to have been understimulated.

Furthermore, evidence is found which suggests conscientiousness to be among the traits affecting training outcomes (Colquitt et al., 2000). Since conscientious individuals are persistent and self-disciplined, conscientiousness is usually related to positive training outcomes (e.g., Tziner et al., 2007; Studer-Lothi et al., 2012b). Since meta-analysis does not support the association between conscientiousness and skill acquisition, it is often assumed that motivation plays a major role in the relation of conscientiousness and training outcome (Colquitt et al., 2000). Still, there seems to be no straightforward interpretation of this relation, and the interaction of conscientiousness and training interventions is likely to be somewhat complex.

The aim of this study

In the current study, we investigated whether individual differences in fluid reasoning, conscientiousness and neuroticism, previously found related to cognitive outcomes (e.g., Eysenck et al., 2007), might contribute to cognitive training efficacy. More specifically, we aimed to investigate the contribution of these traits in different phases of the training program. Training dose, as already mentioned, is a potentially important moderator of training outcomes, and studying variables with a differential effect in different phases of training or skill acquisition might shed additional light on the complex interplay of training dose and its outcomes.

We employed a training procedure (same training task and format) previously validated on a sample of psychology students (Tkalčević, 2013; Tkalčević & Vranić, 2013). The efficacy was investigated on the measures of fluid reasoning (CFT-3), attention (d2 test, Brickenkamp, 1999) and spatial memory (TPP, Vranić, 2011) and it showed posttest improvements for all measures which were maintained at the 6 months follow-up, although to a somewhat lesser extent. We hypothesized that fluid reasoning ability will have a beneficiary effect on the training outcome, providing a linear, “rich-get-richer” increase throughout the training (Lövden et al., 2012). Furthermore, we hypothesized that neuroticism will have a larger (negative) impact in the initial training stages, while its effect should decrease over time. It seems that more experience with the situation and task requirements, as well as the feedback about progress – all of which is provided by repetitive exercise in adaptive training – reduces the influence which neuroticism might exert on training outcomes (e.g. Eysenck et al., 2007). Further along, we wanted to investigate whether conscientiousness will exhibit a differential effect on training efficacy.

Method

Participants

Undergraduate psychology students at the University of Zagreb participated in this study. A total of 47 students with an average age of 19 years (range 18–23 years) were included. Most participants were female (89%). All participants provided their written informed consent to participate in the study.

Training task

Dual n-back task (Jaeggi et al., 2010). An adaptive updating task (described in Jaeggi et al., 2010). The task consists of two simultaneous *n*-back tasks (spatial and verbal). In a spatial task a single blue square is sequentially presented at one of eight different screen locations. In a verbal variant, a one among eight different letters is presented through headphones. The task requires a response (key press) only when the currently presented stimulus (square location and/or letter) are the same as the ones *n* positions back in the sequence. Each session included 20 blocks consisting of 20+*n* trials (total of 20 min). Feedback is given after each block and the *n* level changes depending on participants' performance: If participant's performance was above 90% accurate, *n*-level increased for *n* = 1, and if performance was less than 75% accurate, *n*-level decreased for *n* = 1.

Tests

Operational span task (OSPAN; Unsworth et al., 2005). OSPAN is a WM task in which participants try to remember sequentially presented letters in their correct order while simultaneously solving simple math equations. A feedback is provided on the number of correctly recalled letters, as well

as the percentage of correctly solved mathematical equations. Participants had to be correct in at least 85% of the equations. The task consists of sets of 3-7 equations and letters, and a set of each length is presented three times (75 sets in total). The order of sets was randomized across participants. OSPAN result represented the sum of the letters in the accurately recalled sequence. For example, if a participant correctly recalls 2 letters in a 2-letter sequence, 3 letters in a 3-letter sequence and 3 letters in a 4-letter sequence, his result is $2 + 3 + 0 = 5$.

Cattell's Culture Fair Intelligence Test, Scale 3 (CFT-3). CFT-3 is a non-verbal test that measures fluid reasoning, excluding the influence of verbal ability, culture and educational attainment. The test consists of 4 subtests: 1) Series (13 tasks), 2) Classifications (14 tasks), 3) Matrices (13 tasks), and 4) Conditions (10 tasks). It takes approximately 15min to complete. Based on the answers to all subtests, a single total score is formed that can be directly transformed into the intelligence quotient score (IQ score). In this sample, average IQ score was 126.6 ($SD = 12.40$; range: 103–157).

International Personality Item Pool - 100 (IPIP-100, Goldberg, 1999). IPIP-100 is a cross-cultural personality questionnaire constructed according to the Big-Five model. It consists of 100 items represented by short statements and the task is to indicate, on a scale from 1 ("Very Inaccurate") to 5 ("Very Accurate"), how much each statement describes the participant himself. IPIP-100 covers five personality traits: extraversion, agreeableness, emotional stability, conscientiousness, and intellect. Each trait is measured by 20 statements. In this research, we used only items referring to emotional stability and conscientiousness. Cronbach alpha in this study showed good internal consistency (emotional stability $\alpha = 0.94$; conscientiousness $\alpha = 0.93$).

Procedure

Participants took part in the dual n -back WM training. The training consisted of 15 individual sessions of 20 minutes, extended over every other day for 4 weeks. Prior to the training participants filled in the IPIP-100, CFT-3, and have completed the OSPAN task. During the training, three additional measurement points were conducted at which the performance in OSPAN task was measured (after 5th, 10th, and the last (15th) training session). Participants were given class credits for their participation. Ethical approval was obtained from the faculty ethics committee.

Table 1

Descriptive data for OSPAN (at four measurement points), CFT-3 score, neuroticism and consciousness for ($N = 47$)

Variable	<i>M</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.
1. t_1 OSPAN	41.96	14.503							
2. t_2 OSPAN	46.17	14.344	.62**						
3. t_3 OSPAN	50.45	14.336	.47**	.46**					
4. t_4 OSPAN	55.40	13.322	.67**	.42**	.63**				
5. CFT-3	126.55	12.395	.25	.22	.21	.32*			
6. Neuroticism	56.87	13.760	-.21	-.26	-.04	-.11	-.08		
7. Consciousness	63.68	13.239	-.30*	-.27	-.07	-.16	-.09	-.14	

* $p < .05$; ** $p < .01$

Statistical analysis

We analysed the data within the multilevel modelling approach. This analysis is appropriate for hierarchically nested data in which a lower level unit of analysis is nested within a higher level of analysis.

Our data had a two-level structure: repeated measurements over time (level-1) and participants (level-2). Measurement occasion was introduced as a level-1 predictor, and participant's CFT-3 score, level of neuroticism and conscientiousness were introduced as level-2 predictors. All predictors were centred by subtracting the mean from each value of the measured variable.

As the first step in the analysis, the intercept-only model was computed (model 1). This model gives the estimate of the intraclass correlation ρ which indicates the proportion of the variance explained by the grouping structure in the population (Hox, 2010). As a second step, time variable was added as a predictor in the model. In this step we tested models with linear (model 2a) and quadratic change (model 2b) of the dependent variable over the time. As a third step, the slope of the time variable was allowed to vary across individuals (model 3). In the last step, we entered the predictor variables at the second level (model 4).

For parameter estimation, we used the maximum likelihood estimation (ML). Models were nested, therefore chi-square difference test based on the deviance statistic was used to compare models. All hypotheses were tested at the 5% alpha error rate.

Results

Descriptive statistic and the correlation matrix of dependent variables and the predictors are given in Table 1. Moderate to large positive correlations were found for OSPAN results at different time points, and moderate positive correlation was found for CFT-3 score and OSPAN score at the last time point. Other correlations were non-significant. Mean personality scores fitted the mid-range values. A peculiarity of the sample is a rather high Gf score (95th percentile).

The results of MLM analysis are shown in Table 2. The first model was used to calculate intraclass correlation (ICC) that shows the proportion of between-persons variance in the total variance. In our sample, the ICC was 0.44 indicating that there is within-person change over time (56% of total

Table 2

Change of OSPAN scores over time as predicted by CFT-3 and personality traits ($N = 47$ participants; 188 observations)

Fixed effects	Model 1	Model 2a	Model 2b	Model 3	Model 4a	Model 4b
Predictor	Coefficient (SE)					
Intercept	48.5 (1.66)**	48.5 (1.66)**	48.3 (1.88)**	48.5 (1.66)**	48.5 (1.51)**	41.8 (1.78)**
Time		4.5 (0.62)**	4.5 (0.62)**	4.5 (0.63)**	4.5 (0.62)**	4.5 (0.62)**
Time ²			0.2 (0.69)			
CFT-3					0.3 (0.12)*	0.3 (0.12)*
Neuroticism					-0.2 (0.12)	-0.2 (0.12)
Conscientiousness					-0.2 (0.11)	-0.2 (0.11)
Random effects						
σ^2_e	123.41	89.11	89.11	87.83	89.23	89.23
σ^2_{u0}	98.01	106.31	106.32	105.76	82.59	82.59
Deviance	1504.80	1460.22	1460.15	1459.97	1450.64	1450.64

* $p < .05$; ** $p < .01$

variance), but it is also worth grouping by persons to explain the variation in change between-persons (44% of total variance; Allerhand, 2016).

We tested linear and nonlinear effects separately in order to define the shape of the growth trajectory in model 2. An examination of the fixed effects of model 2a in Table 2 indicates that there is a significant linear relationship between time and OSPAN scores. Compared to model 1, this model showed better fit to data, $\chi^2(1) = 44.57$, $p < 0.01$. Model 2b with quadratic nonlinear trend showed non-significant change from Model 2a, $\chi^2(1) < 1$, $p = 0.79$, and quadratic effect of time was not a significant predictor. Therefore, in all presented models trajectories of time-related OSPAN score were modelled through linear change.

In the next step we tested whether model with random slopes (model 3) fits the data better than the model with the fixed slope. Although, a somewhat better fit was found for model 3, the results of chi-square showed non-significant

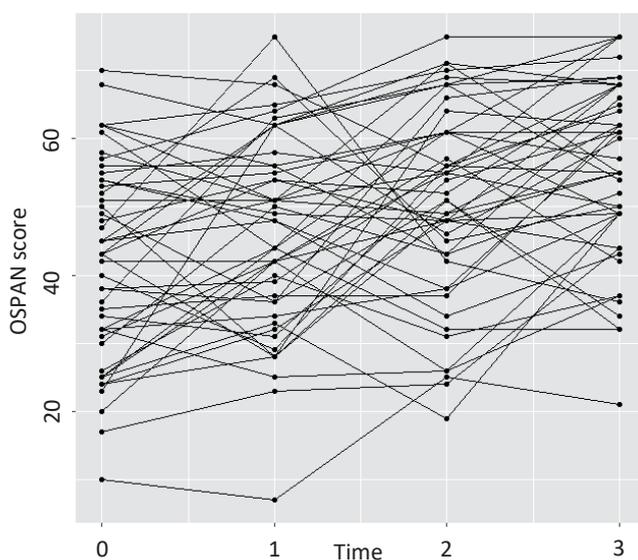
change from model 2a, $\chi^2(2) < 1$, $p = 0.88$, suggesting the same pattern of change between individuals OSPAN scores over the time (Figure 1).

In our next step we added time-invariant predictors (level 2 predictors; model 4a). In this model we tried to explain between-person variation in the dependent variable (the between-person differences in the intercept). This model showed significantly better fit compared to model 2a, $\chi^2(3) = 9.59$, $p < 0.05$, but the only significant predictor was CFT-3. Adding the level-2 predictors, we explained 22.3% of between-subject variance compared to model 2 that was used here as a baseline model. In order to enable the interpretation of intercept as the baseline measurement, in model 4b the initial testing (1st measurement point) was set as zero. This model predicts the initial OSPAN score of 41.8, which increases by 4.5 points at each succeeding measurement point. With each scale point higher on the CFT-3 test, OSPAN score increased by 0.3.

The results of this analysis suggest that OSPAN score grew linearly over time and the trajectory is similar between participants. Although the growth is similar for all participants, there are differences in intercepts, which might be partly explained by the differences in fluid reasoning, but not with the personality traits of conscientiousness and neuroticism.

Figure 1

Individual changes in OSPAN scores during training at four measurement points ($N = 47$)



Discussion

This study was set with the aim to further elucidate the differential contribution of some individual characteristics to the benefits of cognitive training. More specifically, we opted to investigate the contribution of fluid reasoning, conscientiousness and neuroticism to cognitive training efficacy in the different phases of the 15-session WM updating training. Based on the literature, we hypothesized that fluid reasoning will have a linear beneficiary effect in different phases of training, in line with the rich-get-richer effect (Lövdén et al, 2012), and that neuroticism will have a larger impact in the initial training stages. Due to ambiguity of findings on relation of conscientiousness and training outcomes (e.g., Eysenck et al., 2007), no directional prediction was set in this case.

Our data was modelled using a multilevel approach, with repeated measurements over time (4 time points; level 1) and participants (level 2).

We used the OSPAN task (Unsworth et al., 2005), a complex WM span task, as a criterion of training efficacy in each of the three phases, following the initial testing. Our results partly support the set hypothesis. We found that OSPAN scores show a comparable linear growth across different training phases, and that the differences found in intercepts can be partly explained by initial differences in fluid reasoning. Our hypothesis regarding the differential contribution of conscientiousness and neuroticism to the training efficacy were not confirmed. Although not as overwhelming in terms of expectations met, our findings cast some interesting light on the widely discussed relation of fluid reasoning, *n*-back task and span tasks and WM ability. We will address this topic in later sections of the paper, and we will first comment on the results regarding personality traits.

Personality as a predictor of training outcomes

Insofar, only a handful of studies have investigated the relative contribution of non-cognitive factors, such as personality, to training effectiveness and outcomes, although there is an upward trend in the number of such studies (e.g. Colquitt et al., 2000; Jaeggi et al., 2014; Studer-Luethi et al., 2012a, 2012b). The importance of these studies is that they could enable the design of individualized programs, better suited for each individual, which in turn could boost the efficacy of such personalized procedures. Our expectation of neuroticism to be relevant in determining the (in)efficacy of the training was based on the findings of anxiety and depressive symptoms being negatively related to training efficacy (Backman et al., 1996; Studer-Luethi et al., 2012a). Emotional instability and anxious behavior are often used to describe individuals with high neuroticism scores. Almost contrary to this, individuals high on conscientiousness, which are described as reliable, persistent and diligent, could be expected to profit from the training, although the motivation might play a critical role in determining the direction of this relation.

Studer-Luethi and colleagues (2012a; 2012b) have found a detrimental effect of *neuroticism* in dual *n*-back training, such as the one employed in this study. The same study, however, found that participants with high neuroticism exhibited higher gains in fluid reasoning measures after the single *n*-back training. These results have offered a speculative account of the relation of neuroticism and training efficacy, in line with the processing efficiency theory (Eysenck & Calvo, 1992). It is suggested that anxiety – and neuroticism, by the same token – can be a useful feature in solving simple tasks due to the increased activation that helps sustain vigilance. But when the task is more demanding, the same feature imposes cognitive load on further processing, and in turn reduces processing capacity within WM. Anxiety is often present in novel situations, which can provoke lower self-efficacy and less efficient resource mobilization, while repetitive exercise within adaptive training provides

opportunity to get acquainted with the task setting which consequently weakens the impact of neuroticism. However, in our study neuroticism did not have a differential effect on dual *n*-back training efficacy.

Conscientious individuals are usually motivated to improve and excel in their skills and performance (Komarraju & Karau, 2005), which leads to the assumption that they will be equally eager and hard-working in the training conditions. Previous studies have partly confirmed this notion, with participants high in conscientiousness having greater training success in the single *n*-back compared to dual *n*-back condition (Studer-Luethi et al., 2012). Authors argue that these individuals might have preferred lower task complexity because it provides higher chances for excellent performance. Following on that, dual *n*-back condition, such as the one in our study, could have reduced their resources via increased self-attention, and the tendency of conscientious individuals to be self-deceptive (Martocchio & Judge, 1997; Studer-Luethi et al., 2012b). However, results of our study which regard complex dual *n*-back condition did not reveal a significant contribution of conscientiousness to the training outcomes, neither in positive, nor in negative manner.

Although relatively small, our sample was twice the size of the dual *n*-back group in the study of Studer-Luethi and colleagues (2012a), and it could be considered as more robust. However, given a body of literature on the differentially impaired cognitive processing in neuroticism, much research is needed to give a firm conclusion on the interplay of neuroticism and cognitive training outcomes. By the same token, the relation of conscientiousness and training outcomes needs a more systematic research approach, with varying levels of task complexity.

Fluid reasoning as a predictor of training outcomes

We now come to the discussion of the findings on the interrelation of cognitive variables within this study: *Gf*, *n*-back and OSPAN. Fluid reasoning can briefly be defined as the ability of successful reasoning with novel problems. It vastly relies on working memory capacity (WMC; e.g., Kyllonen & Christal, 1990). WMC accounts for a significant portion of the variance of general cognitive ability. WMC tasks – the so-called span tasks, such as OSPAN – have enjoyed strong support for being stable, and with a minimal contribution of error (e.g. Conway et al., 2005). Performance on WM span tasks correlates with a number of higher order cognitive skills, such as language comprehension (Daneman & Carpenter, 1983), reasoning (e.g. Barrouillet, 1996), and complex-task learning (Kyllonen & Stephens, 1990).

Although evidence suggests that *n*-back and WM span are weakly correlated (Gajewski et al., 2018), in our study *n*-back training has resulted in gains on OSPAN task. Moreover, OSPAN scores at different phases of training were found related to initial scores in fluid reasoning (*Gf*). Insofar, these findings provide support for the notion that: 1) WM shares considerable variance with *Gf*, and 2) even though weakly correlated as expected, *n*-back and OSPAN might share some processing components that, in turn, could enable transfer

between these tasks. In light of the recent critiques of *n*-back as a measure of WM, our results provide support for some common ground of *n*-back and WMC. Thus, future studies should not discard *n*-back as a WM measure, rather focus on the specific context in which WM updating might prove beneficial for a more general cognitive functioning. Also, keeping in mind the relatively high initial *Gf* score of our participants, further support for the interrelation of *n*-back and span tasks is needed from the samples with lower mean scores.

Previous studies have found contrary findings about the relationship between initial abilities and training gains which can be explained by the two opposite effects: “rich get richer” account (Lövdén et al., 2012) and compensatory effect (Karbach and Spengler, 2012). The findings of our study do not confirm the existence of any of those effects. Although fluid reasoning is associated with WMC, it does not explain changes in the training gain. It is possible that other individual characteristics, such as training motivation, have a more important role. Incorporating these variables within the design of future studies might shed light on the training gain differences.

Implications and limitations

There are limitations to this study, which need to be addressed. Most limitations result from the convenience sampling employed. The first and utmost limitation comes from a small sample size, thus, some of the insignificances or lack of effects could be attributed to a lack of statistical power. Another problem is the uneven gender distribution; thus, gender effects could not be controlled. Third, specific personality (i.e. high/low conscientiousness, high/low neuroticism) was not an “inclusion” factor; thus, the effects might have been stronger had we included participants with the personality profile in the upper and lower part of the trait spectrum. Finally, although previous studies have shown effectiveness of this training procedure in enhancing WM in young adults (Tkalčević, 2013; Tkalčević and Vranić, 2013), adding a control group could separate the effect of the intervention from multiple testing effects.

Furthermore, more often than not computerized cognitive training (CCT) in older adults suggests there is a maximal dose for CCT, after which other factors (such as fatigue) come into play (Lampit et al., 2014). Comparative studies in younger adults, as in this study, link training schedules distributed over time with a greater training efficacy (Penner et al., 2012). The dosage in this study and its predecessors was decided upon to closely mirror the evidenced-based efficient training conducted by Jaeggi and colleagues (2011), our procedure was mirrored to our previous control study (Tkalčević and Vranić, 2013). However, it remains likely that our results might change given a different dosage or training schedule.

Although with some limitations, this study holds a strong point in the multilevel modeling approach (Kwok et al., 2008). Considering the studied constructs, this type of analysis has enabled the calculation of cross-level interaction effects and explanation of between-subject variance in growth curves by subject-level predictor (fluid reasoning). Also, MLM treats regression parameters from

all the individual growth models (intercepts) as random effects for estimations thus reducing type I error in statistical inference.

Future research should further consider possible moderators, which requires adequate sample sizes to allow the adequate detection of effects. Knowledge of such moderators should ultimately allow the design of interventions aimed at particular cognitive skills on an individual level, and with the ultimate goal to increase the efficacy of the intervention (Buschkuhl et al., 2012).

References

- Allerhand, M. (2016). *Longitudinal data analysis*. Centre for Cognitive Ageing and Cognitive Epidemiology. http://www.ccace.ed.ac.uk/sites/default/files/longitudinal-2up_1.pdf
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, *115*, 602–639.
- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review*, *22*, 366–377.
- Backman, L., Hill R.D. & Forsell, Y. (1996). The influence of depressive symptomatology on episodic memory functioning among clinically nondepressed older adults. *Journal of Abnormal Psychology*, *105*, 97–105.
- Barrouillet, P. (1996). Transitive inferences from set-inclusion relations and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1408–1422.
- Brickenkamp, R. (1999). *Test opterećenja pažnje: Priručnik za test d2* [Test of attention: Manual]. Naklada Slap.
- Buschkuhl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience*, *2*, S167–S179.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, *85*, 678–707.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769–786.
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 561–584.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168.
- Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 170–189.
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, *6*, 409–434.

- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336–353.
- Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the n-back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology, 9*, Article 2208.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, (Vol. 7, pp. 7–28). Tilburg University Press.
- Grimley, M., Dahraei, H., & Riding, R. J. (2008). The relationship between anxiety-stability, working memory and cognitive style. *Educational studies, 34*, 213–223.
- Hox, J. J. (2010). *Quantitative methodology series: Multilevel analysis: Techniques and applications (2nd ed.)*. Routledge/Taylor & Francis Group.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 108*, 10081–10086.
- Jaeggi, S. M., Buschkuhl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition, 42*, 464–480.
- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence, 38*, 625–635.
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 615–622.
- Karbach, J. & Spengler, M. (2012). Who benefits the most? Individual differences in the transfer of executive control training in younger and older adults. In A. Bröder et al. (Eds.), *Abstracts of the 54th Meeting for Experimental Psychologists (TEAP)*, S. 64. Pabst.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*, 352–358.
- Komarraju, M., & Karau, S. J. (2005). The relationship between the big five personality traits and academic motivation. *Personality and Individual Differences, 39*, 557–567.
- Kwok, O. M., Underhill, A. T., Berry, J. W., Luo, W., Elliott, T. R., & Yoon, M. (2008). Analyzing longitudinal data with multilevel models: An example with individuals living with lower extremity intra-articular fractures. *Rehabilitation Psychology, 53*, 370–386.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*, 389–433.
- Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2*, 129–160.
- Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: A systematic review and meta-analysis of effect modifiers. *PLoS Medicine, 11*(11), e1001756.
- Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science, 20*, 240–245.
- Lövdén, M., Brehmer, Y., Li, S. C., & Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Frontiers in Human Neuroscience, 6*, Article 141.
- Martocchio, J. J., & Judge, T. A. (1997). Relationship between conscientiousness and learning in employee training: Mediating influences of self-deception and self-efficacy. *Journal of Applied Psychology, 82*, 764–773.
- Matysiak, O., Kroemeke, A., & Brzezicka, A. (2019). Working memory capacity as a predictor of cognitive training efficacy in the elderly population. *Frontiers in Aging Neuroscience, 11*, 126.
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270–291.
- 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. *Cognitive Psychology, 41*, 49–100.
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111–121.
- Naquin, S. S., & Holton, E. F. III. (2002). The effects of personality, affectivity, and work commitment on motivation to improve work through learning. *Human Resource Development Quarterly, 13*, 357–376.
- Pauwels, L., Chalavi, S., & Swinnen, S. P. (2018). Aging and brain plasticity. *Aging, 10*, 1789–1790.
- Penner, I-K, Vogt, A., Stöcklin, M, Gschwing, L, Opwis, K., & Calabrese, P. (2012). Computerised working memory training in healthy adults: A comparison of two different training schedules. *Neuropsychological Rehabilitation, 22*, 716–33.
- Redick, T. S., Shipstead, Z., Wiemers, E. A., Melby-Lervåg, M., & Hulme, C. (2015). What’s working in working memory training? An educational perspective. *Educational Psychology Review, 27*, 617–633.
- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2009). On the relation of mean reaction time and intraindividual reaction time variability. *Psychology and Aging, 24*, 841–857.
- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). Younger adults show long-term effects of cognitive training on broad cognitive abilities over 2 years. *Developmental Psychology, 50*, 2304–2310.
- Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*, 138–166.
- Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*, 1077–1096.

- Studer-Luethi, B., Bauer, C. E., & Perrig, W. J. (2012a). Neuroticism affects working memory and training performance in regularly developed school children. *International Journal for Cross-Disciplinary Subjects in Education*, 3, 640–647.
- Studer-Luethi, B., Jaeggi, S. M., Buschkuhl, M., & Perrig, W. J. (2012b). Influence of neuroticism and conscientiousness on working memory training outcome. *Personality and Individual Differences*, 53, 44–49.
- Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, 78, 852–868.
- Tkalčević, B. (2013). *Working memory training: Testing the effects of transfer and maintenance* [Master's thesis]. Faculty of Humanities and Social Sciences.
- Tkalčević, B., & Vranić, A. (2013). Working memory training: Transfer and maintenance effects. In G. Kuterovac Jagodić, I. Erceg Jugović, & A. Huić (Eds.), *21. days of Ramiro and Zoran Bujas: Book of abstracts* (pp. 203–203). Faculty of Humanities and Social Sciences, University of Zagreb.
- Tziner, A., Fisher, M., Senior, T., & Weisberg, J. (2007). Effects of trainee characteristics on training effectiveness. *International Journal of Selection and Assessment*, 15, 167–174.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498–505.
- von Bastian, C. C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research*, 78, 803–820.
- Vranić, A. (2011). *Test prostornog slijeda*. Centar za psihodijagnostičke instrumente.
- Wiemers, E.A., Redick, T.S., & Morrison, A.B. (2019). The influence of individual differences in cognitive ability on working memory training gains. *Journal of Cognitive Enhancement*, 3, 174–185.
- Zatorre, R. J., Douglas Fields, R., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience*, 15, 528–536.