Influence of neuroticism and conscientiousness on working memory training outcome

Barbara Studer-Luethi a,b,*, Susanne M. Jaeggi c, Martin Buschkuehl c,d, Walter J. Perrig a,b

a Department of Psychology, University of Bern, Switzerland
b Research Center for Cognition, Learning, and Memory, University of Bern, Switzerland
c Department of Psychology, University of Maryland, USA
d Department of Psychology, University of Michigan, USA

1. Introduction

Cognitive training has become increasingly popular in research. Growing evidence suggests that training on working memory (WM) can lead to increased performance in non-trained tasks (for reviews, see e.g. Buschkuehl & Jaeggi, 2010; Morrison & Chein, 2011). For example, we trained college students on either an adaptive single or dual n-back task, and as a result, both groups improved performance in non-trained matrix reasoning tasks to a comparable extent (Jaeggi et al., 2010). However, we have repeatedly observed that some participants are positively challenged and demonstrate large training gains whereas others feel overwhelmed and hardly improve or even regress (Jaeggi, Buschkuehl, Jonides, & Shah, 2011). Therefore, considering the role of individual differences seems to be crucial when evaluating the efficacy of training. By doing so, the current study answers numerous calls for more work examining the impact of individual differences on training (Colquitt, LePine, & Noe, 2000; Martocchio & Judge, 1997; Mount & Barrick, 1998). For example, there is evidence suggesting that the personality traits conscientiousness and neuroticism affect training outcomes (Colquitt et al., 2000). Furthermore, the finding that individual differences affects how people react to more or less complex tasks (e.g. Walsh, Wilding, & Eysenck, 1994) suggests that the relationship between personality and training outcome might depend on the complexity of the training task. Thus, in the current study, we investigate whether conscientiousness and neuroticism might determine cognitive training performance and transfer to non-trained tasks by using two training tasks that differed in the degree of complexity.

1.1. Personality traits and training outcomes

There are only few studies available that investigated the influence of personality characteristics on cognitive training and the results revealed that anxiety and depressive symptoms seem to be consistent negative predictors of training outcomes (Backman, Hill, & Rosell, 1996; Yesavage & Jacobs, 1984). These two personality characteristics are related to a personality factor commonly referred to as neuroticism. Subjects high in neuroticism are described as anxious and emotionally unstable and usually obtain the least benefit from trainings (e.g. Naquin & Holton, 2002; Yesavage, 1989; for a meta-analysis, see Judge & Ilies, 2002). In contrast, conscientiousness seems to be the personality trait with the most positive influence on training performance (e.g. Barrick, Stewart, & Piotrowski, 2002; Tziner, Fisher, Senior, & Weisberg, 2007). Since subjects with high levels of conscientiousness are described as persistent, hardworking
and self-disciplined, most assumptions emphasize the role of motivation (Colquitt et al., 2000). However, the relationship with training proficiency is complex as conscientiousness is not positively associated with either declarative knowledge or skill acquisition during training (Dean, Conte, & Blankenhorn, 2006; for a meta-analysis, see Colquitt et al., 2000).

Thus, it seems that there is a relationship between personality traits and training outcomes; however, the underlying mechanisms of this association are not clear. Apart from motivational factors, previous research suggests that psychophysiological correlates and task demands might drive some of these interactions, as discussed in the following paragraphs.

1.2. Neuroticism, conscientiousness and task demand

The processing efficiency theory, later redefined as the attentional control theory, explains the detrimental influence of trait anxiety, which is closely related to neuroticism, with intrusive thoughts and worry. These intrusive thoughts interfere with cognitive performance by detracting from the resources available to control attention (Eysenck, Derakshan, Santos, & Calvo, 2007). This interference results in poor processing and limited storage recruitment of prefrontal attentional control mechanisms of subjects scoring high in neuroticism when performing a WM task (Bishop, 2009; Gray et al., 2005).

Another approach was proposed by Eysenck (1967), who suggested a physiological basis for neuroticism, namely its positive association with autonomic arousal. This so-called activation level and its effect on performance can be described by an inverted U-shaped curve: A moderate level of activation facilitates the best performance, whereas under- or overactivation impairs performance (Yerkes & Dodson, 1908). In addition, as task difficulty increases, the optimal level of activation decreases (Eysenck, 1967). Based on these models, numerous studies demonstrated that neuroticism and related traits are negatively related to performance in demanding and stressful tasks, but positively related to performance in relatively simple and monotonous tasks (e.g. Corr, 2003; Oswald, Hambrick, & Jones, 2007; Poposki, Oswald, & Chen, 2009).

Regarding conscientiousness, there is some evidence that conscientious individuals overestimate the importance of their performance and show heightened levels of evaluation apprehension, which make them perceive a challenging task all the more difficult (Thompson, Duxbury, & Behrend, 2008). They furthermore tend to be self-deceptive, which in turn decreases learning (Martocchio & Judge, 1997).

1.3. The current study

Here, we investigate whether individual differences in neuroticism and conscientiousness determine cognitive training outcomes and whether these individual differences interact with the complexity of the training task. We used a single n-back task, consisting of a single stream of stimuli, and a dual n-back task that combines two streams of stimuli that have to be processed independently. As such, the latter produces a substantial amount of interference, since both task components compete for common resources and mechanisms in WM. As in other multitasking situations, it is likely that participants experience failure and higher stress while doing a dual n-back task compared to performing a single n-back task.

Based on the research reviewed above, we predict the following outcomes:

1. Since the WM training tasks largely rely on attentional control (e.g. interference), we assume that overall training outcomes are adversely affected by neuroticism. Regarding the training benefit, we expected an interaction between neuroticism and training task: In the context of the complex dual training task, high levels of neuroticism disrupt the training process which results in less transfer and training enjoyment. However, if the processing demands of the training task are low enough, such as in the single n-back task, participants with high levels of neuroticism might show better training and transfer performance because of their higher basic activation level, which may support sustained attention.

2. With regard to conscientiousness, we predict a positive association with training performance and enjoyment. However, since prior research indicated that the relationship between conscientiousness and training proficiency is not straightforward, we avoid making directional predictions about the association with transfer.

Although we are well aware that other personality traits might also influence training and transfer, for this study, we only focus on neuroticism and conscientiousness. Please note that the same data set used here was used in a previous publication that focused on transfer to reasoning performance not taking into account personality traits (Jaeggi et al., 2010).

2. Methods

2.1. Participants

One hundred and twelve native Chinese speaking undergraduates from the National Taiwan Normal University volunteered to take part in the study (85 females; mean age = 19.5 years, SD = 1.5). Forty-seven participants (36 females) were assigned to a four-week WM training, whereas 43 students were assigned to a control group (34 females; mean age = 19.4 years, SD = 1.0). Nine participants of the control group who did not complete the post tests and 13 subjects who only completed the personality assessment were included in baseline analyses in order to increase statistical power.

In return for participation, course credits were offered, and participants of the training group received an additional NT$ 600 (~US$ 20). Based on demographic variables and pre-test performance (age, gender and baseline reasoning performance), the training group was divided into two comparable groups using the software ‘Match’ (Van Casteren & Davis, 2007). One of these groups was assigned to the single n-back training (n = 21, 17 females; mean age = 19.1 years, SD = 1.5), while the other group was assigned to the dual n-back training. One subject from the dual training group dropped out after a few training sessions, leading to a final dual n-back group of 25 participants (18 females; mean age = 19.1 years, SD = 1.2). Hence, we included 112 subjects for reliability analyses, 99 subjects for analyses of baseline performance, and 89 subjects for analyses of training outcomes.

2.2. Material

2.2.1. Training tasks

2.2.1.1. Dual n-back task. We used an adaptive dual n-back task as previously described (Jaeggi et al., 2010). The task consisted of a sequential presentation of single blue squares at one of eight different locations on the computer screen (stimulus length: 500 ms; interstimulus interval: 2500 ms). At the same time, a series of eight letters was presented through headphones. The task required responding by pressing a key only when the current
stimulus (square location and/or letter) was the same as the one \( n \) positions back in the sequence. Each training session included 15 blocks consisting of 20 + \( n \) trials. After each block, performance feedback was provided and the level of \( n \) was changed depending on the participants’ performance: If accuracy was 90% or more, the level of \( n \) was increased by one, and it decreased by one if accuracy was 70% or lower. Mean training performance was determined as the average \( n \)-back level for all training sessions combined. Performance gain in training was determined as the mean \( n \)-back level in the two last training sessions minus the level in the two first training sessions.

2.2.1.2. Single \( n \)-back task. As a second intervention, we used a modified version of the dual \( n \)-back task which only required processing of visuospatial stimuli, i.e. there was no auditory stimulus stream. Everything else (stimulus rates, number of blocks, adaptivity) was the same as in the dual \( n \)-back version.

2.2.2. Transfer tasks

2.2.2.1. \( n \)-Back. Two \( n \)-back tasks versions were used to assess near transfer (cf. Jaeggi et al., 2010). These non-adaptive versions consisted of three \( n \)-back levels, namely 2-, 3- and 4-back (3 blocks each), presented in this order. We used a dual \( n \)-back task with the same stimulus material as in the training task, and a single \( n \)-back task with eight random shapes as stimuli, which were centrally presented in a yellow color on a black background. The stimulus rate was the same as in the training tasks. We assessed near transfer performance as mean accuracy (\( P_c \); proportion hits minus false alarms) over all \( n \)-back levels in both, the single and the dual \( n \)-back task.

2.2.2.2. Fluid intelligence (Gf). We used two standard tests to assess matrix reasoning, the short version of the Bochumer Matrizentest (BOMAT; Hossiep, Turck, & Hasella, 1999), and Raven’s Advanced Progressive Matrices (RAPM, set II; Raven, 1990), with a time limit of 16 and 11 min, respectively (cf. Jaeggi et al., 2010). Parallel versions of both tests were given with counterbalanced order. The dependent measure (Gf score) was the number of correct solutions given within the time limit. We calculated standardized baseline Gf scores as the average of standardized scores in both matrix reasoning tests (pre-test score divided by the SD). To assess far transfer, we calculated transfer using standardized gain scores (i.e., the average standardized gain of both Gf tests, that is, Gf post – Gf pre, divided by the common SD of the pretest; cf. Jaeggi et al., 2011).

2.2.3. Questionnaires

2.2.3.1. Personality inventory. Personality was assessed with the Mandarin version of the Mini-Marker Set (Saucier, 1994). The inventory has been shown to be psychometrically reliable (\( z \) coefficients range from .75 to .83; Saucier, 1994). It is based on the Five-Factor Model which proved its generalizability in many different cultures, including Taiwan (Wu, Lindsted, Tsai, & Lee, 2008). Participants were asked to indicate their level of agreement with each of 40 single adjective personality descriptors (e.g. moody, shy) on a 5-point Likert scale. The five factor structure of the original questionnaire was replicated in the presented sample (\( z \) coefficients range from .70 to .86; see Table 1), however, for the present study, we only analyzed neuroticism and conscientiousness.

2.2.3.2. Post-training evaluation. In order to obtain a qualitative feedback concerning the training experience, participants were asked to indicate their training enjoyment by answering the question “How much did you enjoy the training?” on a 5-point Likert scale.

### Table 1

<table>
<thead>
<tr>
<th>Baseline performance</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluid intelligence</td>
<td>.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single ( n )-back</td>
<td>.43*</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual ( n )-back</td>
<td>.37*</td>
<td>.58*</td>
<td>.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personality scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>-.05</td>
<td>-.07</td>
<td>-.01</td>
<td>.80</td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-.02</td>
<td>-.06</td>
<td>-.03</td>
<td>-.30*</td>
<td>.86</td>
</tr>
</tbody>
</table>

Note: \( N = 99 \); values on the diagonal represent test–retest reliabilities of the control group and Cronbach’s alpha for the five personality traits; Fluid intelligence = mean of standardized scores in the two matrix reasoning tasks.

* \( p < .05 \).
** \( p < .01 \).

2.3. Procedure and design

One week before the intervention, all participants completed the personality inventory. Three days before the beginning of training as well as 3 days after training completion, participants were tested on the measures of \( n \)-back and matrix reasoning. Between pre- and post-testing, all participants of the experimental groups completed 20 training sessions within a period of 4 weeks. Training took place in small groups of 10–15 participants and one experimenter, 5 days a week for approximately 20 min per training session.

3. Data analyses

For data analyses, we used SPSS 18. All statistic tests were based on a significance level of \( \alpha = .05 \). First, we analyzed zero-order correlations in the whole group and in the three subgroups. However, correlation analyses in different subgroups cannot necessarily be attributed to actual modulating effects, but rather to differences in variance of the quantitative variable (e.g. Whisman & McClelland, 2005). Therefore, to investigate prediction strength of the different factors on training outcome, we ran hierarchical analyses with the effect weighted intervention variable (single vs. dual \( n \)-back training vs. no-contact control) in the first step, the centered personality scores in the second step and the intervention \( \times \) personality interaction in the third step (cf. Cohen, Cohen, West, & Aiken, 2003).

4. Results

4.1. Personality and performance in cognitive tasks

Correlations among the cognitive ability tasks and the personality variables are displayed in Table 1, along with their reliabilities. The overall baseline \( n \)-back accuracy was related to Gf baseline performance, but none of the cognitive measures was reliably correlated with either neuroticism or conscientiousness.

4.2. Personality and training outcome

Correlations between personality variables and training outcomes are presented in Table 3.

4.2.1. Training scores

There was a modest negative correlation of neuroticism and overall training mean score, but not with training gain (see Table 3). In contrast, conscientiousness was positively related to training mean and gain scores, which were mainly driven by the single \( n \)-back intervention (see Fig. 1A and Table 2). The regression analysis revealed a significant prediction of training mean score by...
personality traits, but the intervention \times neuroticism interaction significantly predicts improvement in Gf (see Table 3). For far transfer performance, these interaction effects are visualized in Fig. 2. As can be seen, interaction effects were carried by the fact that a combination of low neuroticism and dual n-back training as well as high neuroticism and single n-back training yielded the highest gain on Gf.

4.2.4. Post-training feedback

Neuroticism was negatively correlated with overall training enjoyment, which was mainly driven by the dual n-back training group. Concerning conscientiousness, there was a positive correlation to overall training enjoyment, particularly in the single training group. Regression analysis only revealed a trend for an intervention \times neuroticism interaction ($\beta = .27$, $t(44) = 1.56$, $p = .06$).

### Table 2

Correlations between personality traits and training, transfer and post-training feedback presented for all training participants, as well as separately for the dual and single n-back training group.

<table>
<thead>
<tr>
<th>Personality trait</th>
<th>Neuroticism</th>
<th>Conscientiousness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training group</td>
<td>All</td>
<td>Single</td>
</tr>
<tr>
<td>Mean score</td>
<td>-.24</td>
<td>-.14</td>
</tr>
<tr>
<td>Gain n-back</td>
<td>-.11</td>
<td>.40</td>
</tr>
<tr>
<td>Feedback</td>
<td>-.28</td>
<td>-.09</td>
</tr>
</tbody>
</table>

Note: $N = 46$. $p < .05$.

### Table 3

Regression model including the intervention (single vs. dual), the personality scores as well as the intervention \times personality interaction as predictors of near transfer (n-back tasks) and far transfer (matrix reasoning).

<table>
<thead>
<tr>
<th>Transfer</th>
<th>$R^2$ change</th>
<th>B</th>
<th>SE B</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near transfer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>Intervention</td>
<td>-.08</td>
<td>.02</td>
<td>-.48**</td>
</tr>
<tr>
<td>Step 2</td>
<td>Neuroticism</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>(.C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td>$I \times N$</td>
<td>.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I \times C$</td>
<td>.01</td>
<td></td>
<td></td>
<td>.36**</td>
</tr>
<tr>
<td>Far transfer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>Intervention</td>
<td>.01</td>
<td></td>
<td>-.01</td>
</tr>
<tr>
<td>Step 2</td>
<td>Neuroticism</td>
<td>.07</td>
<td></td>
<td>-.02</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>(.C)</td>
<td></td>
<td></td>
<td>-.28</td>
</tr>
<tr>
<td>Step 3</td>
<td>$I \times N$</td>
<td>.11</td>
<td></td>
<td>-.04</td>
</tr>
<tr>
<td>$I \times C$</td>
<td>.04</td>
<td></td>
<td></td>
<td>.40*</td>
</tr>
</tbody>
</table>

Note: $N = 90$. $^* p < .05$. $^{**} p < .01$.
5. Discussion

The primary goal of the present study was to investigate whether individual differences in neuroticism and conscientiousness determine performance in cognitive training and transfer, and whether personality interacts with two types of interventions, i.e. a single and a dual n-back training, to predict these training outcomes. Our results demonstrate that this is indeed the case.

It is important to note that we observed no differences in baseline performances in matrix reasoning or the n-back task as a function of personality. Therefore, none of the described effects of personality on training and transfer can be attributed to initial performance differences.

Our results largely support our first hypothesis: High neuroticism was associated with lower training scores overall, however, the group difference did not reach significance. This is in line with the statement that neuroticism does not necessarily have a negative impact on effectiveness (quality of performance) but rather on efficiency (performance effectiveness divided by effort; Eysenck & Calvo, 1992). However, we observed an interaction of neuroticism and training task for near and far transfer performance. Further analysis within the dual n-back group revealed that participants with high scores on neuroticism showed lower gains in near and far transfer tasks and they also reported lower training enjoyment than participants with low neuroticism levels. In contrast, in the single n-back group, it was the participants with high neuroticism who showed more gains in far transfer. In line with the activation model in connection to task difficulty, our results indicate that the high demand of the dual n-back training task led subjects with high levels of neuroticism in a suboptimal activation state which derailed complex cognitive transfer processes.

In line with the processing efficiency theory, a speculative interpretation of this result suggests that the cognitive load imposed by neuroticism (e.g. worry) reduced the processing capacity of the WM system. Subjects high in neuroticism then increased their on-task effort as a compensatory strategy in order to reach a good training level. However, in the challenging dual n-back condition, the increased activity of emotional and cognitive processes did not leave sufficient resources for transfer processes to higher cognitive abilities, which resulted in lower training proficiency. In contrast, the higher baseline activation of participants high in neuroticism may have allowed them to sustain vigilant attention in the less demanding single n-back condition which then led to higher transfer performance compared to their understimulated counterparts.

Our results also supported the second hypothesis in that participants with high scores on conscientiousness showed better overall training performance (Fig. 1A), which was also accompanied with higher training enjoyment. Because conscientious participants tend to be competitive and motivated to excel and improve their skills (Komarraju & Karau, 2005), they might have been highly committed to the training and might have diligently pursued the goal of possible cognitive profit. However, the success of conscientious people in the training as well as training enjoyment seem to be more strongly tied to the single n-back task, suggesting that they might have preferred the lower complexity level of this task which goes along with a higher possibility to excel in the task. Concerning transfer effects, high conscientiousness was found to predict larger near transfer, which can be related to their higher training performance in the single training. However, despite these positive training results, conscientiousness predicted smaller far transfer effects in both training conditions (Fig. 1B). This rather counterintuitive finding is in line with prior findings which revealed that some attributes of conscientious subjects derail training proficiency (Martocchio & Judge, 1997). On one side, the conscientious participants may have developed effective but highly task-specific strategies, which brought them success in the training and the near transfer tasks, but did not foster far transfer process. On the other side, the increased self-attention of conscientious subjects, namely their higher evaluation apprehension, the overestimation of the importance of their performance as well as their tendency to be self-deceptive, may have detracted mental resources, which would have been necessary for an efficient transfer process to higher cognitive abilities.

5.1. Implications and limitations

There are some limitations that have to be considered. First of all, our sample size was rather small, thus, some of the null-effects might have resulted due to a lack of power. Further, gender was not evenly distributed in the present sample and it was therefore impossible to control potential gender effects. Additionally, the Mini-Marker Set is a self-report questionnaires and therefore prone to response bias. Personality was not systematically included as a factor, thus, the effects might have been stronger if only participants with personality profiles in the upper or lower end of the
spectrum would have been included. Finally, regarding the generalization of the results, cultural differences need to be considered, even though there is no reason to believe that the described underlying processes would be different in a Western culture (cf. McCrae, Terracciano, et al., 2005).

Nevertheless, the current data contribute to the growing body of literature linking personality traits with cognitive performance. We provide further evidence that neuroticism and conscientiousness interact with training outcomes. To conclude, our results have important implications in that they can serve as guidelines for the selection of an optimal training regimen based on individual personality profiles. Based on the level in neuroticism, conclusions regarding subjects’ available mental resources can be made and the appropriate training task can be chosen. Our results furthermore extend the findings that conscientiousness is one of the best predictors for high training performance, which, however, can be detrimental for transfer.

References


