Predictors of Near Transfer Gains in Working Memory Training

Reidar Nervander, 1801697
Master's thesis in Psychology
Supervisors: Matti Laine, Daniel Fellman & Liisa Ritakallio
Faculty of Arts, Psychology and Theology
Åbo Akademi University
2020

ÅBO AKADEMI UNIVERSITY – FACULTY OF ARTS, PSYCHOLOGY AND THEOLOGY

Subject: Psychology		
Author: Reidar Nervander		
Title: Predictors of Near Transfer Gains in Working Memory Training		
Supervisor: Matti Laine	**Supervisor:** Daniel Fellman	**Supervisor:** Liisa Ritakallio

Abstract:

Background: After more than a decade of studies regarding working memory (WM) training programs, the evidence for their effectiveness is limited. The group-level transfer gains observed are mostly restricted to task-specific near transfer, that is, improvements in untrained pre-post tasks structurally similar to the trained one. At the same time, large inter-individual differences in improvement levels have been observed. Understanding the mechanisms behind this phenomenon is essential for the future of WM training programs, as this possibly allows for more effective, individually tailored interventions, which could make the training program benefits more prominent.
Methods: Eight individual characteristics (age, education, gender, WM baseline performance level, strategy use, motivation, alertness and cognition-related beliefs) were analyzed for their predictive role in task-specific near transfer gains using multiple regression analyses. The data was obtained from a study by Fellman et al. (2020), where 113 young adults underwent a 12-session long WM training program with a digit n-back task over the course of four weeks.
Results: Out of the eight variables, only WM baseline performance level was found to significantly predict task-specific near transfer gains, so that a lower baseline level in the beginning of training resulted in larger transfer gains.
Conclusion: The present results support the existence of a compensation effect, where cognitively low-performing individuals show larger transfer gains, presumably because they have more room to improve. In practice, this would indicate that WM training programs are possibly most beneficial for those who have a weaker WM performance level to start with.

Keywords: working memory training; task-specific near transfer; individual differences; compensation effect

Date: 19.05.2020	Page count: 41

ÅBO AKADEMI – FAKULTETEN FÖR HUMANIORA, PSYKOLOGI OCH TEOLOGI

<table>
<tr><td colspan="3">Ämne: Psykologi</td></tr>
<tr><td colspan="3">Författare: Reidar Nervander</td></tr>
<tr><td colspan="3">Avhandlingens titel: Prediktorvariabler för nära transfer i arbetsminnesträning</td></tr>
<tr><td>Handledare: Matti Laine</td><td>Handledare: Daniel Fellman</td><td>Handledare: Liisa Ritakallio</td></tr>
</table>

Abstrakt:

Bakgrund: Bevis från över ett årtionde av studier gällande arbetsminnesträningsprogram har visat att programmens effektivitet är begränsad. De förbättringar som konstaterats på gruppnivå innefattar till största del uppgiftsspecifik nära transfer, vilket inkluderar förbättringar i otränade uppgifter som strukturellt är likadana som den tränade uppgiften. Samtidigt har man noterat stora interindividuella skillnader i mängden förbättring. En förståelse för mekanismerna bakom detta fenomen är centralt eftersom det kan bidra till utvecklandet av skräddarsydda och mera effektiva träningsprogram.
Metoder: Åtta individuella karaktärsdrag (ålder, utbildning, kön, grundnivå på arbetsminnesprestation, strategianvändning, motivation, vakenhetsnivå och kognitionsrelaterade åsikter) analyserades som eventuella prediktorvariabler för uppgiftsspecifik nära transfer-förbättring med hjälp av multipla regressionsanalyser. Datat som analyserades erhölls från en studie av Fellman et al. (2020), där 113 unga vuxna genomgick ett arbetsminnesträningsprogram på 12 sessioner under loppet av fyra veckor. Träningsuppgiften var en sifferversion av n-back testet.
Resultat: Av alla åtta prediktorvariabler så var det endast grundnivån på arbetsminnesprestation som signifikant förutspådde den mängd uppgiftspecifik nära transfer deltagarna uppvisade. En lägre grundnivå på arbetsminnesprestation förutspådde mera transfer.
Slutsats: Resultaten stöder förekomsten av en kompensationseffekt, där kognitivt svagpresterande individer uppvisar mera transfer, eventuellt eftersom de har mera utrymme för förbättring. I praktiken innebär detta att arbetsminnesträningsprogram möjligen är mest gynnsamma för de individer som från början har en svagare arbetsminnesförmåga.

Nyckelord: arbetsminnesträning; uppgiftsspecifik nära transfer; individuella skillnader; kompensationseffekt

Datum: 19.05.2020	Sidantal: 41

Table of contents

ACKNOWLEDGEMENTS

First and foremost, thank you to Daniel Fellman and colleagues for allowing me to analyze your data. The whole project with all its sophisticated and intellectual details was a privilege to take part of. A massive thank you goes out to my supervisors Matti Laine and Daniel Fellman. Even though you have had a lot on your own plates during this time of crisis, both of you always took the time to help me throughout this process. A big thanks to Liisa Ritakallio and Matti Laine, who helped me receive a financial grant from Mensa Finland for the publication rights of this thesis. An additional thank you to Liisa, who always so eagerly offered her assistance no matter what. A final acknowledgement goes out to all fellow students and friends who in any way helped me finish this project. Thank you.

Turku, April 2020,

Reidar Nervander

1. Introduction

Working memory (WM) is commonly described as one's ability to temporarily store, as well as manipulate information, and it is considered to play a significant role in many everyday cognitive activities (Baddeley, 1992; Miyake & Shah, 1999). It has also been linked to abilities such as reading comprehension (Daneman & Carpenter, 1980), language processing (Baddeley, 2003) and reasoning (Baddeley, 1992). Ever since Jaeggi et al. (2008) reported that fluid intelligence could be improved through WM training, numerous studies have been published with the aim of exploring possible benefits of WM training regimes on cognitive abilities (Karbach & Verhaeghen, 2014; Melby-Lervåg et al., 2016; Soveri, et al., 2017). However, recent meta-analyses show that *far transfer*[1] from the trained cognitive domain to another domain is either weak (Au et al., 2014; Karbach & Verhaeghen, 2014) or non-existent (Melby-Lervåg et al., 2016; Soveri et al., 2017). Redick (2019) labels this process in the research field as the *hype cycle* of WM training, where the initial hype from some studies showing positive results has been replaced by recently published meta-analyses, holding almost no evidence for any kind of improvements on cognitive domains different from the one being trained. What seem to be present, however, are moderate to strong improvements on the trained task. Also, performance enhancement on tasks supposedly measuring the same cognitive domain has been found, referred to as *near transfer*[2]. In other words, the potential benefits of WM training seem not to consist of far transfer to other cognitive domains, but of improvements on the trained task and its close variants (Fellman et al., 2020; Holmes et al., 2019; Laine et al., 2018; Soveri et al., 2017). These limited transfer effects and their underpinnings are at focus in the present study.

Even though moderate to large improvements on near transfer tasks have been seen at group level, individual differences on WM training outcomes are considerable (Karbach & Verhaeghen, 2014; Melby-Lervåg et al., 2016; Soveri et al., 2017). Despite this, there is a surprisingly big gap in the literature regarding factors related to these individual differences. Most studies have focused on group-level changes and only a few WM studies have, to a varying degree, considered the large inter-individual differences seen in near transfer gains (Anguera et al., 2012; Borella et al., 2017; Bürki et al., 2014; Foster et al., 2017; Hunt et al., 2014; Salminen et al., 2015; Studer-Luethi et al., 2012; Zinke et al., 2014). Although methodological flaws in the research setups

---

[1] *Far transfer* refers to improvements seen in another cognitive domain, other than the trained task, e.g., WM training would increase performance in tasks measuring intelligence (Soveri et al., 2017).

[2] *Near transfer* includes improved performance in tasks intended to measure the same cognitive domain as the trained task, e.g., by being structurally different or having the same task paradigm as the trained task but differing in stimuli (Fellman e al., 2020).

and/or publication bias may in part account for the discrepancies regarding the effectiveness of WM training (Melby-Lervåg et al., 2016), individual characteristics might help explain another part of the variance between study results (Guye et al., 2017). It is important to understand the limitations of these training programs in order to provide realistic and useful suggestions of their use (Redick, 2019).

In the present study, the aim was to analyze the predictive value of an extensive set of predictor variables (age, gender, education, WM baseline performance level, strategy use, motivation, alertness, and cognition-related beliefs) for near transfer gains in WM training. Understanding the potential effects of individual differences on training outcomes could help to better understand the large inter-individual variability in WM training.

1.1 Transfer gains in WM training

The concept of transfer in learning is not new, as research on the matter dates back a century. Transfer refers to a learned skill being applied and adapted to novel situations, and it is essential in all education and training programs. However, a century of studies has shown that significant far transfer effects are hard to find when measuring any type of cognitive or educational capacities. The strongest transfer effects have been found on near transfer tasks, where the participants' familiarity with relevant task context as well as their underlying cognitive capacity predict success (Barnett & Ceci, 2002). Yet, the initial research regarding WM training was mainly done in the spirit of the *Capacity Theory*, which presumes that intensive, adaptive training at individual performance limits can expand WM capacity (Engle & Kane, 2004). The assumption was that such repeated training elicits long-term plasticity in WM-related brain regions, which consequently can produce significant improvement on the numerous cognitive functions that rely on WM-related brain regions (Constantinidis & Klingberg, 2016). However, as noted above, it seems to be the case that WM training can only produce small to large effect-sizes on near transfer measures (Melby-Lervåg, 2016; Soveri et al., 2017). This is in accordance with Barnett and Ceci's (2002) conclusion from a century of studies on the matter of transfer.

An alternative framework to account for WM training improvements has been provided by the *Strategy Mediation hypothesis* (Fellman et al., 2020; Laine et al., 2018; Peng & Fuchs, 2015; Soveri et al., 2017), that considers WM to be a rather limited and fixed cognitive capacity. According to this hypothesis, training can improve WM performance (but not capacity) by eliciting a more efficient spontaneous use of strategies during repeated exposure to the trained tasks.

The level of performance on WM tasks is seen as a result of how the limited WM capacity is being used. Efficient use of strategies can therefore enhance performance by freeing up cognitive capacity, which can be used to boost performance even further (e.g., Peng & Fuchs, 2015). Some researchers have recently argued that the Strategy Mediation hypothesis is a more suitable explanation for the limited results from WM training (Gathercole et al., 2019, Fellman et al., 2020; Laine et al., 2018; Soveri et al., 2017). They argue that strategy use is one important reason why participants improve on the trained tasks and their untrained variants, but not so much on other WM task paradigms or far transfer tasks. The reason would be that strategies adopted for the trained task or tasks can be used on structurally similar task paradigms but not necessarily on structurally dissimilar tasks.

Based on recent studies, the concept of near transfer in WM training appears to be too broad and should additionally be divided into task-general and task-specific near transfer (e.g., Fellman et al., 2020; Laine et al., 2018; Soveri et al., 2017). *Task-specific near transfer* entails transfer to WM tasks sharing the same task paradigm with the trained task(s), while *task-general near transfer* pertains to WM tasks that are structurally dissimilar to the trained task (Fellman et al., 2020). Soveri et al. (2017) criticized previous meta-analyses for not making this distinction and hence, overestimating the transfer effects within the WM domain. In their own meta-analysis, the only more substantial near transfer gain they found was task-specific near transfer. They noted that individuals who trained with n-back tasks (a standard measure of WM, which requires participants to maintain and update series of stimuli in order to decide whether each new item matches the one presented n items ago), showed a relatively strong transfer effect ($g = 0,62$) for untrained versions of the n-back, and only a small effect ($g = 0,24$) for other WM transfer tasks. These results allowed Soveri et al. (2017) to conclude that more emphasis should be put on task-specific aspects of transfer when trying to understand the mechanisms of WM training. The next section will consider previous findings regarding predictors of near transfer, while separating between task-specific and task-general near transfer whenever possible.

1.2 The role of individual characteristics

Due to the inconsistent findings and great inter-individual variability regarding transfer gains from WM training, several researchers are beginning to turn their focus on the role of individual characteristics on training outcomes (e.g., Borella et al., 2017; Zinke et al., 2014). The literature regarding the topic is still very limited, even though researchers like Jaeggi et al. (2012) pointed out the importance of investigating potential predictors of WM training programs relatively early on. Some researchers have touched the subject of individual predictors of near transfer a few years earlier, though only with narrow sets of characteristics (e.g., Bürki et al., 2014; Zinke et al., 2014). Guye et al. (2017) were amongst the first to analyze the impact of a broad range of individual difference variables on WM training outcomes, albeit they focused only on training task gains and not on transfer effects. Some researchers have, however, found positive correlations between training and transfer gains (Chein & Morrison 2010; Zinke et al., 2014). If the Strategy Mediation hypothesis is correct (as suggested by e.g. Laine et al., 2018), a rapidly increasing individual learning curve on the trained task(s) could indicate that a new strategy is being learned and applied (Chein & Schneider, 2012). Consequently, this could predict the emergence of task-specific near transfer as well (Chein & Morrison 2010; Zinke et al., 2014). Note however that not everyone agrees with this conclusion. Researchers like Redick (2019), as well as Tidwell et al. (2014) highlight the need to regard the correlation between training and transfer gains with caution, as not all studies have been able to replicate the existence of such links (Au et al., 2014).

In the next section, previous studies addressing the individual predictors targeted in the current study (age, gender, education, strategy use, WM baseline performance level, motivation, alertness and cognition-related beliefs) are briefly reviewed. Studies focusing on predictors of specifically near transfer are summarized in Table 1.

1.3 Individual predictors

WM baseline performance level at pre-test seems to moderate near transfer gains, as seen in the studies summarized in Table 1. However, inconsistent results have been obtained regarding the direction of this. Some researchers have observed a *magnification effect* (Foster et al., 2017), some have provided evidence for a *compensation effect* (Hunt et al., 2014) and others have observed both (Borella et al., 2017). A magnification effect indicates that individuals with a higher WM baseline performance level at the beginning of training show greater improvements, assumedly because they have more cognitive resources and can therefore learn new cognitive routines more easily. A compensation effect, on the other hand, indicates that low-performing individuals benefit more from training, possibly because they have more room for improvement than high-performing individuals, who are already at the top of their performance (e.g., Borella et al., 2017). The only meta-analysis on this matter by Melby-Lervåg et al. (2016), found that individuals with a condition

Table 1

Characteristics found to predict either task-general or task specific near transfer in individual studies

Predictors	**Participants** Older (60+) Younger (18-59), Children (<18)	**Sample size** (n)	**Level of analysis** (group vs. individual)	**Type of near transfer**	**Predictive power**
WM baseline level					
Hunt et al. (2014)	Children (ADHD)	62	Individual level	Task-specific	Found (-)
Borella et al. (2017)	Older	148	Group-level	Task-general	Found (+ & -)
Foster et al. (2017)	Younger	116	Group-level	Task-general	Found (+)
Age					
Borella et al. (2017)	Older	148	Group-level	Task-general	Found (-)
Zinke et al. (2014)	Older	80	Individual level	Task-general	Found (-)
Dahlin et al. (2008)	Younger & older	64	Group-level	Task-general	Found (-)
Bürki et al. (2014)	Younger & older	128	Group-level	Task-specific	Not found
Li et al. (2008)	Younger & older	87	Group-level	Task-specific	Not found
Hunt et al. (2014)	Children (ADHD)	62	Individual level	Task-specific	Not found
Salminen et al. (2015)	Younger & older	95	Group-level	Task-general	Not found
Personality					
Studer-Luethi et al. (2012)	Younger	85	Individual level	Task-specific	Found
Motivation					
Anguera et al. (2012)	Younger	28	Group-level	Task-specific	Not found
Gender					
Hunt et al. (2014)	Children (ADHD)	62	Individual level	Task-specific	Not found

Level of analysis refers to whether the individual differences were analyzed by 1. splitting the individuals by a dichotomous variable, often by a median-split design of high and low, and comparing the two subgroups using e.g., ANOVA (group-level analysis) or 2. employing the predictors as continuous individual difference variables in a regression analysis (Individual level analysis).
Plus (+) or minus (-) refers to the moderating effect of the significant predictor. Plus indicates that the individuals higher in that predictor (e.g., higher age) improved more and minus that individuals lower in that predictor (e.g., lower WM baseline level) improved more.

associated with impaired WM showed less near transfer than healthy adults. However, no significant difference was found in near transfer gains between high- vs. low-performing healthy individuals.

Age as a predictor has received inconsistent support, as demonstrated in Table 1. Additionally, only one meta-analysis (Melby-Lervåg, 2016) found age to predict near transfer so that children improved less than younger and older adults. Among adults, no differences were found based on age, a result supported by two other meta-analyses that included age as a moderator (Karbach & Verhaeghen, 2014; Soveri et al., 2017).

The connection between WM baseline performance level and age is a factor that makes it harder to determine how these two variables predict near transfer gains. This is because some normal (non-pathological) cognitive decline occurs in adult aging. The domains most affected are processing speed, reasoning, memory and executive functions (Deary et al., 2009). Gajewski et al. (2018) found that these cognitive functions are important for the overall functionality of WM, and accordingly, that also WM performance decreases with age. This has led to the standpoint that these two predictors are interlinked (e.g., Melby-Lervåg et al., 2016). Foster et al. (2017) noted that individuals with a low WM capacity were also generally older than those having a higher WM capacity. Borella et al. (2017) found support for this association, as significant gains on tasks tapping task-general near transfer were moderated by both age and WM baseline level. On more "active" WM tasks (resembling the n-back task that is of particular interest in the present study), participants with a higher initial performance and younger age gained more from the training regime, in line with the magnification effect. On more "passive" WM tasks like the Forward Digit Span task (recalling series of digits in the same order they were presented), however, participants with a lower WM baseline and an older age benefited more from training, thus indicating a compensation effect. Taken together, the predictors varied in their moderating effect on task-general near transfer depending on the task at hand. Zinke et al. (2014) also found that the older the individuals were, the less task-general near transfer gains and training task gains they showed in their sample of 80 older adults (65 - 95 years of age). However, Zinke et al. (2014) found a lower WM baseline performance level to result in larger training task gains, regardless of age.

Personality features and their relationships with WM training and transfer outcomes have also been explored in a few studies. Studer-Luethi et al. (2012) found that neuroticism and conscientiousness (as measured by the Mini-Marker Set [Saucier, 1994], which is based on the Five-Factor Model) predicted task-specific near transfer in their sample of 85 undergraduates. Individuals high in conscientiousness or low in neuroticism showed higher task-specific near transfer. Those high in neuroticism also reported less training enjoyment, suggesting that high neuroticism (e.g., worrying) reduces the cognitive capacity available and leads to less near transfer (Studer-Luethi et al., 2012). Urbánek and Marček (2015), on the other hand, did not find any significant effects of personality traits (as measured by The Personality Styles and Disorders Inventory [Kuhl & Kazén, 1997]) on far transfer when analyzing data compiled from three different cognitive training groups (two versions of the n-back task and one mental rotation task). This was supported by Guye et al. (2017), who also failed to find any significant associations between training task gains and personality traits of the Big Five.

Motivation is another variable that has been speculated to moderate training outcomes (Jaeggi et al., 2013), though several studies have shown that this might not be the case (Anguera et al. 2012; Guye et al., 2017; Jaeggi et al., 2013). Anguera et al. (2012) found their participants showing equally large task-specific near transfer during post-test, despite some participants losing some of their self-assessed motivation and engagement along the way. Guye et al. (2017) suggested that motivation accounted for daily fluctuations in training performance, but not for the overall training task gains. Jaeggi et al. (2013) found evidence that motivational factors predict whether individuals completed the training program, rather than altering the learning curve or the degree of transfer. They also found that intrinsic motivation seems to be more beneficial for WM training programs, whereas extrinsic rewards, such as monetary compensation, can reduce intrinsic motivation (Deci et al., 1999) and consequently, reduce training and transfer gains. According to Jaeggi et al. (2013), most researchers extrinsically compensate participants for participating in WM training programs, which might result in participants lacking the motivation to commit to training. This, they argue, could account for some studies lacking significant transfer effects.

Cognition-related beliefs (e.g., whether an individual consider it possible to improve intelligence) have been argued to affect WM training results, though no study so far has focused specifically on their effect on near transfer. Guye et al. (2017) found that subjects with beliefs in the malleability of intelligence showed less improvement in the trained task following WM training. Jaeggi et al. (2013), on the other hand, noted that individuals who believed intelligence to be malleable, showed greater far transfer gains than those who believed intelligence to be fixed. Jaeggi et al. (2013) speculated that this improvement could be modulated by a placebo effect. In a small study with 25 participants by Foroughi et al. (2016), a similar relationship was found. The experiment group was intentionally induced with a placebo-effect through a suggestive advertisement. They were then compared to a control group, recruited by a nonsuggestive advertisement. Only the placebo group showed significant improvement on fluid intelligence tasks after one hour of training with a dual n-back task, supporting the notion of a strong placebo effect in WM training. The results highlighted the importance of active controls. Soveri et al. (2017), on the other hand, found no differences in the degree of transfer in studies using an active vs. passive control group, making it hard to interpret why cognition-related beliefs would significantly predict transfer gains following WM training.

Strategy use is another, more recently highlighted predictor that can moderate the effect of near transfer. As previously mentioned, the Strategy Mediation hypothesis postulates that the more advanced the used strategies are, the better the outcome of WM training regimes both on the trained task and on tasks tapping task-specific transfer (Fellman et al., 2020; Gathercole et al., 2019; Laine et al., 2018). No previous study has yet examined whether strategy use already prior to intervention predicts subsequent near transfer gains. The aim of the studies by Fellman et al. (2020) and Laine et al. (2018) was instead to explicitly analyze how strategies are being developed during training and how they then moderate WM training outcomes. According to Gajewski et al. (2018), younger adults rely more heavily on general executive functions when performing the n-back task, while older adults tend to rely more on attention, verbal memory and updating in order to compensate for age-related cognitive decline. Younger individuals might have a more suitable cognitive capacity to adopt sophisticated strategies and have more resources to spare after the task becomes automated (Chein, & Schneider, 2012). This could explain why young adults seem to improve faster on a trained task than older ones, as they have been found to adopt effective strategies more quickly (Bürki et al., 2014). Older adults might simply take longer to adopt these strategies, as they can rely less on general executive functions (Gajewski et al., 2018).

Some other variables have also been hypothesized to predict transfer and/or training task gains, but the evidence is lacking. For instance, **gender** has not been found to moderate any training task gains (Guye et al., 2017; Matysiak et al., 2019) or near transfer (Hunt et al., 2014). Variables such as **leisure activities**, **computer literacy** and **previous cognitive training** have not been found to predict training task gains (Guye et al., 2017), and the **need for cognition** (the amount of enjoyment from difficult cognitive tasks) was not connected to the level of far transfer shown (Jaeggi et al., 2014).

In summary, the limited literature indicates that the most consistent predictor of transfer effects is WM baseline performance level. For the other individual characteristics, except age, there is only limited support for them being reliable predictors of near transfer, with evidence usually stemming just from a single study. One should also note that there are large discrepancies in both the methodology and the results between these studies. In some studies, researchers have focused only on a single predictor like age (e.g., Salminen et al., 2015) or WM baseline (Foster et al., 2017), ignoring the fact that these predictors can overlap (Melby-Lervåg et al., 2016). In other studies, the samples have been quite small (e.g., Anguera et al., 2012; Hunt et al., 2014), making it hard to determine whether the results are reliable or not (Karbach & Verhaeghen, 2014). In some cases, task-specific and task-general near transfer effects are intertwined, blurring the findings (Soveri et al., 2017). A distinction should also be made between studies basing their conclusions on group-level comparisons or on regression models where the predictors are employed as continuous variables. The first ones are often conducted as post hoc comparisons where a continuous variable is split into a dichotomous variable, often by a median-split. This leads to a loss of power and information, as well as a reduction of effect sizes, partly because of regression towards the mean (Moreau et al., 2016). Analyzing predictive variables at an individual level in regression models allows for more sensitive data analysis, with less statistical distortion in the results (Guye et al., 2017). More research, especially focusing on multiple predictors on an individual level, is needed before any firm conclusions can be drawn on predictors of near transfer.

2. Research questions

Based on previous literature, the aim of the present study was to examine whether individual differences predict transfer gains on untrained tasks that are structurally similar to the trained task (task-specific near transfer) following WM training. The individual characteristics analyzed were age, education, gender, WM baseline performance level, strategy use, motivation, alertness and cognition-related beliefs. Due to scarcity of relevant previous studies and their discrepant results, the present study was exploratory, and no hypotheses were put forth prior to analysis. The overall purpose was to add to the large knowledge gap regarding individual predictors of near transfer in WM training, shedding light on the heavily debated question as to why near transfer gains vary so much between individuals. This study followed the guidelines by Soveri et al. (2017), who highlighted the importance of separating training task gains and transfer gains, as well as task-general and task-specific near transfer gains. Here the focus was narrowed down to task-specific near transfer, as this type of transfer has recently been shown to be the only more substantial form of transfer from WM training (Soveri et al., 2017).

3. Method

3.1 Procedure

The data used in the present study stemmed from a pre-registered fully online-administered randomized controlled trial by Fellman et al. (2020) who tested the Strategy Mediation hypothesis in WM training in detail. The original study was approved by the Institutional Review Board of the Department of Psychology and Logopedics, Åbo Akademi University in accordance with the Helsinki Declaration. Prior to the data collection, the study protocol was pre-registered at AsPredicted.org (https://aspredicted.org/r7qs9.pdf). All testing and training sessions were performed on an in-house developed customizable Java-based experiment platform called SOILE. This platform enables creation, distribution and management of psychological experiments and cognitive interventions via the Internet. In the Fellman et al. (2020) study, a total of 419 participants, varying in age from 18-50 years, were recruited through the crowdworking site Prolific Academic (https://www.prolific.co/). The participants remained anonymous during the whole study, as they were only identified using a Prolific ID (a string of random characters assigned to them by Prolific). Before starting the study on the SOILE platform, they had to agree to an informed consent. They were informed about the study setup, as well as about how the forthcoming monetary compensation would be provided. They were also informed about their right to withdraw from the study if/whenever they wanted. As the Prolific Academic participant pool was not screened for all the exclusion criteria, potential participants completed a few minutes' pre-screening survey using the SOILE platform. The inclusion criteria were as follows: English native speaker, no serious psychiatric or neurological illnesses, no current use of CNS medication and no current psychotropic drug use. Out of 419 participants initially invited, 9 were excluded for not meeting the inclusion criteria, 78 were excluded for not completing the pre-test, 30 withdrew after completing only the pre-test, 25 withdrew during training, 13 were excluded for using external memory aids during training and 6 for being multivariate outliers. All in all, 161 participants were excluded for either not meeting the inclusion criteria or not completing the study, leaving a total of 258 participants.

The participants completed a pre-test in week 1, including background questionnaires and a test battery consisting of 10 WM tasks. The pre-test took approximately 2 hours (reimbursement £10). After this, the participants were randomly divided into three groups: a strategy training group (*n* = 73), a traditional training group (*n* = 118) and a passive control group (*n* = 67). Prior to randomization, the participants were told that they would be assigned to either a training or a control group. The ratio of group selection was 2:1:1 to provide a 50 % chance of being allocated in the traditional training group, and 25 % chance to join the other two groups, respectively. The participants in the traditional training group had a typical WM training regime where they practiced with an adaptive n-back task without any strategy guidance. The strategy training group trained with the same task but was provided with external strategy instructions aimed to boost the use of an effective visualization strategy. The passive control group did not receive any training and participated only in the pre-test, the intermediate test and the post-test. In addition to being compensated 10£ for the pre-test, the monetary compensation for the rest of the study was 50£ for the participants in the training group and 20£ for the ones in the control group.

Figure 1

The study setup in Fellman et al.'s (2020) study

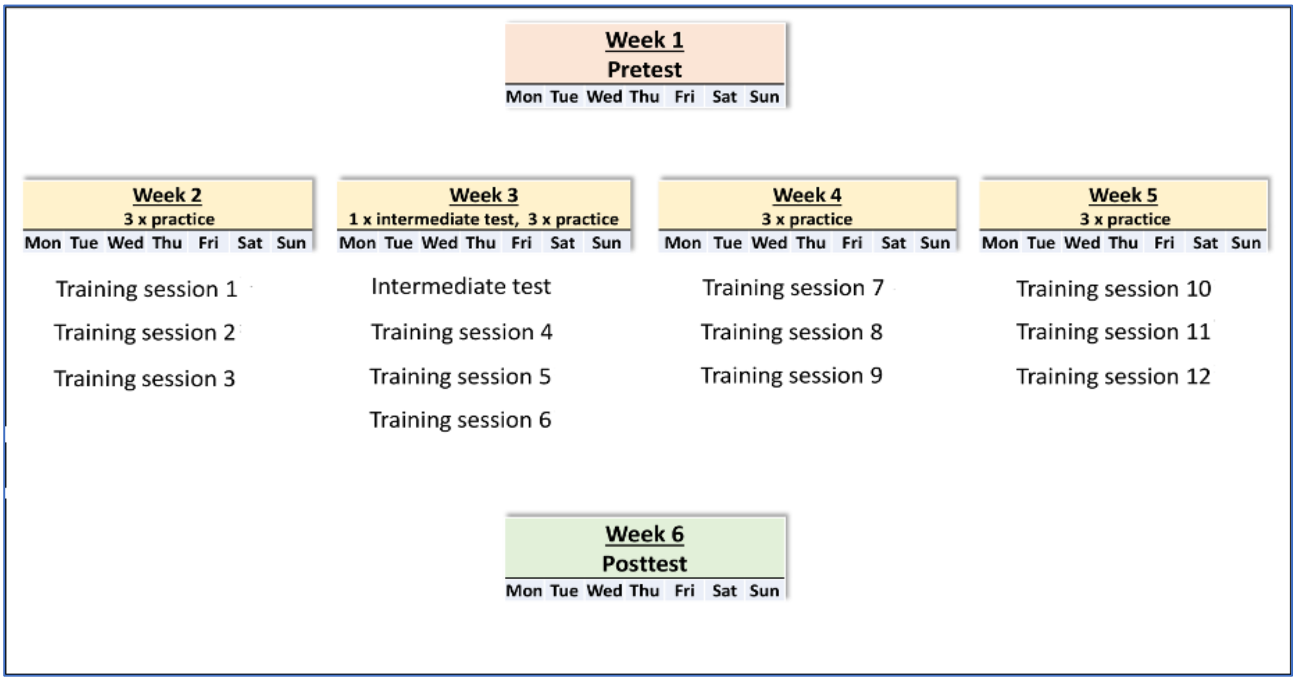
Out of the three groups, the intervention in the traditional training group (training without given strategy) was similar to a considerable portion of previous WM training studies (Soveri et al., 2017) and hence, chosen as the target group in the present study. This allowed for a better comparison with previous studies on predictor variables. An important prerequisite for the present study was that the group-level analyses in Fellman et al. (2020) revealed statistically significant task-specific near transfer effects on all three untrained variants of the n-back task in both training groups. This allowed for further investigation of potential individual predictors for transfer. Prior to the current analysis, an additional five participants in the traditional training group were removed due to missing values caused by technical errors during data gathering. The missing values appeared random and no imputation was conducted.

3.2 The traditional training group

After the pre-test, the participants in the traditional training group underwent a four-week training program, including twelve 30-minute WM training sessions. The intermediate test was conducted during week 3 following three initial training sessions, and the post-test during week 6 when the participants had completed the whole training program. As compared to the pre-test, both the intermediate and the post-test took approximately 2 hours to complete and included the same cognitive tasks. The timetable of the 4-week long WM training program and the assessment tests is presented in Figure 2.

Figure 2

Timetable for the assessment tests and the WM training program in the traditional training group



3.2.1 Training task

The sole training task in all of the 12 training sessions was an adaptive n-back task with digits as stimuli. In this task, a sequence of digits ranging from 1 to 9 was presented, one digit at a time. The participants were instructed to respond via computer keyboard (one button for yes and one for no) whether or not the current digit was the same as the one presented n items back. Each training session consisted of 20 blocks, with each block containing 20 + n trials. Out of the 20 trials in a block, six were targets and 14 non-targets. Four of the non-targets were so-called lures, that is, they were identical to the target items but in adjacent n ± 1 positions. The lures were included to discourage decisions based on mere familiarity of the items. Each sequence in a block was structured in the same way. First a blank screen was presented for 450 ms, then a stimulus for 1500 ms, followed by another blank screen for 450 ms, and then the next stimulus appeared. This pattern continued until the end of the block. The level of the n-back was automatically determined by the participants' success rate, rendering the task easier or more difficult (i.e., adaptive) depending on how the participant performed. The level of the n-back varied between 1 and 15, starting with a 1-back task in the first training session. If the participants answered correctly on 18-20 trials in a block, they advanced one level. If one succeeded with 15-17 trials, the n-back level stayed the same and anything less than correct 15 trials resulted in a decrease of n by one. The starting level of each new training session followed the n – 2 principle, i.e., two levels below the highest n-back level reached in the previous training session.

3.2.2 Task-specific near transfer tasks

Three WM tasks sharing the same task paradigm as the trained task were used as measures of task-specific near transfer. They were administered at pre-test, intermediate test and post-test. These were thus untrained variants of the n-back task, having the same task structure but different stimuli. They included n-back with letters (NBL), n-back with colors (NBC) and n-back with boxes (NBB). They all had the same number of blocks, targets and non-targets (including lures), as well as the same stimulus exposure/response time. The NBL used the nine first letters of the English alphabet (A, B, C, D, E, F, G, H, I) and the NBC used nine colors (purple, black, pink, yellow, red, green, blue, grey, orange). The NBB used a 3 x 3 matrix where a target (cell marked with a contrast color) moved from one spatial location to another. The participants completed 12 blocks of each task and the average level of n-back reached was used as the dependent variable.

3.2.3 Predictor variables

In the study by Fellman et al. (2020), information on several individual differences variables was gathered at pre-test. Demographic variables (**age**, **years of education, gender**) were registered with background questionnaires. Means and standard deviations for the demographic variables, as well as gender distribution, are presented in Table 2.

Table 2

Background characteristics of the participants.

Sample size (n)	113
Gender (F/M)	72/41
Age (M, SD)	35.12 (8.75)
Years of education (M, SD)	15.80 (3.22)

WM baseline performance level (see Table 3) was a composite score of the three task-specific near transfer n-back tasks (n-back with letters, colors and blocks) in the pre-test battery. All three task scores were standardized before they were modulated further. The average standardized level of n-back reached in each of the three tasks were added together and divided by three, to create a mean of the average level of n-back reached for each participant. This mean was used as a composite score for WM baseline performance level in the analysis.

Strategy use was measured separately for each WM task at pre-test, starting with a yes/no question as to whether the participant used any kind of strategy. If the participant confirmed to be using one, the next question prompted to describe the self-generated strategy in their own words in the comment field, and to rate how consistently the described strategy was employed on a 10-point Likert scale (1 = *Highly inconsistently*, 10 = *Highly consistently*). The written strategy reports were coded for two variables, Strategy Type and Level of Detail. The former indicated what type of strategy the individual used (e.g., rehearsal, grouping of stimuli etc.) and the latter reflected the level of sophistication of the strategy by determining the number of details given in the strategy report. The latter variable, introduced by Laine et al. (2018), served as the measure of strategy use in this study, ranging from 0 to 4 depending on how advanced the strategy was. This allowed individuals to be compared to each other by rank and the rank numbers were suitable for multiple regression analysis. Zero points denote the absence of any strategy in the strategy report. Highly vague strategies were given one point (e.g., "I tried to keep up but found it quite difficult"). Rather non-specific strategies were given two points (e.g., "Just tried to concentrate on the numbers"). A general strategy description with at most one detail was given three points (e.g., "I tried to repeat the previous digits each time a new one appeared"). Four points were given to descriptions with two or more details (e.g., "Memorized the four letters as one sequence and replaced each letter as a new one came up"). Fellman et al. (2020) used two evaluators to code the strategy reports independently and any discrepancies were solved through consensus. The agreement between the two independent evaluators was examined using a linearly weighted kappa analysis (κw) that revealed good-toperfect agreement for the Level of Detail classifications for the WM tasks at pre-test (κw range 0.86 – 0.96). In the present study, the predictor variable strategy use (see Table 3) was created by adding the strategy points from the three task-specific near transfer tasks and dividing the sum by three, thus creating a mean score for strategy use.

Motivation and **alertness** (see Table 3) were respectively assessed after the last WM task of the pre-test using a Likert scale ranging from 1 to 5. Motivation was measured by asking the participants to rate their level of motivation towards completing the tasks with the help of a 5-point Likert scale (1 = *Not at all motivated*, 5 = *Very motivated*). Alertness was measured using a similar Likert scale, as the participants rated their energy level (1 = *Very tired*, 5 = *Very alert*).

Cognition-related beliefs (see Table 3) were assessed using the Working Memory Questionnaire (WMQ), developed by Vallat-Azouvi et al., (2012). The original scale was developed for clinical use to assess the subjective consequences of WM deficits in everyday life. The WMQ is a self-administered scale probing short-term storage, attention and executive control in daily life. It consists of 30 questions (e.g., "Do you feel that fatigue excessively reduces your attention?", "When you are carrying out an activity, if you realize you are making a mistake, do you find it difficult to change strategy?"), each rated on a 5-point Likert scale (0 = *Not at all*, 4 = *Extremely*) (Vallat-Azouvi et al., 2012). In their sample of 313 healthy participants, Vallat-Azouvi et al. (2012) reported good reliability (Cronbach's alpha; .89) and good internal validity. Fellman et al. (2020) adopted the WMQ as an indicator of how individuals perceive the functionality of their own WM. In this study, the total sum score (maximum: 120 points) of the questions was used as the predictor variable. Lower points indicated that the individual perceived himself/herself as having a highly functioning WM, while higher points indicated the opposite.

Table 3

Descriptive statistics for the predictor variables

Predictor	*M*	*SD*	Min	Max
Pre-test WM performance level	0.00	0.86	-1.68	2.10
Strategy use	0.96	1.21	0	4
Motivation	4.18	0.93	1	5
Alertness	3.89	1.04	1	5
WMQ	51.04	16.30	30	110

WMQ = Working Memory Questionnaire. Pre-test WM performance level used standardized values.

3.2.4 Dependent variable

The dependent variable was the total amount of task-specific near transfer seen in the three untrained versions of the n-back task (n-back with letters, colors and blocks). Although all task-specific near transfer tasks encompassed the same task paradigm, score range and setup, they were standardized to z-scores. This was done to avoid the potential risk of some n-back task with a higher transfer effect to exert a disproportionally strong influence on the total task-specific near transfer score. The transfer score was attained as follows; the average standardized level of n-back reached at post-test was subtracted with the average standardized level of n-back reached at pre-test, respectively for each task. This yielded a separate task-specific near transfer score for each of the three tasks. These three scores were added together for each participant after which the total sum was divided by three. Thus, a standardized mean variable was created for the average level of n-back reached in all three untrained n-back tasks (M = 0.00, SD = 0.71, Min = -1.76, Max = 1.79). A positive score indicated that the participant's average performance had improved following WM training, meaning task-specific near transfer gains had occurred. A negative score indicated that the participant's performance level in the task-specific near transfer tasks had worsened following WM training. Zero points indicated no change.

3.3 Analytical approach

The statistical analysis was done using IBM SPSS Statistics 26. A preliminary analysis investigated Pearson's correlations between all the independent variables, as well as the dependent variable. Multiple regression analyses were then conducted to explore the predictive values of the individual difference variables for task-specific near transfer gains. Alpha level was set to $p < .05$ in all analyses. Due to the large number of independent predictors in relation to the sample size, the multiple regression analyses were separated into three sets to avoid model overfitting. The five statistical assumptions for a multiple regression analysis (Williams et al., 2013) were investigated in all three models. By testing for these assumptions, the risk for type 1 errors (falsely accepting the findings as significant when they in fact occurred by chance) and type 2 errors (falsely reporting no significant effect when there actually is one) were reduced (Osborne & Waters, 2002). First, the models were tested for the normality of residuals in order to meet the *statistical assumption of normally distributed residuals*. This was done by graphically applying a histogram and a probability-probability plot (P-P plot) with the observed standardized residuals. The histogram and P-P plot were used to observe whether the standardized residuals were closely related to the expected standardized residuals or if any deviations existed in the data. If the residuals were normally distributed, they would appear evenly distributed on both sides of the standardized zero value in the histogram and conform to the diagonal normality line in the P-P plot. Second, a scatter plot, with the standardized residuals on the x-axis and standardized predicted values on the y-axis, was applied to check for the *statistical assumption of a homoscedastic distribution* of the residuals. This was done to make sure that the variance (residuals) around the regression line was the same for all levels of the predictor variable. If the assumption of a homoscedastic distribution was not met, the residuals in the scatter plot would bunch together at some values and be further apart at other values. Third, the *statistical assumption of linearity* refers to a straight-line relationship between the predictor variables and the dependent variable and this assumption was automatically met if the two above-mentioned assumptions (normally distributed and homoscedastic residuals) were met. Fourth, the *statistical assumption of independent errors* was tested for by applying the Durbin-Watson test (values ranging from 0 to 4). This was conducted in order to make sure that the residuals did not correlate with each other. Values close to 2 indicated that the assumption of independent errors was met, as there was no problem with correlating residuals. Fifth, all predictors within each model were tested for the *statistical assumption of no multicollinearity* by running collinearity diagnostics (Tolerance levels & VIF-values). This was done to guard against highly correlated predictors, as that would cause problems in determining which predictors in the model are important, even though it would not reduce the predictive power of the whole model per se. Tolerance levels above 0,1 (preferably over 0,2) and VIF-values below 10 indicated that the assumption of no multicollinearity was met.

Due to the exploratory nature of the study, all variables included in a set were added to the regression analysis simultaneously. This was justified, as these variables either lack or have inconsistent evidence of their influence on near transfer gains. In other words, analyzing the potential effect of each predictor in the model was more important than finding the best possible model. The use of three separate analyses followed a rule of thumb for multiple regression analyses, that recommends at least 20 participants per variable (Field, 2018). In the present study a sample size of 160 participants would have been required, if all eight variables had been included in the same model. Thus, with 113 participants, an exclusion of three of the eight predictors would have been needed to meet this demand. As there are no robust previous study results regarding predictors of near transfer, an exclusion would have been almost completely on random and hence, not warranted. Instead, by chunking the predictors that were naturally related, the statistical assumptions for multiple regression were met and a wider set of predictors could be analyzed. The three predictor sets are presented in Table 4.

Table 4

The predictor sets applied in the multiple regression analyses

Regression model	**Set 1**	**Set 2**	**Set 3**
Type of set	Demographic:	Cognitive:	Subjective:
Individual Predictors included	Age Education Gender	Pre-test WM Strategy use	Motivation Alertness WMQ

WMQ = Working Memory Questionnaire. Pre-test WM = pre-test WM performance level.

4. Results

4.1 Descriptive data

The final sample consisted of 113 individuals. The sample data met parametric test demands, allowing for further analysis. The Pearson's correlation analysis (see Table 5) with the predictor variables (age, education, pre-test WM performance level, strategy use, motivation, alertness and WMQ) and the dependent variable (task-specific near transfer gains) revealed a few significant correlations. Gender was not included due to its dichotomous nature.

The demographic variables age and education, included in the first multiple regression analysis, did not show a statistically significant correlation with transfer gains and they did not correlate with each other. The two cognitive measures (pre-test WM performance level and strategy use), included in the second multiple regression analysis, showed a moderate positive intercorrelation. Put differently, the higher the WM performance level, the more sophisticated strategies were used. However, only pre-test WM performance level showed a significant bivariate correlation with transfer gains. This correlation was moderate and negative, meaning that a lower pre-test WM performance level was connected to more transfer gains. None of the subjective measures (motivation, alertness and WMQ) included in the third multiple regression analysis did show statistically significant correlations with transfer gains. They did, however, show moderate to strong intercorrelations between each other. A positive strong correlation was found between motivation and alertness. If the participants were motivated, they were also likely alert and vice versa. WMQ showed a moderate negative correlation with alertness and motivation. In other words, the more alert and motivated the participants were, the less WM-related problems they reported having. Additionally, weak but significant positive correlations were found between strategy use and education, as well as between motivation and age. A weak but significant negative correlation was found between WMQ and age, meaning that higher age resulted in less self-perceived everyday problems with WM.

Table 5

Pearson's product moment correlations between the independent and dependent variables

Variable	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**
1. Transfer gains							
2. Pre-test WM	**-.36****						
3. Strategy use	-.01	**.41****					
4. Motivation	-.18	.02	.04				
5. Alertness	-.14	.08	.05	**.64****			
6. WMQ	.00	-.04	-.12	**-.44****	**-.42****		
7. Age	-.14	.15	.14	**.32***	.12	**-.25****	
8. Education in years	.04	.09	**.20***	-.02	-.04	-.06	.10

WMQ = Working Memory Questionnaire. The correlation matrix applied standardized scores for task-specific near transfer gains and pre-test WM baseline performance level.
** Indicates p < .05; ** indicates p < .01*

4.2 Predictors of task-specific near transfer gains

The predictive value of the three sets of variables (**demographic:** age, education, gender; **cognitive:** pre-test WM performance level and pre-test strategy use; **subjective:** motivation, alertness, WMQ) was examined using three separate multiple regression analyses. In the first analysis with the demographic variables (see Table 6), all statistical assumptions were met, allowing for further analysis. The model did not, however, reach statistical significance ($R^2 = .02$, $F[3, 109] = .81$, $p = .49$), indicating that the combined effect of age, education and gender was not related to the task-specific near transfer gains.

Table 6

Multiple regression analysis with demographic variables as predictors of gain in task-specific near transfer measures

		Model	
Variable	**B**	**SE B**	**β**
Age	-.01	.01	-.15
Education	.01	.02	.05
Gender	-.06	.14	-.04
R^2		.03	
p-value		.41	

** Indicates p < .05; ** indicates p < .01*

The second multiple regression model, with pre-test WM performance level and strategy use included, met all statistical assumptions, allowing for further analysis. The results from the analysis (see Table 7) showed a statistically significant regression equation ($R^2 = .15$, $F[2, 110] = 9.66$, $p = .00$). In other words, the model explained 15 % of the variance in task-specific near transfer gain scores. The possibility to generalize the results from the regression model to the population was decent, as the adjusted R square (adjusted $R^2 = .13$) was close to the determination coefficient. Out of the two variables, only pre-test WM performance level reached significance. The relationship between pre-test WM performance level and task-specific near transfer gains was negative, such that lower pre-test WM performance level resulted in more transfer gains. This connection is graphically presented as a partial regression plot in Figure 3, demonstrating the negative linear relationship between pre-test WM performance level and transfer gains.

Table 7

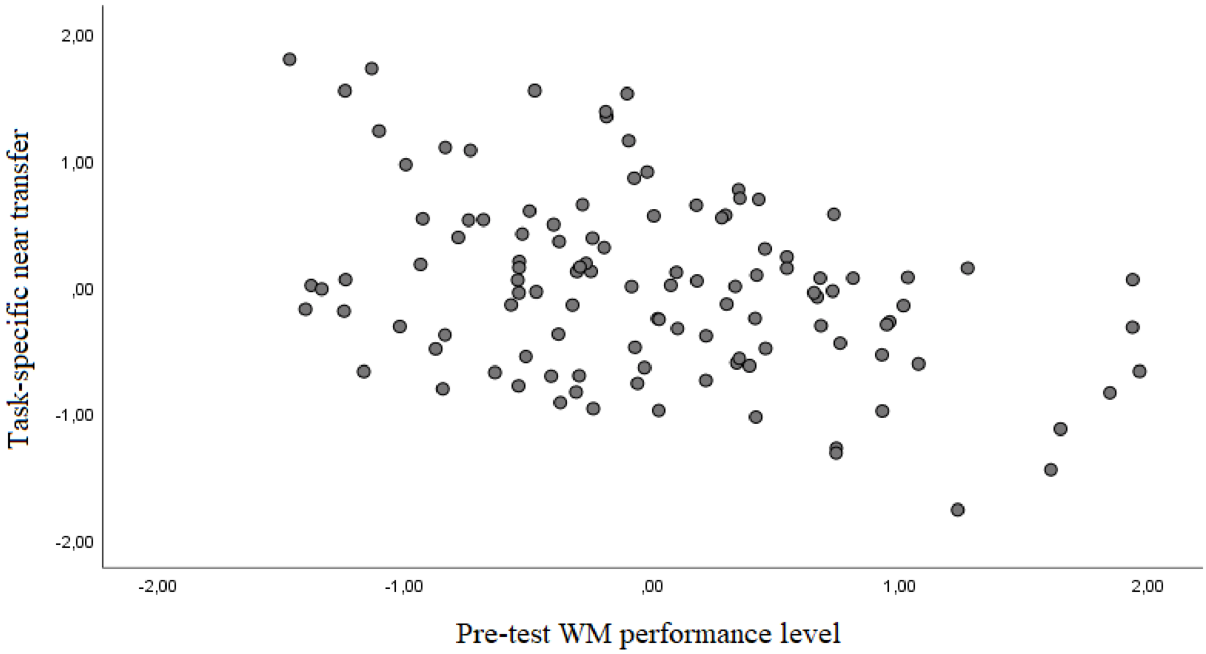
Multiple regression analysis with cognitive measures as predictors of gain in task-specific near transfer measures

		Model	
Variable	**B**	**SE B**	**β**
Pre-test WM	-.35	.08	**-.42****
Strategy use	.10	.06	.16
R^2		**.00****	
p-value		.15	

** Indicates p < .05; ** indicates p < .01.*

Figure 3

Partial regression plot with pre-test WM performance level as the independent variable (x-axis) and task-specific near transfer gains as the dependent variable (y-axis)



In the third multiple regression analysis, the subjective measures motivation, alertness and WMQ were tested as predictors of task-specific near transfer gains. The regression model met all statistical assumptions allowing for further analysis. However, the model did not reach statistical significance ($R^2 = .04$, $F[3, 109] = 1.57$, $p = .20$) (see Table 8), indicating that the combined effect of the participants' motivation- and alertness levels, as well as their perceived WM functionality, was not related to task-specific near transfer gains.

Table 8

Multiple regression analysis with subjective measures as predictors of gain in task-specific near transfer measures

		Model	
Variable	**B**	**SE B**	**β**
Motivation	-.13	.10	-.17
Alertness	-.05	.09	-.08
WMQ	-.01	.01	-.11
R^2		.04	
p-value		.20	

** Indicates p < .05; ** indicates p < .01. WMQ = Working Memory Questionnaire.*

5. Discussion

5.1 Main findings

The aim of the study was to examine if certain individual characteristics contribute to explaining previous inconsistent results regarding near transfer effects in WM training programs. Despite the importance of considering individual characteristics that might help in explaining the large inter-individual differences found in WM training outcome, only a few studies have considered this issue. Most WM training studies have focused on training improvements at the group level, ignoring the large differences in training outcomes at the individual level (Guye et al., 2017). In this study, eight variables (age, education, gender, WM baseline performance level, strategy use, motivation, alertness and cognition-related beliefs) were considered as predictors of task-specific near transfer gains. This type of transfer concerned improvements on untrained WM tasks (n-back with letters, colors and blocks) that were structurally similar to the trained task (n-back with digits). To the author´s knowledge, no previous study has examined task-specific near transfer effects in relation to an equally broad array of potentially relevant predictors.

Using multiple regression, the eight variables included in this study were organized into three distinct sets of predictors that were fed into separate models: demographic variables (age, education, gender), cognitive measures (WM baseline performance level, strategy use) and subjective measures (motivation, alertness, cognition-related beliefs). Out of the three regression models, only the second model, with the cognitive measures WM baseline performance level and strategy use as predictors, reached significance (discussed in the next paragraph). Neither the model with demographic variables nor the model with subjective measures reached significance. Put differently, none of these six predictors (age, gender, education, motivation, alertness and cognition-related beliefs) was found to predict task-specific near transfer. It is possible that the predictors in the set with subjective measures might have overlapped, as they were found to have moderate to strong intercorrelations. Especially motivation and alertness were strongly related and thus, including them both in the same set might have removed some of their unique variance. Out of the three demographic variables and the three subjective measures, only motivation, gender and age have previously been studied in relation to task-specific near transfer. The absence of motivation as a predictor was in line with the results by Anguera et al. (2012), and the absence of a predicting effect of gender was in accordance with the results by Hunt et al. (2014). The absence of age as a predictor of task-specific near transfer was also in accordance with most previous studies (Karbach & Verhaeghen, 2014; Soveri et al., 2017). Some studies have shown that younger age is related to higher transfer gains, but these have been conducted in relation to task-general near transfer and with wider age gaps between participants, as they included individuals up to 80+ years (Borella et al., 2017; Dahlin et al., 2008; Zinke et al., 2014). In this study, the age of the participants only ranged from 18-50. Normal age-related cognitive decline (Deary et al., 2009) would have been more evident if the sample had included older adults (60+ years), who are more likely to have experienced more significant cognitive decline in relation to younger adults. In other words, with a larger age gap between participants, age might have predicted near transfer gains, as was seen in the study by Dahlin et al. (2008).

5.2 Cognitive measures as predictors of task-specific near transfer

In the statistically significant regression model with the cognitive measures included, only WM baseline performance level was found to significantly predict task-specific near transfer gains, while strategy use showed a trend of statistical significance (p = .10). This matches the fact that the former showed a significant bivariate correlation with transfer gains, while the latter did not. The connection between WM baseline performance level and transfer gains in the multiple regression model was negative so that lower WM baseline performance resulted in more task-specific near transfer gains achieved. These results support the existence of a compensation effect, indicating that low-performing individuals benefit more from training, possibly because they have more room for improvement than high-performing individuals, who are already at the top of their performance (Borella et al., 2017). This is in line with some studies on near transfer, where such a compensation has been found in relation to both task-general (Borella et al., 2017) and task-specific near transfer (Hunt et al., 2014). These studies were, however, conducted with selective subgroups consisting of either older adults or children. When a young adult age group was investigated, Foster et al. (2017) found an opposite magnification effect of WM baseline performance level positively predicting task-general near transfer. Thus, cognitively high-performing individuals improved more than low-performing individuals. The results by Foster et al. (2017) are possibly affected by the fact that they used group level analyses of high vs. low WM baseline performance level and not a more sensitive regression analysis with individual performance levels. Their results might have been distorted due to a regression towards the mean (Moreau et al., 2016), if their sample included several individuals with either really high or really low WM baseline performance levels. However, a magnification effect has also been more prominent, as compared to a compensation effect, in studies on training task gains with healthy young adults (Foster et al., 2017; Guye et al., 2017; Wiemers et al., 2018). One potential explanation for the opposite effect found in all aforementioned studies could be that a magnification effect requires more training sessions to become evident, and a compensation effect or null results are more likely with less amount of training. In all the above-mentioned studies, at least 20 training sessions were applied as compared to the 12-session WM training program in this study. One meta-analysis did find the amount of training sessions to moderate near transfer gains so that more training led to more prominent near transfer (Weicker et al., 2016), albeit such a link has been refuted by two other meta-analyses (Schwaighofer et al., 2015; Soveri et al., 2017). Nonetheless, it is possible that low-performing individuals might show larger near transfer improvements during the initial phase of the WM training program, but this improvement might level out and dissipate further on if these participants are faced with their cognitive barriers. High-performing individuals on the other hand, might continue improving and enhancing their performance on near transfer tasks the more training they receive. However, such a phenomenon does not seem to apply for training task gains, as cognitively high-performing individuals also have been shown to improve faster than individuals lower in cognitive ability (Guye et al., 2017; Matysiak et al., 2019). This does not automatically mean that the same applies for task-specific near transfer, as many researchers warns against relying too strongly on a connection between training task gains and transfer gains (Redick, 2019; Tidwell et al., 2014). It is also important to remember that not all studies have been able to detect either a compensation or a magnification effect (Melby-Lervåg et al., 2016). This adds to the controversy regarding the predictive power of WM baseline performance level. Therefore, more investigation is needed in order to understand whether and how a compensation/magnification effect occurs in relation to task-specific near transfer in WM training.

The fact that strategy use did not reach significance in the multiple regression model is surprising, as the Strategy Mediation hypothesis is receiving more and more support (Fellman et al., 2020; Gathercole et al., 2019; Laine et al., 2018; Soveri et al., 2017). This might partly be due to the fact that strategy use and WM baseline performance level share some of their explained variance, leading to their unique contribution disappearing when fed into one and the same model. This is supported by the significant positive bivariate correlation ($r = .41$) found between these measures. However, when only analyzing the predictive value of strategy use post hoc, using a simple linear regression analysis, the variable was not even near to reaching significance. This would indicate that WM baseline performance level was the carrying force in making the multiple regression analysis model with the two cognitive predictors statistically significant. Thus, the statistically significant positive correlation found between WM baseline performance level and strategy use would not account for the non-significant beta value of the strategy predictor. To sum up, based on these results, implicit strategies do not appear to play a significant role, suggesting that most individuals would start from the same clean slate in terms of strategy development at the beginning of a WM training program. Instead strategy development and strategy use seem to obtain a more central role following repeated practice with a WM task (i.e., the actual WM training), as strategies are developed and enhanced (Fellman et al., 2020; Laine et al., 2018). In other words, individuals do not appear to have pre-existing strategies obtained from real-life situations that would be directly suitable to adapt to the novel tasks in WM training. Gathercole et al. (2019) support this conclusion, as they argue for the acquisition of new routines following WM training with unfamiliar tasks such as the n-back, and they compare the process to learning a new skill.

5.3 Implications

The results in this study provide crucial information regarding individual predictors of near transfer, and more specifically, task-specific near transfer. Due to the large discrepancies regarding the effects of WM training and poor generalizability to other cognitive domains (Soveri et al., 2017), researchers are beginning to lose interest in WM training programs (Redick, 2019). Before completely abandoning the concept, it is necessary to thoroughly investigate whether WM programs can be tailored to individually suit participants, who have the largest potential to improve their cognitive functioning following WM training. This could justify the commercial use of WM training programs for certain target groups (Redick, 2019). WM baseline performance level was found to significantly predict task-specific transfer gains and thus, add to a slightly more optimistic view regarding WM training programs. In theory, the compensation effect found could potentially mean that WM training programs can be applied for rehabilitation purposes (e.g. for individuals with an intellectual or a learning disability), as individuals with a weaker WM performance level appear to benefit more from training. This is a rather optimistic view, however, as the gains are only task-specific and appear to be a result of strategy use, rather than anything else (Fellman et al., 2020). Put differently, the problem with generalization to real-life settings still exists, making any commercial or rehabilitating use of WM training programs questionable (Redick, 2019). Additionally, with only WM baseline performance reaching significance, another issue emerges in relation to any commercial implications. This variable is problematic as the mapping of a baseline level would require the use of a battery of cognitive tests. The question then arises as to whether the potential benefits of the WM training programs are worth the large-scale mapping of cognitive abilities in order to choose suitable participants. The predictors in this study that could have easily been used to determine inclusion criteria (age, education and gender), did not reach significance.

Neither does it seem relevant to ask the participants for their subjective opinions regarding WM programs as none of the subjective measures (motivation, alertness and cognition-related beliefs) reached significance.

5.4 Limitations

Despite several strengths in the present study, there are some limitations to consider. First, Fellman et al. (2020) conducted an intermediate test in the middle of the training period, in addition to a pre- and a post-test. The negative consequence of this for the present study is an increased risk for a so-called *test-retest effect*, which leads to an overestimation of the amount of transfer found. A test-retest effect (also called practice effect) refers to the fact that individuals tend to improve their performance following repeated exposure to the same task (Lemay et al., 2004). This problem is already evident with only a pre-test and a post-test but increases even further with an intermediate test. This is not problematic when analyzing the overall effectivity of a WM training program at a group-level, as the control group removes the problem with a test-retest effect (Soveri et al., 2017) However, when focusing on individual predictors using a within-group multiple regression analysis, the problem exists. This means that a part of the task-specific near transfer found in the training group had nothing to do with the training regime, but simply with the fact that the test battery was familiar at the time of the post-test (Tidwell et al., 2014). One way to tackle this would be to compare the present kind of training group analysis to a similar one conducted with a control group. For the latter group, any significant predictors of pre-post gain would reflect only the test-retest effect.

A second limitation stems from the fact that the research sample comes from a fully online-administered trial and not a controlled laboratory setup. This makes it impossible to control for environmental factors (e.g. loud background) or to ensure that participants perform the tasks the way they were intended. However, Germine et al. (2012) found that studies conducted online need not systematically differ from traditional lab-based studies in terms of a reduction in data quality, even for demanding cognitive and perceptual experiments. Additionally, apart from the problem of not being able to control for all relevant factors following online-administered trials, studies have shown that near transfer gains might be reduced if the study instructions are not provided by a supervisor in a real-life setting (Schwaighofer et al., 2015). It is possible that more task-specific near transfer gains could have been achieved amongst the participants, increasing the power to detect significant predictors, if the study had been conducted in a laboratory-like setting with a real supervisor. Whether this is the case is debatable, however, as one meta-analysis found no difference in training outcomes between conducting a WM training program at home or in a laboratory (Au et al., 2014).

A third limitation concerns the way some predictor variables were obtained in the present study. WM baseline performance level was selectively determined based on the average baseline performance level on the three task-specific near transfer tasks at focus here, and not on all the available WM tasks in the pre-test. It is possible that the inclusion of more WM baseline task scores to the predictor variable composite would have led to different results in the multiple regression analysis [3]. Moreover, the strategy use measure was determined by having the participants write down whether they used any strategies and what type of strategy they used. For this variable to be accurate, it required the participants to have the energy and will to thoroughly record this even though it was possible to proceed with the study without doing so. As the pre-test was rather long (approximately 2 hours) and the questions on strategy use appeared multiple times, some commitment to the study might have been required in order to answer in a detailed way every time. Mccaul et al. (1987) found a link between commitment and performance, as they noticed that commitment to a task or a goal has motivational properties that increase effort. Since the participants received monetary compensation, which has been shown to reduce intrinsic motivation and be less beneficial for the WM training (Deci et al., 1999; Jaeggi et al., 2013), some participants might have taken the easy way out and answered in an undetailed fashion. Apart from strategy use, the subjective measures motivation and alertness can also be criticized, as they were assessed with quite minimal methods. Both variables were based on data stemming from only one pre-test Likert scale question. With only one question to map the variables, no reliability or internal consistency could be determined. Additionally, when using only one Likert scale question, the risk for participants who "sit on the fence" increases. These are individuals who find it hard to determine their opinion and thus, choose the option in the middle (Brown, 2000). Having the scales ranging from one to five in the present study allowed for a so called "neutral" answer, as number three appeared exactly in the middle. An alternative way to assess motivation and alertness could be to use for example, the Multidimensional Crowdworker Motivation Scale (MCMS; Posch et al., 2019) and the Karolinska Sleepiness Scale (KSS; Shahid et al., 2011). The MCMS was developed specifically to measure subjective motivation in the context of crowdworking and is rather short, as

[3] Indeed, in an unpublished study performed in parallel and independently of the present work, Fellman and colleagues found a magnification effect of WM baseline performance level on task-specific near transfer when analyzing the same data set as in the present study. This is possibly related to the fact that they used the average baseline score of all 10 WM tasks when compiling the predictor variable, thus not only including n-back tasks that are typically deemed as more difficult ones. Moreover, they used a different analysis method and also included the intermediate test in their statistical model, making any direct comparison between these two studies more difficult.

the scale only contains 18 items. The MCMS has been demonstrated to be both valid and reliable, as well as comparable across countries and income groups (Posch et al., 2019). The KSS measures the subjective level of sleepiness and only takes five minutes to complete (Shahid et al., 2011). The KSS has been proven to be a valid and reliable scale (Kaida et al., 2006) and could be adopted to measure the alertness level of participants in WM training programs.

5.5 Directions for future studies

It is by now obvious that the effectiveness and usefulness of WM training programs are far from the initial hype (Redick, 2019). The amount of studies on WM transfer effects is already sufficient enough to conclude that the benefits of these programs are narrow (Melby-Lervåg et al., 2016), as they mostly seem to be limited to improvements on untrained tasks structurally similar to the trained task (Soveri et al., 2017). However, as some studies have found larger transfer gains than others, the next thing to do is to explore why this might be the case. Based on the results from over a decade of WM training studies, the focus should now be turned away from far transfer and task-general near transfer, and instead zoom in on task-specific near transfer (Soveri et al., 2017). As this type of transfer seems to exist, investigating the underlying mechanisms and potential predictors would be essential for the future of all WM training programs. When doing so, controlling for the test-retest effect is crucial in order to obtain a realistic view of the task-specific near transfer occurring in WM training. Researchers should also try to avoid the inaccurate median-split designs and instead use more sensitive data analysis methods in order to minimize the loss of power and a reduction of effect sizes (Moreau et al., 2016). It might also be that predictors of task-specific near transfer effects work differently depending on the type of WM tasks applied. One meta-analysis found correlations between the n-back task and other WM tasks to be very low (Redick & Lindsey, 2013), while another found n-back tasks to strongly correlate with other WM tasks at a latent level (Schmiedek et al., 2014). By increasing the number of studies applying different task paradigms in relation to predictors of task-specific near transfer, it would be possible to rule out whether there might be any significant differences between different WM tasks.

In this study, WM baseline performance level was the only predictor to reach statistical significance. As the role of this predictor is still disputed, both in relation to its predictive power on near transfer (Melby-Lervåg et al., 2016) and whether a compensation effect (e.g., Hunt et al., 2014) or a magnification effect (e.g., Foster et al., 2017) is more prominent, more research regarding this predictor is needed. When doing so, it would be important to consider the potential role of training dose, as the direction of the predictive power might vary depending on the amount of training received. It could also be beneficial to control for other intelligence measures (e.g., fluid intelligence), as they might interact with WM baseline performance level in relation to near transfer. One possibility is that fluid intelligence works as a third variable in the WM training equation by moderating both pre-test WM performance level and task-specific near transfer gains.

Other than WM baseline performance level, no other individual characteristics included in this study were found to predict task-specific near transfer gains. Despite that, more research on predictors of near transfer is needed for any conclusions at a meta-analytic level. As more studies are grouped together and the sample pool increases, the risk for both type 1 and type 2 errors is reduced. A meta-analytic approach also minimizes the risk for methodological flaws diluting the results by the principle of "one's weakness becoming another one's strength". Some of the predictors (i.e., motivation and alertness) should be improved using more sophisticated data gathering methods (e.g. the MCMS for motivation and the KSS for alertness) as this might increase the reliability and validity of the variables.

5.6 Conclusions

To the best of the author´s knowledge, no previous study has analyzed an equally wide set of individual characteristics in relation to task-specific near transfer in WM training. The predictor variables included in the analysis were age, education, gender, WM baseline performance level, strategy use, motivation, alertness and cognition-related beliefs. Out of all eight variables, only WM baseline performance level reached significance and predicted task-specific near transfer gains. Individuals with a lower pre-test WM performance score showed larger task-specific near transfer gains at post-test than those with a higher pre-test score. The present study supports the existence of a compensation effect, where cognitively low-performing individuals show larger transfer gains, as they might have more room to improve as compared to cognitively high-performing individuals, who are already performing at the top of their ability. In theory, this would highlight the need for WM training programs to be tailored to suit individuals who are struggling cognitively (e.g. individuals with an intellectual or a learning disability), as they might benefit the most from WM training.

6. Swedish summary – Svensk sammanfattning

Prediktorvariabler för nära transfer i arbetsminnesträning

Inledning

Arbetsminnet beskrivs vanligen som förmågan att tillfälligt lagra, samt manipulera information. Denna förmåga anses spela en central roll i många av vardagens kognitiva processer (Baddeley, 1992; Miyake & Shah, 1999). Forskare har under det senaste decenniet aktivt debatterat över huruvida det är möjligt att med hjälp av arbetsminnesträning förbättra individers arbetsminneskapacitet. Målet har varit att hitta s.k. *transfer effekter*, där en förbättring i den tränade uppgiften även leder till en förbättring i andra, otränade uppgifter. I praktiken skulle detta betyda att man generellt kunde förbättra människors kognitiva kapacitet (Redick, 2019). Dock har man i övergripande meta-analyser kunnat konstatera att arbetsminnesträningsprogram inte är så effektiva som man ursprungligen antagit (Melby-Lervåg et al., 2016; Soveri et al., 2017). En signifikant förbättring har endast enhetligt noterats i så kallade *uppgiftsspecifika nära transfer uppgifter*, vars struktur är likadan som den uppgift man tränat med (Soveri et al., 2017). Trots att arbetsminnesträning verkar leda till en signifikant förbättring i en begränsad mängd otränade uppgifter så har man inom denna förbättring observerat stora interindividuella skillnader. Med andra ord verkar vissa individer dra mera nytta av arbetsminnesträningen än andra (Karbach & Verhaeghen, 2014; Melby-Lervåg et al., 2016; Soveri et al., 2017). Även om metodologiska brister i studieupplägget kan ansvara för en del av skillnaden, så kan möjligen deltagarnas *individuella karaktärsdrag* förklara en annan del av variansen (Guye et al., 2017). Trots arbetsminnesträningsforskningens popularitet, så har väldigt få studier beaktat individuella karaktärsdrag som en eventuell förklaringsmodell för diskrepansen i tidigare studieresultat.

Ett par karaktärsdrag som i tidigare studier relativt frekvent beaktats i förhållande till arbetsminnesträning är *grundnivå på arbetsminnesprestation* (t.ex. Borella et al., 2017) och *ålder* (t.ex. Melby-Lervåg, 2016). Basnivå på arbetsminnesprestation har relativt robust bevisgrund som prediktorvariabel för nära transfer, trots att forskare är delade om det rör sig om en *kompensationseffekt* (Hunt et al., 2014) eller en *magnifikationseffekt* (Foster et al., 2017). En kompensationseffekt betyder att de individer som är kognitivt lågpresterande får mera effekt ut av träningen eftersom de möjligen har mera utrymme för förbättring. En magnifikationseffekt tyder däremot på att de som redan är kognitivt högpresterande blir ännu starkare då de möjligen använder sin kognitiva förmåga mera fördelaktigt (Borella et al., 2017). Dock verkar det även som att ålder partiellt hänger ihop med grundnivån på arbetsminnesprestation. Forskare har nämligen konstaterat att en normal kognitiv nedgång sker i samband med ökad ålder och att denna kognitiva nedgång delvis berör arbetsminnets funktioner (Deary et al., 2009). I enstaka arbetsminnesträningsstudier har man också observerat en magnifikationseffekt, där både en yngre ålder och en högre grundnivå på arbetsminnesprestation resulterat i mera transfer (Borella et al., 2017). Dock är ålder som en ensamstående prediktorvariabel tveksam eftersom man i majoriteten av studier inte funnit att denna variabel skulle förutspå prestationen i transferuppgifter (Melby- Lervåg, 2016; Soveri et al., 2017).

Andra prediktorvariabler som spekulerats kunna ansvara för variansen i arbetsminnesträningsresultat är *personlighet* (Studer-Luethi et al., 2012; Urbánek & Marček, 2015), *motivation* (Anguera et al., 2012), *kognitionsrelaterade åsikter* (Jaeggi et al., 2013), *kön* (Matysiak et al., 2019), samt *fritidsaktiviteter*, *datorkunskap* och *tidigare kognitiv träning* (Guye et al., 2017). Dock har dessa variabler bristfälligt stöd från enstaka studier med relativt små sampel eller metodologiska brister. En variabel som nyligen observerats ha en central roll i arbetsminnesträning är *strategianvändning* (Fellman et al., 2020; Gathercole et al., 2019; Laine et al., 2018). Trots att strategianvändning konstaterats vara viktig för den förbättring som sker under själva träningen så har inga tidigare studier analyserat denna variabel som en eventuell prediktorvariabel för mängden uppvisad nära transfer.

I denna studie analyserades åtta individuella karaktärsdrag (ålder, utbildning, kön, grundnivå på arbetsminnesprestation, strategianvändning, motivation, vakenhetsnivå och kognitionsrelaterade åsikter) som eventuella prediktorvariabler för mängden uppgiftsspecifik nära transfer, efterföljt ett arbetsminnesträningsprogram. På grund av begränsad tidigare forskning angående ämnet så var denna studie explorativ ur den synvinkel att inga hypoteser lades fram. Huvudmålet med studien var att bidra till det stora glappet i litteraturen gällande den potentiellt signifikanta rollen individuella karaktärsdrag har i arbetsminnesträning.

Metod

Det analyserade datat i studien var en del av en nätbaserad och randomiserad, samt kontrollerad studie av Fellman et al. (2020), där strategianvändning i arbetsminnesträning undersöktes. Den ursprungliga studien godkändes av den etiska nämnden för psykologi och logopedi vid Åbo Akademi i enlighet med Helsingforsdeklarationen. I den ursprungliga studien deltog 258 deltagare men deltagarna delades efter pretestet in i två olika träningsgrupper och en kontrollgrupp. Denna studie fokuserade endast på deltagarna i en av grupperna, dvs. den traditionella träningsgruppen, vars träningsupplägg starkast påminde om upplägget i andra arbetsminnesträningsstudier (Soveri et al., 2017).

Samplet i den traditionella träningsgruppen bestod av 113 individer som alla efter pretestet utförde en tolv sessioner lång träningsperiod under loppet av fyra veckor, ett mellantest i mitten av träningsperioden, samt ett posttest i slutet. Träningsuppgiften bestod av en sifferversion av n-back testet (ett standardmått på arbetsminne, där man ska lagra och uppdatera stimulusserier). Uppgiftsspecifik nära transfer mättes med hjälp av en bokstavs-, en färg- och en spatialversion av n-back testet. Det standardiserade medeltalet av dessa tre transferuppgifter utgjorde den beroende variabeln.

Den statistiska analysen utfördes med hjälp av IBM SPSS Statistics 26. En preliminär Pearsons korrelationsanalys genomfördes med prediktorvariablerna och den beroende variabeln inkluderade. De åtta prediktorvariablernas förmåga att förutspå uppgiftsspecifik nära transfer analyserades med hjälp av multipla regressionsanalyser. En alfa-nivå på $p < .05$ applicerades i alla analyser. För att minska risken för variabelöverlapp så delades variablerna in i tre olika grupper; demografiska variabler (ålder, utbildning, kön) kognitiva variabler (grundnivå på arbetsminneprestation, strategianvändning) och subjektiva variabler (motivation, vakenhetsnivå och kognitionsrelaterade åsikter). Alla regressionsmodeller granskades mot de statistiskt viktiga antagandena för en multipel regressionsanalys innan resultaten granskades (Williams et al., 2013).

Resultat

Resultaten från Pearsons korrelationsanalys (se tabell 5) uppvisade några statistiskt signifikanta korrelationer. Av alla prediktorvariabler så korrelerade endast grundnivån på arbetsminnesprestation med uppgiftsspecifik nära transfer på så sätt att en lägre grundnivå hängde ihop med en större grad transfer. Grundnivån på arbetsminnesprestation korrelerade även medelstarkt med strategianvändning och sambandet var positivt. Motivation och vakenhetsnivå uppvisade en stark positiv interkorrelation och dessa två prediktorvariabler uppvisade även en medelstark negativ korrelation med kognitionsrelaterade åsikter. Med andra ord, ju mera motiverade och vakna deltagarna kände sig, desto mindre kognitionsrelaterade problem beskrev de sig ha. Andra svaga korrelationer existerade mellan ålder och motivation (positiv), utbildning och strategianvändning (positiv), samt ålder och kognitionsrelaterade åsikter (negativ).

Av de tre multipla regressionsanalyserna så var det endast modellen med de kognitiva prediktorvariablerna (grundnivån på arbetsminnesprestation och strategianvändning) som statistiskt signifikant förutspådde uppgiftsspecifik nära transfer (se tabell 7). Varken modellen med de demografiska (ålder, utbildning & kön; se tabell 6) eller de subjektiva (motivation, vakenhetsnivå & kognitionsrelaterade åsikter; se tabell 8) prediktorvariablerna kunde förutspå prestationsförbättring på uppgiftsspecifika nära transfer uppgifter. Modellen med de kognitiva prediktorvariablerna kunde däremot signifikant förklara 15 % av variansen. Inom modellen nådde bara grundnivån på arbetsminnesprestation signifikans och variabelns relation med uppgiftsspecifik nära transfer var negativ så att en lägre grundnivå på arbetsminnesprestation förutspådde större förbättring.

Diskussion

Målet med studien var att beakta huruvida individuella karaktärsdrag bidrar till att förklara de inkonsekventa studieresultat som nåtts gällande nära transfer. Åtta karaktärsdrag (ålder, utbildning, kön, grundnivå på arbetsminnesprestation, strategianvändning, motivation, vakenhetsnivå och kognitionsrelaterade åsikter) analyserades som potentiella prediktorvariabler för uppgiftsspecifik nära transfer med hjälp av multipla regressionsanalyser. Denna typ av transfer utgjordes av otränade varianter av n-back testet (n-back med bokstäver, färger och block) som till sin struktur påminde om träningsuppgiften (n-back med siffror).

Prediktorvariablerna analyserades som tre skilda variabelgrupper i tre separata modeller: demografiska variabler (ålder, utbildning, kön), kognitiva variabler (grundnivån på arbetsminnesprestation, strategianvändning) och subjektiva variabler (motivation, aktivitetsnivå, kognitionsrelaterade åsikter). Endast modellen med de kognitiva variablerna nådde signifikans och kunde förutspå uppgiftsspecifik transfer. Varken modellen med de demografiska eller de subjektiva variablerna förutspådde träningsframgång, vilket överensstämmer med existerande litteratur (Anguera et al., 2012; Hunt et al., 2014; Karbach & Verhaeghen, 2014; Soveri et al., 2017). I modellen med de kognitiva variablerna inkluderade, nådde endast grundnivån på arbetsminnesprestation signifikans på så sätt att en lägre grundnivå resulterade i mera uppgiftsspecifik transfer. Detta överensstämmer med den kompensationseffekt som observerats i flertalet studier (Borella et al., 2017; Hunt et al., 2014), där kognitivt lågpresterande individer uppvisat mera nära transfer, möjligen eftersom de har mera utrymme för förbättring (Borella et al., 2017). Strategianvändning kunde däremot inte förutspå transfer. Detta indikerar på att den bevisligen centrala rollen strategianvändning har i arbetsminnesträning verkar begränsa sig till själva träningen, där strategianvändning starkt ihopkopplats med förbättring i träningsuppgiften (Fellman et al., 2020; Gathercole et al., 2019; Laine et al., 2018; Soveri et al., 2017). Med andra ord så verkar individer stå på samma grundnivå gällande, för uppgifterna gynnsam, strategianvändning i början av träningsprogrammet. Detta överensstämmer med konstaterandet om att arbetsminnesträning involverar inlärning av en helt ny förmåga (Gathercole et al., 2019).

De praktiska implikationerna av resultaten i studien kunde möjligen innefatta användandet av arbetsminnesträning som rehabiliteringsprogram för kognitivt lågpresterande (t.ex. individer med intellektuell nedsättning eller inlärningssvårigheter), eftersom individer med lägre arbetsminnesprestation verkar dra mera nytta av träningen. Dock anses den allmänna generaliseringen från träningsfärdighet i arbetsminnesuppgifter till vardagslivet vara bristfällig, vilket skapar tveksamhet gällande arbetsminnesträningens nytta och effektivitet (Redick, 2019). Problemet med det faktum att basnivån på arbetsminnesprestation verkar vara den enda av de åtta prediktorvariabler som förutspår träningsframgång, är att denna variabel kräver ett testbatteri för att kunna fastslås. Därmed uppstår frågan huruvida det är värt att utföra en storskalig kognitiv kartläggning över människors arbetsminnesprestation för att sedan enbart välja en subgrupp individer för själva arbetsminnesträningen, vars arbetsminnesprestation är under en viss nivå.

Trots flertalet styrkor i studien så existerar det även ett fåtal brister. Först och främst kontrollerade man inte för test-retest effekten (träningseffekten), vilket leder till en överestimering av mängden uppgiftsspecifik nära transfer bland deltagarna. En del av förbättringen kan således härledas till att testbatteriet vid posttestet var bekant från tidigare (Lemay et al., 2004). En annan brist härstammar från att forskningsprojektet utfördes som en nätbaserad studie, vilket gjorde det omöjligt att i praktiken kontrollera för eventuella omgivningsfaktoriella brister som möjligen påverkat prestationen (t.ex. ljudnivå i bakgrunden). En tredje brist att beakta berör sättet hur viss variabeldata samlades in. Strategianvändningsvariabeln baserades på subjektiva beskrivningar av vald strategi, där kvaliteten på beskrivningarna möjligen påverkats av deltagarnas engagemang. Mccaul et al. (1987) observerade att engagemang verkar hänga ihop med prestation så att mera engagemang leder till en bättre prestation, vilket i denna studie möjligen betytt att vissa individer inte haft energi att ge en detaljerad och sanningsenlig beskrivning. Insamlingsmetoden för prediktorvariablerna motivation och vakenhetsnivå kan även kritiseras för att ha genomförts med relativt minimala metoder. Bägge två baserade sig enbart på en separat fråga gällande hur motiverad och hur vaken individen kände sig.

Framtida studier kunde fortsätta fokusera på olika individuella karaktärsdrag i förhållande till uppgiftsspecifik nära transfer, med särskilt fokus på grundnivån på arbetsminnesprestation. Detta skulle möjliggöra skapandet av robusta meta-analytiska slutsatser gällande den potentiella roll individuella karaktärsdrag spelar i arbetsminnesträning. På det här sättet kunde man nå fram till en fundamental förståelse för arbetsminnesträningens alla mekanismer och således dra välinformerade beslut gällande deras potentiella användningsområde i praktiken.

7. References

Anguera, J., Bernard, J., Jaeggi, S., Buschkuehl, M., Benson, B., Jennett, S., Humfleet, J., Reuter-Lorenz, P., Jonides, J., & Seidler R. (2012). The effects of working memory resource depletion and training on sensorimotor adaptation. *Behavioural Brain Research*, *228*(1), 107-115. doi: 10.1016/j.bbr.2011.11.040

Au, J., Sheehan, E., Tsai, N., Duncan, G., Buschkuehl, M., & Jaeggi, S. (2014). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *22*(2), pp.366-377.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*(3), pp.189-208.

Baddeley A. (1992). Working memory. *Science 255*, 556–559. doi: 10.1126/science.1736359.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*(4), 612–637. doi: 10.1037/0033-2909.128.4.612

Borella, E., Carbone, E., Pastore, M., Beni, R. D., & Carretti, B. (2017). Working memory training for healthy older adults: The role of individual characteristics in explaining short- and long-term gains. *Frontiers in Human Neuroscience*, *11*. doi: 10.3389/fnhum.2017.00099

Brown, J. (2000). What issues affect Likert-Scale Questionnaire formats? *Shiken: JALT Testing & Evaluation SIG Newsletter*, *4*(1), 27-30.

Bürki, C., Ludwig, C., Chicherio, C., & de Ribaupierre, A. (2014). Individual differences in cognitive plasticity: an investigation of training curves in younger and older adults. *Psychological Research*, *78*(6), 821-835. doi: 10.1007/s00426-014-0559-3

Chein, J., & Morrison, A. (2010). Expanding the minds workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review*, *17*(2), 193–199. doi: 10.3758/pbr.17.2.193

Chein, J., & Schneider, W. (2012). The brain’s learning and control architecture. *Current Directions in Psychological Science*, *21*(2), 78-84. doi: 10.1177/0963721411434977

Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, *17*(7), pp.438-449.

Dahlin, E., Nyberg, L., Bäckman, L., & Neely, A. (2008). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging*, *23*(4), 720-730. doi: 10.1037/a0014296

Deci, E., Koestner, R., & Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627-668. doi: 10.1037/0033-2909.125.6.627

Deary, I., Corley, J., Gow, A., Harris, S., Houlihan, L., Marioni, R., Penke L., Rafnsson, S., & Starr, J. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135-152. doi: 10.1093/bmb/ldp033

Engle, R., & Kane, M. (2004). *Executive attention, working memory capacity, and a two-factor theory of cognitive control*. In Ross, B. (Ed.), The psychology of learning and motivation, 145–199. New York, NY: Academic Press.

Fellman, D., Jylkkä, J., Waris, O., Soveri, A., Ritakallio, L., Haga, S., Salmi, J., Nyman, T., & Laine, M. (2020). The role of strategy use in working memory training outcomes. *Journal of Memory and Language*, *110*, 104064. doi: 10.1016/j.jml.2019.104064

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics*. London: SAGE Publications.

Foroughi, C., Monfort, S., Paczynski, M., McKnight, P., & Greenwood, P. (2016). Placebo effects in cognitive training. *Proceedings of The National Academy of Sciences*, *113*(27), 7470-7474. doi: 10.1073/pnas.1601243113

Foster, J., Harrison, T., Hicks, K., Draheim, C., Redick, T., & Engle, R. (2017). Do the effects of working memory training depend on baseline ability level? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1677–1689. doi: 10.1037/xlm0000426

Gajewski, P., 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*. doi: 10.3389/fpsyg.2018.02208

Gathercole, S., Dunning, D., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, *105*, 19-42. doi: 10.1016/j.jml.2018.10.003

Germine, L., Nakayama, K., Duchaine, B., Chabris, C., Chatterjee, G., & Wilmer, J. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847–857. doi: 10.3758/s13423-012-0296-9

Guye, S., De Simoni, C., & von Bastian, C. (2017). Do individual differences predict change in cognitive training performance? A latent growth curve modeling approach. *Journal of Cognitive Enhancement*, *1*(4), 374-393. doi: 10.1007/s41465-017-0049-9

Holmes, J., Woolgar, F., Hampshire, A., & Gathercole, S. (2019). Are working memory training effects paradigm-specific? *Frontiers in Psychology*, *10*. doi: 10.3389/fpsyg.2019.01103

Hunt, A., Kronenberger, W., Dunn, D., Gibson, B., & Gondoli, D. (2014). Predictors of change in short-term memory span following working memory training. *Acta Neuropsychologica*, *12*(4), 445-458.

Jaeggi, S., Buschkuehl, M., Jonides, J. & Perrig, W. (2008). Improving fluid intelligence with training on working memory. *PNAS*, *105*(19), 6829-6833.

Jaeggi, S., Buschkuehl, M., Jonides, J., & Shah, P. (2012). Cogmed and working memory training—Current challenges and the search for underlying mechanisms. *Journal of Applied Research in Memory and Cognition*, *1*(3), 211-213. doi: 10.1016/j.jarmac.2012.07.002

Jaeggi, S., Buschkuehl, M., Shah, P., & Jonides, J. (2013). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, *42*(3), 464–480. doi: 10.3758/s13421-013-0364-z

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*, *117*(7), 1574-1581. doi: 10.1016/j.clinph.2006.03.011

Karbach, J. & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive control and working memory training in younger adults. *Psychological Science*, *25*(11), pp.2027-2037.

Kuhl, J., & Kazén, M. (1997). *Persönlichkeits-Stil-und Störungs-Inventar (PSSI) [Personality Styles and Disorders Inventory]*. Göttingen, Germany: Hogrefe.

Laine, M., Fellman, D., Waris, O., & Nyman, T. (2018). The early effects of external and internal strategies on working memory updating training. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-018-22396-5

Lemay, S., Bédard, M.-A., Rouleau, I., & Tremblay, P.-L. (2004). Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. *The Clinical Neuropsychologist*, *18*(2), 284–302. doi: 10.1080/13854040490501718

Li, S., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, *23*(4), 731-742. doi: 10.1037/a0014343

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*. doi: 10.3389/fnagi.2019.00126

Mccaul, K., Hinsz, V., & Mccaul, H. (1987). The effects of commitment to performance goals on effort. *Journal of Applied Social Psychology*, *17*(5), 437–452. doi: 10.1111/j.1559-1816.1987.tb00323.x

Melby-Lervåg, M., Redick, T. & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer". *Perspectives on Psychological Science*, *11*(4), pp.512-534.

Miyake, A., & Shah, P. (1999). *Models of working memory – Mechanisms of active maintenance and executive control*. Cambridge, Cambridge University Press.

Moreau, D., Kirk, I., & Waldie, K. (2016). Seven pervasive statistical flaws in cognitive training interventions. *Frontiers in Human Neuroscience*, *10*. doi: 10.3389/fnhum.2016.00153

Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, *8*(2). doi: https://doi.org/10.7275/r222-hv23

Peng, P., & Fuchs, D. (2015). A randomized control trial of working memory training with and without strategy instruction. *Journal of Learning Disabilities*, *50*(1), 62-80. doi: 10.1177/0022219415594609

Posch, L., Bleier, A., Lechner, C. M., Danner, D., Flöck, F., & Strohmaier, M. (2019). Measuring motivations of crowdworkers. *ACM Transactions on Social Computing*, *2*(2), 1–34. doi: 10.1145/3335081

Redick, T. (2019). The hype cycle of working memory training. *Current Directions in Psychological Science*, *28*(5), pp.423-429.

Redick, T., & Lindsey, D. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin & Review*, *20*(6), 1102-1113. doi: 10.3758/s13423-013-0453-9

Salminen, T., Frensch, P., Strobach, T., & Schubert, T. (2015). Age-specific differences of dual n-back training. *Aging, Neuropsychology, And Cognition*, *23*(1), 18-39. doi: 10.1080/13825585.2015.1031723

Saucier, G. (1994). Mini-Markers: A brief version of Goldbergs unipolar Big-Five markers. *Journal of Personality Assessment*, *63*(3), 506–516. doi: 10.1207/s15327752jpa6303_8

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01475

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, *50*(2), 138–166. doi: 10.1080/00461520.2015.1036274

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. (2011). Karolinska Sleepiness Scale (KSS). *STOP, THAT and One Hundred Other Sleep Scales*, 209–210. doi: 10.1007/978-1-4419-9893-4_47

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*(4), 1077–1096. doi: 10.3758/s13423-016-1217-0

Studer-Luethi, B., Jaeggi, S., Buschkuehl, M., & Perrig, W. (2012). Influence of neuroticism and conscientiousness on working memory training outcome. *Personality and Individual Differences*, *53*(1), 44-49. doi: 10.1016/j.paid.2012.02.012

Tidwell, J., Dougherty, M., Chrabaszcz, J., Thomas, R., & Mendoza, J. (2014). What counts as evidence for working memory training? Problems with correlated gains and dichotomization. *Psychonomic Bulletin & Review*, *21*(3), 620–628. doi: 10.3758/s13423-013-0560-7

Urbánek, T., & Marček, V. (2015). Investigating the effectiveness of working memory training in the context of Personality Systems Interaction theory. *Psychological Research*, *80*(5), 877-888. doi: 10.1007/s00426-015-0687-4

Vallat-Azouvi, C., Pradat-Diehl, P., & Azouvi, P. (2012). The Working Memory Questionnaire: A scale to assess everyday life problems related to deficits of working memory in brain injured patients. *Neuropsychological Rehabilitation*, *22*(4), 634-649. doi: 10.1080/09602011.2012.681110

Zinke, K., Zeintl, M., Rose, N., Putzmann, J., Pydde, A., & Kliegel, M. (2014). Working memory training and transfer in older adults: Effects of age, baseline performance, and training gains. *Developmental Psychology*, *50*(1), 304-315. doi: 10.1037/a0032982

Weicker, J., Villringer, A., & Thöne-Otto, A. (2016). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology*, *30*(2), 190–212. doi: 10.1037/neu0000227

Wiemers, E., Redick, T., & Morrison, A. (2018). The influence of individual differences in cognitive ability on working memory training gains. *Journal of Cognitive Enhancement*, *3*(2), 174–185. doi: 10.1007/s41465-018-0111-2

Williams, M., Grajales, C., & Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research, and Evaluation, 18*(11). doi: https://doi.org/10.7275/55hn-wk47

PRESSMEDDELANDE

Människors arbetsminnesförmåga verkar förutspå framgång i arbetsminnesträning,
Pro gradu-avhandling i psykologi
Fakulteten för humaniora, psykologi och teologi, Åbo Akademi

Resultaten från en pro-gradu avhandling i psykologi vid Åbo Akademi tyder på att grundnivån på människors arbetsminnesprestation kan förutspå vem som drar mest nytta av arbetsminnesträning. Det verkar som att individer med lägre arbetsminnesprestation uppnår en förbättring i större grad, jämfört med de som redan presterar på hög nivå före träningsprogrammet. Även andra karaktärsdrag beaktades men de visade sig inte förutspå träningsframgång. Dessa karaktärsdrag var ålder, utbildning, kön, strategianvändning, motivation, aktivitetsnivå och kognitionsrelaterade åsikter. I studien deltog 113 vuxna individer, vars prestation mättes före och efter ett internetbaserat arbetsminnesträningsprogram på tolv sessioner. Träningsframgången mättes som mängden förbättring på otränade uppgifter, som till strukturen är identiska med den tränade uppgiften (även kallat nära transfer). Resultatet från avhandlingen bidrar till den hektiska debatten angående arbetsminnesträningens nytta och effektivitet.

Avhandlingen utfördes av Reidar Nervander under handledning av professor Matti Laine (Åbo Akademi), forskare Daniel Fellman (Umeå Universitet) och doktorand Liisa Ritakallio (Åbo Akademi).

Ytterligare information fås av: Reidar Nervander
Email: reidar.nervander@abo.fi