

論文要約

Thesis Summary

Experience-based Changes in Human Cognition: Positive and Negative Effects of

Working Memory Training

(経験による認知機能の変化:ワーキングメモリ・トレーニングによる正と負の効果)

NI NAN

The present thesis aimed to investigate how working memory (WM) training can lead to improvements as well as negative impacts on human cognitive functions, and to explore the mechanisms underlying these training outcomes. Chapter 1 provides a general introduction to this topic, featuring a comprehensive literature review of previous findings and theories of WM training and transfer effects.

WM broadly refers to the temporary retention and manipulation of a limited amount of information to support ongoing cognitive activities (Logie et al., 2020). Over the past two decades, the central role of WM in human cognition has motivated a large body of intervention studies aimed at inducing general enhancements in complex cognition by improving WM function through intensive training. However, although initial findings demonstrated promising transfer effects to untrained domains such as fluid intelligence (i.e., far transfer; Jaeggi et al., 2008), later methodology-strengthened studies and recent meta-analyses showed that training-induced improvements are largely restricted to the trained tasks and their untrained variants (i.e., near transfer, Melby-Lervåg et al., 2016; Tsubomi et al., 2019). This narrow pattern of transfer suggests that WM training rarely increases WM capacity but instead improves the efficiency with which the existing capacity is utilized (von Bastian et al., 2022).

Recent studies have explained this enhanced efficiency in WM from the perspective of cognitive skill learning. The cognitive routine framework (Gathercole et al., 2019) conceptualizes training-induced changes as the acquisition of novel cognitive routines, similar to learning a new skill. Routines are cognitive procedures that coordinate existing processes in a new sequence to meet novel task demands. These newly learned skills facilitate performance on trained tasks and may transfer to structurally similar untrained tasks. A similar view of

strategy development during training, namely the strategy mediation hypothesis, also suggests that WM training stimulates the generation of efficient task-specific strategies that contribute to training-related improvements (e.g., Fellman et al., 2020). The present research is based on these accounts of cognitive skill learning during training.

Chapter 2 elaborates on two novel aspects of the present thesis: the theoretical inference of negative transfer and the use of a two-phase training paradigm. Based on the perspective of cognitive skill learning, we further inferred that WM training might not always generate positive outcomes, because previously acquired routines may affect subsequent cognitive activities in various ways. Therefore, transfer effects should be evaluated on a continuum ranging from positive to null to negative. However, as negative transfer was not the primary focus of most training interventions, this possibility has rarely been investigated. To examine potential negative effects, we employed a novel two-phase training paradigm instead of the traditional pretest-posttest design used in previous studies. This paradigm involves an initial training phase followed by a re-training phase, and it may encourage the application of newly learned skills to subsequent cognitive activities, thus offering a more sensitive and ecologically valid approach to examine transfer effects. Chapter 2 thus provides the foundation for the following five studies. Studies 1 and 2 (Chapters 3 and 4) established the phenomena of positive and negative transfer in two-phase WM training, while Studies 3–5 (Chapters 5–7) addressed the mechanisms behind these effects from different perspectives.

Chapter 3 presents the first study (Study 1), which aimed to demonstrate positive and negative effects of WM training via two experiments. Both experiments involved two sets of adaptive training (Phase 1 and 2) on serial order backward recall tasks with verbal or spatial

stimuli. In Experiment 1, participants were trained on either the backward digit span, color change detection (active control), or backward circle span task in Phase 1. In Phase 2, all participants were trained on the backward letter span task. We conducted Experiment 1 online using with three training sessions per phase, and successfully replicated the key findings of Gathercole and Norris (in prep.) that backward circle training transferred negatively to subsequent backward letter training. We conducted Experiment 2 using a reversed task order design corresponding to Experiment 1. The results revealed that transfer from backward letter training to backward circle training was not negative, but rather weakly positive, suggesting that the direction of negative transfer effect is asymmetric. This asymmetric effect indicates that backward circle and backward letter tasks require different optimal routines, and that the locus of negative transfer might be the acquisition process of such optimal routines. Routines already established for backward circle may hinder the development of optimal routines for backward letter, but not vice versa. Overall, these experiments confirmed that a negative transfer effect can indeed occur under certain WM training designs, and provided critical support for the cognitive routine framework.

Chapter 4 reports the second study (Study 2), which aimed to replicate and extend the cross-phase near transfer effects observed in the first study with minimal training durations. We shortened training duration because a more compact two-phase design may offer an efficient yet effective way to investigate various experimental manipulations, thus facilitating its future application. Two experiments employed the same training design as in Experiment 1 but with different training durations. In Experiment 3, participants completed a single session of 100 trials in both training phases, but no evidence of transfer was observed. In Experiment

4, we extended the first training phase to three sessions with a total of 240 trials, and found that backward digit training positively transferred to subsequent backward letter training. In contrast, backward circle training did not enhance subsequent backward letter training, instead showing a trend toward negative transfer. These results suggest that transfer effects are unstable with minimal training durations. Therefore, we recommend a minimum of three to five training sessions per phase, as implemented in Experiment 1, to ensure reliable acquisition and transfer of cognitive skills. In summary, these findings highlight the potential of the two-phase training paradigm and provide practical guidance for developing more effective training protocols in future research.

Having demonstrated positive and negative transfer effects in two-phase WM training, the thesis then turns to investigate why these effects occur. Chapter 5 describes the third study (Study 3), focusing on the connection between the task relationships (e.g., correlations) and transfer outcomes. Using data from Experiments 1, 3 and 4, we investigated whether Phase 1 performance predicts Phase 2 performance and whether this association determines the direction of transfer. The initial optimism that WM training may induce wide-reaching cognitive benefits was driven by strong positive pre-training correlations between WM and complex cognition. However, across three experiments, we showed that such relationships do not guarantee that improvements in one task will positively transfer to another. Results revealed that Phase 1 performance on both the backward digit and circle span significantly predicted Phase 2 performance on the backward letter span. However, while backward digit training led to positive transfer to Phase 2 backward letter span performance, backward circle training resulted in negative transfer. In contrast, Phase 1 performance on color change detection

training showed a weaker association and minimal transfer to Phase 2, due to distinct task features. These findings suggest that although positive task relationships reflect shared cognitive components across tasks, those components enable the transfer of newly acquired cognitive skills that may be either beneficial or detrimental. This challenges the assumption that high pre-training correlations predict beneficial outcomes, but aligns with broader discussions that WM training promotes the learning of new cognitive skills, rather than improving WM function itself.

Moreover, Chapter 6 presents the fourth study (Study 4) investigating the role of cognitive strategy use in two-phase training. We analyzed self-reported strategy use questionnaires from Experiments 3 and 4, with a novel focus on the combined use of multiple strategies. Cluster analyses of Phase 2 backward letter training revealed distinct patterns (i.e., clusters) in strategy use combinations consistently across two experiments, and that these patterns were associated with objective training performance. A combination of most of the strategies was related to the highest performance, whereas reliance on visuospatial or reversal strategies but without grouping strategies produced the lowest performance. Moreover, in Experiment 4, which employed a longer initial training phase, strategy use in Phase 2 was influenced by different training tasks in Phase 1, suggesting a cross-phase transfer of strategy use. In Experiment 3 with a shorter training duration, we also observed significant differences in baseline performance between clusters, which reflects the potential influence of participants' initial abilities on effective strategy development. Overall, these findings highlight the important role of strategy combinations, rather than single strategies, in understanding WM training mechanisms. In line with cognitive skill learning perspective, the acquisition and transfer of

specific strategy combinations may underlie both positive and negative training outcomes.

In Chapter 7, the final study (Study 5) addressed the fundamental differences in representing and processing sequences within WM domains, which may contribute to the transfer effects observed in two-phase training. Using baseline data from Experiments 1 and 2, we compared performance on serial recall tasks across two domains (verbal and spatial) and two recall directions (forward and backward). The results suggested that verbal serial recall was overall slower and less accurate in the backward direction than in the forward direction, whereas in the spatial domain, backward recall was about as difficult as forward recall. In addition, there was a more marked recency effect in verbal backward recall than in spatial backward recall. These results suggest that verbal WM favors the forward-going direction during encoding and retrieval. On the contrary, spatial WM may not inherently encode sequences in the forward direction, making retrieval possible in either direction (Norris et al., 2019). This partly explains the negative transfer observed in the above studies. That is, Phase 1 training on a spatial backward recall task may result in the acquisition of skills that are not fine-tuned for the backward direction (e.g., reduced use of effective grouping strategies), thus rendering the routines suboptimal for the subsequent verbal backward recall task. Therefore, these findings connect the phenomenon of negative transfer to more basic cognitive mechanisms, advancing our understanding of memory for serial order.

Finally, Chapter 8 summarizes the main findings, theoretical and practical implications, limitations, future directions and conclusions. By using the two-phase WM training paradigm to study positive and negative transfer effects, the present thesis contributes to our understanding of changes induced by WM training and their underlying mechanisms. Our study

supports the perspective of WM training as skill acquisition, and provides further implications for recent frameworks of cognitive training and related theories such as skill learning, expertise and metacognitive control. At the same time, we offer practical implications for the design of cognitive training programs, particularly with regard to the training duration and strategy use in the two-phase design, as well as for applied contexts such as educational and clinical psychology. More importantly, we call for more careful implementation of cognitive training interventions, as the existence of negative transfer may reflect decreased cognitive flexibility caused by intensive training. All of these are important for better understanding the experience-based plasticity of human cognitive functions. Nevertheless, major limitations in our study include the short training duration, as well as the restricted scope of WM tasks, as we focused primarily on backward recall tasks. Future research should employ longer training periods and a broader range of WM tasks, and further elucidate the role of strategy use to provide deeper and more robust insights into the mechanisms of transfer in the novel two-phase training design.

In conclusion, our main finding is that training-induced changes are primarily driven by task-specific cognitive skills acquired during WM training. These skills may evoke beneficial, neutral, or even detrimental transfer effects depending on whether they are optimal or suboptimal when applied to subsequent cognitive tasks. These insights demonstrate how humans can flexibly, but sometimes cannot, adapt to cognitively challenging situations. In this way, the present research opens future avenues for developing effective and efficient training interventions that can foster not only improved task performance but also beneficial skills that genuinely transfer beyond the training context.