Psychology and Aging
2016, Vol. 31, No. 4, 390–408

0882-7974/16/$12.00 http://dx.doi.org/10.1037/pag0000081

Multi-Domain Training Enhances Attentional Control

Julia C. Binder, Mike Martin, Jacqueline Zöllig, Christina Röcke, Susan Mérillat, Anne Eschen, and Lutz Jäncke
University of Zurich

Yee Lee Shing
University of Stirling and Max Planck Institute for Human Development, Berlin, Germany

Multi-domain training potentially increases the likelihood of overlap in processing components with transfer tasks and everyday life, and hence is a promising training approach for older adults. To empirically test this, 84 healthy older adults aged 64 to 75 years were randomly assigned to one of three single-domain training conditions (inhibition, visuomotor function, spatial navigation) or to the simultaneous training of all three cognitive functions (multi-domain training condition). All participants trained on an iPad at home for 50 training sessions. Before and after the training, and at a 6-month follow-up measurement, cognitive functioning and training transfer were assessed with a neuropsychological test battery including tests targeting the trained functions (near transfer) and transfer to executive functions (far transfer: attentional control, working memory, speed). Participants in all four training groups showed a linear increase in training performance over the 50 training sessions. Using a latent difference score model, the multi-domain training group, compared with the single-domain training groups, showed more improvement on the far transfer attentional control composite. Individuals with initially lower baseline performance showed higher training-related improvements, indicating that training compensated for lower initial cognitive performance. At the 6-month follow-up, performance on the cognitive test battery remained stable. This is one of the first studies to investigate systematically multi-domain training including comparable single-domain training conditions. Our findings suggest that multi-domain training enhances attentional control involved in handling several different tasks at the same time, an aspect in everyday life that is particularly challenging for older people.

Keywords: cognitive training, healthy old age, iPad, multi-domain training, transfer

Supplemental materials: http://dx.doi.org/10.1037/pag0000081.supp

Julia C. Binder and Mike Martin, Department of Psychology, International Normal Aging and Plasticity Imaging Center (INAPIC), and University Research Priority Program (URPP) "Dynamics of Healthy Aging," University of Zurich; Jacqueline Zöllig, Department of Psychology and International Normal Aging and Plasticity Imaging Center (INAPIC), University of Zurich; Christina Röcke, Susan Mérillat, and Anne Eschen, International Normal Aging and Plasticity Imaging Center (INAPIC) and University Research Priority Program (URPP) "Dynamics of Healthy Aging," University of Zurich; Lutz Jäncke, Department of Psychology, International Normal Aging and Plasticity Imaging Center (INAPIC), University Research Priority Program (URPP) "Dynamics of Healthy Aging" Center for Integrative Human Physiology, University of Zurich; Yee Lee Shing, Department of Psychology, University of Stirling and Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.

This work was supported by the Velux Stiftung (project No. 369), the Jacobs Foundation, the LIFE program that supported Julia C. Binder's doctoral studies, and the University Research Priority Program (URPP) "Dynamics of Healthy Aging" of the University of Zurich, Switzerland. Hotel Plastisse is one of the core projects at the International Normal Aging and Plasticity Imaging Center (INAPIC) and the University Research Priority Program (URPP) "Dynamics of Healthy Aging" of the University of Zurich. We thank the following people who were involved in study enrollment: Ladina Costa-Bezzola, Serena Fiacco, Anna Lena Gillessen, Vicky Krack, Lisa Ronchetti, Sarah Schoch, and Brigitte Sonderegger. We thank Mathias Allemand for his support and the stimulating discussions about the methodological approach. We thank the Zurich University of the Arts, Specialization in Game Design (ZHdK) for their fruitful contributions in developing the concept of the training software and design, and implementing the training software for iPad: René Bauer, Michael Burgdorfer, Florian Faller, Ulrich Götz, Daniel Hug, Mela Kocher, Konradin Kuhn, Bruno Meilick, and Cornelius Müller. Furthermore, we thank Bruno Meilick for his technical support during the training phase. Special thanks go to our participants who participated in the extensive training study.

Julia C. Binder and Jacqueline Zöllig planned the empirical study, Julia C. Binder was responsible for and supervised data collection. Julia C. Binder analyzed the data and wrote the manuscript supervised by Yee Lee Shing. Yee Lee Shing kindly offered Julia C. Binder a research exchange to work together on the structural equations models. Anne Eschen, Susan Mérillat, and Christina Röcke commented on the manuscript. Anne Eschen, Lutz Jäncke, Mike Martin, Susan Mérillat, Christina Röcke, and Jacqueline Zöllig were involved in developing the original research idea and initiating the collaboration with the ZHdK to develop the training software. Anne Eschen and Jacqueline Zöllig had the project lead of the software development. Jacqueline Zöllig conceptually prepared the empirical study. Lutz Jäncke and Mike Martin raised the funding and supervised the project.

Correspondence concerning this article should be addressed to Julia C. Binder, Department of Psychology (Gerontopsychology and Gerontology), University of Zurich, Binzmühlestrasse 14, 8050 Zürich, Switzerland. E-mail: j.binder@psychologie.uzh.ch

With the increasing number of people living very long lives (Cauley, 2012), identifying effective training interventions to counteract the typical decline of cognitive abilities, such as executive functions, processing speed, reasoning, and episodic memory across the adult life span (for reviews see, e.g., Salthouse, 2010; Schaie, 2012), is highly relevant for individuals as well as societies. However, the overall picture of training older adults' cognition is mixed (see Ballesteros, Kraft, Santana, & Tziraki, 2015, for a recent, comprehensive review). Recent meta-analyses on computerized cognitive and video game training revealed at least small effect sizes of near and far transfer (Karbach & Verhaeghen, 2014; Kelly et al., 2014; Lampit, Hallock, & Valenzuela, 2014; Toril, Reales, & Ballesteros, 2014). However, null findings have also been reported (e.g., Melby-Lervåg & Hulme, 2013; Owen et al., 2010). Hence, there is an ongoing debate on the extent to which cognitive training generalizes to untrained domains and real life. The attempt to understand the mechanisms of cognitive training is complicated by the fact that studies differ widely with regard to the cognitive functions trained, the assessed transfer measures, and design factors, such as type of control groups or training duration (Noack, Lövdén, & Schmiedek, 2014; Noack, Lövdén, Schmiedek, & Lindenberger, 2009; Shipstead, Redick, & Engle, 2012).

It is assumed that working memory and executive functions are relevant for a broad range of cognitive functions and even for the daily functioning of older adults (e.g., Tomaszewski Farias et al., 2009). A recent meta-analysis showed reliable transfer effects of working memory and executive function training in older adults, with effects being larger for near than far transfer measures (Karbach & Verhaeghen, 2014). In contrast to targeting these functions directly, such as with classic working memory tasks, multi-domain training interventions require the handling of several tasks simultaneously or sequentially (Strobach, Frensch, & Schubert, 2012; Strobach, Salminen, Karbach, & Schubert, 2014). The simultaneous coordination of multiple training domains demands higher order executive functions (Strobach et al., 2012; Strobach et al., 2014). Hence, simultaneous multi-domain training potentially trains each training domain and, in addition, executive functions demanded by the concurrent orchestration of these domains. Based on the overlap hypothesis of training and transfer (Buschkuehl, Jaeggi, & Jonides, 2012; Dahlin, Neely, Larsson, Bäckman, & Nyberg, 2008; Jonides, 2004; Kuwajima & Sawaguchi, 2010; Lustig, Shah, Seidler, & Reuter-Lorenz, 2009; Taatgen, 2013), increasing training breadth by training multiple domains should theoretically increase the likelihood of such an overlap with transfer tasks. Based on this assumption, recent multi-domain training studies combined different cognitive domains with social stimulation, physical training, health advice, or nutritional guidance (e.g., the FINGER trial, Kivipelto et al., 2013). For example, a training study that aimed at older adults' memory, goal management, and psychosocial well-being increased all targeted areas by an intervention of 12 weeks with the sequential administration of each training module for four weeks (Craik et al., 2007; Levine et al., 2007; Stuss et al., 2007; Winocur et al., 2007). Positive synergistic effects have also been reported by the combination of physical and cognitive training (Bamidis et al., 2014; Theill, Schumacher, Adelsberger, Martin, & Jäncke, 2013).

In the present study, we focus on the simultaneous combination of inhibition, spatial navigation, and visuomotor function training. A prominent view of cognitive aging puts forward inhibitory deficits as the driving factor of working memory declines during aging (Hasher et al., 2007; Hasher & Zacks, 1988). Spatial navigation performance has a high ecological validity for everyday life functioning, but declines with age (Moffat, 2009). In addition, from a brain aging perspective, lateral prefrontal cortex and mediotemporal lobe are particularly affected by structural deterioration during aging (Raz et al., 2005; Raz & Rodrigue, 2006). The inhibition training targets frontal lobe functioning, specifically the right inferior frontal gyrus (Chambers, Garavan, & Bellgrove, 2009). The spatial navigation training targets hippocampal functioning (Moffat, 2009; Wolbers & Hegarty, 2010). The aging hippocampus is one of the few regions that has persistently shown to undergo shrinkage (Raz, Ghisletta, Rodrigue, Kennedy, & Lindenberger, 2010). However, attempts to investigate how training possibly counters hippocampal deterioration is sparse (see, e.g., Lövdén, Schaefer, et al., 2012). The choice of a motor component was based on the dedifferentiation hypothesis, suggesting that cognitive and motor processes are less separable during aging (Baltes & Lindenberger, 1997; Lindenberger & Baltes, 1994). According to this hypothesis, sensorimotor functioning, as a marker of physical integrity of the aging brain, is a prominent source of individual differences in cognitive aging. We specifically compare the training of each of these three domains, single-domain training of inhibition, spatial navigation, and visuomotor function, to the training of their simultaneous combination (multi-domain training) with regard to near and far transfer. We refer to near transfer for improvements on a task different from the training tasks measuring the cognitive function under training, while we refer to far transfer for improvements on a task measuring another cognitive function (cf. Karbach & Verhaeghen, 2014; for a general discussion see, e.g., Noack et al., 2014; Noack et al., 2009).

Different Training Approaches to Train Several Cognitive Domains Simultaneously

When designing a training intervention targeting several cognitive functions simultaneously, researchers have to consider a trade-off between experimental control over the trained function and complexity. Classic dual-task and task switching training allows fine-grained manipulation and close experimental control. However, the training tasks are not very complex. Dual-task training of same and different modality discrimination has shown near transfer to similar tasks with different stimuli in older adults (Bherer et al., 2005, 2008). Far transfer to executive functions and fluid intelligence has been shown by task switching training when compared to the training of each of the two tasks separately (Karbach & Kray, 2009). In contrast to dual-task and task-switching training, video game training is more complex and provides a motivating training environment (Anguera & Gazzaley, 2015; Green & Bavelier, 2008), an aspect that is increasingly recognized as critical in the training literature. Video game training has been successful in improving older adults' overall cognitive functioning, memory, attention, and reaction time (RT) when compared with active and passive control groups (for a meta-analysis of video game training with older adults see Toril et al., 2014). However, video game training does not allow a direct inference about which cognitive functions are trained (Karbach, 2014; Karbach & Verhaeghen, 2014), thereby making good and informed predictions for transfer difficult (Noack et al., 2014). Furthermore, finding appropriate control conditions for video game training is difficult. Hence training and control conditions within a study usually differ substantially. Across studies, different training regimes vary greatly, in turn hampering comparisons (Anguera & Gazzaley, 2015; Toril et al., 2014).

In the present study, we compared multi-domain and single-domain training with the Hotel Plastisse training program that was specifically designed to combine the advantages of dual-task training and video game training regimes (Binder et al., 2015). Hotel Plastisse uses game-based elements to create an interesting training environment, while, at the same time, the cognitive functions under training are clearly defined. Furthermore, the single-domain and multi-domain training conditions are comparable with regard to important context-dependent variables (training environment, cover story, number of games per training session, difficulty adaption, type of feedback). This tight comparison is an important advancement over video game training studies that typically compare different types of game that differ vastly in many dimensions. There is one recent video game training study that also succeeded in including comparable control conditions. Anguera et al. (2013) designed the video game Neuroracer to compare a dual-task training to the sequential training of both single tasks. In the dual-task training condition, participants had to drive a car along a road and simultaneously react to a signal detection task. This condition was compared with the training of each task component for half of the total training time (sequential training) and a passive control group. After a total of 12 1-hr training sessions, participants in the dual-task training condition improved more on working memory and sustained attention compared to participants in the sequential training and the passive control groups. Compared with Neuroracer, the Hotel Plastisse multi-domain training goes a step further by combining three different training domains. We compare the *simultaneous* training of inhibition, visuomotor function, and spatial navigation to the *separate* training of each single-domain. A multi-domain task that trains three different cognitive functions allows several task switches between the three functions (e.g., inhibition–visuomotor function, inhibition–spatial navigation, spatial navigation–visuomotor function, and vice versa). This is qualitatively different from switching between two tasks only (6 possibilities vs. 2 possibilities of switching) and requires more flexibility than the multi-domain training by Anguera et al. (2013) and dual-tasking training (Bherer et al., 2005; Bherer et al., 2008).

Taken together, the present study design allows us to investigate to what extent multi-domain training might lead to broader (far) transfer, while at the same time it possibly leads to smaller near transfer compared with single-domain training because each component function is trained less extensively. We hypothesize that the simultaneous training shows far transfer by improving higher order executive functions in addition to improvements in each component function (near transfer), while the single-domain training should increase performance on the trained domain (near transfer) without transferring to executive functions (far transfer). We therefore assessed a cognitive test battery of tasks measuring performance on inhibition, visuomotor function, spatial navigation (near transfer), and executive functions (far transfer) at baseline, post-training, and 6-month follow-up.

Interindividual Differences in Cognitive Training Effects

Older adults usually show substantial interindividual differences in cognitive performance. Baseline performance has shown to be related to training gains and transfer (Lövdén, Brehmer, Li, & Lindenberger, 2012; Zinke et al., 2014). Lower baseline performance in working memory training tasks has been associated with higher training gains (Zinke, Zeintl, Eschen, Herzog, & Kliegel, 2012; Zinke et al., 2014). Furthermore, higher training gains have been associated with higher transfer effects (Bürki, Ludwig, Chicherio, & de Ribaupierre, 2014; Zinke et al., 2014). The opposite pattern has also been found, such that better performing individuals benefitted more from memory strategy training (Verhaeghen & Marcoen, 1996). Depending on whether training induces plastic changes or draws on flexibility, Lövdén, Brehmer, et al. (2012) predicted compensation or magnification effects in a memory training paradigm (Brehmer, Li, Müller, von Oertzen, & Lindenberger, 2007). It is postulated that training that draws on flexibility refers to optimization within the available cognitive resources and should lead to training-induced compensation effects, such that lower performing individuals improve more through training compared to higher performing individuals. In contrast, training that taps on plasticity implies plastic changes and hence an expansion of currently available cognitive resources often accompanied by structural brain changes (see, e.g., Kühn, Gleich, Lorenz, Lindenberger, & Gallinat, 2014). Plastic changes are assumed to be bigger for higher performing individuals because they already use available cognitive resources optimally and cannot further increase performance by flexible adaption. Therefore, they would have to expand on their resources (Lövdén, Brehmer, et al., 2012). To investigate individual differences and how they are related to training effectiveness, we used a structural equation modeling approach with a latent difference score model to analyze training-related change in performance across the various cognitive tasks. Only a few training studies have analyzed training-related improvements with structural equation modeling so far (Bellander et al., 2015; Lövdén, Brehmer, et al., 2012; Schmiedek, Lövdén, & Lindenberger, 2010; Zelinski, Peters, Hindin, Petway, & Kennison, 2014). Hence, this study contributes to the training literature by both a unique study design that incorporates broad assessment of near and far transfer, and a latent difference score model approach to explicitly test individual differences in training-related changes.

Long-Term Effects of Cognitive Training

Although some impressive long-term effects of cognitive training in healthy old age have been shown (e.g., training-related maintenance of up to 10 years: Ball et al., 2002; Rebok et al., 2014; Willis et al., 2006), there are hardly any multi-domain training studies assessing maintenance effects. In the above-mentioned study by Anguera et al. (2013), participants of the dual-task training condition maintained performance on the training task five months after training. However, maintenance on the transfer test battery was not reported. With regard to transfer, training of the video game Space Fortress with strategy instructions to change the focus on particular game aspects from time to time did not result in maintenance of training-related improvements of an executive control task at the 3-month follow-up (Stern et al., 2011). In contrast, the sequential multi-domain training of several cognitive functions (reasoning, memory, problem solving, visuospatial map reading, handcraft, and physical exercise) as compared with the single-domain training of reasoning only resulted in maintained reasoning at the 12-month follow-up (Cheng et al., 2012). Interestingly, both intervention groups showed improvements on reasoning compared with a passive control group immediately after training and at the 6-month follow-up, but only the multi-domain training group maintained reasoning performance one year after training.

To our knowledge, there is no theoretical model to explain how training-related improvements and transfer are maintained. Furthermore, predictions about which training conditions enable maintenance is hampered by the scarce empirical basis with only a few studies including follow-up measurements. In the present study, we assessed performance on the cognitive test battery six months after training. If multi-domain training increased the likelihood of an overlap with transfer measures and even demands of daily life, then the trained abilities would have a higher probability of being used during the six months after training termination. In this case, we would expect participants of the multi-domain training to show better maintenance of training-related improvements and transfer than individuals in the single-domain training groups.

The Present Study

In summary, the present study introduces an iPad-based training specifically designed to have comparable multi-domain and single-domain training conditions. This training regime provides healthy older adults with a motivating learning environment including a cover story and detailed feedback about training performance (Binder et al., 2015). Eighty-four healthy participants aged 64 to 75 years were randomly assigned to one of four training conditions, namely training inhibition, visuomotor function, spatial navigation, or their simultaneous combination (multi-domain training) over 50 training sessions with adaptive difficulty. The cognitive transfer test battery was very different from the training tasks. We expected the simultaneous multi-domain training to have a higher chance of overlapping with transfer tasks and daily demands. Hence, we hypothesized multi-domain training to transfer to executive functions (far transfer) and to show maintenance of training-related improvements at the 6-month follow-up.

Method

Participants

Participants were recruited for a "cognitive training study" through study advertisements in local newspapers and magazines for seniors, lectures for senior citizens at the University of Zurich, and the participant database of the International Normal Aging and Plasticity Imaging Center (INAPIC) of the University of Zurich. They were first screened for eligibility in a telephone interview. Inclusion criteria included age between 64 and 75 years, retirement, right-handedness, speaking German fluently, neurologically and psychiatrically healthy, no severe vision or hearing impairments, and no participation in a cognitive training study within the last two years. If these inclusion criteria were met based on self-reports in the telephone interview, individuals were scheduled for a baseline session. At the beginning of the baseline session, participants provided written informed consent and completed further health questionnaires and tests to finally decide on study admission. Participants were required to score at least 27 points or higher (of a maximum of 30 points) in the Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975). All participants self-reported that they had not suffered from a depression within the last three years, and participants were screened for current depressive symptoms with the Geriatric Depression Scale (GDS; Gauggel & Birkner, 1999; Yesavage et al., 1982). In addition to participants' self-report of being right-handed, we assessed handedness with the questionnaire by L. J. Chapman and Chapman (1987). Three participants self-reported that they had been retrained to write with the right hand during school (which was a common practice for this generation), but were included in the study. If participants were admitted to the study, we randomized them to one of the four training conditions. For participation in the training including pre- and posttest, participants were reimbursed 60 CHF (approximately 60 USD). When they attended the 6-month follow-up, they were paid an additional 50 CHF (approximately 50 USD). The study was approved by the ethics committee of the Faculty of Arts of the University of Zurich.

At baseline, we excluded two participants (one participant had severe vision impairments, one participant scored low in the MMSE and additionally had impaired color vision). An additional 10 participants were excluded from all analyses (for excluded participants' characteristics see supplementary information Table A1): Three participants were admitted to the study but never started with the cognitive training, six participants withdrew from the study during training and did not come back for the posttest and follow-up assessments, and one participant was excluded from all analyses because she was diagnosed with a psychiatric condition after training. The final sample consisted of 84 participants (see Table 1 for demographics). Three participants did not complete the 50 training sessions, but were included in the analyses because they took part in all pre-, post-, and follow-up assessments (2 participants in the inhibition group quit at training sessions 32 and 44, one participant in the visuomotor function training quit at training session 42). The remaining 81 participants completed all 50 training sessions. Eight of them did not take part in the 6-month follow-up assessment but were included in all other analyses (see supplementary information Table A1).

The four training groups did not differ with respect to the ratio of male to female participants, $\chi^2(3) = .76$, $p = .858$, age, $F(3, 80) = 1.63$, $p = .189$, MMSE, $F(3, 80) = .37$, $p = .776$, depressive symptoms (GDS; $F(3, 80) = .05$, $p = .986$), handedness, $F(3, 80) = 1.65$, $p = .185$, years of school education, $F(3, 80) = .62$, $p = .602$, and vocabulary knowledge (MWT-B; Lehrl, 2005), $F(3, 80) = .43$, $p = .730$. The age of the whole sample ranged from 64 to 75 years at baseline ($M = 69.90$, $SD = 2.80$) with more female than male participants (58.33% females). All but 2 participants had a computer at home, all but 5 people indicated that they were familiar with the Internet, 13 participants possessed an iPad, and 33 participants possessed a smartphone.

We also collected data from a sample of no-contact control participants comparable to the training study sample about a year after the training study took place. Participants of the no-contact control group performed the cognitive test battery twice with an interval of about 10 weeks in-between, similar to the training participants' baseline and posttest sessions (see supplementary

Table 1
Study Characteristics of the Whole Sample and for Each Training Group Separately

Demographics	All	Training group: Inhibition	Visuomotor function	Spatial navigation	Multi-domain
Sample size (f, m)	84 (49, 35)	22 (14, 8)	21 (11, 10)	20 (11, 9)	21 (13, 8)
Age	69.49 (2.83)	70.50 (3.05)	68.81 (2.48)	68.95 (2.76)	69.62 (2.85)
MMSE	28.93 (.85)	28.86 (.71)	29.10 (.83)	28.85 (.99)	28.90 (.89)
Depression	1.08 (1.47)	1.00 (1.75)	1.14 (1.62)	1.05 (1.40)	1.14 (1.15)
Handedness	12.96 (2.40)	12.91 (1.60)	12.57 (1.33)	13.95 (4.20)	12.48 (1.12)
School education	10.02 (1.99)	10.36 (2.23)	10.12 (2.12)	10.03 (1.98)	9.55 (1.61)
Vocabulary	32.86 (2.11)	32.73 (2.41)	33.24 (1.87)	32.95 (2.11)	32.52 (2.09)

Note. Means and standard deviations (in parentheses) are indicated. Age: Age at baseline in years; MMSE: exclusion if score was below 27 points; depression (GDS) with 15 items; handedness (12 questions): 12–17 points – right-handedness, 18–31 – ambidexterity, 32–36 points – left-handedness; school education in years; vocabulary (MWT-B): mean of 32 points indicates high average crystallized intelligence.

information Table A2 for the no-contact control group's sample characteristics and Table A3 for descriptive performance on the cognitive test battery). Data from this group allowed us to estimate retest effects. Our focus was to compare single- versus multi-domain training, and we thought that the single-domain training groups function as a very strict active control condition for the multi-domain training. Hence, we used the no-contact control group only for additional analysis to compare training-related improvements against retest effects.

Apparatus

Training took place individually at home with an iPad (versions 1, 2, 3) by Apple Inc. Participants were handed out an iPad at the end of the baseline session. Because of a limited number of iPads, participants were divided into two waves. As soon as a participant brought back an iPad, we could hand it out to another trainee. Individual cognitive testing in the laboratory consisted of paper-pencil and computer-based tests administered on a PC with a 22-inch monitor using the keyboard, the mouse, and special button boxes.

Training Procedure

The three single-domain training groups trained inhibition, visuomotor function, or spatial navigation exclusively, while the multi-domain training group trained these three cognitive functions *simultaneously*. Each training condition consisted of five different training tasks called minigames. A training session included the completion of all five minigames in a fixed, quasi-randomized order. Each minigame took about 6 to 10 min to complete, which resulted in a total session time of about 45 to 60 min including instructions and feedback. All training conditions encompassed 50 training sessions with adaptive task difficulty (about 5 training sessions per week, one training session per day). The training parameters and the training setting were as comparable as possible between the multi-domain training and the single-domain training conditions (for a detailed description of the training software see Binder et al., 2015). The level of the current training session depended on the performance of the previous training session: A score of 80% or higher resulted in a level increase for the subsequent training session, a score below 60% resulted in a level decrease, and a score between 60% to 80% resulted in level maintenance. Training score protocols were transferred to a data server immediately after training completion to enable supervision of training progress by the study team.

Inhibition training. The inhibition training consisted of five different minigames with go/no-go tasks (washday, labeling, fruit salad, dishwashing, chasing mice). In all five minigames, participants were presented with a continuous stream of go and no-go stimuli. They were supposed to react to go stimuli and inhibit their reaction to no-go stimuli (the whole screen registered taps independent of the tapping location). For example, participants sorted laundry in the washday minigame. The clothes were blown out of the dryer at the top of the screen and fell toward two baskets. Go stimuli were pieces of clothing labeled with the hotel logo, no-go stimuli were pieces of clothing without the hotel logo. Hence, reacting to a go stimulus shifted the baskets such that the particular piece of clothing was sorted to the basket with the hotel logo. Upon a no-go stimulus, no response was required. The delay between two stimuli was shortened with increasing level across the training sessions (e.g., washday: level 1 with 173 go and 36 no-go stimuli and a delay of 1.72 s between the stimuli; level 50: 747 go and 153 no-go stimuli with a delay of 0.40 s between the stimuli). The percentage of correct responses to the total of all responses (correct and incorrect reactions to go and no-go stimuli) determined the level for the subsequent training session (increase, decrease, maintenance). Level and percentage of correct responses were the dependent variables of training performance.

Visuomotor function training. The visuomotor function training consisted of five minigames to practice eye-hand coordination with unimanual or bimanual hand and finger movements (paw prints, darts, rolling fruits, marble box, model aircraft). In all five minigames, participants were presented with a continuous stream of visuomotor targets that had to be aimed at as precisely as possible. Depending on the minigame, difficulty increased across levels by the parameters speed or the size of the targets. For example, participants had to sink colored marbles in the marble box minigame. The color of the target marble was indicated by a colored ring around the hole in the middle of the screen where the marbles had to be sunk. Participants could move the marbles by tilting the iPad (bimanual control). The number of marbles increased with increasing training level (level 1: 2 marbles, level 50: 12 marbles). Furthermore, the speed of the marbles was gradually increased (game-specific metric of speed at level 1: 1.01; level 50:

1.60). The percentage of hits (e.g., correctly sunk target marbles) to the total of all reactions (hits and misses; correctly sunk target marbles and incorrectly sunk marbles) determined the level for the next training session (increase, decrease, maintenance). Level and percentage of correct responses were the dependent variables of training performance.

Spatial navigation training. The spatial navigation training consisted of five minigames that required participants to memorize and recall different paths in labyrinths (hedge labyrinth, pantry, wine cellar, room service, odyssey). All tasks consisted of an encoding and a retrieval phase. During encoding, either 2D-maps (bird's-eye perspective, time-unlimited encoding) or 3D-videos of labyrinths (landmark perspective, time-limited encoding) were presented. Retrieval always required recalling the memorized path in a 3D-labyrinth by deciding on the correct direction at every crossroad. The decision at the crossroads was either time-unlimited by choosing an arrow (unimanual control), or time-limited by tilting the iPad to the left, to the right, or no tilting to keep straight on (bimanual control). For example, participants' task in the hedge labyrinth minigame was to find lost items. During the encoding phase, participants were walked through the hedge labyrinth by a video animation (landmark perspective, time-limited encoding). During the retrieval phase, participants walked through the same labyrinth again. At every crossroads, the animation stopped and the participants had to decide on the correct direction by pressing the respective arrow (unimanual control, time unlimited). The animation time between two crossroads was 4 s. The labyrinths at level 1 consisted of three crossroads, whereas the labyrinths at level 50 consisted of 12 crossroads. Across training sessions, difficulty increased by the number of crossroads of a labyrinth and the complexity of the labyrinths. The percentage of correct responses to the total of all responses (correct and incorrect decisions at the crossroads) determined the level for the subsequent training session (increase, decrease, maintenance). Level and percentage of correct responses were the dependent variables for training performance.

Multi-domain training. The multi-domain training required participants to *simultaneously* handle an inhibition task, a visuomotor function task, and a spatial navigation task (raking leaves, pipe burst, wine tasting, vacuum cleaner, model car racing). Therefore, the five multi-domain training tasks consisted of two parts accommodating requirements for the spatial navigation task by an encoding and a retrieval phase. During the retrieval phase of the spatial navigation task, participants had to simultaneously perform an inhibition and a visuomotor task.

During encoding, 2D-maps (bird's-eye perspective, time-unlimited encoding) or 3D-videos of labyrinths (landmark perspective, time-limited encoding) were presented. Retrieval always required recalling the memorized path in a 3D-labyrinth by deciding on the correct direction at every crossroad (spatial navigation component; unimanual or bimanual control). The decision was always time-limited and the animation did not stop. Between two crossroads, participants were presented with a continuous stream of go and no-go stimuli that they had to reach or ignore, respectively (inhibition component). In addition, the go stimuli served as visuomotor function targets: these targets had to be hit as precisely as possible (unimanual or bimanual control; it was always the same control mode as the spatial navigation component required for retrieval). Whereas the timing of the reactions was critical for the inhibition component, the precision was critical to the visuomotor function component. For example, in the raking leaves minigame, the participants' task was to rake leaves in the hedge labyrinth. During the encoding phase, participants were walked through the hedge labyrinth (landmark perspective, time-limited encoding). During the retrieval phase, participants were walked through the animated labyrinth again. At every crossroads, participants had to decide on the correct direction by pointing to the respective arrow (spatial navigation component). Between the crossroads, participants had to pick up leaves (go stimuli of the inhibition component) and ignore garbage items (no-go stimuli of the inhibition component). At the same time, participants had to aim at the leaves as precisely as possible (visuomotor component; unimanual control).

Across training sessions, difficulty increased by the number of crossroads of a labyrinth, the complexity of the labyrinths (spatial navigation component; from 3 to 12 crossroads), and increasingly shorter delays between go and no-go stimuli (inhibition and visuomotor components; the raking leaves minigame started with 144 go and 38 no-go stimuli with a delay of 1.97 s between stimuli and ended with 569 go and 151 no-go stimuli with a delay of 0.50 s on level 50). The mean percentage of the three training components (correct responses to the total of all responses) determined the level for the subsequent training session (increase, decrease, maintenance). Level and percentage of correct responses were the dependent variable of training performance.

Maintenance of performance on trained tasks. To evaluate to what extent participants maintained performance on the trained tasks, they played the five minigames again at the 6-month follow-up. To compare performance with training session 50, they worked on the minigames at their individually reached end level of training session 50. Therefore, the percentage of performance was directly comparable within individuals. The follow-up data of the minigames of two participants had to be excluded because of errors in level setting. The data of three participants who did not complete all 50 training sessions were excluded.

The Cognitive Transfer Test Battery

According to the four training conditions, we created composite scores for inhibition, visuomotor function, and spatial navigation to evaluate the effects of the single-domain training (near transfer). In addition, composite scores for executive control functions were calculated for working memory, speed, and attentional control (far transfer). For each composite score, an average score was calculated across the tasks that made up the domain. For all tests, we first gave the instructions, made sure that participants had understood the task with examples, and practiced the task when a practice run was available.

Inhibition composite (near transfer). The inhibition composite consisted of two RT tasks, a stop signal task and a Stroop task.

Stop signal task. This task from the Vienna Test System assessed motor response inhibition (Kaiser, Aschenbrenner, Pfüller, Roesch-Ely, & Weisbrod, 2012). Participants had to sort arrows pointing to the left and the right side of the computer screen by pressing two keyboard buttons. Whenever they heard an acoustic signal (stop signal) after an arrow, they were instructed not to respond. The task consisted of two parts directly following each other. Each part consisted of 100 arrows presented for 1 s with an interstimulus interval of 1 s. Succeeding 24 of the 100 arrows, a tone of 1000 Hz with a duration of 100 ms was presented as the stop signal. This stop signal had a variable delay that increased when participants correctly inhibited their reaction and decreased when they did not inhibit their reaction (range 50–350 ms, a longer delay indicates better performance and requires more inhibitory control). The dependent variable entered to calculate the composite score was the stop signal RT, which was the mean RT minus the delay of the stop signal (main variable for inhibition performance as described in the Vienna Test System).

Stroop task. This task is a measure of response inhibition as it requires suppressing the dominant response of reading to correctly name the color of words (MacLeod, 1991). We programmed the task with the E-prime version 2.0 (Schneider, Eschman, & Zuccolotto, 2002a, 2002b). In each trial, a stimulus appeared in red, blue, green, or black ink and participants were instructed to react to the ink color and ignore the semantic meaning of the stimulus by pressing one of four keys on the keyboard. In congruent trials, the semantic meaning of the word matched the ink color, while in incongruent trials, the semantic meaning of the word did not match the ink color. There were 28 congruent and 84 incongruent trials that were presented in a fixed pseudorandom order. Stimuli remained on the screen until the participant gave a response or until 2000 ms had passed. The interstimulus interval was 500 ms. The dependent variable used to calculate the composite score was the Stroop effect based on the median RTs for incongruent minus congruent trials (correct responses). The data of three participants were not available (for one at baseline and two at posttest due to technical reasons and color discrimination difficulties).

Visuomotor function composite (near transfer). We used the short version of the motor performance series ("motorische Leistungsserie"; MLS) from the Vienna Test System (Neuwirth & Benesch, 2011; Schoppe, 1974; Sturm & Büssing, 1985) including the four subtests steadiness, line tracing, aiming, and tapping. The MLS work panel had touch-sensitive contact surfaces and holes. Each test was administered twice: once with the dominant right hand, once with the left hand. The visuomotor function composite consisted of the mean performance score of both hands for the subtests steadiness, line tracing, and aiming. Tapping was left out because of poor correlation with the other three tasks, likely due to its emphasis on speed rather than acuity. One participant did not complete the tests at posttest.

Steadiness. As a measure of arm or hand unrest and tremor, participants were required to hold a thin pen in a hole with a diameter of 5.8 mm without touching the rim or the bottom. The board was positioned vertically. Testing lasted 32 s. The dependent variable for the composite score was the number of touches, which were counted as errors.

Line tracing. Participants were required to trace a groove in the work panel with a thin pen as quickly and as precisely as possible. The board was positioned horizontally. The dependent variable for the composite score was rim touches, which were counted as errors.

Aiming. Participants had to touch a series of 20 circles positioned in a line with a thin pen as quickly as possible (contact points of the work panel with a diameter of 5 mm separated by a gap of 4 mm). The board was positioned horizontally. The dependent variable for the composite score was the total time in seconds for task completion.

Spatial navigation composite (near transfer). The spatial navigation composite consisted of the Corsi block forward, a mental rotation task, and a map learning task.

Corsi block forward. This task was originally developed by Corsi (1972) and is a measure for visuospatial short-term memory (STM). We used the subtest from the Wechsler Memory Scale Revised (Härting et al., 2000). The task consisted of a board containing nine blocks at fixed positions. The experimenter tapped several blocks in a predefined order with a speed of 1 s per block and the participant had to recall this order by tapping the presented sequence in the same order. The presented block sequences gradually increased in difficulty with sequences of two blocks at the beginning to sequences of a maximum of seven blocks. Two different sequences of the same length were always presented subsequently. The task was terminated as soon as two sequences of the same length were not correctly reproduced. The dependent variable for the composite score was the total number of correctly reproduced sequences (0–12).

City map path learning. This subtest of the Berlin Intelligence Structure Test (Jäger, Süβ, & Beauducel, 1997) assessed visuospatial STM. Participants were shown a city map on which a path from one house to another was drawn. Participants memorized this path for 30 s and were then asked to redraw the presented path on an empty map. Recall time was time-limited to 30 s. The dependent variable for the composite score was the number of correctly recalled segments redrawn on the empty map.

3D spatial orientation. This test from the Vienna Test System measured spatial perception and spatial rotation abilities (Bratfisch & Hagmann, 2012). A target figure composed of several blocks was presented at the top of the computer screen. An arrow pointed to the figure from a particular direction. The participants had to imagine what the figure looked like from this perspective. At the bottom, there were four different figures of which the correctly rotated figure had to be identified. The test consisted of 30 items, but there was a time limit of 3 min to solve as many items as possible. The dependent variable was the total of correctly solved items. One participant had problems with three-dimensional thinking and did not complete the test at baseline.

Working memory composite (far transfer). The working memory composite consisted of a 2-back task, the Corsi block backward, and the digit span backward.

2-back task with two-digit numbers. The 2-back task is a measure of working memory by requiring online monitoring, updating, and manipulating remembered information (Owen, McMillan, Laird, & Bullmore, 2005). We used a 2-back test version from the test battery of attentional performance by Zimmermann and Fimm (2002a). Participants were shown a sequence of visually presented two-digit numbers. They had to press a button whenever the current number was the same as the one presented two positions before (target). The task consisted of 100 two-digit numbers presented with a rate of 3 s. Fifteen numbers were targets. The total duration of the task was 5 min without the practice trial and instructions. The dependent variable for the composite score was the sum of the number of errors (commissions) and the number of omissions. The data of two participants were not available because of technical problems (one dataset at baseline, one dataset at posttest) and one participant did not understand the task at baseline.

Corsi block backward. This task is basically the same as the Corsi block forward but measures visuospatial working memory as it requires to recall the series backward (Corsi, 1972). We used the subtest from the Wechsler Memory Scale Revised (Härting et al., 2000). The experimenter tapped several blocks in a predefined order with a speed of 1 s per block and the participant had to recall this order by tapping the presented sequence in the reverse order. The presented block sequences gradually increased in difficulty with sequences of two blocks at the beginning to sequences of a maximum of seven blocks. Two different sequences of the same length were always presented subsequently. The task was terminated as soon as two sequences of the same length were not correctly reproduced. The dependent variable for the composite score was the total number of correctly reproduced sequences (0–12).

Digit span backward. This task is the verbal version of the Corsi span backward and hence measures verbal working memory. We used the subtest from the Wechsler Memory Scale Revised (Härting et al., 2000). The experimenter read a series of one-digit numbers aloud with a speed of one number per second. At the end of the series, participants had to repeat the series of numbers in the reverse order. The sequences gradually increased in difficulty with sequences of two numbers at the beginning to sequences of a maximum of seven numbers. Two different sequences of the same length were always presented subsequently. The task was terminated as soon as two sequences of the same length were not correctly reproduced. The dependent variable for the composite score was the number of correctly reproduced sequences (0–12). Digit span forward was not included in any of the composite scores since it did not fit in any of the composites from a theoretical perspective.

Speed composite (far transfer). The processing speed composite consisted of the trail making test (part A) and the digit substitution task.

Trail making test part A. Part A of the trail making test assessed visual search and motor speed skills (Bowie & Harvey, 2006). There were 25 circles containing numbers distributed on a sheet. Participants drew a line to connect the circles in ascending numerical order (1–25) as quickly as possible. Whenever an error was committed, the experimenter stopped the subjects and returned them to the last correct response for continuation. The dependent variable for the composite score was the total time for completion in seconds.

Digit substitution test. The digit substitution test measured processing speed. It was administered as a paper-and-pencil test (Härting et al., 2000; Wechsler, 1981) that consisted of a code table at the top of the page that paired nine numbers with a distinct symbol. Below, participants were presented a series of numbers in a quasi-random order and were required to fill in the respective symbols as shown in the code. The code was presented during the whole test. First, 6 number-symbol pairs were completed as practice trials followed by 94 test items of which as many number-symbol pairs had to be completed in a 90-s time interval. The dependent variable for the composite score was the total of correctly filled-in symbols.

Attentional control composite (far transfer). The attentional control composite consisted of four tests: the Test D2 for focused attention, two tests for divided attention (divided attention, trail making test part B), and a test of flexibility or set shifting.

Test D2. This test was a measure of sustained and focused attention (Brickenkamp, Schmidt-Atzert, & Liepmann, 2010). For a total of 14 lines on a page, subjects had to identify a target among several distractors (each line contained 21–22 distractors and 25 to 26 targets). Participants were required to start at the beginning of each line and work sequentially through the items by marking as many targets as possible. For each line, a time limit of 20 s was set. Even if they did not finish a line, they had to continue with the next line when 20 s had passed. For analysis, the first and the last line were discarded. The dependent variable for the composite score was a "concentration score" calculated by subtracting the sum of errors and omissions from the number of correctly identified targets. One participant did not complete the test at baseline because of vision problems.

Divided attention. In this task from the test battery of attentional performance by Zimmermann and Fimm (2002c), participants performed an auditory and visual task simultaneously. In both tasks, they had to detect target stimuli and respond as fast as possible on a response button (the same for auditory and visual targets). The visual task required participants to identify when moving crosses on a grid formed a rectangle. The auditory task required participants to react when two tones of the same pitch followed each other. The whole task took 3 min 25 s. A total of 100 visual stimuli including 17 targets were presented with a stimulus presentation time of 2 s. Simultaneously, 200 auditory stimuli were presented including 16 targets with a stimulus presentation time of 433 ms and an interstimulus interval of 1 s. The dependent variable for the composite score was the median of the RTs for both visual and auditory stimuli. The data of one participant at baseline were not available due to technical problems and one participant did not understand the task at posttest.

Flexibility/Set shifting. We used the nonverbal set shifting task called flexibility from the test battery of attentional performance by Zimmermann and Fimm (2002b). Each trial consisted of two figures, an angular and a round figure, one presented on the right and the other on the left side of the computer screen. The participants had two response buttons, one on the left and one on the right side. Every trial, the target changed and participants had to alternate with focusing on the angular or round figure by pressing the button on the respective side. One hundred trials were presented. There was no time limit for a trial. The next trial was presented as soon as a response had occurred. If an error was committed, participants got an auditory signal and were shown the next correct response. The dependent variable for the composite score was a general performance index calculated by the test program in which the RTs and the number of errors were included. A high index indicates good performance (fast reactions, few errors), and a low index indicates bad performance (slow reactions, many errors).

Trail making test part B. Part B of the trail making test assessed visual search, motor speed skills, and executive control, such as set shifting and working memory (Bowie & Harvey, 2006; Sánchez-Cubillo et al., 2009). As in part A, there were 25 circles distributed on the sheet. In part B, half of the circles contained numbers (1–13) and half of the circles letters (A–L). Participants drew a line to connect the circles in ascending order as quickly as possible. However, they had to alternate between the numbers and the letters. Errors were not scored directly, however, the experimenter stopped the subjects whenever an error was committed and returned them to the last correct response for continuation. The dependent variable for the composite score was the total time for completion in seconds.

Data Analyses

Analyses were conducted using SPSS 22 and AMOS 22 (http://www.spss.com). We used MATLAB R2012a (Mathworks Inc., Natick, MA; http://www.mathworks.com) for data organization, creating figures of training data, and computing composite scores of the dependent variables.

Training data. Upon completion of each minigame in each training session, a high score protocol with the participant's code was uploaded to a data server containing all the relevant training scores (level, percentage). Missing data of training sessions resulting from technical problems were not imputed, however such missingness was rare. At the 6-month follow-up, participants played the five minigames again at the last level they had reached. Thereby, we could compare how much performance on the trained tasks declined by comparing percentage of performance in the last training session and at follow-up (assessed at the same level).

Transfer test battery. First, distribution of the raw scores of the dependent variables of the transfer test battery were visually inspected and transformed with the natural logarithm when very skewed. Next, outlier values outside the range of mean ±4 standard deviations were replaced by the mean ±4 standard deviations. We decided on this liberal procedure to keep the data as close as possible to the original data. We repeated all analyses with the whole data set including the outlier values and results did not change. Second, we rescaled all values such that higher values meant better performance (e.g., RTs and errors were inverted by multiplying them with −1). Third, to get the same metric, we *z*-standardized all dependent values based on the mean of the baseline score and the pooled standard deviation of the three measurement points (baseline, posttest, and follow-up). Finally, we computed the composite scores by calculating the mean of the *z*-standardized dependent variables for each measurement time point. The variables that formed a composite score intercorrelated well except for the inhibition composite (see supplementary information Table A4). Consequently, we could not build an inhibition composite score based on the Stroop and the stop signal task. Therefore, we built up the models with each inhibition variable (Stroop effect, stop signal RT).

To evaluate training-related changes at posttest and follow-up, we used multigroup structural equation modeling. Because of the small sample size, we could not establish latent factors for the dependent variables; instead, we set up the measurement model with the composite score for each time point (baseline, posttest, follow-up). We then estimated a latent change score for the difference from baseline to posttest and from posttest to follow-up (see Figure 1).

We started with the just identified model with free parameters across groups and then subsequently constrained the means, the variances, and covariances across groups. If the model fit dropped significantly upon a constraint as evaluated with the likelihood ratio test (difference in χ^2; $\Delta\chi^2$), we freed the respective parameter and continued by constraining the subsequent parameters in the model. Model fit was evaluated using the χ^2 exact fit test, the comparative fit index (CFI), and the root mean square of approximation (RMSEA). In general, CFI above .95 and RMSEA values below .06 indicate that a model is adequately parameterized and reflect good model fit. Values for CFI above .90 and for RMSEA of below .08 are also acceptable (Browne & Cudeck, 1992; Hu & Bentler, 1999). The model fit of the final model for the Stroop effect variable was not acceptable, $\chi^2(7) = 9.43$; CFI = .64; RMSEA = .09 (.00–.23). We therefore do not report any results on the Stroop effect.

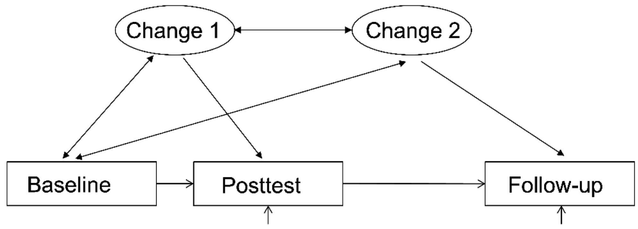

Figure 1. Latent difference score model to investigate training-related change on the composite scores. Rectangles represent the composite scores for baseline, posttest, and follow-up, circles represent estimated latent change scores. The two small arrows pointing to the posttest and follow-up boxes indicate error terms.

We additionally ran traditional repeated measures ANOVAs with the between-group factor Training group and the within-group factor Time (baseline, posttest) to make analyses comparable with other studies (see supplementary information Tables A5–A10). The results did not differ between the two approaches.

Retest analysis. In addition to our main structural equation models, we ran latent difference score models including the no-contact control group. Because we only have retest data on two measurement time points with an interval of about 10 weeks in-between, we reduced the models to one change score from baseline to posttest (the original analyses also include the follow-up time point). When there was a significant difference in the change score between the training groups, we compared the training groups separately to the no-contact control group (e.g., for the attentional control composite, multi-domain vs. no-contact control group, single-domain vs. no-contact control group). If there was no group difference in the change score, we collapsed across the training groups and compared them against the no-contact control group.

Effect sizes. Alpha level was set to $p < .05$ for all analyses. Effect sizes of analyses of variance were partial eta-square values and categorized according to the following conventions: small effect: $\eta_p^2 = .01$; medium effect: $\eta_p^2 = .06$; large effect: $\eta_p^2 = .14$ (Lakens, 2013). Effect sizes for the change scores of the structural equation models were calculated as Cohen's *d* (Cohen, 1992) by dividing the change score from baseline to posttest by the standard deviation at baseline (variances were always equal across groups) and the change score from posttest to follow-up by the standard deviation of the change score of baseline to posttest (variance for speed differed between groups). Cohen's *d* to quantify differential training improvements were only calculated when there were significant differences in change (difference in change score divided by the standard deviation). Effect sizes were classified according to the following conventions: small effect: $d = .20$; medium effect: $d = .50$; large effect: $d = .80$.

Results

Our main interest was in the comparison between multi-domain and single-domain training with respect to training-related transfer and maintenance. We expected the simultaneous multi-domain training to have a higher chance of overlapping with far transfer tasks and therefore hypothesized that multi-domain training transferred to executive functions. Furthermore, we were interested to what extent gains from training followed a compensation or magnification pattern of individual differences.

Training-Related Improvements on the Trained Tasks

Performance increased over the course of training in all training groups, as indicated by the increasing level of difficulty (see Figure 2). A simple linear regression was calculated to predict training level based on training session for each training group separately. The regression equations were highly significant (all $ps < .001$, all $R^2 > .94$) with highly significant linear slopes for each group (all $\beta s > .97$, $ps < .001$). With increasing level, the training task became more difficult to challenge individual performance levels. Increased difficulty was reflected in a decreasing percentage of performance over the training course. Percentage of performance of each training session determined the level of the next training session, such that the difficulty level could increase, decrease, or stay the same. Means and standard deviations of the level and percentage of performance for the mean of all five minigames of the last training session are shown in Table 2.

Table 2
Group Means of Level and Percentage of Performance for Training Session 50 and Follow-Up

	Level	Percentage of performance	
	Session 50	Session 50	Follow-up
Training group	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Inhibition	39.30 (1.72)	71.69 (3.49)	59.14 (5.35)
Visuomotor function	44.09 (2.06)	85.28 (3.33)	74.52 (5.29)
Spatial navigation	43.54 (6.03)	85.69 (5.64)	78.75 (6.30)
Multi-domain	42.12 (2.92)	77.88 (3.11)	66.27 (6.29)

Note. The group means are based on each participant's mean over all five minigames. See also Figure 2 caption for further information.

Training-Related Improvements on the Transfer Tasks: Comparing Multi-Domain to Single-Domain Training

There were no baseline differences for the composite scores across the four groups (ANOVAs with the factor group) and for the comparisons of interest (t-tests for independent samples), nor for individual variables of the composite scores for the comparisons of interest (t-tests for independent samples; for descriptives of the composite scores and individual variables for the group comparisons of interest see Table 3; for descriptives of the composite scores and individual variables of each group see supplementary information Table A11).

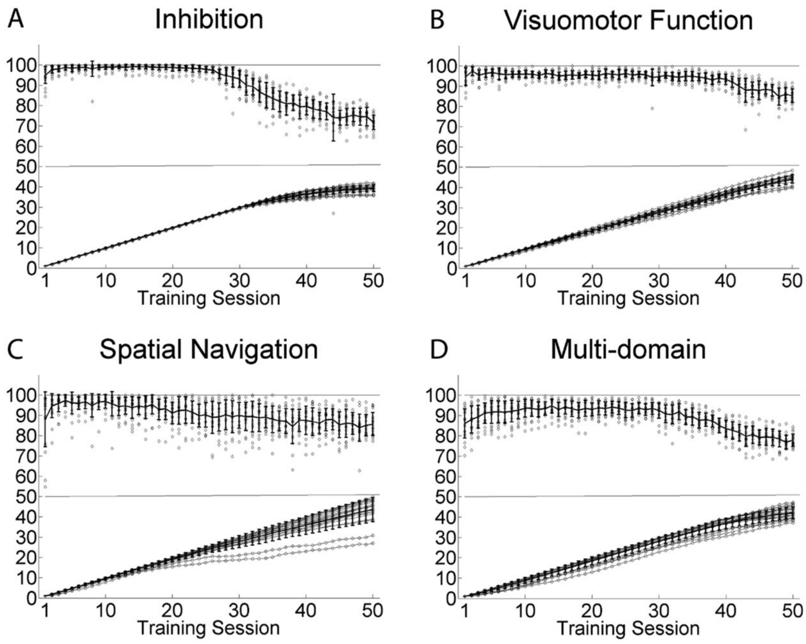

Figure 2. Training curves for the four training groups. Group means and standard deviations of level (lower curve) and percent correct (upper curve) are displayed for each participant's mean of all five minigames per training session. Level ranged from 1–50, all participants started with level 1. Participants could increase, decrease, or maintain the level in the subsequent training session based on performance (percent correct) of the previous training session. Percentage correct of performance ranged from 0–100%.

Table 3
Means and Standard Deviations for the Composite Scores and for the Individual Tests of Each Composite Score for Baseline, Posttest, and Follow-Up Measurements

	Multi-domain training			Single-domain training		
Composite/Test	Baseline	Posttest	Follow-up	Baseline	Posttest	Follow-up
Attention	.00 (.75)	.56 (.65)	.72 (.63)	.01 (.77)	.32 (.74)	.48 (.79)
Trail making B	*−.13 (1.50)*	*.45 (.71)*	*.65 (.59)*	*.04 (1.16)*	*.24 (.91)*	*.28 (.96)*
D2	*.15 (.98)*	*.83 (.77)*	*1.11 (.94)*	*−.05 (1.05)*	*.39 (.99)*	*.81 (1.03)*
Divided attention	*.09 (.86)*	*.49 (.71)*	*.47 (.72)*	*−.03 (1.05)*	*.31 (.98)*	*.19 (1.16)*
Flexibility	*−.09 (1.01)*	*.50 (1.34)*	*.65 (1.01)*	*.03 (.97)*	*.34 (.93)*	*.64 (.98)*
Working memory	−.05 (.51)	.28 (.64)	.10 (.62)	.02 (.69)	.16 (.67)	.22 (.78)
2-back	*−.15 (1.04)*	*.15 (1.07)*	*−.05 (.99)*	*.05 (.95)*	*.26 (1.02)*	*.28 (1.00)*
Digit span backward	*−.19 (.75)*	*.05 (.94)*	*.06 (1.26)*	*.06 (.99)*	*−.08 (.91)*	*.14 (1.13)*
Corsi block backward	*.20 (.89)*	*.64 (1.01)*	*.28 (1.21)*	*−.07 (.89)*	*.24 (.95)*	*.24 (1.13)*
Speed	.03 (.68)	.25 (.51)	.53 (.78)	−.01 (.92)	.27 (.86)	.35 (.94)
Trail making A	*.06 (1.11)*	*.28 (.67)*	*.50 (.86)*	*−.02 (1.09)*	*.16 (.94)*	*.25 (1.09)*
Digit symbol	*−.01 (.84)*	*.23 (.73)*	*.56 (.84)*	*.00 (1.05)*	*.39 (1.08)*	*.44 (1.05)*
	Multi-domain training			Inhibition training		
	Baseline	Posttest	Follow-up	Baseline	Posttest	Follow-up
Inhibition						
Stop signal	*.17 (.91)*	*.38 (1.06)*	*.81 (.57)*	*.12 (1.05)*	*1.30 (.82)*	*1.10 (.72)*
Stroop	*.07 (1.02)*	*.42 (1.08)*	*−.37 (.89)*	*−.36 (1.17)*	*−.09 (.93)*	*.03 (.91)*
	Multi-domain training			Visuomotor function training		
	Baseline	Posttest	Follow-up	Baseline	Posttest	Follow-up
Visuomotor function	−.02 (.61)	.23 (.75)	.38 (.86)	.05 (.60)	.11 (.66)	.16 (.72)
Aiming	*−.05 (.88)*	*.42 (1.06)*	*.53 (1.05)*	*.08 (.69)*	*.32 (1.24)*	*.58 (1.11)*
Steadiness	*−.21 (.75)*	*.20 (1.07*	*.26 (1.07)*	*.09 (.81)*	*−.01 (.82)*	*.03 (1.10)*
Line drawing	*.22 (.79)*	*.06 (.87)*	*.36 (1.17)*	*.00 (1.06)*	*.02 (.72)*	*−.13 (.87)*
	Multi-domain training			Spatial navigation training		
	Baseline	Posttest	Follow-up	Baseline	Posttest	Follow-up
Spatial navigation	.10 (.86)	.29 (.82)	.65 (.72)	.03 (.65)	.15 (.78)	.26 (.71)
Mental rotation	*.33 (1.12)*	*.60 (1.30)*	*.96 (1.21)*	*−.04 (.73)*	*.32 (.81)*	*.44 (.85)*
Map learning	*−.09 (1.18)*	*.06 (1.01)*	*.56 (1.01)*	*.11 (1.06)*	*.29 (1.23)*	*.00 (.98)*
Corsi block forward	*.05 (1.06)*	*.23 (1.00)*	*.41 (1.10)*	*.02 (1.01)*	*−.16 (1.17)*	*.36 (.79)*

Note. Standardized scores for the composites and the individual variables (tests) of each composite (italic). Standard deviations are in parentheses. Single-domain training refers to the three training groups inhibition, visuomotor function, and spatial navigation.

We evaluated whether training resulted in group differences in change from baseline to posttest with a latent difference score model in a sequential manner, starting with the just identified model and moving to a series of nested models with constrained means, variances, and covariances across groups unless a constraint significantly reduced model fit. Model fits of the final models for each composite score are shown in Table 4, parameter estimates of the mean change scores and the correlations are shown in Table 5. This set of analysis answered the specific question of whether multi-domain training shows more or less benefits than single-domain training both in terms of near and far transfer.

Multi-domain versus single-domain training for the trained domains (near transfer). We first tested whether multi-domain training resulted in different performance gains on the trained domains compared to the single-domain training group that trained the particular function exclusively (e.g., difference in the change score for the multi-domain vs. visuomotor function training groups on the visuomotor function composite).

Constraining the change score of the stop signal inhibition test from baseline to posttest resulted in a significant reduction of model fit ($\Delta\chi^2 = 8.88, p < .01$). Only the inhibition training group showed improved performance (change score of the inhibition training group: $M = 1.16$, $SE = .25$, $p < .001$, $d = 1.21$; change score of multi-domain training group: $M = .24$, $SE = .25$, $p = .341$, $d = .25$; effect size for the group difference in change: $d = .96$). Follow-up analyses including the no-contact control group revealed a baseline difference indicating that the no-contact control group performed significantly better on this test (see supplementary information Tables A3, A12–A13). Consequently, we do not interpret the parameters. In contrast, we did not find any group differences in performance change from baseline to posttest for the other two near transfer measures. Hence, constraining the change score to be equal across groups did not result in significant reductions of model fit: The visuomotor function and the multi-domain training group improved similarly on visuomotor function (change score of visuomotor function independent of group: $M = .15$,

Table 4
Model Fits for the Final Models After Constraining All Parameters Across Groups That Did Not Result in a Significant Reduction of Model Fit

Final model	χ^2	*df*	CFI	RMSEA (90% CI)
Attention	4.12	8	1.00	.00 (.00–.07)
Working memory	6.29	9	1.00	.00 (.00–.09)
Speed	11.55	7	.97	.09 (.00–.18)
Stop signal (inhibition test)	4.22	5	1.00	.00 (.00–.20)
Visuomotor function	6.64	9	1.00	.00 (.00–.14)
Spatial navigation	5.46	9	1.00	.00 (.00–.12)

Note. CFI values above .95 and RMSEA values below .06 indicate that a model is adequately parameterized and reflect good model fit. Values for CFI above .90 and for RMSEA below .08 are also acceptable.

SE = .07, p = .039, d = .25). Likewise, the spatial navigation and the multi-domain training group showed a statistical trend for improvement on spatial navigation (change score of spatial navigation independent of group: M = .16, SE = .09, p = .091, d = .21). These group-independent changes from baseline to posttest did not differ from the changes in the no-contact control group (no significant reduction of model fit when constraining the change score from baseline to posttest across the two training groups and the no-contact control group for the visuomotor function composite: $\Delta\chi^2$ = 0.54 and for the spatial navigation composite $\Delta\chi^2$ = 1.36; for model fits and parameters see supplementary information Tables A12 and A13).

Multi-domain versus single-domain training for far transfer. Next, we tested whether multi-domain training resulted in greater performance gains on executive transfer tasks compared to the mean of the three single domain trainings as reflected in a higher change score in attentional control, working memory, and processing speed. Constraining the change score of the attentional control composite from baseline to posttest to be equal across groups resulted in a significant reduction of model fit ($\Delta\chi^2$ = 6.11, $p < .05$). The multi-domain training group showed higher performance increases on the attentional control composite (M = .55, SE = .08, $p < .001$, d = .74) compared with the single-domain training groups (M = .31, SE = .05, $p < .001$, d = .42; effect size for the group difference in change: d = .32). Follow-up analyses considering the no-contact control group revealed the same result pattern (see supplementary information Tables A12–A13). The multi-domain training group showed significantly higher performance increases on the attentional control composite (significant reduction of model fit when constraining the change score from baseline to posttest; $\Delta\chi^2$ = 5.49, $p < .05$). However, the change score of the single-domain training groups did not differ from the change score of the no-contact control group ($\Delta\chi^2$ = 0.22). With regard to the composites of working memory and speed, we did not find any group differences in performance change from baseline to posttest. All groups showed similar performance increases after training of small effect sizes (working memory: M = .19, SE = .07, p = .005, d = .30; speed: M = .27, SE = .06, $p < .001$, d = .31). These group-independent performance increases did not differ from the increases of the no-contact control group (no significant decreases of model fit when constraining the change score across the training groups and the no-contact control group for speed: $\Delta\chi^2$ = 2.22; and for working memory: $\Delta\chi^2$ = 0.02; see supplementary information Tables A12–A13).

Stability of Performance Six Months After Training

According to the hypothesis that multi-domain training has a higher probability of a functional overlap between training and transfer, we expected the multi-domain training group to show better maintenance based on the assumption that the trained processes may be applied to everyday life during the six months after training.

Stability of performance on training tasks. In a 2 × 2 mixed ANOVA with the within-group factor Time (training Session 50, follow-up) and the between-group factor Training (multi-domain training, single-domain training), percentage of performance at indi-

Table 5
Model Parameters for the Means of the Change Scores and the Correlations

Composite	Training group	Mean change 1 *E.* (*SE*)	Mean change 2 *E.* (*SE*)	Corr. T1–change 1	Corr. T1–change 2	Corr. change 1–change 2
Attention	Multi-domain	.55 (.08)***	.12 (.04)**	−.36**	.04	−.31*
	Single-domain	.31 (.05)***				
Working memory	Multi-domain	.19 (.07)**	−.01 (.07)	−.45***	.15	−.53***
	Single-domain					
Speed	Multi-domain	.27 (.06)***	.10 (.07)	−.80***	.14	−.23**
	Single-domain			−.41**		
Stop signal (inhibition test)	Multi-domain	.24 (.25)	.44 (.20)*	−.70***	.11	−.67***
	Inhibition	1.16 (.25)***	−.22 (.19)		−.31	−.25
Visuomotor function	Multi-domain	.15 (.07)*	.14 (.07)*	−.15	−.02	−.13
	Visuomotor function					
Spatial navigation	Multi-domain	.16 (.09)†	.21 (.09)*	−.32†	−.12	−.48**
	Spatial navigation					

Note. *E.* = estimate; *SE* = standard error; Corr. = correlation (standardized covariance); single-domain = mean across inhibition, visuomotor function, and spatial navigation training; T1 = baseline; change 1 = change from baseline to posttest; change 2 = change from posttest to follow-up. Parameter estimates are provided for the final models. When groups differed significantly, parameters are provided for both groups, otherwise parameters are constrained across training groups. Correlation coefficients differed in value (not in significance) when the variances were not the same in both groups (e.g., speed). In those cases we report only the correlation for the multi-domain group.
† ≤ .09. * $p < .05$. ** $p < .01$. *** $p < .001$.

vidual training end level decreased in all groups as indicated by a main effect of Time, $F(1, 69) = 119.40$, $p < .001$, $\eta_p^2 = .63$. There was no interaction effect, $F(1, 69) = .64$, $p = .428$, $\eta_p^2 = .01$. Hence, performance in the multi-domain training group did not decrease less than performance in the single-domain training groups (mean performance difference multi-domain: 11.93; single-domain: 10.30). A 2 × 4 mixed ANOVA with all training groups as between-group factor indicated a statistical trend for an interaction of Time × Training group, such that the spatial navigation group showed the smallest performance decrease, $F(3, 67) = 2.42$, $p = .073$, $\eta_p^2 = .10$ (for means of training performance see Table 2).

Stability of performance on transfer tasks. To assess stability of improvements six months after training, we tested to what extent performance changed from posttest to follow-up. Constraining the change score of posttest to follow-up to be equal across groups did not result in a significant reduction of model fit in any of the above described models except for the stop signal inhibition task. With regard to this test, constraining the change score from posttest to follow-up to be equal across the inhibition and the multi-domain training group significantly reduced model fit ($\Delta\chi^2 = 4.68$, $p < .05$). The multi-domain training group improved significantly from posttest to follow-up ($M = .44$, $SE = .20$, $p = .029$, $d = .34$), whereas the inhibition training group remained stable ($M = -.22$, $SE = .19$, $p = .261$, $d = -.17$; effect size for group difference in change: $d = .51$). In contrast, there were no differential group effects for the other two near transfer composites. We found a significant change of visuomotor function performance, such that both the visuomotor function and the multi-domain group increased performance from posttest to follow-up equally (change score of visuomotor function performance independent of group: $M = .14$, $SE = .07$, $p = .039$, $d = .31$). A similar pattern was found for spatial navigation, indicating that the spatial navigation and multi-domain training groups significantly increased spatial navigation performance from posttest to follow-up ($M = .21$, $SE = .09$, $p = .015$, $d = .36$). With regard to far transfer, there was only one significant group-independent change on the attentional control composite ($M = .12$, $SE = .04$, $p = .006$, $d = .28$), while performance on the working memory ($M = -.01$, $SE = .07$, $p = .898$, $d = -.01$) and the speed composite ($M = .10$, $SE = .07$, $p = .134$, $d = .14$) did not change from posttest to follow-up.

Individual Differences in Baseline Performance and Training-Related Change

In our structural equation models, we found significant interindividual differences indicated by significant variances at baseline, for the estimated latent difference from baseline to posttest, and the latent difference from posttest to follow-up. This pattern held true for all composite measures independent of group (exception: variance of speed for the change score from baseline to posttest revealed higher variability in the multi-domain than in the single-domain groups, $\Delta\chi^2 = 6.83$, $p < .05$). Furthermore, we found a consistent pattern such that participants with lower baseline performance improved more through training as indicated by significant negative correlations of baseline performance with the change score from baseline to posttest (see Table 5). There were two exceptions: the negative correlations did not reach significance in the visuomotor function and spatial navigation models. Furthermore, there was a significant group difference in correlations between the multi-domain and the single-domain training groups for the speed composite ($\Delta\chi^2 = 9.05$, $p < .01$; multi-domain training group: $r = -.80$, $p < .001$; single-domain training groups: $r = -.41$, $p = .002$). Consequently, initially lower performing individuals in the multi-domain and the single-domain training could increase their speed performance more through training, and this pattern was significantly stronger in the multi-domain group. Moving to the correlation of the two change scores, constraining the correlations of the stop signal inhibition change score from baseline to posttest with the one from posttest to follow-up to be equal across the inhibition and the multi-domain training groups resulted in a significant reduction of model fit ($\Delta\chi^2 = 5.28$, $p < .05$). The correlation was not significant in the inhibition group, $r = -.25$, $p = .210$, although it was significant in the multi-domain training group, $r = -.67$, $p < .001$. This indicated that the greater the improvement from baseline to posttest, the smaller the change from posttest to follow-up. In addition, there was a significant group difference in the correlations from baseline to the change scores from posttest to follow-up ($\Delta\chi^2 = 4.01$, $p < .05$), although the correlations in both groups (inhibition, multi-domain) did not reach significance (see Table 5).

Discussion

In the present study, we showed that simultaneous multi-domain training of cognitive domains that are key ingredients of cognitive functioning, namely inhibition, visuomotor function, and spatial navigation, showed *far transfer* to quite different cognitive tasks tapping into attentional control. *Near transfer* effects in terms of increases of performance on the trained functions were group-independent, however, and did not exceed retest effects assessed with an additional no-contact control group. An exception was the inhibition training group who increased performance on the stop signal inhibition task compared to the multi-domain training group. Furthermore, there was evidence for reliable interindividual differences in intraindividual transfer gains in that participants with lower initial performance generally improved more through training. At the 6-month follow-up, there were no other differential maintenance effects, both the multi-domain and the single-domain training groups maintained performance to comparable degrees. The only exception was the stop signal inhibition task where we found a group difference in change: The multi-domain training group improved from posttest to follow-up, whereas the inhibition training group remained stable.

Identifying the Processes Underlying Multi-Domain Training Interventions

To our knowledge, this is the first study that systematically compared the effects of a simultaneous multi-domain training of three different cognitive functions to the training of each individual function (single-domain training). We assumed that training three domains is qualitatively different from training two domains with respect to the imposed flexibility demands. Whereas a multi-domain training targeting two cognitive functions simultaneously allows only two possibilities for switching back and forth (e.g., switching back and forth between the visual tracking and signal detection task; Anguera et al., 2013), the simultaneous combination of three cognitive functions allows six possibilities for switching back and forth. The mechanism for the far transfer to attentional control induced by the present multi-domain training regime might well be explained by its increased flexibility demands. The multi-domain training participants had to switch between inhibition, spatial navigation, and visuomotor function. Previous multi-domain training studies with video game training, for example, did not allow inference about the exact training content. Hence, the mechanisms of transfer were hardly identifiable although these studies were promising with respect to cognitive improvements in older adults (for a meta-analysis, see Toril et al., 2014). An exception was the training study with the custom-designed video game Neuroracer targeting visuomotor tracking and signal detection (Anguera et al., 2013).

With regard to the cognitive functions targeted by the training, we selected inhibition based on the known deficits during aging and its key function in working memory (Hasher et al., 2007; Hasher & Zacks, 1988). The selection of spatial navigation was based on its importance in everyday life functioning and dependency on hippocampal functioning (Moffat, 2009; Wolbers & Hegarty, 2010), and the selection of visuomotor function on the dedifferentiation hypothesis (Baltes & Lindenberger, 1997; Lindenberger & Baltes, 1994). Only the inhibition training group showed near transfer to the stop signal inhibition task. Research on inhibition training in old age is sparse (Buitenweg, Murre, & Ridderinkhof, 2012; Strobach et al., 2014), and it has been difficult to show transfer. Our results should be taken with caution because we could not build an inhibition composite. The absence of other near transfer effects raises the question to what extent training the orchestration of several cognitive functions is independent of the particular cognitive functions trained. Future studies combining different cognitive functions in a way that they are still identifiable will further shed light on multi-domain transfer mechanisms. Furthermore, intensively training individual cognitive functions might not be the most promising approach for older adults. Because cognitive aging is a complex process including declines and maintenance of various cognitive functions (de Frias, Lövdén, Lindenberger, & Nilsson, 2007; Hedden & Gabrieli, 2004; Park & Reuter-Lorenz, 2009), the ability to orchestrate these functions flexibly might be a key for stable mental functioning. This orchestration can consist of switching, sequencing, coordinating, or synchronizing.

Interindividual Differences in Intraindividual Training Effects

The structural equation modeling approach allowed us to take into account individual differences in baseline performance and relate them to training-related changes in the cognitive functions assessed with the transfer test battery. We found a pattern that fitted the compensation account proposed by Lövdén, Brehmer et al. (2012): Initially lower performing participants showed higher performance improvements through training. According to this account, the compensation pattern emerges when training fosters flexibility (optimization within available cognitive resources) rather than inducing plastic changes (expansion of currently available cognitive resources). As shown in other studies, plastic changes could have been expected considering the intensity of our training regime (see, e.g., S. B. Chapman et al., 2015; Kühn et al., 2014; Lövdén, Schaefer, et al., 2012). However, we cannot draw conclusions about plastic brain changes since we did not include neuroimaging to assess structural brain changes. The multi-domain training condition might well have fostered flexibility by demanding the simultaneous administration of three tasks that had to be kept in mind and required quick task set shifts rather than maximizing only one cognitive function. This is supported by the transfer to the attentional control composite. Furthermore, magnification effects have rarely been reported and pertained mainly to the memory domain (Lövdén, Brehmer, et al., 2012; Verhaeghen & Marcoen, 1996). It is possible that such a pattern only emerges when training demands high cognitive effort from the beginning, thereby putting individuals with lower cognitive ability at a disadvantage. The participants in our study were highly functioning with a good cognitive and health status, high average crystallized intelligence, and high levels of education. Our somewhat selective sample of participants probably entered the study with a high level of cognitive resources, making it more difficult to create the "demand-supply mismatch" necessary for the induction of plastic changes (Lövdén, Bäckman, Lindenberger, Schaefer, & Schmiedek, 2010; Lövdén, Brehmer, et al., 2012). An adaptive level to start training based on baseline performance or steeper adjustments could have further increased training demands, thereby bringing high-performing participants to their individual performance limits faster (Kliegl, Smith, & Baltes, 1989).

Maintenance of Training Effects

At the 6-month follow-up, there were no differential training effects on the transfer test battery (except for the stop signal inhibition test) and we do not have retest data for this third measurement time point. Independent of the training conditions, all groups showed maintained performance and sometimes even improved performance from posttest to follow-up. Interestingly, the multi-domain group did not differ from the single-domain training groups on the attentional control composite at the 6-month follow-up, which could have been expected if multi-domain training transferred to everyday life because of the overlapping demands of multitasking. However, termination after multi-domain training did not appear to differentially facilitate maintenance of these improvements. One could speculate that the training was not sufficiently applicable to, or did not imitate the demands of everyday life. This is in line with findings from other studies. Direct transfer to everyday life has hardly ever been shown (but see, e.g., Ball, Edwards, & Ross, 2007; for a meta-analysis, see Kelly et al., 2014).

What Would Be the Ideal Multi-Domain Training Setup?

An important factor to consider for the construction of comparable multi-domain and single-domain training is the complexity and controllability of the trained functions for a better understanding of the processes underlying the observed training and transfer effects. There is usually a trade-off between the amount of training spent on each domain in a multi-domain training condition and the number of training trials for each domain (Strobach et al., 2014). Comparing multi-domain and single-domain training and thereby holding the total amount of training time constant across these conditions, single-domain training trains the targeted function more intensely (e.g., see simultaneous vs. sequential dual-tasking; Anguera et al., 2013; or sequential multi-domain vs. single-domain training; Cheng et al., 2012). This can (partly) be overcome by simultaneous multi-domain training, although pure simultaneous conditions are difficult to construct. The advantage of simultaneous training of several cognitive functions is the additional training of higher order executive functions needed to coordinate the different individual tasks (Strobach et al., 2012; Strobach et al., 2014). Although it is assumed that the simultaneous training does not necessarily improve the single-domains maximally, but rather improves the single-domains equally and coordination skills in addition, it has been proposed that a maximal training effect can be achieved by a combination of dual-task training and training of each single task component (for a discussion see Strobach et al., 2014). Adapting the Hotel Plastisse training, this could potentially be investigated by combining multi-domain and single-domain training tasks. Furthermore, a training regime that allows a parametric modulation of the number of cognitive functions combined could possibly give insights into this matter. Another follow-up question is whether combining certain cognitive functions leads to interaction effects. Are there particular combinations of cognitive functions that facilitate or hamper transfer? Because multi-domain training has targeted very different cognitive functions and most of them do not allow inference about the particular cognitive functions trained (e.g., video game training), it is largely unknown to which training aspect transfer can be attributed (see discussion in Binder et al., 2015; Karbach, 2014; Winocur et al., 2007).

Our training regime with 50 training sessions of about 45 to 60 min was intense. Recent meta-analyses (Karbach & Verhaeghen, 2014; Lampit, Ebster, & Valenzuela, 2014; Toril et al., 2014) have found mixed results concerning optimal training duration. Although it is assumed that a sustained demand-supply mismatch is required for training-induced plastic changes in the brain (Lövdén et al., 2010), a meta-analysis by Karbach and Verhaeghen (2014) did not find a dose-response relationship of working memory and executive function training duration and transfer. Similarly, a meta-analysis of physical and cognitive training in older adults did not find treatment effects to be associated with treatment duration, session duration, and session frequency (Karr, Areshenkoff, Rast, & Garcia-Barrera, 2014). In contrast, Toril et al. (2014) found shorter video game training studies to be more effective. Future studies should provide insights into the progression of plasticity by manipulating training duration, the duration of each single training session, and optimal spacing (see also Lampit, Hallock, et al., 2014).

Limitations

Including several control conditions demands large sample sizes. There is often a trade-off between the number of training and control conditions to disentangle the mechanisms of training and the effort, time, and costs to recruit and support an adequate number of trainees. Our primary interest was in the comparison of multi-domain and single-domain training to investigate differential training effects. These comparisons were quite conservative because all training groups underwent an intensive training regime. However, we thought that these comparisons best control for training-unspecific effects, such as participants' expectations (Green, Strobach, & Schubert, 2014). Nevertheless, we also assessed retest data with a comparable no-contact control group that performed on the cognitive test battery twice with an interval comparable to the training regime. This no-contact control group did not do any control activities during this interval and was not originally randomized in the training study. The small sample size of approximately 20 subjects per training condition restricted power. Given that we found a training-related group difference on the attentional control composite, the effect size that we found was likely in the lower boundary. However, we possibly lack power to detect other effects, especially for effects at the 6-month follow-up because of additional dropouts. In addition, a bigger sample size would have allowed an estimation of transfer at the latent level, a step that is important for studying cognitive training (Noack et al., 2014). Unfortunately, when estimating our transfer abilities at a latent level, our latent difference score model estimations were not reliable with only 20 participants per group. Consequently, our composite scores were not error-free and we could not test for measurement invariance across time (Bellander et al., 2015; Miyake & Friedman, 2012; Schmiedek et al., 2010). In future cognitive training studies, larger sample sizes are needed to allow for examination of transfer constructs at a latent level. Examining not only transfer at a latent level, but also training progress would allow the investigation of how intraindividual training trajectories relate to interindividual differences in transfer (Könen & Karbach, 2015; Schmiedek et al., 2010; Zelinski et al., 2014). Moderators such as motivation, emotion, personality, or health variables could then also be included to unveil possible mechanisms of transfer (e.g., Jaeggi, Buschkuehl, Shah, & Jonides, 2014).

Conclusion

Our results suggest that multi-domain training enhances functions that involve handling several different tasks at the same time, which closely mimics typical everyday challenges, especially for older people. We extended the literature of existing multi-domain training studies using video game training by applying a training regime that offers more control over the trained functions. Hence, we could better relate training to transfer based on theoretically involved underlying processes. More studies are needed to systematically investigate how multi-domain training in healthy old age relates to transfer, and neuroimaging can further shed light on the mechanisms of the relationship between training and transfer.

References

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., . . . Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature, 501,* 97–101. http://dx.doi.org/10.1038/nature12486

Anguera, J. A., & Gazzaley, A. (2015). Video games, cognitive exercises, and the enhancement of cognitive abilities. *Current Opinion in Behavioral Sciences, 4,* 160–165. http://dx.doi.org/10.1016/j.cobeha.2015.06.002

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . . for the ACTIVE Study Group. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *JAMA: Journal of the American Medical Association, 288,* 2271–2281. http://dx.doi.org/10.1001/jama.288.18.2271

Ball, K., Edwards, J. D., & Ross, L. A. (2007). The impact of speed of processing training on cognitive and everyday functions. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 62B,* 19–31. http://dx.doi.org/10.1093/geronb/62.special_issue_1.19

Ballesteros, S., Kraft, E., Santana, S., & Tziraki, C. (2015). Maintaining older brain functionality: A targeted review. *Neuroscience and Biobehavioral Reviews, 55,* 453–477. http://dx.doi.org/10.1016/j.neubiorev.2015.06.008

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 12,* 12–21. http://dx.doi.org/10.1037/0882-7974.12.1.12

Bamidis, P. D., Vivas, A. B., Styliadis, C., Frantzidis, C., Klados, M., Schlee, W., . . . Papageorgiou, S. G. (2014). A review of physical and cognitive interventions in aging. *Neuroscience and Biobehavioral Reviews, 44,* 206–220. http://dx.doi.org/10.1016/j.neubiorev.2014.03.019

Bellander, M., Bäckman, L., Liu, T., Schjeide, B. M., Bertram, L., Schmiedek, F., . . . Lövdén, M. (2015). Lower baseline performance but greater plasticity of working memory for carriers of the val allele of the COMT Val^{158}Met polymorphism. *Neuropsychology, 29,* 247–254. http://dx.doi.org/10.1037/neu0000088

Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Becic, E. (2005). Training effects on dual-task performance: Are there age-related differences in plasticity of attentional control? *Psychology and Aging, 20,* 695–709. http://dx.doi.org/10.1037/0882-7974.20.4.695

Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Becic, E. (2008). Transfer effects in task-set cost and dual-task cost after dual-task training in older and younger adults: Further evidence for cognitive plasticity in attentional control in late adulthood. *Experimental Aging Research, 34,* 188–219. http://dx.doi.org/10.1080/03610730802070068

Binder, J. C., Zöllig, J., Eschen, A., Mérillat, S., Röcke, C., Schoch, S. F., . . . Martin, M. (2015). Multi-domain training in healthy old age: Hotel Plastisse as an iPad-based serious game to systematically compare multi-domain and single-domain training. *Frontiers in Aging Neuroscience, 7,* 137. http://dx.doi.org/10.3389/fnagi.2015.00137

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols, 1,* 2277–2281. http://dx.doi.org/10.1038/nprot.2006.390

Bratfisch, O., & Hagmann, E. (2012). *Räumliches Orientierungsvermögen (3D). Wiener Testsystem (WTS)* [3D spatial orientation: Vienna Test System]. Mödling, Austria: Schuhfried.

Brehmer, Y., Li, S.-C., Müller, V., von Oertzen, T., & Lindenberger, U. (2007). Memory plasticity across the life span: Uncovering children's latent potential. *Developmental Psychology, 43,* 465–478. http://dx.doi.org/10.1037/0012-1649.43.2.465

Brickenkamp, R., Schmidt-Atzert, L., & Liepmann, D. (2010). *Test d2-R. Aufmerksamkeits-Belastungs-Test* [d2-R test of attention]. Göttingen, Germany: Hogrefe.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21,* 230–258. http://dx.doi.org/10.1177/0049124192021002005

Buitenweg, J. I. V., Murre, J. M. J., & Ridderinkhof, K. R. (2012). Brain training in progress: A review of trainability in healthy seniors. *Frontiers in Human Neuroscience, 6,* 183. http://dx.doi.org/10.3389/fnhum.2012.00183

Bürki, C. N., 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,* 821–835. http://dx.doi.org/10.1007/s00426-014-0559-3

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience, 2,* S167–S179. http://dx.doi.org/10.1016/j.dcn.2011.10.001

Cauley, J. A. (2012). The demography of aging. In A. B. Newman & J. A. Cauley (Eds.), *The epidemiology of aging* (pp. 3–14). Dordrecht, The Netherlands: Springer. http://dx.doi.org/10.1007/978-94-007-5061-6_1

Chambers, C. D., Garavan, H., & Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neuroscience and Biobehavioral Reviews, 33,* 631–646. http://dx.doi.org/10.1016/j.neubiorev.2008.08.016

Chapman, L. J., & Chapman, J. P. (1987). The measurement of handedness. *Brain and Cognition, 6,* 175–183. http://dx.doi.org/10.1016/0278-2626(87)90118-7

Chapman, S. B., Aslan, S., Spence, J. S., Hart, J. J., Jr., Bartz, E. K., Didehbani, N., . . . Lu, H. (2015). Neural mechanisms of brain plasticity with complex cognitive training in healthy seniors. *Cerebral Cortex, 25,* 396–405. http://dx.doi.org/10.1093/cercor/bht234

Cheng, Y., Wu, W., Feng, W., Wang, J., Chen, Y., Shen, Y., . . . Li, C. (2012). The effects of multi-domain versus single-domain cognitive training in non-demented older people: A randomized controlled trial. *BMC Medicine, 10,* 30. http://dx.doi.org/10.1186/1741-7015-10-30

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34*(02), 891B.

Craik, F. I. M., Winocur, G., Palmer, H., Binns, M. A., Edwards, M., Bridges, K., . . . Stuss, D. T. (2007). Cognitive rehabilitation in the elderly: Effects on memory. *Journal of the International Neuropsychological Society, 13,* 132–142. http://dx.doi.org/10.1017/S1355617707070166

Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320,* 1510–1512. http://dx.doi.org/10.1126/science.1155466

de Frias, C. M., Lövdén, M., Lindenberger, U., & Nilsson, L.-G. (2007). Revisiting the dedifferentiation hypothesis with longitudinal multicohort data. *Intelligence, 35,* 381–392. http://dx.doi.org/10.1016/j.intell.2006.07.011

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Gauggel, S., & Birkner, B. (1999). Validität und Reliabilität einer deutschen Version der Geriatrischen Depressionsskala (GDS) [Validity and reliability of a German version of the Geriatric Depression Scale (GDS)]. *Zeitschrift für Klinische Psychologie, 28,* 18–27. http://dx.doi.org/10.1026//0084-5345.28.1.18

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging, 23,* 692–701. http://dx.doi.org/10.1037/a0014345

Green, C. S., Strobach, T., & Schubert, T. (2014). On methodological standards in training and transfer experiments. *Psychological Research, 78,* 756–772. http://dx.doi.org/10.1007/s00426-013-0535-3

Härting, C., Markowitsch, H., Neufeld, H., Calabrese, P., Deisinger, K., & Kessler, J. (2000). *WMS-R Wechsler Gedächtnistest - Revidierte Fassung* [WMS-R Wechsler Memory Scale - Revised]. Bern, Switzerland: Hans Huber.

Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York, NY: Oxford University Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York, NY: Academic Press. http://dx.doi.org/10.1016/S0079-7421(08)60041-9

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience, 5,* 87–96. http://dx.doi.org/10.1038/nrn1323

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition, 42,* 464–480. http://dx.doi.org/10.3758/s13421-013-0364-z

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test: BIS-Test; Form 4; mit einem separat verwendbaren Kurztest der allgemeinen Intelligenz und der Verarbeitungskapazität* [Berlin intelligence structure test: BIS-Test. Form 4 with a separate short test for general intelligence and processing capacity]. Göttingen, Germany: Hogrefe.

Jonides, J. (2004). How does practice makes perfect? *Nature Neuroscience, 7,* 10–11. http://dx.doi.org/10.1038/nn0104-10

Kaiser, S., Aschenbrenner, S., Pfüller, U., Roesch-Ely, D., & Weisbrod, M. (2012). *Response inhibition (INHIB). Wiener Testsystem (WTS).* Mödling, Austria: Schuhfried.

Karbach, J. (2014). Game-based cognitive training for the aging brain. *Frontiers in Psychology, 5,* 1100. http://dx.doi.org/10.3389/fpsyg.2014.01100

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12,* 978–990. http://dx.doi.org/10.1111/j.1467-7687.2009.00846.x

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological Science, 25,* 2027–2037. http://dx.doi.org/10.1177/0956797614548725

Karr, J. E., Areshenkoff, C. N., Rast, P., & Garcia-Barrera, M. A. (2014). An empirical comparison of the therapeutic benefits of physical exercise and cognitive training on the executive functions of older adults: A meta-analysis of controlled trials. *Neuropsychology, 28,* 829–845. http://dx.doi.org/10.1037/neu0000101

Kelly, M. E., Loughrey, D., Lawlor, B. A., Robertson, I. H., Walsh, C., & Brennan, S. (2014). The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: A systematic review and meta-analysis. *Ageing Research Reviews, 15,* 28–43. http://dx.doi.org/10.1016/j.arr.2014.02.004

Kivipelto, M., Solomon, A., Ahtiluoto, S., Ngandu, T., Lehtisalo, J., Antikainen, R., . . . Soininen, H. (2013). The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): Study design and progress. *Alzheimer's & Dementia, 9,* 657–665. http://dx.doi.org/10.1016/j.jalz.2012.09.012

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology, 25,* 247–256. http://dx.doi.org/10.1037/0012-1649.25.2.247

Könen, T., & Karbach, J. (2015). The benefits of looking at intraindividual dynamics in cognitive training data. *Frontiers in Psychology, 6,* 615.

Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., & Gallinat, J. (2014). Playing Super Mario induces structural brain plasticity: Gray matter changes resulting from training with a commercial video game. *Molecular Psychiatry, 19,* 265–271. http://dx.doi.org/10.1038/mp.2013.120

Kuwajima, M., & Sawaguchi, T. (2010). Similar prefrontal cortical activities between general fluid intelligence and visuospatial working memory tasks in preschool children as revealed by optical topography. *Experimental Brain Research, 206,* 381–397. http://dx.doi.org/10.1007/s00221-010-2415-z

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4,* 863. http://dx.doi.org/10.3389/fpsyg.2013.00863

Lampit, A., Ebster, C., & Valenzuela, M. (2014). Multi-domain computerized cognitive training program improves performance of bookkeeping tasks: A matched-sampling active-controlled trial. *Frontiers in Psychology, 5,* 794. http://dx.doi.org/10.3389/fpsyg.2014.00794

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: A systematic review and meta-analysis of effect modifiers. *PLoS Medicine, 11,* e1001756. http://dx.doi.org/10.1371/journal.pmed.1001756

Lehrl, S. (2005). Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B [Multiple choice vocabulary test of intelligence] (5th ed.). Balingen, Germany: Spitta.

Levine, B., Stuss, D. T., Winocur, G., Binns, M. A., Fahy, L., Mandic, M., . . . Robertson, I. H. (2007). Cognitive rehabilitation in the elderly: Effects on strategic behavior in relation to goal management. *Journal of the International Neuropsychological Society, 13,* 143–152. http://dx.doi.org/10.1017/S1355617707070178

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9,* 339–355. http://dx.doi.org/10.1037/0882-7974.9.3.339

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin, 136,* 659–676. http://dx.doi.org/10.1037/a0020080

Lövdén, M., Brehmer, Y., Li, S.-C., & Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Frontiers in Human Neuroscience, 6,* 141. http://dx.doi.org/10.3389/fnhum.2012.00141

Lövdén, M., Schaefer, S., Noack, H., Bodammer, N. C., Kühn, S., Heinze, H.-J., . . . Lindenberger, U. (2012). Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. *Neurobiology of Aging, 33,* 620.e9–620.e22. http://dx.doi.org/10.1016/j.neurobiolaging.2011.02.013

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review, 19,* 504–522. http://dx.doi.org/10.1007/s11065-009-9119-9

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203. http://dx.doi.org/10.1037/0033-2909.109.2.163

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49,* 270–291. http://dx.doi.org/10.1037/a0028228

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21,* 8–14. http://dx.doi.org/10.1177/0963721411429458

Moffat, S. D. (2009). Aging and spatial navigation: What do we know and where do we go? *Neuropsychology Review, 19,* 478–489. http://dx.doi.org/10.1007/s11065-009-9120-3

Neuwirth, W., & Benesch, M. (2011). *Motorische Leistungsserie. Wiener Testsystem (WTS)* [Motor performance series: Vienna Test System]. Mödling, Austria: Schuhfried.

Noack, H., Lövdén, M., & Schmiedek, F. (2014). On the validity and generality of transfer effects in cognitive training research. *Psychological Research, 78,* 773–789. http://dx.doi.org/10.1007/s00426-014-0564-6

Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: Gauging the generality of cognitive intervention effects. *Restorative Neurology and Neuroscience, 27,* 435–453.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., . . . Ballard, C. G. (2010). Putting brain training to the test. *Nature, 465,* 775–778. http://dx.doi.org/10.1038/nature09042

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25,* 46–59. http://dx.doi.org/10.1002/hbm.20131

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology, 60,* 173–196. http://dx.doi.org/10.1146/annurev.psych.59.103006.093656

Raz, N., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. *NeuroImage, 51,* 501–511. http://dx.doi.org/10.1016/j.neuroimage.2010.03.020

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., . . . Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15,* 1676–1689. http://dx.doi.org/10.1093/cercor/bhi044

Raz, N., & Rodrigue, K. M. (2006). Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience and Biobehavioral Reviews, 30,* 730–748. http://dx.doi.org/10.1016/j.neubiorev.2006.07.001

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., . . . the ACTIVE Study Group. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatrics Society, 62,* 16–24. http://dx.doi.org/10.1111/jgs.12607

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16,* 754–760. http://dx.doi.org/10.1017/S1355617710000706

Sánchez-Cubillo, I., Periáñez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., & Barceló, F. (2009). Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society, 15,* 438–450. http://dx.doi.org/10.1017/S1355617709090626

Schaie, K. W. (2012). *Developmental influences on adult intelligence: The Seattle Longitudinal Study*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:osobl/9780195386134.001.0001

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience, 2,* 27. http://dx.doi.org/10.3389/fnagi.2010.00027

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.

Schoppe, K. P. (1974). Das MLS-Gerät: Ein neuer Testapparat zur Messung feinmotorischer Leistungen [MLS device: A new device for measurement of fine motor abilities]. *Diagnostica, 20,* 43–47.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138,* 628–654. http://dx.doi.org/10.1037/a0027473

Stern, Y., Blumen, H. M., Rich, L. W., Richards, A., Herzberg, G., & Gopher, D. (2011). Space Fortress game training and executive control in older adults: A pilot intervention. *Aging, Neuropsychology, and Cognition, 18,* 653–677. http://dx.doi.org/10.1080/13825585.2011.613450

Strobach, T., Frensch, P. A., & Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychologica, 140,* 13–24. http://dx.doi.org/10.1016/j.actpsy.2012.02.001

Strobach, T., Salminen, T., Karbach, J., & Schubert, T. (2014). Practice-related optimization and transfer of executive functions: A general review and a specific realization of their mechanisms in dual tasks. *Psychological Research, 78,* 836–851. http://dx.doi.org/10.1007/s00426-014-0563-7

Sturm, W., & Büssing, A. (1985). Ergänzende Normierungsdaten und Retest-Reliabilitätskoeffizienten zur Motorischen Leistungsserie (MLS) nach Schoppe [Supplementary standarization and retest reliability coefficents of the Motorische Leistungsserie (MLS) by Schoppe]. *Diagnostica, 31,* 234–245.

Stuss, D. T., Robertson, I. H., Craik, F. I. M., Levine, B., Alexander, M. P., Black, S., . . . Winocur, G. (2007). Cognitive rehabilitation in the elderly: A randomized trial to evaluate a new protocol. *Journal of the International Neuropsychological Society, 13,* 120–131. http://dx.doi.org/10.1017/S1355617707070154

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120,* 439–471. http://dx.doi.org/10.1037/a0033138

Theill, N., Schumacher, V., Adelsberger, R., Martin, M., & Jäncke, L. (2013). Effects of simultaneously performed cognitive and physical training in older adults. *BMC Neuroscience, 14,* 103. http://dx.doi.org/10.1186/1471-2202-14-103

Tomaszewski Farias, S., Cahn-Weiner, D. A., Harvey, D. J., Reed, B. R., Mungas, D., Kramer, J. H., & Chui, H. (2009). Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. *The Clinical Neuropsychologist, 23,* 446–461. http://dx.doi.org/10.1080/13854040802360558

Toril, P., Reales, J. M., & Ballesteros, S. (2014). Video game training enhances cognition of older adults: A meta-analytic study. *Psychology and Aging, 29,* 706–716. http://dx.doi.org/10.1037/a0037507

Verhaeghen, P., & Marcoen, A. (1996). On the mechanisms of plasticity in young and older adults after instruction in the method of loci: Evidence for an amplification model. *Psychology and Aging, 11,* 164–178. http://dx.doi.org/10.1037/0882-7974.11.1.164

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–revised*. New York, NY: Psychological Corporation.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., . . . the ACTIVE Study Group. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA: Journal of the American Medical Association, 296,* 2805–2814. http://dx.doi.org/10.1001/jama.296.23.2805

Winocur, G., Craik, F. I. M., Levine, B., Robertson, I. H., Binns, M. A., Alexander, M., . . . Stuss, D. T. (2007). Cognitive rehabilitation in the elderly: Overview and future directions. *Journal of the International Neuropsychological Society, 13,* 166–171. http://dx.doi.org/10.1017/S1355617707070191

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences, 14,* 138–146. http://dx.doi.org/10.1016/j.tics.2010.01.001

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37–49. http://dx.doi.org/10.1016/0022-3956(82)90033-4

Zelinski, E. M., Peters, K. D., Hindin, S., Petway, K. T., II, & Kennison, R. F. (2014). Evaluating the relationship between change in performance on training tasks and on untrained outcomes. *Frontiers in Human Neuroscience, 8,* 617. http://dx.doi.org/10.3389/fnhum.2014.00617

Zimmermann, P., & Fimm, B. (2002a). *Arbeitsgedächtnis. Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 2.2)* [Working memory: Test of Attentional Performance (TAP, Version 2.2)]. Herzogenrath, Germany: Psytest.

Zimmermann, P., & Fimm, B. (2002b). *Flexibilität. Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 2.2)* [Flexibility: Test of Attentional Performance (TAP, Version 2.2)]. Herzogenrath, Germany: Psytest.

Zimmermann, P., & Fimm, B. (2002c). *Geteilte Aufmerksamkeit. Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 2.2)* [Divided attention: Test of Attentional Performance (TAP, Version 2.2)]. Herzogenrath, Germany: Psytest.

Zinke, K., Zeintl, M., Eschen, A., Herzog, C., & Kliegel, M. (2012). Potentials and limits of plasticity induced by working memory training in old-old age. *Gerontology, 58,* 79–87. http://dx.doi.org/10.1159/000324240

Zinke, K., Zeintl, M., Rose, N. S., 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,* 304–315. http://dx.doi.org/10.1037/a0032982

Received July 22, 2015
Revision received December 18, 2015
Accepted January 27, 2016 ■