



EFFECTS OF GAME AND GAME-LIKE TRAINING ON NEUROCOGNITIVE PLASTICITY

EDITED BY: Guido P. H. Band, Chandramallika Basak, Heleen A. Slagter
and Michelle W. Voss

PUBLISHED IN: Frontiers in Human Neuroscience





frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-840-5

DOI 10.3389/978-2-88919-840-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

EFFECTS OF GAME AND GAME-LIKE TRAINING ON NEUROCOGNITIVE PLASTICITY

Topic Editors:

Guido P. H. Band, Leiden University, Netherlands

Chandramallika Basak, University of Texas at Dallas, USA

Heleen A. Slagter, University of Amsterdam, Netherlands

Michelle W. Voss, University of Iowa, USA

Cognitive training is not always effective. This is also the case for the form of cognitive training that this Research Topic focuses on: prolonged performance on game-like cognitive tasks. The ultimate goal of this cognitive training is to improve ecologically-valid target functions. For example, cognitive training should help children with ADHD to stay focused at school, or help older adults to manage the complexity of daily life. However, so far this goal has proven too ambitious. Transfer from trained to non-trained tasks is not even guaranteed in a laboratory, so there is a strong need for understanding how, when and for how long cognitive training has effect. Which cognitive functions are amenable to game training, for whom, and how? Are there mediating factors for success, such as motivation, attention, or age? Are the improvements real, or can they be attributed to nonspecific factors, such as outcome expectancy or demand characteristics? Are there better strategies to improve cognitive functions through game training?

This Research Topic of *Frontiers in Human Neuroscience* charts current insights in the determinants of success of game training.

Citation: Band, G. P. H., Basak, C., Slagter, H. A., Voss, M. W., eds. (2016). *Effects of Game and Game-like Training on Neurocognitive Plasticity*. Lausanne: Frontiers Media.
doi: 10.3389/978-2-88919-840-5

Table of Contents

- 04 Editorial: Effects of Game and Game-Like Training on Neurocognitive Plasticity**
Guido P. H. Band, Chandramallika Basak, Heleen A. Slagter and Michelle W. Voss
- 07 Videogame interventions and spatial ability interactions**
Thomas S. Redick and Sean B. Webster
- 12 Evaluating the relationship between change in performance on training tasks and on untrained outcomes**
Elizabeth M. Zelinski, Kelly D. Peters, Shoshana Hindin, Kevin T. Petway II and Robert F. Kennison
- 26 Harnessing the neuroplastic potential of the human brain & the future of cognitive rehabilitation**
Jyoti Mishra and Adam Gazzaley
- 30 Spaced cognitive training promotes training transfer**
Zuowei Wang, Renlai Zhou and Priti Shah
- 38 Differential effect of motivational features on training improvements in school-based cognitive training**
Benjamin Katz, Susanne Jaeggi, Martin Buschkuhl, Alyse Stegman and Priti Shah
- 48 Video game training and the reward system**
Robert C. Lorenz, Tobias Gleich, Jürgen Gallinat and Simone Kühn
- 57 The impact of binaural beats on creativity**
Susan A. Reedijk, Anne Bolders and Bernhard Hommel
- 64 Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer**
Vessela Stamenova, Janine M. Jennings, Shaun P. Cook, Lisa A. S. Walker, Andra M. Smith and Patrick S. R. Davidson
- 77 Corrigendum: Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer**
Vessela Stamenova, Janine M. Jennings, Shaun P. Cook, Lisa A. S. Walker, Andra M. Smith and Patrick S. R. Davidson
- 80 Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches**
Erin C. Connors, Elizabeth R. Chrastil, Jaime Sánchez and Lotfi B. Merabet
- 93 Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task**
Aki Nikolaidis, Michelle W. Voss, Hyunkyu Lee, Loan T. K. Vo and Arthur F. Kramer



Editorial: Effects of Game and Game-Like Training on Neurocognitive Plasticity

Guido P. H. Band^{1*}, Chandramallika Basak², Heleen A. Slagter³ and Michelle W. Voss⁴

¹ Leiden Institute for Brain and Cognition, Leiden, Netherlands, ² School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, TX, USA, ³ Department of Psychology, University of Amsterdam, Amsterdam, Netherlands, ⁴ Department of Psychological and Brain Sciences, University of Iowa, Iowa City, IA, USA

Keywords: videogame, learning, training effects, cognitive control, prefrontal cortex, neuromodulators, working memory, plasticity

The Editorial on the Research Topic

Effects of Game and Game-Like Training on Neurocognitive Plasticity

Playing videogames is a widely practiced leisure activity. Because of the time spent playing and the interactive content, game-guided learning could potentially supplement or replace more traditional mediums for learning and rehabilitation. There are many shades of gray between pure cognitive training, where the participant pursues the goal of a cognitive benefit directly, and game-based training, where training is contained in activities with their own intrinsic motivational value. Finding the optimal approach requires a better understanding of the underlying mechanisms (cf. Karbach and Schubert, 2013). We therefore conducted a research topic entitled “Effects of game and game-like training on neurocognitive plasticity,” composed of eight empirical papers, an opinion paper and a review that contribute to the quest for effective ways to employ game-guided learning.

Several studies have shown benefits from videogames for cognitive functions such as visual attention (Green and Bavelier, 2003, 2007), and multitasking (Basak et al., 2008; Anguera et al., 2013). However, transfer of cognitive training to criterion tasks is not always observed. For a reliable picture of game effects, methodological rigor is necessary; to recalibrate positive findings (Redick and Webster) and to give way to publication of reliable null-results (Boot et al., 2011). Optimizing methodology is an ongoing process, as Zelinski et al. demonstrated using structural equation modeling to validly evaluate transfer.

How do games make training more effective? Games can help overcome cognitive limitations by employing adaptive difficulty, informative feedback, and a sufficient dosage (Mishra and Gazzaley). But the positive role that is often attributed to increased motivation (Habgood and Ainsworth, 2011), prolonged immersion and fun inherent in games may be overestimated. Wang et al. tested the optimal way to schedule a working memory training intervention. They found that transfer from training to fluid intelligence scores occurred only after spreading training sessions over 20 days, and not after any of the denser schedules. Thus, long game sessions may not be the most efficient way to learn. Katz et al. performed a systematic study to identify motivational game features that could boost children’s working memory. Surprisingly, game elements such as real-time score displays modulated learning only negatively. This illustrates that mere addition of game elements to disguise the burden of active training works counterproductive. Rather, game-guided training should comply with the same

OPEN ACCESS

Edited and reviewed by:

Srikantan S. Nagarajan,
University of California, San Francisco,
USA

Reviewed by:

Ian Spicer Ramsay,
San Francisco VA Medical Center,
USA

*Correspondence:

Guido P. H. Band
band@fsw.leidenuniv.nl

Received: 04 February 2016

Accepted: 07 March 2016

Published: 30 March 2016

Citation:

Band GPH, Basak C, Slagter HA and
Voss MW (2016) Editorial: Effects of
Game and Game-Like Training on
Neurocognitive Plasticity.
Front. Hum. Neurosci. 10:123.
doi: 10.3389/fnhum.2016.00123

evidence-based principles as traditional forms of training (e.g., Schmidt and Bjork, 1992; Pashler et al., 2007).

What are the cognitive mechanisms that make games beneficial for training? Both game designers and trainers acknowledge the difficulty of merging gaming with training features. When there is too much focus on cognitive benefit, game elements may fail to raise motivation. Conversely, when game elements dominate, they exert a cognitive load at the expense of capacity available for target training (cf. Sweller et al., 1998), or at least draw attention away from the target training. It is fair to assume that the amount of learning is a function of the intensity and the focus of attention for relevant learning material. Indeed, Nikolaidis et al. demonstrated that individual differences in the working memory related brain response to game training are predictive of post-intervention changes on the Sternberg memory search task.

The influence of attentional focus and intensity in games can be translated to game effects on neuromodulation (e.g., Deveau et al.). Reward is a common game ingredient, associated with the release of dopamine (Koepp et al., 1998) and consequently, new habit formation. Gaming elements can help preserve reward responsiveness as measured in ventral striatum after training (Lorenz et al.). Furthermore, games often trigger arousal, characterized by phasic norepinephrine release, which facilitates attention and memory encoding (Tully and Bolshakov, 2010). Note, however, that trainees vary in neuromodulatory activity across age groups and between trainees with and without psychopathology (e.g., schizophrenia, ADHD). Thus, game mechanics should ideally adapt to individual differences in optimal levels of reward, reward responsiveness (cf. Gray, 1982) and arousal.

A different form of brain stimulation targets brain activity at EEG spectral bands associated with attention, working memory and control. Reedijk et al. used binaural beats to entrain target frequencies alpha (10 Hz) or gamma (40 Hz), which led to improved divergent thinking scores, at least for participants with lower dopamine levels. The same reasoning, that entrainment of the brain activity spectrum can improve cognitive performance, is followed in applying neurofeedback and transcranial AC stimulation. Mishra and Gazzaley argued that integrating such entrainment techniques into games, the closed-loop approach, may be able to boost game efficacy.

Finally, games can target qualitative changes in task performance. Bavelier et al. (2012) argued that transfer occurs if games help in “learning to learn,” which involves better selection of relevant information and recognition of a problem structure. Stamenova et al. found data challenging this principle, however. They trained older adults to better distinguish targets from foils in a recollection training task. Although false alarm rates decreased, there was no transfer to non-trained memory tests, possibly because the stimulus set used for training was too limited. Generalization of learning increases as variability in input and training tasks increases (Schmidt and Bjork, 1992). High variability at the level of the learning material likely fosters learning at a more abstract level of representation (Green and Bavelier, 2008; Slagter, 2012). Another example of a qualitative cognitive change by game-guided training is the observation of Connors et al. who trained blind participants navigation through a building, using either audio- or game-guided training. Although standard performance improved equally for both groups, the game-guided training group could apply the same spatial knowledge faster and more flexibly, suggesting deeper or more implicit learning.

In sum, game elements have clear potential to improve training via multiple mechanisms. Perhaps the largest challenge in optimizing the efficacy of game-guided training is to gear interventions to individual differences. Variable brain anatomy, connectivity, baseline performance, age, neuromodulation, and strategy use can all modulate effects of game training parameters. As more is known about the neurocognitive properties of the trainee, games can be adapted accordingly. Ultimately, this approach would merge game design with insights from cognitive neuroscience and educational neuroscience.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

The third author was supported by a VIDI award of the Netherlands Organisation for Scientific Research.

REFERENCES

- Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., et al. (2013). Video game training enhances cognitive control in older adults. *Nature* 501, 97–101. doi: 10.1038/nature12486
- Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy videogame attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765–777. doi: 10.1037/a0013494
- Bavelier, D., Green, C. S., Pouget, A., and Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Ann. Rev. Neurosci.* 35, 391–416. doi: 10.1146/annurev-neuro-060909-152832
- Boot, W. R., Blakely, D. P., and Simons, D. J. (2011). Do action video games improve perception and cognition? *Front. Psychol.* 2:226. doi: 10.3389/fpsyg.2011.00226
- Gray, J. A. (1982). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford: Oxford University Press.
- Green, C. S., and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature* 423, 534–537. doi: 10.1038/nature01647
- Green, C. S., and Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychol. Sci.* 18, 88–94. doi: 10.1111/j.1467-9280.2007.01853.x
- Green, C. S., and Bavelier, D. (2008). Exercising your brain: a review of human brain plasticity and training-induced learning. *Psychol. Aging* 23, 692–701. doi: 10.1037/a0014345
- Habgood, M. P. J., and Ainsworth, S. E. (2011). Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. *J. Learn. Sci.* 20, 169–206. doi: 10.1080/10508406.2010.508029

- Karbach, J., and Schubert, T. (2013). Training-induced cognitive and neural plasticity. *Front. Hum. Neurosci.* 7:48. doi: 10.3389/fnhum.2013.00048
- Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., et al. (1998). Evidence for striatal dopamine release during a video game. *Nature* 393, 266–268. doi: 10.1038/30498
- Pashler, H., Rohrer, D., Cepeda, N. J., and Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: choices and consequences. *Psychon. Bull. Rev.* 14, 187–193. doi: 10.3758/BF03194050
- Schmidt, R. A., and Bjork, R. A. (1992). New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. *Psychol. Sci.* 3, 207–217. doi: 10.1111/j.1467-9280.1992.tb00029.x
- Slagter, H. A. (2012). Conventional working memory training may not enhance intelligence. *Trends Cogn. Sci.* 16, 582–583. doi: 10.1016/j.tics.2012.10.001
- Sweller, J., Van Merriënboer, J. J. G., and Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10, 251–296. doi: 10.1023/A:1022193728205
- Tully, K., and Bolshakov, V. Y. (2010). Emotional enhancement of memory: how epinephrine enables synaptic plasticity. *Mol. Brain* 3:15. doi: 10.1186/1756-6606-3-15
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Band, Basak, Slagter and Voss. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Videogame interventions and spatial ability interactions

Thomas S. Redick* and Sean B. Webster

Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

Edited by:

Michelle W. Voss, University of Iowa, USA

Reviewed by:

Michael Dougherty, University of Maryland at College Park, USA
Walter R. Boot, Florida State University, USA

***Correspondence:**

Thomas S. Redick, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA
e-mail: tredick@purdue.edu

Numerous research studies have been conducted on the use of videogames as tools to improve one's cognitive abilities. While meta-analyses and qualitative reviews have provided evidence that some aspects of cognition such as spatial imagery are modified after exposure to videogames, other evidence has shown that matrix reasoning measures of fluid intelligence do not show evidence of transfer from videogame training. In the current work, we investigate the available evidence for transfer specifically to nonverbal intelligence and spatial ability measures, given recent research that these abilities may be most sensitive to training on cognitive and working memory tasks. Accordingly, we highlight a few studies that on the surface provide evidence for transfer to spatial abilities, but a closer look at the pattern of data does not reveal a clean interpretation of the results. We discuss the implications of these results in relation to research design and statistical analysis practices.

Keywords: video games, training, transfer, intelligence, cognitive interventions

VIDEOGAME INTERVENTIONS AND SPATIAL ABILITY INTERACTIONS

In the past 10 years, there has been substantial interest in the idea that playing videogames may serve to improve certain cognitive functions. A recent meta-analysis (Powers et al., 2013) provided a quantitative summary of the numerous studies in which a videogame-playing group was compared against a control group that did not receive the videogame "treatment" of interest. As is typically done in meta-analyses, Powers et al. (2013) used broad operational definitions of cognitive outcomes, such as combining together various outcome measures of "executive functions", which included multitasking, inhibition, task-switching, short-term/working memory, and intelligence. Overall, videogame training effects on executive functions was statistically significant ($d = 0.16$), although the effect would be classified as small according to Cohen (1992). Notably, transfer to intelligence was not significant, $d = 0.06$, and inhibition was the only executive function that was significantly improved by videogame training. On the other hand, tests of spatial imagery, such as mental rotation tasks, exhibited stronger meta-analytic effects, $d = 0.43$. Our current work investigates specifically the effect of videogame training on spatial ability transfer outcomes, given recent research that argues these spatial ability tests may be most sensitive to cognitive training (Colom et al., 2013).

Basak et al. (2008) provide an illustrative example of almost ideal intelligence transfer results, as a function of videogame training. In their study, older adults in the training group played a videogame (*Rise of Nations*) during 15 sessions. Raven Advanced Progressive Matrices, a matrix reasoning test commonly used to measure fluid or nonverbal intelligence, was among the battery of tests administered during pre-test and post-test transfer sessions. As seen in **Figure 1A**, the training group improved on Raven scores from pre- to post-test, whereas the control group that

did not do anything between pre- and post-test (a *no-contact control group*) showed no improvement on Raven. In addition, the Raven mean pre-test scores were similar for the two groups. The interaction pattern in Basak et al. (2008) is straightforward to interpret, and the transfer data easily support the argument that the training "worked" for those subjects.

The pattern of transfer results observed in Basak et al. (2008) is what one would predict if the pre- to post-test change for the training group is what drives the observed significant interaction. However, videogame training studies such as Maillot et al. (2012), van Muijden et al. (2012), and Cherney (2008) are more difficult to interpret, because the pattern of transfer results is not similar to the Basak et al. (2008) example above. As outlined in the summaries below, each of these studies has been used to provide support for the efficacy of videogame interventions upon spatial abilities, yet a closer examination of the results warrants a more cautious interpretation. Note that our discussion in the current manuscript focuses on interactions and how strongly one can interpret the results as support for the efficacy of videogame interventions; recent articles have discussed many other methodological and measurement issues in videogame and brain training studies (e.g., Boot et al., 2011, 2013; Shipstead et al., 2012; Green et al., 2013).

VAN MUIJDEN ET AL. (2012)

Two groups of participants completed the study: the video game group ($n = 53$) and the documentary group ($n = 19$). Participants in the videogame training group played five custom-built games (cf. **Figure 1**; van Muijden et al., 2012). Participants in the documentary control group watched documentaries of about 30 min in length followed by a three-to-five multiple-choice question quiz about the documentary. Participants in both the videogame and documentary control groups were instructed to

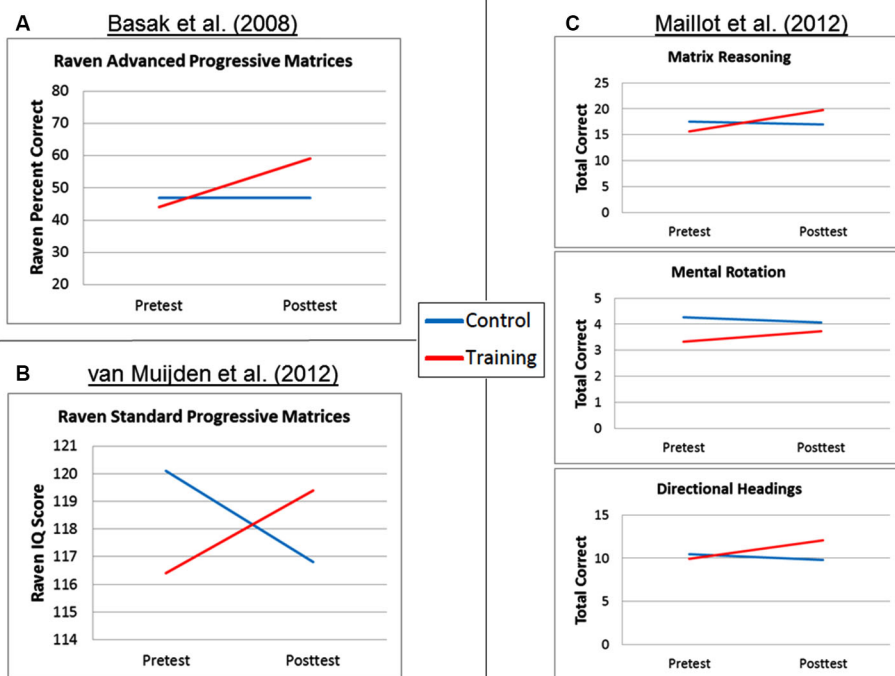


FIGURE 1 | Spatial ability transfer results as a function of group for (A) Basak et al. (2008), (B) van Muijden et al. (2012), and (C) Maillot et al. (2012).

complete one 30-min session daily for 7 weeks, and participants that completed the study completed on average over 21 h of the prescribed activity. Before and after the videogame and documentary sessions, all participants completed a cognitive test battery that consisted of nine cognitive tests, including Raven standard progressive matrices. Raw scores from Raven were converted to IQ scores. van Muijden et al. (2012) concluded that “the results from the present study suggest that modest improvements of inductive reasoning can also be achieved by means of playing cognitive training games” (p. 10).

The group \times session interaction on Raven was significant ($\eta_p^2 = 0.068$), and is displayed in **Figure 1B**. However, this pattern is quite different from the Raven interaction pattern shown by Basak et al. (2008). The pre-test Raven IQ mean score for the videogame group was 116.4 and the post-test mean score was 119.4, an increase of 3.0 IQ points. The pre-test Raven IQ mean score for the documentary group was 120.1 and the post-test mean score was 116.8, a decrease of 3.3 IQ points. Although the between-groups comparison of the pre-test scores reported by van Muijden et al. (2012) was not significant (“ $p > 0.05$ ”, p. 4), the effect size for the pre-test scores indicated a small-to-medium difference (Cohen’s $d = 0.38$). van Muijden et al. (2012) conducted subsequent simple main effect analyses to examine the pre- to post-test change within each group separately. Within the videogame group only, the pre- to post-test Raven increase was considered marginally significant ($p = 0.05$), whereas the pre- to post-test Raven decrease for the documentary group was not significant ($p > 0.10$). Critically, however, there was also a large difference in the group sample sizes ($n = 53$ and $n = 19$ for the videogame training group and the documentary control group,

respectively). The sample size difference is extremely important as it pertains to the aforementioned follow-up tests—the reported effect size for the pre- to post-test Raven change score was actually larger for the nonsignificant control group ($\eta_p^2 = 0.128$) than it was for the significant training group ($\eta_p^2 = 0.071$). Thus, despite the significant interaction and associated effect size, we view the crossover pattern of Raven results as relatively weak evidence for the efficacy of videogame training to improve spatial abilities.

MAILLOT ET AL. (2012)

Two groups of older adults (between the ages of 65 and 78 years old) completed the study: one group ($n = 15$) was assigned to the exergame training condition and the other ($n = 15$) was assigned to the no-training control group. The exergame training group completed two 1-h exergame sessions per week for 12 weeks for a total training time of 24 h. During these sessions participants played *Wii Sports*, *Wii Fit*, and *Mario and Sonic on Olympic Games*. In pre- and post-test sessions, all participants completed a battery of cognitive assessment tasks—we focus here on the matrix reasoning test (a subtest of the Wechsler Abbreviated Scale of Intelligence), a mental rotation task, and a directional headings test, all measures of spatial ability. Maillot et al. (2012) concluded “exergame training, which combines cognitive and physical demands in an intrinsically attractive activity, might be an effective way to promote physical and cognitive improvements among older adults” (p. 597).

In the published article, only the pre-test/post-test change scores were reported for each test, but the pre- and post-test values were provided upon request (P. Maillot, personal communication,

10/9/13), and are shown in **Figure 1C**. First, Maillot et al. (2012) reported that the mental rotation task change scores were not significantly different between the exergame and control groups ($p = 0.24$, $\eta_p^2 = 0.019$), so this result is not discussed further. Maillot et al. (2012) reported that the matrix reasoning test change scores were significantly different between the exergame and control groups ($p < 0.01$) with a very large effect size of $\eta_p^2 = 0.531$. For the directional headings test, Maillot et al. (2012) reported that the change scores were significantly different between the exergame and control groups ($p = 0.02$), again with a large effect size of $\eta_p^2 = 0.149$. Critically, across all three dependent variables, the control group shows a numerical decrease from pre-test to post-test on each test, and the control group pre-test scores were numerically larger than the training group. Given the small sample sizes, between-groups comparisons of the pre-test scores were not significant (all p 's > 0.13), yet the effect sizes for the pre-test scores indicated small-to-medium differences (Cohen's $d = 0.14$ – 0.59). Although the matrix reasoning and directional headings difference scores were significantly different between the training and control groups, examination of the pre- and post-test scores instead of only the change scores reveals an interaction pattern that complicates a strong interpretation of the exergame training efficacy. Combined with the absence of exergame effects on the mental rotation task and the use of a no-contact control group, we view the spatial ability transfer results here as modest at best, which contrasts with the large effect sizes reported in the article.

CHERNEY (2008)

There has been substantial interest in the idea that videogames could reduce or eliminate gender effects in spatial ability, following the study by Feng et al. (2007). Cherney (2008) investigated gender differences in videogame effects upon spatial abilities, reported in a paper titled “Mom, let me play more computer games: They improve my mental rotation skills”. Separate groups of male ($n = 30$) and female ($n = 31$) undergraduate students completed the study. Participants were randomly assigned to one of the three conditions: 3D videogame training ($n = 20$), 2D videogame training ($n = 21$), or control ($n = 20$) group, such that there were either 10 or 11 participants of each gender in each group (personal communication, I. Cherney, 11/25/2013). The 3D group played *Antz Extreme Racing*, the 2D group played *Tetris*, and the remaining control participants played paper-and-pencil logic games. Thirty-six participants completed the training during 2 weeks and 25 participants completed the training during 1 week. The participants all completed mental rotation and card rotation spatial ability tests in pre- and post-test sessions before and after the training period, respectively. Cherney (2008) concluded that “the results suggest that even a very brief practice (4 h) in computer game play does improve performance on mental rotation measures” (p. 783), and specifically, “practice with the *Antz* game, a 3-D computer game, seemed particularly beneficial for women's Vandenberg mental rotation test (VMRT) scores” (p. 785).

Figure 2 displays the results for each transfer test as a function of group and gender (mean and standard deviations provided by I. Cherney, personal communication, 10/17/13). Collapsing

across training and control groups, Cherney (2008) reported that there were significant improvements on the mental rotation test scores for women ($p < 0.001$) but not for men, and that card rotation test scores improved for both men and women ($p < 0.001$). Unfortunately, the necessary statistical comparisons testing whether the male and female training and control groups differentially improved from pre- to post-test on each dependent variable were not reported in Cherney (2008). As seen in **Figure 2**, the results are not clear. For the mental rotation test, men playing the 2D *Tetris* game showed almost the same amount of decrease in scores from pre- to post-test (2.1 items) that the women increased in scores from pre- to post-test (2.6 items). In addition, for the mental rotation test, the effect of 3D and 2D training for the female groups was to bring their post-test scores up to the level of the pre-test mental rotation score for the female control group. For card rotation, Cherney reported significant pre- to post-test increases for each of the six groups tested, but no comparison of differences in the gain scores across groups. Given the small sample sizes, between-groups comparisons of the mental rotation and card rotation pre-test scores were not significant (all p 's > 0.07), yet the effect sizes for the pre-test scores indicated small-to-large differences (Cohen's $d = 0.11$ – 0.81). Again, we do not find the pattern of results presented in **Figure 2** as providing compelling evidence for the efficacy of videogame training to improve spatial abilities, for either men or women.

CONCLUSION

The studies reviewed here do not represent all of the published evidence in support of the efficacy of videogame training to improve spatial abilities (see Powers et al., 2013, for review). However, the highlighted studies do illustrate the numerous challenges both scientists and laypersons face when trying to interpret the available research. In the three studies reviewed here, the sample sizes are smaller than the recommended minimum number of observations per cell for at least one of the comparison groups (Simmons et al., 2011), and the use of small samples leads to many problems. First, and most obvious, the use of small samples leaves the study underpowered to quantify the true effect of videogame training, and even more problematically leads to an increased likelihood of producing an inflated effect size (Button et al., 2013). Although meta-analyses are more informative than any individual study in quantifying the exact magnitude of the effect of videogame training, the meta-analysis is not a panacea if studies with small sample sizes produce large effect sizes that are then included in the meta-analysis. For example, Powers et al. (2013) provided Cohen's d estimates of the effect sizes obtained in Maillot et al. (2012): $d = 2.056$ and $d = 0.810$ for matrix reasoning and directional headings, respectively. As reviewed above, the pattern of results for both the training and control groups were not straightforward, and inclusion of these large effect sizes in the meta-analysis will influence the overall meta-analytic estimate of the training effect size.

Second, although random sampling and assignment should eliminate pre-existing differences between the training and control groups, smaller samples provide less accurate estimates of the population values and will be more strongly influenced by an outlier value, and as such pre-test differences between training

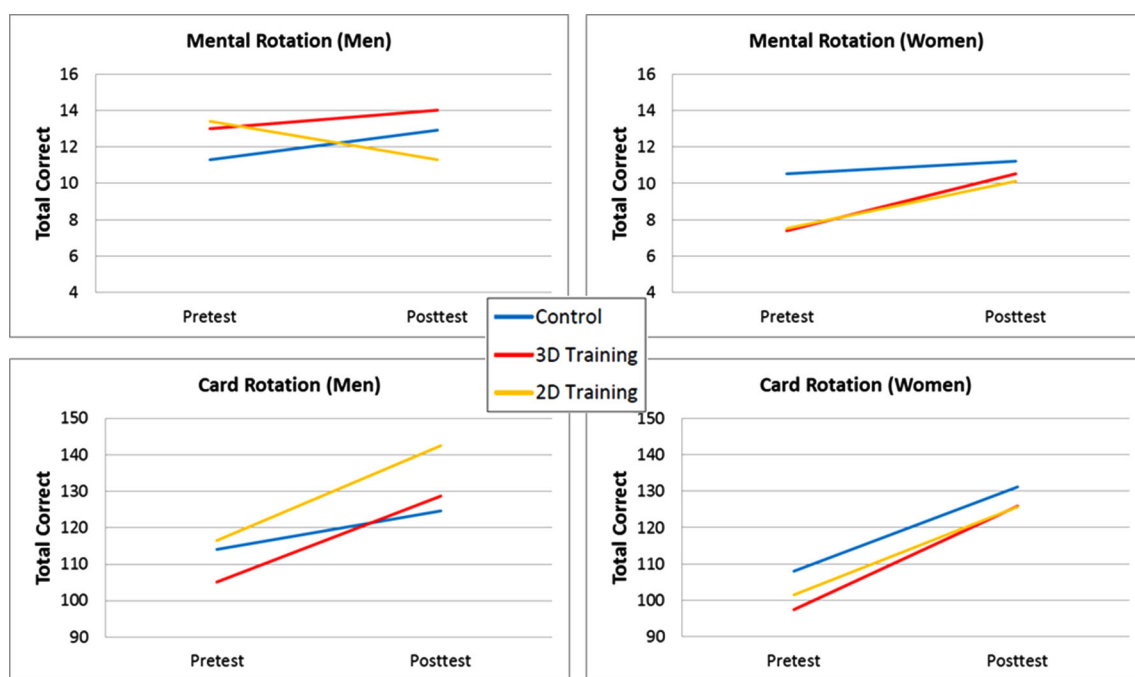


FIGURE 2 | Spatial ability transfer results as a function of gender and group for Cherney (2008).

and control groups may be more likely. However, the use of small sample sizes also means that statistical tests of pre-test values are likely to be non-significant, allowing researchers to declare the training and control groups did not differ at pre-test (" $p > 0.05$ ") despite the numerical differences between the groups. Note also that researchers are relying on failure to reject the null hypothesis as evidence for no difference between the training and control groups at pre-test (see Redick et al., 2013, for related discussion). As seen in the studies reviewed here (Figures 1 and 2), the control groups' pre-test values were numerically but not significantly higher across several of the transfer tests, complicating interpretation of results when compared with the videogame training groups.

In closing, we offer a few suggestions for future videogame training studies. First, as noted in other recent reviews (Boot et al., 2011, 2013; Shipstead et al., 2012; Green et al., 2013), we strongly encourage the use of appropriate experimental control procedures that measure and counteract placebo effects. Second, for the reasons outlined above, we advocate the use of much larger sample sizes for training and control groups. Third, we strongly recommend presenting the pre- and post-test values for transfer tests, instead of or in addition to only reporting pre- to post-test change scores. Presenting the pre- and post-test values will allow the reader to determine if the pattern of significant results allows a strong conclusion or if there is ambiguity in the transfer results. A further suggestion is to provide the full data for each participant as supplemental material. Having the full data will be beneficial for interested readers and researchers conducting future meta-analyses, as they will be able to conduct both between- and within-subject analyses and additionally make comparisons not presented in the article by the authors. Finally, we note there is

considerable debate about how best to statistically assess change in such intervention studies, with various authors pointing out limitations with independent-samples *t*-tests of gain scores, group by session interactions in ANOVA, and using pre-test scores as covariates in an ANCOVA (Lord, 1967; Huck and McLean, 1975; Miller and Chapman, 2001; Wright, 2006). Sampling techniques that lead to similar pre-test values for training and control groups will help minimize differences among the statistical analyses, whether that is achieved via random assignment with larger samples or through some sort of matching technique (for pros and cons of matching, see Green et al., 2013). In addition, incorporation of Bayesian analyses to either supplement or replace null-hypothesis significance-testing may more accurately quantify the intervention effect in a particular study (for a recent example of the use of Bayes Factors in cognitive training research, see Sprenger et al., 2013).

Above all, we hope that researchers will not focus so much on obtaining a significant *p*-value that they fail to examine the pattern of results to understand the cause of the significant result.

AUTHOR CONTRIBUTIONS

Thomas S. Redick and Sean B. Webster contributed to the literature review. Sean B. Webster contacted authors for necessary additional information. Thomas S. Redick and Sean B. Webster created the figures, drafted the manuscript, and approved the final version for submission.

ACKNOWLEDGMENTS

While writing this article, Thomas S. Redick was supported by the Office of Naval Research (Award # N00014-12-1-1011). The

authors thank Tyler Harrison for constructive feedback on an earlier draft.

REFERENCES

- Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765–777. doi: 10.1037/a0013494
- Boot, W. R., Blakely, D. P., and Simons, D. J. (2011). Do action video games improve perception and cognition? *Front. Psychol.* 2:226. doi: 10.3389/fpsyg.2011.00226
- Boot, W. R., Simons, D. J., Stothart, C., and Stutts, C. (2013). The pervasive problem with placebos in psychology: why active control groups are not sufficient to rule out placebo effects. *Perspect. Psychol. Sci.* 8, 445–454. doi: 10.1177/1745691613491271
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., et al. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376. doi: 10.1038/nrn3475
- Cherney, I. D. (2008). Mom, let me play more computer games: they improve my mental rotation skills. *Sex Roles* 59, 776–786. doi: 10.1007/s11199-008-9498-z
- Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155
- Colom, R., Román, F. J., Abad, F. J., Shih, P. C., Privado, J., Froufe, M., et al. (2013). Adaptive n-back training does not improve general ability at the construct level; gains on individual tests suggest training may enhance visuospatial processing. *Intelligence* 41, 712–727. doi: 10.1016/j.intell.2013.09.002
- Feng, J., Spence, I., and Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychol. Sci.* 18, 850–855. doi: 10.1111/j.1467-9280.2007.01990.x
- Green, C. S., Strobach, T., and Schubert, T. (2013). On methodological standards in training and transfer experiments. *Psychol. Res.* doi: 10.1007/s00426-013-0535-3. [Epub ahead of print].
- Huck, S. W., and McLean, R. A. (1975). Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: a potentially confusing task. *Psychol. Bull.* 82, 511–518. doi: 10.1037/h0076767
- Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychol. Bull.* 68, 304–305. doi: 10.1037/h0025105
- Maillot, P., Perrot, A., and Hartley, A. (2012). Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. *Psychol. Aging* 27, 589–600. doi: 10.1037/a0026268
- Miller, G. A., and Chapman, J. P. (2001). Misunderstanding analysis of covariance. *J. Abnorm. Psychol.* 110, 40–48. doi: 10.1037//0021-843x.110.1.40
- Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A., and Alfieri, L. (2013). Effects of video-game play on information processing: a meta-analytic investigation. *Psychon. Bull. Rev.* 20, 1055–1079. doi: 10.3758/s13423-013-0418-z
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *J. Exp. Psychol. Gen.* 142, 359–379. doi: 10.1037/a0029082
- Shipstead, Z., Redick, T. S., and Engle, R. W. (2012). Is working memory training effective? *Psychol. Bull.* 138, 628–654.
- Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22, 1359–1366. doi: 10.1177/0956797611417632
- Sprenger, A. M., Atkins, S. M., Bolger, D. J., Harbison, J. I., Novick, J. M., Chrabaszcz, J. S., et al. (2013). Training working memory: limits of transfer. *Intelligence* 41, 638–663. doi: 10.1016/j.intell.2013.07.013
- van Muijden, J., Band, G. P. H., and Hommel, B. (2012). Online games training aging brains: limited transfer to cognitive control functions. *Front. Hum. Neurosci.* 6:221. doi: 10.3389/fnhum.2012.00221
- Wright, D. B. (2006). Comparing groups in a before-after design: when t test and ANCOVA produce different results. *Br. J. Educ. Psychol.* 76, 663–675. doi: 10.1348/000709905x52210

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 January 2014; accepted: 12 March 2014; published online: 26 March 2014.
Citation: Redick TS and Webster SB (2014) Videogame interventions and spatial ability interactions. *Front. Hum. Neurosci.* 8:183. doi: 10.3389/fnhum.2014.00183
This article was submitted to the journal *Frontiers in Human Neuroscience*.
Copyright © 2014 Redick and Webster. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Evaluating the relationship between change in performance on training tasks and on untrained outcomes

Elizabeth M. Zelinski^{1*}, Kelly D. Peters², Shoshana Hindin¹, Kevin T. Petway II² and Robert F. Kennison^{1,3}

¹ Zelinski Laboratory, Center for Digital Aging, Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

² Psychology Department, University of Southern California, Los Angeles, CA, USA

³ Psychology Department, California State University, Los Angeles, CA, USA

Edited by:

Michelle W. Voss, University of Iowa, USA

Reviewed by:

Erika K. Hussey, University of Illinois, USA

Sobanawartiny Wijekumar, University of Iowa, USA

*Correspondence:

Elizabeth M. Zelinski, Zelinski Laboratory, Center for Digital Aging, Davis School of Gerontology, University of Southern California, 3715 S. McClintock St., Los Angeles, CA 90089-0191, USA
e-mail: zelinski@usc.edu

Training interventions for older adults are designed to remediate performance on trained tasks and to generalize, or transfer, to untrained tasks. Evidence for transfer is typically based on the trained group showing greater improvement than controls on untrained tasks, or on a correlation between gains in training and in transfer tasks. However, this ignores potential correlational relationships between trained and untrained tasks that exist before training. By accounting for crossed (trained and untrained) and lagged (pre-training and post-training) and cross-lagged relationships between trained and untrained scores in structural equation models, the training-transfer gain relationship can be independently estimated. Transfer is confirmed if only the trained but not control participants' gain correlation is significant. Modeling data from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study (Smith et al., 2009), transfer from speeded auditory discrimination and syllable span to list and text memory and to working memory was demonstrated in 487 adults aged 65–93. Evaluation of age, sex, and education on pretest scores and on change did not alter this. The overlap of the training with transfer measures was also investigated to evaluate the hypothesis that performance gains in a non-verbal speeded auditory discrimination task may be associated with gains on fewer tasks than gains in a verbal working memory task. Gains in speeded processing were associated with gains on one list memory measure. Syllable span gains were associated with improvement in difficult list recall, story recall, and working memory factor scores. Findings confirmed that more overlap with task demands was associated with gains to more of the tasks assessed, suggesting that transfer effects are related to task overlap in multimodal training.

Keywords: cognitive training, aging, untrained outcomes, multimodal training

INTRODUCTION

Longitudinal declines in many cognitive processes, including memory, attention, working memory, and speed of processing, are normative in aging (e.g., Zelinski et al., 2011a). This has led to concerns that declines may negatively impact quality of life and increase the risk of losing independence, as cognition plays an important role in many activities of daily living including financial management (e.g., Jobe et al., 2001). At the same time, it has become increasingly clear that individual differences in healthy older adults' cognitive performance is associated with a wide range of potentially enriching experiences, including education, healthy lifestyle practices, engagement in cognitively challenging activities, social involvement, avoidance of stress, and positive attitudes that promote psychological well-being (Hertzog et al., 2009). Interventions to enhance cognition have also shown benefits; many of these involve training on tasks thought to benefit processes that decline with aging. An important indicator of the effectiveness of interventions designed to improve cognitive performance in older adults is whether training benefits generalize to

tasks or cognitive activities that were not trained (e.g., Jobe et al., 2001). It is well established that training of specific strategies, such as mnemonics, does not produce transfer in older adults (e.g., Park et al., 2007). This approach to training holds little promise for reducing risk of decline or even supporting the maintenance of cognitive ability, possibly because older adults often do not apply strategies to new tasks. This may occur because older people experience difficulties in engaging such strategies (Zelinski, 2009), have greater willingness to use suboptimal strategies (Hertzog et al., 2007), or have poor memory self-concept (West et al., 2008).

However, extended practice of tasks such as dual-tasking or N-back, can transfer to untrained tasks (Zelinski, 2009). Game play that involves repetitive practice of cognitive skills that involve multitasking also can produce transfer (e.g., Basak et al., 2008; Anguera et al., 2013). A recent meta-analysis directly evaluated effects of extended practice cognitive training on untrained tasks. These interventions significantly improved older adults' performance on untrained cognitive tasks, with an estimated mean effect size of 0.32 after accounting for practice in the experimental

and control groups (Hindin and Zelinski, 2012). All of the 25 extended practice studies in the meta-analysis evaluated improvements in untrained outcomes by comparing pre-post differences between experimental and control groups. None examined how individuals' performance was affected by training. Yet if transfer has occurred, those in the experimental group who gain more on the training task should improve more on the untrained task because the training should generalize to other tasks with common components (e.g., Persson et al., 2007; Lövdén et al., 2010) or at least the same task-specific demands (e.g., Buschkuhl et al., 2012). Several studies published subsequently to the Hindin and Zelinski meta-analysis have examined correlations between improvements on trained and untrained tasks in older adults, reporting significant correlations in the experimental group (e.g., Anguera et al., 2013; Stepankova et al., 2014).

McArdle and Prindle (2008) suggested that it is necessary to test for transfer with a more sophisticated modeling approach than the use of *t*-tests, ANOVA, or bivariate correlations. They argued that if trained and untrained tasks invoke similar constructs, these should be correlated at baseline as well as after training. This suggests that in order to assess transfer, existing relationships between performance on trained and untrained tasks at baseline should be accounted for, so that the independent relationship between baseline and posttest training and transfer task performance relationships can be ascertained. Relationships between the initial baseline and posttraining scores should also be accounted for, as individual differences in the construct measured may be related to performance gains (see also von Bastian et al., 2013). Therefore, the strongest test of whether training produces transfer is that those who received the training intervention show a significantly stronger relationship between changes in trained and untrained task performance after training than those in the control group after all other possible relationships between trained and untrained tasks prior to, as well as subsequent to, training in each group have been accounted for. It would also be expected that demographic covariates should not affect transfer if a clear interpretation of training benefits is to be made. Otherwise, interactions between the characteristics of participants and training might confound transfer.

McArdle and Prindle (2008) evaluated a series of structural equation models accounting for relationships between near (trained) and far (untrained) cognitive tasks that compared 699 participants trained over 10 h to improve reasoning with 698 members of a no-contact control group. Data were from the initial phase of the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) trial (Ball et al., 2002), a randomized controlled single-blind study of three interventions examining whether older adults' cognitive abilities and everyday functioning could be improved over 2 years. The trained group had a higher latent change mean than the untrained group on the reasoning measures, as they had in the study, showing that training improved performance on the trained measure. The models also indicated that at baseline, relationships were significant and positive between the trained and untrained measures. There was also a significant and positive relationship between the trained and untrained latent change measures, but this relationship did not vary differentially for the trained and control group participants.

Thus, this study showed no relationship between change in training and in transfer in the experimental group participants. However, no group effects of transfer had been observed in the main study (Ball et al., 2002), and the elegant structural analysis of McArdle and Prindle did not produce any new findings to support the existence of training-related transfer in the trained group. The present analyses extended the modeling approach of McArdle and Prindle to a different dataset that had produced transfer effects at the group level for the trained participants.

HYPOTHESES

Data were from The Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study (Smith et al., 2009). The training protocol of the IMPACT study is based on a conceptualization of age declines in memory that are associated with negative neuroplasticity. Mahncke et al. (2006) suggested that deficits associated with cognitive aging are due to reduced frequency of engaging in cognitively demanding activities with age, declines in the integrity of perceptual experience due to sensory deficits that lead to reduced signal to noise ratios in information processing, reduced neuromodulation of the attention-reward system due to reduced cognitive stimulation, negative learning, and coping with reduced stimulation by reducing cognitively engaging behaviors further, creating a negative spiral of increasing decline in cognitive functioning. This can be reversed by undoing the activities that cause negative neuroplasticity and engaging in activities that cause positive neuroplasticity: frequent intense practice of cognitively challenging tasks requiring fine sensory discrimination, rapid processing of sensory information, deep attention, and novelty (Mahncke et al., 2006). The training program, described below, was adaptive, so as to remain cognitively demanding, it improved the signal to noise ratio by training discrimination of increasingly finer differences between stimuli while reducing the stimulus presentation rate with sound compression, and included feedback and rewards to maintain deep attention. Stimuli ranged from sound sweeps, non-word syllables (phonemes), syllables, and verbal instructions, to stories. The primary training measure was performance on the simplest training task, time-ordered sound sweep discrimination, measured as the duration of the sound sweeps needed for high accuracy in performance.

The training program was multimodal in that multiple processes involving rapid auditory discrimination were trained. For example, the training tasks included discriminating easily confused phonemes, remembering them in order, remembering their locations in a matrix, remembering and following increasingly complex sets of instructions to move objects in particular sequences (e.g., *move the dog next to the girl with the black hat, then move the police officer to the front of the bank*), and remembering facts from stories. It was possible that the primary training measure of sound sweep discrimination might be differentially associated with outcome changes than another measure that had also been collected, syllable span. By assessing relationships of change in the two trained tasks in the IMPACT study, the issue of what changes are measured comes to the forefront. Most multimodal training studies do not include pre and post training measures of all aspects of the training, so it is difficult to

determine what aspects of training gain are associated with transfer gains. In the present analysis, the transfer measures were not only different from the training tasks in terms of the specific materials used (e.g., numbers, letters) but tested recall where only recognition had been trained, using subtests from widely used clinical neuropsychological tests. They involved episodic memory recall or reorganization of material in working memory. That is, transfer tasks were not closely related to training tasks. The training tasks also differed substantially in terms of overlap with transfer tasks.

It was hypothesized that the complete IMPACT training program would produce transfer because the underlying neuroplastic mechanisms would have been improved, producing perceptual and memory representations with greater fidelity, so that there would be better performance on a range of untrained auditory memory tasks. Gains in neural timing and accuracy of auditory perception with the training used in IMPACT have been confirmed in an independent study of older adults (Anderson et al., 2013). The use of the speeded sound sweep discrimination task as the training measure in the published study (Smith et al., 2009) has relatively few components in common with more complex memory tasks. The speed task has a constant memory load of two sound samples, it is non-verbal, and requires emphasis on perception of the sweeps, which are presented in increasingly shorter durations. In the IMPACT study, data from another training task, the reproduction of sequences of easily confusable syllables (pât and mât), were collected. This task used a span measure, whereby sequences of syllables increased in length as individuals improved in their ability to discriminate and recognize them. Performance was measured as the maximal syllable span at pretest and posttest in the task and can be considered an index of training effect in the expansion of working memory span. This measure was not analyzed in the IMPACT publications, but its analysis allows for a comparison of transfer effects on the outcome measures with transfer associated it and with speed training. The syllable span task is a measure of working memory. It has been suggested that interventions that may be most effective for older adults are those retraining working memory or executive control processes (Lövdén et al., 2010). Training cognitive control such as coordination of information in working memory produces transfer in older adults to similar tasks (e.g., Buschkuhl et al., 2008; Karbach and Kray, 2009). The task in this study required discrimination of easily confused syllables presented for increasingly shorter durations, storing them, and remembering them in order. The number of syllables increased as performance improved. The phonemes are verbalizable, can be rehearsed, and the memory demands increase. Though these tasks were learned in the multimodal context, hypotheses about the relative amount of demand can be derived. In contrast to gains in the speeded time ordered auditory discrimination task, transfer may be more easily observed because of the mapping of relatively similar task demands to the untrained tasks (e.g., Buschkuhl et al., 2012).

Testing transfer from change in syllable span to change in the outcome measures of list and story memory and to working memory would provide an important test of the relationship of assessed training gain to transfer task gain based on task

demand overlap. If similarity of demands is the critical predictor of transfer (e.g., Buschkuhl et al., 2012), training change in syllable span would show the strongest relationship with change in the working memory outcome tasks of backwards digit span and letter-number sequencing. Because working memory is also implicated in verbal memory, it was also hypothesized that transfer would also be observed in the other measures of the IMPACT study, though it was expected that story memory measures would show stronger transfer because reconstructing a story is more closely associated with working memory than is list memory (e.g., Lewis and Zelinski, 2010).

Individual differences that affect baseline performance, such as participants' age, should not be expected to affect transfer (see McArdle and Prindle, 2008). However, surprisingly few aging studies have examined how characteristics like age, sex, or education affect training gains. McArdle and Prindle (2008) found that age had a negative effect on baseline and change scores, that gender had small effects on pretest scores and that education affected only pretest scores. These relationships, however, did not affect transfer. In the present study, effects of age, sex, and education were included as covariates in the final set of analyses. Baseline memory outcome scores were expected to be more negatively affected by age, but positively by female gender and education as seen in other studies of memory in large samples (e.g., Zelinski and Gilewski, 2003). It was expected that being older would reduce training gains because of age-related limits on plasticity (e.g., Hertzog et al., 2009), but not the relationship between gains in training and transfer, following McArdle and Prindle (2008). Effects of gender and education on training gains were exploratory, as little was known about how these differences would affect training outcomes. It was also not clear whether transfer would be affected by those individual differences.

MATERIALS AND METHODS

The IMPACT study had tested the efficacy of a commercially available computerized cognitive training program on the speeded auditory discrimination task and on untrained clinical neuropsychological measures of memory and attention (Smith et al., 2009). The study design was a double blinded randomized controlled trial comparing those who participated in the training, which used principles of brain plasticity, that is, was repetitive, adaptive, and trained perceptual discrimination, with active controls who watched DVDs of "usual treatment" educational television programs. Analyses were intent-to-treat. Participants were 487 healthy, cognitively normal men and women aged 65–93 recruited from communities in northern and southern California, and Minnesota. They were randomized into the training ($N = 242$) or active control ($N = 245$) conditions and given computers to use at home for the trial. Trained participants completed a series of six exercises focused on improving speed and accuracy of auditory memory. Exercises used computer-adaptive algorithms to maintain challenge. The specific exercises were:

High or Low: pairs of frequency-modulated sound sweeps. Participants indicated whether the direction of the sweeps is upward (from low to more high pitched) or downward (from high to more low pitched).

Tell Us Apart: pairs of confusable syllables, such as *bō* and *dō*, are presented on the screen. One syllable was spoken and participants indicated which they heard.

Match it: a matrix of buttons was presented on the screen. Clicking a button revealed a written syllable that was spoken aloud. There were two buttons with the identical syllables in the matrix. Participants found the matched pairs; as they identified them correctly, the buttons disappeared until all were gone.

Sound Replay: Sequences of two, three, or more confusable syllables were presented auditorily. Participants listened to the syllables, then clicked buttons identifying the syllables in the order in which they were presented. There were more buttons on the screen than there were syllables, so the task involved recognition of the syllables as well as memory for their ordering.

Listen and Do: A set of spoken instructions was presented. Participants saw a scene with various characters and structures on it, with instructions to click particular characters or structures or to move the characters. Participants followed the instructions in the order given.

Story Teller: Participants listened to segments of stories and answered multiple-choice questions about them.

Active controls watched educational television program series on their computers and answered questions about the content afterwards. Both groups completed their activities 1 h a day, 5 days a week for 8 week, totaling 40 h of exposure. Computers were removed from participants' homes after they completed their training. The top panel of **Table 1** shows demographic information for the experimental and control groups.

Performance was evaluated at baseline before randomization, within 3 weeks of training completion, and 3 months later. The primary outcome was a composite index score of performance on the auditory tests of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998), a test relatively insensitive to age declines before 65. The RBANS test was developed to detect dementia in older adults but is also used to screen younger adults for impairments in cognitive status. The subtests have two alternate forms. Alternate forms were administered at each test occasion. The subtests included in the analyses are:

List learning: A 10 word list is read to the participant for study/recall over 4 trials.

Immediate List Recall: The total number of words recalled correctly over the trials.

Delayed List Recall: recall of the list after completion of seven other tests.

List Recognition: selection of the 10 study words from a list of 20 read by the examiner.

Story memory: A short story is read aloud and recalled over two trials.

Immediate Story Recall: total number of ideas recalled over the two trials.

Delayed Story Recall: recall of the story after 7 other tests.

Digit Span: digit span forwards.

The primary outcome consisted of a normed index score based on the six subtests. Secondary outcomes included performance on the trained speeded sound sweep discrimination task, and on untrained tasks: an auditory memory and attention index composite of list learning scores from the Rey Auditory and Verbal Learning Test (RAVLT) an age-sensitive and more difficult test than the RBANS (Schmidt, 1996), story memory from the Rivermead Behavioral Memory Test (Wilson et al., 2003), and letter number sequencing and digit span backwards from the Wechsler Memory Scale (Wechsler, 1997). Published findings of the IMPACT study revealed significant Group \times Time interactions shortly after the training ended on the primary outcome, on the secondary composite scores, on the trained task, and on individual test scores including RAVLT list memory and delayed list recall, WMS digits backwards, and letter-number sequencing, with larger posttest gains for the experimental group (Smith et al., 2009). Means and standard deviations of the individual tests for the experimental and control groups are published in Smith et al. (2009). Three months after training was discontinued, gains of the plasticity training group were somewhat reduced, but significant Group \times Time interactions for the trained auditory discrimination task, the secondary composite, and for RAVLT word list recall and WMS letter-number sequencing indicated retention of gains in the trained group (Zelinski et al., 2011b).

MEASUREMENT MODEL OF UNTRAINED OUTCOMES

Data from the pretest and immediate post-training assessments of the IMPACT study were analyzed. The published analyses included primary and secondary experimenter-determined outcomes measures that had not been evaluated empirically for their psychometric characteristics. Initial analyses of all subtests administered were conducted to confirm the structure of the two outcomes of the IMPACT study as latent variables so that transfer to the common construct they represented rather than to specific test scores could be appropriately assessed (see Lövdén et al., 2010; Schmiedek et al., 2010). The data were from all participants at pretest, including those who dropped out during the training phase of the study. Confirmatory factor analyses indicated very poor fit of the individual baseline tests to the published experimenter-defined measurement structure of RBANS auditory memory and to the secondary measures of the auditory memory and attention index measure. A psychometrically sound structural model of the untrained outcomes had to be developed in order to test transfer. Individual test scores were evaluated for their intercorrelations, and those with non-significant

Table 1 | Demographic information of experimental and control groups.

	Experimental	Control
N	242	245
Mean age	75.6 (6.6)	75.0 (6.3)
No of women	140	115
Mean education	15.7 (2.6)	15.6 (2.5)

Standard deviations are in parentheses.

correlations with all other tests were dropped, leaving 11 scores for further analysis.

Measurement models of the outcome variables were next assessed using *R* (*R* Project Homepage: <http://www.R-project.org>). To identify the model that best characterized the structure of the data, exploratory maximum likelihood factor analyses (*R*: psych, version 1.3.10.12), extracted 2, 3, 4, and 5 factors, with each indicator (test score) constrained to load only on one factor. A Promax rotation was used to allow factors to correlate, and no equality constraints were imposed on factor loadings. Each model was compared to an independence null model, in which covariances among all observed variables were constrained to zero. For this analysis, four fit indices to determine goodness of fit were used: *RMSEA* (root mean square error of approximation; Steiger, 1990) with a value <0.08 (Browne and Cudeck, 1992), *SRMR* (standardized root mean square residual; Joreskog and Sorbom, 1988), *TLI* (Tucker-Lewis Index; Tucker and Lewis, 1973), and *BIC* (Bayesian Information Criterion; Schwarz, 1978). Results are shown in the top panel of **Table 2**. For the 2-, 3-, and 4-factor models, fit indices were relatively poor (*RMSEA* > 0.1, *SRMR* ≥ 0.1 for 2- and 3-factor models, *TLI* < 0.9). Fit indices for the 5-factor model were acceptable, and this model produced the lowest *BIC* value out of all the models examined.

Confirmatory factor analysis (*R*: lavaan, version 0.5–15) on both the 4- and 5-factor models was next conducted. Each indicator was constrained to load only on the factor it measured and factor covariances were freely estimated. All available data were included in the maximum likelihood estimation. Four fit indices were used to determine goodness of fit: *RMSEA*, *SRMR*, *TLI*, and *CFI* (Comparative Fit Index; Bentler, 1990). Like the *TLI*, the *CFI* takes into account the χ^2 and df of hypothesized model and null model, with values ≥0.95 indicating good fit (Hu and Bentler, 1999). The χ^2 test itself was not used because the sample size of 487 was relatively large, inflating its values so that it would differ significantly from zero under most circumstances (Marsh et al., 1988).

Results of the confirmatory factor analyses supported a 5-factor model. Fit indices for the 5-factor model indicated acceptable fit, whereas fit indices for the 4-factor model were not as strong (see **Table 2**, lower panel). The 5-factor model consisted of: RBANS list memory, the RBANS list learning, list recall, and list recognition scores; RAVLT list memory, immediate recall and delayed recall measures of the Rey Auditory and Verbal Learning Test; RBANS story memory, the story memory and story recall

measures from the RBANS; RBMT story memory, the immediate and delayed story recall from the Rivermead Behavioral Memory Test, and WMS Working Memory, the Wechsler Memory Scale letter-number sequencing and backwards digit span scores.

The next step was to assess invariance of the 5-factor measurement model between the experimental and control groups at pretest to ascertain that the variables measured the same construct and therefore had the same meaning in both groups (see McArdle and Prindle, 2008). Increasingly stringent measurement invariance was assessed using four models in *R*: lavaan (version 0.5–12): configural, metric, scalar, and structural invariance. Configural invariance indicates that the variables load on the same factors across groups, but the value of the factor loadings may vary. Metric invariance indicates that the factor loadings are identical across groups. Scalar invariance indicates that the item intercepts are identical across groups, and structural invariance indicates that the factor means are identical across groups (see Horn and McArdle, 1992). Indices used to evaluate overall model fit included the normed χ^2 (χ^2/df ; Wheaton et al., 1977). A χ^2/df ratio of 3:1 or less indicates good fit (Carmines and McIver, 1981). *RMSEA* was also included in the invariance analyses. Fit statistics are shown in **Table 4**.

Results from the invariance analyses of the 5-factor model across the experimental and control groups supported the strictest measurement invariance and structural invariance, as fit did not worsen with increasing stringency of invariance tests. The models did not vary in *CFI* (.97) but the structural model resulted in a smaller *RMSEA* of 0.06 compared to the 0.07 of all other models. The χ^2/df ratio for the structural model (1.96) also indicated the best fit relative to metric (2.06), scalar (2.06), and configural (2.26) models.

This indicated that any observed differences between experimental and control groups on the factors could be interpreted as representing differences in the same constructs. **Table 3** shows the factor loadings and communalities for the tests in the five-factor model.

The five between-group invariant factors identified in the measurement model seen in **Table 3** were represented by unit weight factor scores of the tasks that loaded on each factor, that is, the sum of the scores on each of the factors. Factor scores was used instead of latent variable models of each factor because analyses estimating latent factors either did not converge or produced non-positive definite covariance matrices.

STRUCTURAL EQUATION MODELS

Multigroup structural equation models were used to test the hypothesis that latent change in each trained task was associated with latent change in each untrained variable after controlling for crossed, lagged, and cross-lagged relationships between the trained and untrained scores assessed at pre and at posttest in the experimental but not in the active control group. The model is shown in **Figure 1**. Rectangles represent manifest variables and circles latent ones. The triangle is an indicator of the latent change means. Indicators of training effects were the time order judgment sound sweep discrimination task, referred to in the tables as *speed* and the recognition of sequences of confusable syllables,

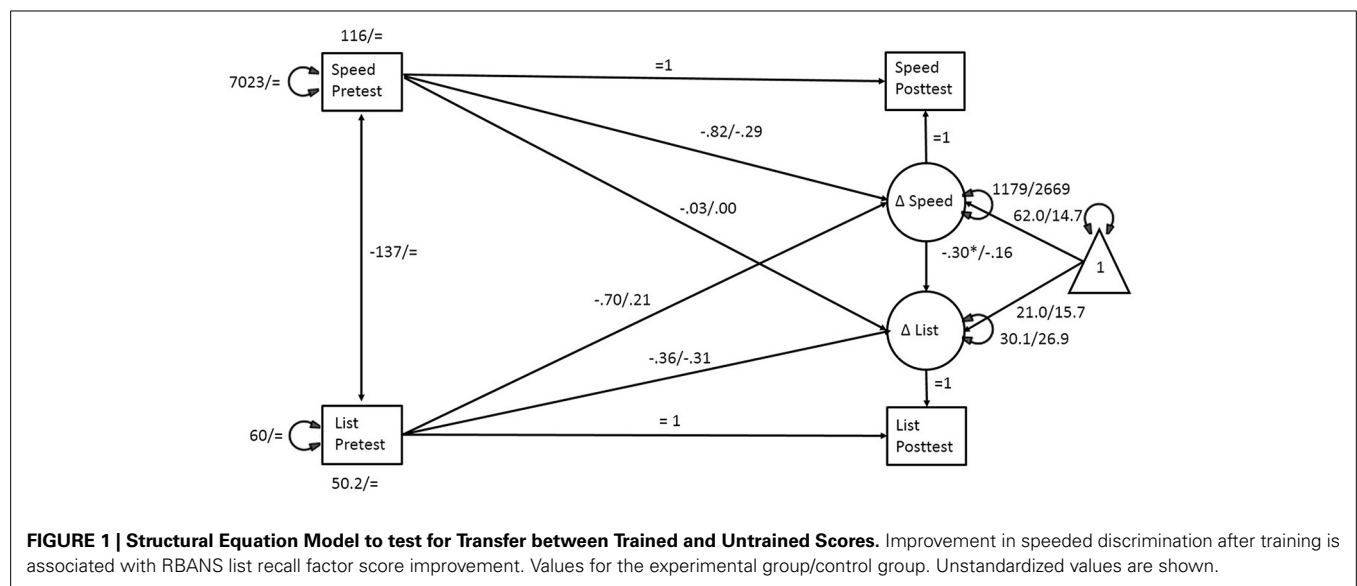
Table 2 | Results of analyses of the measurement model.

Number of Factors	RMSEA (90% CI)	SRMR	TLI	BIC	CFI
EXPLORATORY FACTOR ANALYSES					
2	0.167 (0.153–0.178)	0.15	0.753	306.64	–
3	0.152 (0.136–0.165)	0.10	0.795	163.53	–
4	0.125 (0.107–0.143)	0.07	0.860	47.50	–
5	0.052 (0.025–0.078)	0.04	0.976	–38.74	–
CONFIRMATORY FACTOR ANALYSES					
4	0.101 (0.089–0.114)	0.036	0.909	–	0.937
5	0.074 (0.061–0.088)	0.028	0.951	–	0.969

Table 3 | Standardized factor loadings of the structural invariance model for outcomes.

	RBANS list memory	RAVLT list memory	RBANS story memory	RBMT story memory	WMS working memory	h^2
RBANS						
List learning	0.81					0.60
List recall	0.90					0.93
List recognition	0.80					0.39
RAVLT						
Immediate recall		0.92				1.00
Delayed recall		0.85				0.65
RBANS						
Story recall			0.84			0.95
Story memory			0.82			0.55
RBMT STORY						
Immediate				0.87		1.00
Delayed				0.94		0.77
WMS						
Letter-number sequencing					0.87	0.82
Digits backwards					0.61	0.38

RBANS, Repeatable Battery for the Assessment of Neuropsychological Status; RAVLT, Rey Auditory Verbal Learning Test; RBMT, Rivermead Behavioral Memory Test; WMS, Wechsler Memory Scale.



referred to as *syllable span*. Analyses were conducted separately for each training effect indicator and for each of the five outcome factor scores.

The modeling approach involved estimating the maximum likelihood parameters for the illustrated bivariate change score model and testing whether selected parameters differed between the experimental and control groups. Analyses were conducted separately for each of the five outcomes and the two trained indicators in an intent-to-treat design, so that all available data, including those of the dropouts, were included. For all models, it was assumed that random assignment to groups eliminated baseline differences in test scores so that baseline intercepts for

the trained and untrained variables were set to be equal for both groups. Model 1 was set to be completely invariant over groups, with all parameters constrained to be equal. Model 2 freed the intercepts for the latent change of the training and the outcome indicators across groups with all other parameters constrained to be equal. This tested the hypothesis that training affected the means of the trained and untrained outcomes. Model 3 included the freed intercepts and the regression parameters of the crossed and lagged relationships between pretest and latent change of trained and untrained outcomes across groups. Model 4 additionally freed the variances of the latent changes for trained and untrained outcomes.

RESULTS

Table 4 shows the observed means and standard deviations for the trained and control groups on the pretest and posttest trained measures and untrained factor scores. Latent difference scores, however, were analyzed in the structural equation models.

Table 5 shows the model fit results. Fit indices included the nested $-2 \text{ Log Likelihood}$ ($-2LL$)/number of df test, which subtracts the value of $-2LL$ and df from each successive model, with the $\Delta -2LL/\Delta df$ tested using the χ^2 distribution to determine a significant improvement in fit from the prior model, with a significant $\Delta\chi^2/\Delta df$ indicating improvement in fit. This, together with the smallest AIC , and smallest $RMSEA$, was used to select the best fitting model to characterize the trained and control groups.

Results for sound sweep discrimination training are seen in **Table 6** and **Figure 1** for the training effects of speed on RBANS list memory. The models that best fit the experimental and control groups for each outcome factor score with the training indicator of speed were the ones that freed all tested parameters, indicating that those parameters in the structural model differed across groups. For all outcomes, the fit indices for Model 4 were the smallest of all four models and there were significant reductions in $-2LL$. The critical regression parameter for this study was the path from the latent change in the speed training measure to the latent change in each outcome.

Table 6 shows the unstandardized and standardized parameters for the analyses. The covariance at pretest between speed and each outcome (Speed Pre \leftrightarrow Outcome Pre) in the first row of the middle panel of **Table 6** was significant, indicating a relationship between the two measures before training. The standardized values are their correlations, which were low, ranging from -0.16 to -0.21 for the four memory factor scores and with a moderate value of -0.38 for WMS Working Memory.

Intercepts for latent changes on all of the outcomes ($1 \rightarrow \Delta \text{Outcome}$) differed significantly from zero for the experimental and control groups, suggesting that practice effects were observed in both groups. Pretest speed and pretest outcome performance were negatively associated with their respective latent changes (Speed Pre $\rightarrow \Delta \text{Speed}$; Outcome Pre $\rightarrow \Delta \text{Outcome}$), indicating greater change in those with lower baseline scores and possibly regression to the mean. This was the case for both the experimental and control groups. Crossed and lagged relationships between speed and outcome measures were significant,

Table 4 | Means and Standard Deviations (in parentheses) for the pretest and posttest scores on the trained tasks and untrained task factor scores for the experimental and control groups.

	Pre-test		Post-test	
	Experimental	Control	Experimental	Control
Speed	115.8 (83.8)	116.9 (84.2)	47.7 (38.6)	105.4 (75.8)
Syllable span	3.6 (0.51)	3.6 (0.56)	4.1 (0.57)	3.7 (0.59)
RBANS list memory	50.1 (7.8)	50.3 (7.7)	51.9 (7.6)	51.3 (7.3)
RAVLT list memory	47.0 (12.6)	48.1 (13.4)	48.9 (13.6)	47.8 (12.2)
RBANS story memory	26.1 (5.1)	26.5 (5.0)	27.3 (4.9)	27.6 (5.1)
RBMT story memory	14.3 (6.0)	14.5 (6.4)	15.6 (6.2)	15.8 (6.4)
WMS working memory	17.0 (4.0)	16.8 (4.5)	18.3 (4.2)	17.3 (4.6)

Table 5 | Nested tests of fit for models with speed (top panel) or syllable span (bottom panel) and each of the outcome factor scores testing parameter differences between experimental and active control groups.

	$-2LL$	df	$\Delta 2LL/\Delta df$	AIC	$RMSEA$ (90% CI)
MODELS WITH SPEED					
RBANS List Memory					
Model 1	−8243	14	—	16515	0.28 (0.24–0.31)
Model 2	−8175	16	68/2	16382	0.21 (0.17–0.24)
Model 3	−8125	21	50/5	16292	0.13 (0.09–0.18)
Model 4 ^a	−8107	23	19/2	16260	0.00 (0.00–0.00)
RAVLT List Memory					
Model 1	−8726	14	—	17481	0.28 (0.25–0.31)
Model 2	−8656	16	70/2	17344	0.21 (0.18–0.24)
Model 3	−8605	21	51/5	17252	0.14 (0.10–0.18)
Model 4 ^a	−8587	23	18/2	17220	0.00 (0.00–0.06)
RBANS Story Memory					
Model 1	−7925	14	—	15878	0.28 (0.25–0.30)
Model 2	−7856	16	69/2	15744	0.20 (0.17–0.24)
Model 3	−7808	21	48/5	15659	0.14 (0.09–0.18)
Model 4 ^a	−7790	23	18/2	15627	0.00 (0.00–0.06)
RBMT Story Memory					
Model 1	−8152	14	—	16333	0.28 (0.25–0.30)
Model 2	−8084	16	68/2	16200	0.21 (0.17–0.24)
Model 3	−8037	21	47/5	16115	0.14 (0.10–0.18)
Model 4 ^a	−8019	23	18/2	16083	0.00 (0.00–0.07)
WMS Working Memory					
Model 1	−7607	14	—	15243	0.28 (0.25–0.30)
Model 2	−7534	16	73/2	15101	0.20 (0.17–0.23)
Model 3	−7492	21	42/5	15026	0.15 (0.11–0.19)
Model 4 ^a	−7474	23	18/2	14993	0.03 (0.00–0.10)
MODELS WITH SYLLABLE SPAN					
RBANS List Memory					
Model 1	−3630	14	—	7288	0.18 (0.15–0.21)
Model 2 ^a	−3573	16	57/2	7179	0.03 (0.00–0.07)
Model 3	−3569	21	4/5	7180	0.00 (0.00–0.07)
Model 4	−3567	23	2/2	7180	0.00 (0.00–0.05)
RAVLT List Memory					
Model 1	−4100	14	—	8229	0.19 (0.16–0.22)
Model 2	−4043	16	57/2	8118	0.05 (0.00–0.09)
Model 3 ^a	−4036	21	7/5	8115	0.01 (0.00–0.08)
Model 4	−4035	23	1/2	8116	0.00 (0.00–0.05)
RBANS Story Memory					
Model 1	−3311	14	—	6650	0.19 (0.16–0.22)
Model 2 ^a	−3252	16	59/2	6536	0.05 (0.00–0.09)
Model 3	−3247	21	5/5	6536	0.01 (0.00–0.08)
Model 4	−3244	23	3/2	6536	0.00 (0.00–0.07)
RBMT Story Memory					
Model 1	−3555	14	—	7140	0.18 (0.16–0.21)
Model 2	−3500	16	45/2	7033	0.05 (0.00–0.09)
Model 3 ^a	−3495	21	5/5	7031	0.01 (0.00–0.08)
Model 4	−3493	23	2/2	7032	0.00 (0.00–0.07)

(Continued)

Table 7 | Maximum likelihood estimates and standardized parameters (in parentheses) of the best-fitting bivariate models for effects of syllable span training in the experimental and control groups.

	RBANS list memory		RAVLT list memory		RBANS story memory		RBMT story memory		WMS working memory	
	Exp	Cntl	Exp	Cntl	Exp	Cntl	Exp	Cntl	Exp	Cntl
FIXED PARAMETERS: TRAINED										
1 → Syllable Pre	3.6* (6.8)	=	3.6* (6.8)	=	3.6* (6.8)	=	3.6* (6.8)	=	3.6* (6.8)	=
1 → ΔSyllable	0.89* (2.3)	0.49* (1.3)	0.83* (2.2)	0.84* (2.2)	0.86* (2.3)	0.46* (1.2)	0.92* (2.4)	0.84* (2.2)	0.95* (2.5)	0.96* (2.4)
Syllable Pre → ΔSyllable	-0.20* (-0.28)	=	-0.16* (-0.23)	-0.22* (-0.31)	-0.22* (-0.30)	=	-0.12* (-0.17)	-0.22* (-0.30)	-0.31* (-0.44)	-0.42* (-0.58)
Outcome Pre → ΔSyllable	0.01* (0.13)	=	0.01* (0.17)	0.00 (0.03)	0.02* (0.20)	=	0.00 (0.02)	0.00 (0.05)	0.04* (0.43)	0.04* (0.42)
FIXED PARAMETERS: OUTCOMES										
Syllable Pre ↔ Outcome Pre	1.47* (0.36)	=	2.68* (0.39)	=	0.86* (0.32)	=	0.75* (0.23)	=	1.44* (0.64)	=
1 → Outcome Pre	50.2* (6.5)	=	47.5* (3.7)	=	26.3* (5.2)	=	14.4* (2.3)	=	16.9* (4.0)	=
1 → ΔOutcome	14.9* (2.5)	14.3* (2.4)	-6.8* (-0.70)	9.4* (0.95)	7.6* (1.6)	8.1* (1.7)	3.0 (0.45)	0.08 (0.01)	-3.3* (-1.2)	-0.96 (-0.37)
Outcome Pre → ΔOutcome	-0.35* (-0.45)	=	-0.32* (-0.42)	-0.37* (-0.49)	-0.54* (-0.56)	=	0.65* (-0.59)	-0.49* (-0.48)	-0.50* (-0.71)	-0.28* (-0.45)
Syllable Pre → ΔOutcome	1.04* (0.09)	=	5.9* (0.33)	2.02 (0.11)	1.9* (0.21)	=	1.7* (0.14)	2.2* (0.19)	3.2* (0.59)	1.7* (0.34)
ΔSyllable → ΔOutcome	0.73 (0.05)	=	4.3* (0.17)	0.94 (0.04)	1.2* (0.10)	=	2.7* (0.15)	-0.40 (-0.03)	2.6* (0.34)	-0.07 (-0.01)
RANDOM PARAMETERS: VARIANCES OF LATENT CHANGE										
Outcome	28.9* (.82)	=	77.7* (0.79)	=	170* (0.73)	=	29.7* (0.77)	=	5.92* (0.87)	=
Syllable	0.14* (0.93)	=	0.14* (0.91)	=	0.13* (0.91)	=	0.14* (0.91)	=	0.12* (0.80)	=

**p* < 0.05. Equals signs indicates that the parameter was constrained to be equal for the experimental and control groups.

The differences in parameters for the two RBANS factor scores in Model 2 suggested that the trained group only differed from the control group in the amount of improvement in the model intercepts but not the regression parameters. These did differ for the remaining outcome factor scores for which Model 3 was the best fit. The critical test of the relationship between latent change in the trained and in the untrained scores was significant for RAVLT List Memory, RBMT Story Memory, and WMS Working Memory for the experimental but not the control group, indicating evidence of transfer. In addition, the relationship between the two latent change variables was significant for syllable span and RBANS Story Memory but because the path was constrained to be equal for experimental and control groups in that model, it did not demonstrate transfer of training as defined in the analysis.

The final set of analyses tested whether transfer was associated with individual differences. They included the covariates of age, sex, and education, all of which were associated with baseline training task performance. Bivariate change models tested the baseline and latent change trained and outcome variables regressed on the covariates, with covariate effects fixed across experimental and control groups, because of random assignment. The critical relationships of latent changes in training and outcomes were free to vary. **Table 8** shows the standardized estimates for speed and syllable span, which were identical across outcomes, and **Table 9** the standardized estimates for each of the five outcomes, which were identical across training task analyses, and for the latent change- to- latent change regression coefficients for each training task.

For Speed, the pretest scores only were associated with the covariates. Being younger and male were associated with lower (faster) speed. Paradoxically, having more years of education was associated with slower performance. No correlations were observed for the latent change of speed. Age was negatively associated with syllable span at baseline, with worse performance, and more education was associated with higher scores. For latent change in syllable span, being older was associated with less gain and more highly educated with more gain. There were no sex differences in associations with baseline or latent change syllable span.

The covariates, as expected, had significant relationships with the baseline outcome factor scores, as seen in **Table 9**. Older people had lower baseline scores on all of the outcomes. Women were

Table 8 | Standardized regression parameters for the analyses of the regression of training task variables on age, sex, and education.

	Speed	Syllable span
TRAINING TASKS		
Age → Trained Pre	0.28*	-0.41*
Sex → Trained Pre	0.14*	0.09
Education → Trained Pre	-0.20*	0.17*
Age → ΔTrained	0.03	-0.12*
Sex → ΔTrained	-0.01	-0.04
Education → ΔTrained	-0.05	0.11*

Parameters were constrained to be identical across training groups. **p* < 0.05.

Table 9 | Standardized parameters for analyses with covariates.

	Outcomes									
	RBANS list memory		RAVLT list memory		RBANS story memory		RBMT story memory		WMS working memory	
	Exp	Cntl	Exp	Cntl	Exp	Cntl	Exp	Cntl	Exp	Cntl
FIXED PARAMETERS: TRAINED										
Age → Outcome Pre	−0.36*	=	−0.37*	=	−0.28*	=	−0.18*	=	−0.21*	=
Sex → Outcome Pre	0.27*	=	0.25*	=	0.06	=	−0.06	=	−0.49	=
Education → Outcome Pre	0.14*	=	0.11*	=	0.13*	=	0.20*	=	0.22*	=
Age → Δ Outcome	−0.15*	=	−0.26*	=	−0.22*	=	−0.18*	=	−0.07*	=
Sex → Δ Outcome	0.16*	=	0.07	=	0.03	=	−0.06	=	0.19	=
Education → Δ Outcome	0.06	=	0.08	=	0.02	=	0.04	=	0.10*	=
Δ Speed → Δ Outcome	−0.30*	−0.05	−0.14	−0.08	−0.21	−0.05	−0.11	0.05	−0.00	0.00
Δ Syllable → Δ Outcome	0.04	=	0.14*	0.02	0.08	=	0.12*	−0.04	0.33*	−0.02

* $p < 0.05$. Equals signs indicates that the parameter was constrained to be equal for the experimental and control groups.

better on baseline list memory factor scores, for both the RBANS and RAVLT. More education was associated with better baseline performance on all five factor scores. Age was associated with latent changes in the outcome variables, with less gain for older individuals. Female gender was associated with larger gains on RBANS List Memory, and more education with greater gains on WMS Working Memory.

Despite the relationships of covariates with the outcomes at pretest and for their latent changes, all of the significant latent change training-latent change outcome relationships observed in the main bivariate analyses for the experimental but not the control group remained significant after accounting for covariates. Transfer was therefore independent of the covariates.

DISCUSSION

The goal of cognitive training of older adults is to support them in either maintaining or improving their functioning. Critical to this is the effectiveness of training in producing transfer. It has been suggested that multimodal cognitive training will produce transfer to multiple outcomes (e.g., Basak et al., 2008). However, it is not clear whether transfer is more likely to be observed, in the context of multimodal training, in training tasks that have greater demand overlaps with outcomes, and this was a focus of the present study.

Data modeling included controlling for relationships in performance between trained and untrained tasks not only at baseline, but subsequent to training, in a study dataset that showed improvement in untrained task performance after training at the group level. The data source was the IMPACT study, which involved a design with many strengths, including being the largest multisite randomized controlled double-blind trial of a commercially available cognitive training program with 487 participants over age 65 in experimental and control groups. It included an active control group and was conducted at three different sites. Published results showed interactions between experimental/control group participation and assessment visit, with the trained participants showing better performance, and Cohen's d effect sizes for the interaction ranging from 0.20 to 0.33 (Smith

et al., 2009). However, like most studies in the cognitive training literature, data analyses were only conducted at the group level and only one training effect was reported.

Transfer from a task assessing the speed of discriminating time-ordered sound sweeps was assumed to reflect relatively less task demand overlap with the outcome constructs than transfer from a task assessing expansion of syllable span. Results suggested that transfer to a relatively easy list memory outcome was associated with improvement in the training indicator of speed, and that transfer to relatively difficult list memory, story memory, and working memory outcomes were associated with improvement in the training indicator of syllable span.

Because change in the speeded non-verbal training task was associated only with latent change of one memory task factor score, its utility in the measurement of transfer in this study was limited. Processing speed has long been characterized as a cognitive primitive (e.g., Salthouse, 1996) that underlies age related performance declines in many cognitive tasks, including memory. Perceptual speed was significantly associated with memory for word lists but not for text memory in cross-sectional research (Lewis and Zelinski, 2010). However, perceptual speed training gain in the present study showed transfer only to one factor score from a neuropsychological test that does not differentiate performance at ages under 65 (Randolph, 1998). The task demand explanation would suggest that rapid processing of non-verbal auditory information overlaps only somewhat with skills involved with rapid processing of the relatively low-retrieval demand material of the RBANS list memory factor scores. That score is based on a 10-item 4-trial free recall + delayed recall of the same list. In comparison, the RAVLT list memory factor score is based on a 15-item 5-trial free recall + free recall of an interference list followed by initial list recall, + delayed interference list recall. A lack of transfer was also observed for training on a perceptual speed task and list recall in the ACTIVE trial (Ball et al., 2002) as well. This suggests that improving on a non-verbal training task with a fixed and low memory load has only limited value as an indicator of transfer to gains in verbal memory.

On the other hand, improvement in syllable span was associated with transfer to the more difficult RAVLT list memory, RBMT Story Memory, and WMS Working Memory factor scores. Age declines in working memory performance are well documented, and working memory has been considered to be an important mechanism in word list recall and text recall, as coordinating to-be-remembered information in working memory contributes to retrieval of both item- and discourse-level information (e.g., Lewis and Zelinski, 2010). The largest standardized parameter was observed for the effect of gains in syllable span on gains in the factor score derived from two re-sequencing span measures. It was predicted that the transfer relationship would be stronger for working memory outcomes than for recall outcomes because of similarities in span task demands. This was confirmed. The standardized coefficients for list and story memory transfer, on the other hand, were similar. Pretest correlations were greater for syllable span and the outcomes than for speed and outcomes, suggesting more commonalities of syllable span with transfer measures at baseline. Because those relationships before and after training were covaried in the structural equation model, the relationship of latent changes in training and in transfer was independent of those influences.

If the present analysis had only included the targeted sound sweep discrimination measure, the argument of transfer from the training program would be only weakly supported. By analyzing gains on another training task, the transfer findings suggests extension to more outcomes that tap into similar constructs as those trained. Thus, in general, the findings support an overlapping task demand model of transfer not due to confounding of crossed, lagged, or cross-lagged relationships.

The findings of task-specific transfer are confirmed by several studies reporting limited transfer between different working memory/cognitive control tasks and untrained working memory tasks (Buschkuhl et al., 2008; Li et al., 2008; Karbach and Kray, 2009; Schmiedek et al., 2010). Dahlin et al. (2008) found that, after working memory training, brain activations in young adults increased in the striatum during working memory updating training as well as during transfer tasks. Older adults showed activation during the trained but not the transfer task and showed no evidence of behavioral transfer. Thus, transfer may suggest similarity of functional neural activation patterns between the trained and transfer tasks, but this is not consistently observed (see Buschkuhl et al., 2012).

In the present study, individual differences among participants affected latent change independent of baseline functioning. Increasing age was associated with reduced latent change in all measures except for speed, female sex was associated with more latent change in RBANS List memory, and more years of education with more latent change in syllable span and in WMS Working Memory. This suggests that, as found elsewhere, very elderly adults gain less from training than younger ones, but they do show some benefit (see Buschkuhl et al., 2008; Hertzog et al., 2009; von Bastian et al., 2013). Female gender and more education were associated with better baseline cognitive performance, as is often observed, but this is the first study to demonstrate a benefit for women in list recall and for more years of schooling in training and transfer gains in working memory span tasks. Most critically,

significant transfer in the experimental group only from latent trained change to latent outcome change remained significant.

METHODOLOGICAL IMPLICATIONS

The findings confirm the value of assessing relationships between trained and untrained scores in evaluating transfer. In all cases, there were significant pretraining relationships between the trained task and outcome factor scores for both experimental and control groups. The findings of significant intercepts for latent change in the models for both trained and control participants showed that practice effects were present in both groups. Practice may inflate the apparent training effect size considerably if only the data of experimental groups are included in transfer task effect size computation (see Hindin and Zelinski, 2012). Many training studies only use repeated measures ANOVA of untrained tasks to assess transfer, which accounts for practice, but this study suggests that such findings may be compromised by the complex of pretraining and posttraining relationships between trained and transfer measures.

Recently, theoretical concerns about the interpretation of correlational relationships of gains in trained and transfer variables based on observed strong relationships between baseline task performance measures have been raised (e.g., Redick et al., 2013; Tidwell et al., 2014). It has been assumed that strong baseline relationships indicate that gain score relationships in the trained group reflect a causal change. In the working memory literature, the very strong baseline relationship between working memory and intelligence has been suggested by some as evidence that working memory training can improve intelligence. This has led to the use of analyses that produce misleading results.

Several recent studies that did not report training group differences in transfer used responder analyses to test for training effects (e.g., Jaeggi et al., 2011; Redick et al., 2013; Novick et al., 2014). The idea is that because not all participants improve with training, they should be categorized based on training outcomes, with correlations of change scores for trained and untrained tasks within successful and unsuccessful outcome groups computed. As Tidwell et al. (2014) have shown, this categorization is problematic because of lack of inclusion of control participants, a restriction of range for correlations, and spurious relationships between changes in training and transfer.

In addition, Moreau and Conway (2014) showed that even if training did produce transfer, strong pretest correlations do not guarantee strong gain correlations. Gains on both tasks may be negligibly related, for a number of reasons, but especially if the gain score correlations are computed for manifest variables, which contain error. Negative relationships between pretest trained and untrained scores and their respective changes, possibly because of regression to the mean, have also been observed in training studies (e.g., Whitlock et al., 2012). Shipstead et al. (2010) note that this problem affects outcomes, but is generally ignored. Because of these measurement problems, it is crucial to assess the relationship between training and transfer change independent of all major confounding relationships and to assess latent change, which is free of error. Another issue is that studies in the training literature rarely use intent-to-treat analyses, which include all pretested participants, and any training data, even of

dropouts, to represent all available data, not just that of those self- or experimenter-selected to participate. When maximum likelihood algorithms are used in modeling with all available data, this reduces the possibility that systematic individual differences in dropout characteristics leads to biased findings. One of the serious problems in the training literature is that most published experiments do not include sample sizes adequate for the sophisticated modeling of effects that account for possible confounds as presented here. Many studies are additionally underpowered in terms of sample size and duration of training, thus limiting exposure to the intervention (see Basak et al., 2008 for an example).

We therefore agree that correlational modeling, as practiced in the literature, suffers from interpretive problems, and that unless the complex of interrelationships between trained and transfer measures is assessed and covaried in all participants within the experimental and control groups, latent variables are evaluated, and all available data are modeled, the problems described here lead to interpretive difficulties.

Suggestions have also been made that biases in interpretation of training effects exist because, in effect, competing hypotheses rather than the null hypothesis, are being evaluated. In the working memory training literature, the hypothesis that training transfers to abilities like intelligence assumes that the null is simply the absence of transfer. However, an alternative hypothesis is implied by the intelligence literature, which suggests that the abilities cannot be improved by training (see Tidwell et al., 2014). The Bayesian approach evaluates the likelihood that findings support the null vs. a transfer hypothesis. Following Sprenger et al. (2013), we computed Bayes-factor analysis of the Group \times Time interaction effects observed in the Smith et al. (2009) paper, transforming them to two-sample t 's because there was 1 df in the numerator of the F ratio. We found that one of the seven previously significant interactions on untrained tasks was shown instead to support the null hypothesis with a Bayes factor value of 1.59. A total of 9 untrained task scores (including those that were not significant) was analyzed to compute the median Bayes Factor, which, for all reported outcomes, was 0.79, thus in favor of modest transfer effects.

Hindin and Zelinski (2012) assessed quality of extended practice training studies in their meta-analysis and found that studies with higher quality (measured with respect to random assignment to conditions, reports of attrition, sample size, etc.) had larger effect sizes for transfer tasks. The mean estimated effect size of $d = 0.32$, equivalent to $r = 0.16$, associated with transfer in older adults (Hindin and Zelinski, 2012) may seem inconsequential relative to effect sizes for pre-post change in a trained task. However, many medical interventions become clinical practice with much smaller effect sizes, for example $r = 0.02$ for the effect of aspirin and reduced risk of death by heart attack (Meyer et al., 2001). Provigil (Modafinil), a narcolepsy drug, used off-label to improve working memory and attention, has an estimated mean effect size on working memory and similar tasks of $r = 0.11$ or $d = 0.23$ in young adults (Hindin and Zelinski, 2012). Although expectation of substantial transfer effects, that is, those as large as effects for improvements in pre- to post-task training, may be unrealistic, we note that transfer effects for working memory

interventions, largely in children, as shown by Melby-Lervag and Hulme (2013) are smaller and not different from zero. Older adults may show more transfer from training than young adults on average because their baseline performance is worse due to reduced neuroplasticity, which is re-engaged with training (see Mahncke et al., 2006).

LIMITATIONS

Tidwell et al. (2014) suggest that computation of correlations between trained and transfer tasks are uninformative because it is likely that measurement characteristics of the training task are not invariant as a result of exposure. This is a concern for the current study, but individual item scores were unavailable for differential item functioning analyses before and after training.

Concerns raised in the literature include the observation that training is adaptive whereas active control conditions generally are not, and this was true of the present study. Though this could bias findings because adaptive training promotes performance improvements to a greater extent than standardized training, and because there may be different levels of motivation and strategy use that may affect outcomes in experimental and control groups, the evidence for this potential source of spurious training and transfer effects is quite weak (see Redick et al., 2013).

In the present study, there was a trained group and an active control condition with double blinding. A concern in clinical trials, even with double blinding, is whether the trained group gets more attention from study staff and whether there is an implicit message because of unchallenging sham material that control participants are not getting the experimental treatment, so that they experience less social interaction and expect less improvement, both of which dampen performance. In the present study, there were no differences in the amount of interaction with trainers for the two treatment groups. Participants had been told that after the study was completed, they would receive upon request copies of the training materials that produced better outcomes on the untrained tasks. Some of the control participants requested copies of the DVDs they had watched. This suggests that expectancies of cognitive benefits, which could affect performance, were present in some control participants (see Boot et al., 2013), but this was not systematically assessed so it is unknown whether the majority of those in the control group did expect to improve and to the same degree as those in the training condition on the outcomes.

The study was not informative regarding change in underlying processes compared to overlap in similarities in task characteristics. This could not be evaluated for three reasons. First, the neural basis of overlap was not tested. Second, the multimodal training design could not rule out complex sources of transfer. Third, the speeded auditory discrimination and syllable span tasks differed with respect to whether they were non-verbal or verbal, as well as on their measurement characteristics. Though the findings would suggest that syllable span was more effective for transfer to recall memory than time-ordered sound sweep discrimination, we note that training effects from the four other trained tasks in the program used in the IMPACT study could not be assessed. We also note that all training tasks involved adaptive speeded processing and difficult auditory discrimination training, and that with the

extant design, the specific benefits to transfer within the training program could not be isolated.

We note that what constitutes near and far transfer has not yet been objectively defined, and varies from study to study, so prediction of the amount of transfer that should be observed for a given outcome is difficult. In the present study, the most parsimonious explanation for performance improvements on untrained tasks in older adults is that of overlap in task demands, because training was multimodal. This is an important limitation. However, improvement in untrained tasks rather than broad abilities in older adults may have important implications for public health. The ACTIVE trial showed that training of reasoning and of speed was associated with reductions in risk of dependency 10 years after the study was initiated (Rebok et al., 2014). We agree, though, that elucidating the mechanisms of transfer is a critical goal for the cognitive training literature. Promising approaches for understanding the basis of transfer include testing neural activation patterns during task performance (e.g., Dahlin et al., 2008) and developing targeted tasks that clearly vary process engagement (e.g., Persson et al., 2007).

Other limitations to this study are those of the IMPACT study inclusion and exclusion criteria. This resulted in a convenience sample of very healthy participants, with high fluency in English, and low participation rates by members of ethnic minorities. Participants had committed to engage in the study for a minimum of 6 months. These characteristics suggest that the findings may not be generalizable to the population of older adults.

CONCLUSIONS

The findings have positive implications for the cognitive training of older adults who are healthy and willing to engage in challenging and extensive multimodal training such as that provided in the IMPACT study. The current set of findings suggest that even when individual differences including age are incorporated into models that test transfer independent of other possible within-study influences, the relationship between latent changes in trained and untrained tasks generally remains significant.

ACKNOWLEDGMENTS

This research was supported in part by the Center for Digital Aging of the USC Davis School of Gerontology, grants R01AG10569, P50AG005142, and T32AG00037 from the National Institute on Aging and H133E080024 from the National Institute on Disability and Rehabilitation Research to the University of Southern California, and a grant from Posit Science Corporation.

REFERENCES

- Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related timing delays with training. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4357–4362. doi: 10.1073/pnas.1213551110
- Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., et al. (2013). Video game training enhances cognitive control in older adults. *Nature* 501, 97–101. doi: 10.1038/nature12486
- Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., et al. (2002). Effects of cognitive training interventions with older adults: a randomized controlled trial. *J. Am. Med. Assoc.* 288, 2271–2281. doi: 10.1001/jama.288.18.2271
- Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy videogame attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765–777. doi: 10.1037/a0013494
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychol. Bull.* 107, 238–246. doi: 10.1037/0033-2909.107.2.238
- Boot, W. R., Simons, D. J., Stothart, C., and Stutts, C. (2013). The pervasive problem with placebos in psychology: why active control groups are not sufficient to rule out placebo effects. *Perspect. Psychol. Sci.* 8, 445–454. doi: 10.1177/1745691613491271
- Browne, M. W., and Cudeck, R. (1992). Alternative ways of assessing model fit. *Soc. Methods Res.* 21, 230–258. doi: 10.1177/00491241920021002005
- Buschkuhl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Däpp, C., Müller, M., et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychol. Aging* 23, 743–753. doi: 10.1037/a0014342
- Buschkuhl, M., Jaeggi, S. M., and Jonides, J. (2012). Neuronal effects following working memory training. *Dev. Cogn. Neurosci.* 25, S167–S179. doi: 10.1016/j.dcn.2011.10.001
- Carmine, E. G., and McEver, J. P. (1981). “Analyzing models with unobserved variables: analysis of covariance structures,” in *Social Measurement: Current Issues* eds G. W. Bohrnstedt and E. F. Borgatta (Beverly Hills, CA: Sage Publications), 65–115.
- Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., and Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science* 320, 1510–1512. doi: 10.1126/science.1155466
- Hertzog, C., Kramer, A. F., Wilson, R. S., and Lindenberger, U. (2009). Enrichment effects on adult cognitive development: can the functional capacity of older adults be preserved or enhanced? *Psychol. Sci. Public Interest* 9, 1–65. doi: 10.1111/j.1539-6053.2009.01034.x
- Hertzog, C., Touron, D. R., and Hines, J. C. (2007). Does a time-monitoring deficit influence older adults' delayed retrieval shift during skill acquisition? *Psychol. Aging* 22, 607–624. doi: 10.1037/0882-7974.22.3.607
- Hindin, S., and Zelinski, E. M. (2012). Extended practice and aerobic exercise interventions benefit untrained cognitive outcomes in older adults: a meta-analysis. *J. Am. Geriatr. Soc.* 60, 136–141. doi: 10.1111/j.1532-5415.2011.03761.x
- Horn, J. L., and McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Exp. Aging Res.* 18, 117–144. doi: 10.1080/03610739208253916
- Hu, L. T., and Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., and Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108
- Jobe, J. B., Smith, D. M., Ball, K., Tennstedt, S. L., Marsiske, M., Willis, S. L., et al. (2001). ACTIVE: a cognitive intervention trial to promote independence in older adults. *Control. Clin. Trials* 22, 453–479. doi: 10.1016/S0197-2456(01)00139-8
- Joreskog, K. G., and Sorbom, D. (1988). *LISREL 7 – A Guide to the Program and Applications, 2nd Edn.* Chicago, IL: SPSS.
- Karbach, J., and Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Dev. Sci.* 12, 978–990. doi: 10.1111/j.1467-7687.2009.00846.x
- Lewis, K. L., and Zelinski, E. M. (2010). List and text recall differ in their predictors: Replication over samples and time. *J. Gerontol. Psychol. Sci.* 65B, 449–458. doi: 10.1093/geronb/gbq034
- Li, S., Schmiedek, F., Huxhold, O., Rocke, C., Smith, J., and Lindenberger, U. (2008). Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychol. Aging* 23, 731–742. doi: 10.1037/a0014343
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., and Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychol. Bull.* 136, 659–676. doi: 10.1037/a0020080
- Mahncke, H. W., Bronstone, A., and Merzenich, M. M. (2006). Brain plasticity and functional losses in the aged: scientific bases for a novel intervention. *Prog. Brain Res.* 157, 81–109. doi: 10.1016/S0079-6123(06)57006-2
- Marsh, H. W., Balla, J., and McDonald, R. P. (1988). Goodness of fit indexes in confirmatory factor analysis: the effect of sample size. *Psychol. Bull.* 103, 391–410. doi: 10.1037/0033-2909.103.3.391

- McArdle, J. J., and Prindle, J. J. (2008). A latent change score analysis of a randomized clinical trial in reasoning training. *Psychol. Aging* 23, 702–719. doi: 10.1037/a0014349
- Melby-Lervag, M., and Hulme, C. (2013). Is working memory effective? A meta-analytic review. *Dev. Psychol.* 49, 270–291. doi: 10.1037/a0028228
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., et al. (2001). Psychological testing and psychological assessment: a review of evidence and issues. *Am. Psychol.* 56, 128–165. doi: 10.1037/0003-066X.56.2.128
- Moreau, D., and Conway, A. R. A. (2014). The case for an ecological approach to cognitive training. *Trends Cogn. Sci.* 18, 1364–1366. doi: 10.1016/j.tics.2014.03.009
- Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., and Bunting, M. F. (2014). Clearing the garden-path: improving sentence processing through cognitive control training. *Lang. Cogn. Neurosci.* 29, 186–217. doi: 10.1080/01690965.2012.758297
- Park, D. C., Gutchess, A. H., Meade, M. L., and Stine-Morrow, E. A. L. (2007). Improving cognitive function in older adults: nontraditional approaches. *J. Gerontol. B* 62B, 45–52. doi: 10.1093/geronb/62.special_issue_1.45
- Persson, J., Welsh, K. M., Jonides, J., and Reuter-Lorenz, P. A. (2007). Cognitive fatigue of executive processes: interaction between interference resolution tasks. *Neuropsychologia* 45, 1571–1579. doi: 10.1016/j.neuropsychologia.2006.12.007
- Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status*. San Antonio, TX: Psychological Corporation.
- Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., et al. (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. *J. Am. Geriatr. Soc.* 62, 16–24. doi: 10.1111/jgs.12607
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *J. Exp. Psychol. Gen.* 142, 359–379. doi: 10.1037/a0029082
- R Project Homepage: <http://www.R-project.org/>
- Salthouse, T. A. (1996). The processing-speed theory of adult agedifferences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403
- Schmidt, M. (1996). *Rey Auditory and Verbal Learning Test: A Handbook*. Los Angeles, CA: Western Psychological Services.
- Schmiedek, F., Lövdén, M., and Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study. *Front. Aging Neurosci.* 2:27. doi: 10.3389/fnagi.2010.00027
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136
- Shipstead, Z., Redick, T. S., and Engle, R. W. (2010). Does working memory training generalize? *Psychol. Belg.* 50, 245–276. doi: 10.5334/pb-50-3-4-245
- Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., et al. (2009). A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *J. Am. Geriatr. Soc.* 57, 594–603. doi: 10.1111/j.1532-5415.2008.02167.x
- Sprenger, A. M., Atkins, S. M., Bolger, D. J., Harbison, J. I., Novick, J. M., Chrabaszcz, J. S., et al. (2013). Training working memory: limits of transfer. *Intelligence* 41, 638–663. doi: 10.1016/j.intell.2013.07.013
- Steiger, J. H. (1990). Structural model evaluation and modification. *Multivar. Behav. Res.* 25, 214–212. doi: 10.1207/s15327906mbr2502_4
- Stepankova, H., Lukavsky, J., Buschkuehl, M., Kopecek, M., Ripova, D., and Jaeggi, S. M. (2014). The malleability of working memory and visuospatial skills: a randomized controlled study in older adults. *Dev. Psychol.* 50, 1049–1059. doi: 10.1037/a0034913
- Tidwell, J. W., Dougherty, M. R., Chrabaszcz, J. R., Thomas, R. P., and Mendoza, J. L. (2014). What counts as evidence for working memory training? Problems with correlated gains and dichotomization. *Psychon. Bull. Rev.* 21, 620–628. doi: 10.3758/s13423-013-0560-7
- Tucker, L. R., and Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38, 1–10. doi: 10.1007/BF02291170
- von Bastian, C., Langer, N., Jäncke, L., and Oberauer, K. (2013). Effects of working memory training in young and old adults. *Mem. Cognit.* 41, 611–624. doi: 10.3758/s13421-012-0280-7
- Wechsler, D. (1997). *Wechsler Memory Scale-3rd*. San Antonio, TX: Psychological Corporation.
- West, R. L., Bagwell, D. K., and Dark-Freudeman, A. (2008). Self-efficacy and memory aging: the impact of a memory intervention based on self-efficacy. *Aging Neuropsychol. Cogn.* 15, 302–329. doi: 10.1080/13825580701440510
- Wheaton, B., Muthén, B., Alwin, D., and Summers, G. (1977). “Assessing reliability and stability in panel models,” in *Sociological Methodology* ed D. R. Heise (San Francisco CA: Jossey-Bass, Inc). 84–136.
- Whitlock, L. A., McLaughlin, A. C., and Allaire, J. C. (2012). Individual differences in response to cognitive training: using a multi-modal, attentionally demanding game-based intervention for older adults. *Comput. Hum. Behav.* 28, 1091–1096. doi: 10.1016/j.chb.2012.01.012
- Wilson, B., Cockburn, J., and Baddeley, A. (2003). *The Rivermead Behavioral Memory Test - 2nd Supplement Two*. San Antonio, TX: Harcourt Assessment.
- Zelinski, E. M. (2009). Far transfer in cognitive training of older adults. *Restor. Neurol. Neurosci.* 27, 455–471. doi: 10.3233/RNN-2009-0495
- Zelinski, E. M., Dalton, S. E., and Hindin, S. (2011a). Cognitive changes in healthy older adults. *Generations* 35, 13–20.
- Zelinski, E. M., and Gilewski, M. J. (2003). Effects of demographic and health variables on Rasch scaled cognitive scores. *J. Aging Health* 15, 435–464. doi: 10.1177/0898264303253499
- Zelinski, E. M., Spina, L. M., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., et al. (2011b). Improvement in memory with plasticity-based adaptive cognitive training (impact): results of the 3-month follow-up. *J. Am. Geriatr. Soc.* 59, 258–265. doi: 10.1111/j.1532-5415.2010.03277.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 January 2014; accepted: 23 July 2014; published online: 13 August 2014.
Citation: Zelinski EM, Peters KD, Hindin S, Petway KT II and Kennison RF (2014) Evaluating the relationship between change in performance on training tasks and on untrained outcomes. *Front. Hum. Neurosci.* 8:617. doi: 10.3389/fnhum.2014.00617
This article was submitted to the journal *Frontiers in Human Neuroscience*.
Copyright © 2014 Zelinski, Peters, Hindin, Petway and Kennison. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Harnessing the neuroplastic potential of the human brain & the future of cognitive rehabilitation

Jyoti Mishra* and Adam Gazzaley*

Department of Neurology, Physiology and Psychiatry, University of California, San Francisco, San Francisco, CA, USA

*Correspondence: jyoti.mishra@ucsf.edu; adam.gazzaley@ucsf.edu

Edited by:

Guido P. H. Band, Leiden University, Netherlands

Reviewed by:

Geert Van Boxtel, Tilburg University, Netherlands

Keywords: cognitive training, neurotherapeutics, cognitive control, neuroplasticity, closed loop

Neuroplasticity is the remarkable ability of the brain that allows us to learn and adapt to our environment. Many studies have now shown that plasticity is retained throughout the lifespan from infancy to very old age (Merzenich et al., 1991; Merzenich and DeCharms, 1996; Greenwood and Parasuraman, 2010; May, 2011; Bavelier et al., 2012). Enriching life experiences, including literacy, prolonged engagement in the arts, sciences and music, meditation and aerobic physical activities have all been shown to engender positive neuroplasticity that boosts cognitive function and/or prevents cognitive loss (Vance et al., 2010; Hayes et al., 2013; Matta Mello Portugal et al., 2013; Newberg et al., 2013; Zatorre, 2013). Unfortunately, just as enriching experiences generate positive plasticity, negative plasticity ensues in impoverished settings. For instance, many studies now show that low socioeconomic, resource-poor environments, which are associated with stress, violence and abuse within families and communities, have detrimental effects on cognition and neural function (D'Angiulli et al., 2012; McEwen and Morrison, 2013). As cognitive neuroscientists we observe both positive and negative aspects of plasticity in neural systems, in functional changes of neural activations, neural oscillations and strength of connectivity between brain regions, in structural changes in gray matter volume and white matter integrity, and importantly in the relationship between such neuroplastic changes and concomitant cognitive/behavioral changes. As we come to understand various facets of plasticity, it drives further the quest to develop new activities/interventions that engender

maximal positive plasticity in selectively targeted neural systems; we envision such activities will in turn generate “far transfer of benefit” to generalized cognition and thereby improve the human condition.

In today's modern technological and internet-connected era, individuals are increasingly engaging with cognitive training software to improve cognitive function. In fact over the past 10–15 years, several companies have become established proponents and marketers of such software, transforming it into a multi-million dollar industry with exponential projected future growth. The fact that this technology is easily accessible over the internet to the home-setting, and at low-cost, has facilitated its mass adoption. Scientifically, however, not all “brain training” is made equal. All too often, basic cognitive neuroscience experimental paradigms are embedded in commercial “brain training” approaches with add-on visual graphic skins that attempt to maximize user-engagement; a process known as gamification. Although these experimental paradigms had been originally developed to understand cognition, that does not mean that they are also the best tools to engender positive neuroplasticity. It is no surprise then that some scientific investigations have uncovered that generic brain training approaches yield no positive cognitive outcomes (Owen et al., 2010). However, a blanket statement that all cognitive training is ineffective is also unfair. In recent years, development and evaluation of cognitive training approaches in many labs, including our own, has revealed evidence for positive neuroplasticity, as well as for

transfer of benefit to untrained cognitive abilities (Tallal et al., 1996; Temple et al., 2003; Stevens et al., 2008; Smith et al., 2009; Ball et al., 2010; Berry et al., 2010; Anderson et al., 2013; Anguera et al., 2013; Mishra et al., 2013; Wolinsky et al., 2013). Furthermore, in two of our training studies we find neurobehavioral correlations that relate on-task neuroplasticity to broader improvements in untrained aspects of cognition. Other researchers have also reported positive findings and transfer of training effects to untrained cognitive abilities in the context of custom-designed working memory exercises (Klingberg, 2010; Rutledge et al., 2012), task-switching training (Korbach and Kray, 2009), as well as for a specific genre of commercially available games, i.e., action video games (Bavelier et al., 2012) (although it is difficult to make strong recommendations about many off-the-shelf games given concerns over violent content). From these studies we are coming to understand some of the design principles that may govern the development of effective neuroplasticity-targeted training, as well as the scientific evaluation methods that can be used to provide convincing proof of the efficacy of the training intervention. Here, we summarize some of these principles that have emerged from two of our published training studies that now inform the development and evaluation of our next generation of training tools.

In our first training study in older adults, we simply trained visual perceptual discrimination of Gabor patches that had built-in directed motion animation (Berry et al., 2010). Ten hours

of training improved on-task perception relative to performance changes in a non-training (no-contact) control group. Interestingly, the training also benefitted delayed-recognition working memory of an untrained motion direction task. Not only was working memory performance improved, electroencephalography (EEG) neural recordings showed that training evoked more efficient sensory encoding of the stimuli, which correlated with the working memory performance gains. This finding that 10 h of simple perceptual training engendered transfer of benefits to working memory aligns with recent understanding that perceptual training improves signal to noise contrast, which then leads to refined encoding at multiple neural scales and hence, at least some degree of generalized cognitive benefits (Vinogradov et al., 2012).

We are now gaining an appreciation that the observed gains in our perceptual training study, and in similar studies performed by other labs, some of which have shown long-lasting cognitive benefits (Willis et al., 2006; Rebok et al., 2014), may be mediated by two fundamental design elements that drive neuroplasticity. 1) Training incorporated continuous performance feedback at multiple levels of game play providing repeated cycles of reward to the user 2) Training was adaptive to the trainee's in-the-moment game performance; i.e., adaptivity was incorporated using psychophysical staircase functions that enhance training challenge in response to accurate performance and reduce it for inaccurate performance. The up-down step ratio in such staircases is often chosen to maintain overall task challenge at 75–85%, at which point the user is optimally engaged but not frustrated. Thus, continuous performance feedback rewards and adaptive task challenge uniquely personalize the training to the cognitive capacity of each individual, and allows abilities to improve over time. Overall we have found these features to be critically important in generating positive neuroplasticity and cognitive benefit. Note, it is important to realize that casual game software is often not designed to provide the optimal dose of repetitive rewards nor incorporate adaptive progressions specifically targeted to the cognitive domains that may be deficient in a given

population cohort. These factors, along with the heterogeneity of tested populations, very small training doses on multiple cognitive exercises, and the use of assessment measures that are insensitive to detect training related benefits in the tested population, all may contribute to a failure to observe positive impact (e.g., Owen et al., 2010).

While reward cycles and adaptive progressions are key components of software design, it is equally important to tailor these game mechanics toward improving specific deficits observed in a population cohort. For instance, Anguera et al. (2013) showed that deficient cognitive control abilities, such as working memory and sustained attention, in healthy older adults can be enhanced by specifically training on a multitasking performance-adaptive and rewarding video game, “*Neuroracer*.” “*Neuroracer*” implements visual discrimination training in a go-no-go task for colored shape targets, with the added demand of simultaneously driving on a virtual road. “*Neuroracer*” evidenced extensive gains such that healthy older adults who multi-tasked 175% worse than younger adults on a first assessment, achieved significant post-training performance levels on the game itself that surpassed those of young adults by +44%. Importantly, training on “*Neuroracer*” transferred to untrained measures of sustained attention and working memory in the setting of interference, with EEG-based neural recordings showing that plasticity of midline frontal theta (mf theta) neural oscillations may be a mediator of these cognitive improvements.

While we have tested some aspects of sound game design, as described, other aspects of high-level video games may contribute to their success and we look ahead to assessing these empirically. For example, immersion, fun, real-world features, continuous performance, 3D environments, virtual reality, high-levels of art, story, and music facilitate sustained performance and better compliance, and also deeper engagement that we suspect maximally harnesses plasticity. Evaluation of the influence of these features on training effectiveness requires careful scientific study design. For this, the “*Neuroracer*” study adopted a rigorous three-armed randomized controlled design. In addition to the multitasking

training group, the study included an active single-task training control, as well as a no-contact control group. The single-task training control performed the exact same tasks as the training group of visual discrimination and driving, except that task engagement was not concurrent. This active control directly tested our hypothesis that only training in a setting that stresses cognitive control via a high interference environment would show significant cognitive gains. Outcomes of the “*Neuroracer*” multitasking training were not achieved in the active control group or in the no-contact group, the latter being critical for assessing practice effects due to repeated evaluations. Thus, the “*Neuroracer*” study highlighted that rigorous scientific evaluation of a cognitive training approach requires appropriate control groups, and often more than one control group, especially if we want to understand the underlying mechanisms of training effectiveness. Indeed longitudinal data collection is arduous, but without randomized, controlled and single/double blinded enrollments, we cannot convince ourselves of the significance of the results of new interventions. This is especially appropriate for healthy populations, while single-arm feasibility trials do remain informative as a first pass in cognitively impaired populations. In addition, we should also implement expectation bias measures for all participants, which confirm that all study groups anticipate the same level of influence of their assigned intervention on the outcome measures, thus assuring appropriate placebo control (Boot et al., 2013). Finally, adequately powered large sample size studies and investigations that measure sustainability of the cognitive gains and underlying neuroplasticity in yearly follow-ups are rare and need to be performed more often to address the long-term efficacy of cognitive interventions. Such rigor is convention in pharmaceutical clinical trials, and its adoption for video game testing, along with safety evaluations that detect potential side effects such as game-addiction, would promote a path toward FDA approvals and medical prescription of such technologies.

Equipped with our growing understanding of how to design cognitive training approaches to target plasticity

in specific neural circuits, we are now embarking on the development of the next generation training technologies. We envision these advances to include combining behavior-digital closed loops that link behavioral performance metrics to adaptive modulations of a training task on a digital platform, with neuro-digital closed loops that link neural performance measures to adaptive game mechanics. For example, the “*Neuroracer*” training study discovered that neuroplasticity of midline frontal theta (enhanced mf theta post-training) is a key neural factor that correlates with transferred cognitive gains. In order to test whether mf theta plasticity is truly causal in enabling improved cognition, we are now developing neuro-digital closed loops that directly target mf theta activity. More specifically, technological development is being directed at real-time EEG-based recordings that occur simultaneous with the cognitive task training (Delorme et al., 2011; Makeig et al., 2012; Kothe and Makeig, 2013). The goal is for these measurements to be event-locked to task stimuli, account for ocular and muscle-related artifacts, and use source localization algorithms (Mullen et al., 2013) so that they can be directly integrated in the game environment to guide reward feedback to the user and adaptivity of task challenge in real-time. We hypothesize that using neural performance as the driver for task-adaptivity will generate more rapid, efficient and specific circuit plasticity than is currently obtained using behavior-adaptive cognitive training approaches. This hypothesis is borne out of data, which shows that single-trial behavioral performance is predicted by neural measures such as mf-theta oscillations preceding the behavior. Thus the neuro-digital closed loop offers the potential to selectively train and refine the bottleneck neural processes that govern the final behavioral outcome. Importantly, by directly embedding task-related neural activity in a closed loop, this approach can provide missing causal evidence between neuroplasticity and cognitive benefits. This line of investigation is especially promising in the light of accumulating scientific evidence of the value of conventional neurofeedback approaches (Gruzelier, 2013; Wang and Hsieh, 2013; Arns et al., 2014), which also creates a

neuro-digital closed loop, albeit driven by ongoing scalp EEG oscillations as opposed to task-related neural processes as we envision.

We are aware that unlike traditional cognitive training, a neuro-digital closed loop approach is not feasible as a mainstay in the home setting at present. Yet, with rapid developments of mobile EEG technology (Stopczynski et al., 2013), as well as advances in the real-time computational power available on consumer devices such as laptops and tablets, we expect that deployment in the home environment will be a reality within a few years. Neuro-digital closed loops are also an exciting way to achieve personalized therapeutics, as each feedback loop is customized to the individual user’s neural capacities in the moment. While here we have provided a simplistic example of a closed loop tied to task-related mf theta activation, one can conceive of more sophisticated neural targets, including frontal-posterior effective connectivity based on task interaction dynamics. Further advances in this field are expected as neuroscientists collaborate with neural engineers, who have predominantly focused related efforts on neuroprosthetic development (Borton et al., 2013). Neuro-engineers have designed efficient closed loop decoding algorithms for brain-machine interfaces in animal model systems, and these techniques are now ripe for adoption in humans (Carmena, 2013). Finally, especially beneficial for clinical populations that exhibit weakened neural responsivity, another intriguing step will be the integration of neuro-digital closed loop systems with transcranial electrical current stimulation or even deep brain stimulation technologies (Coffman et al., 2013), which may provide a needed plasticity boost to impaired brain regions.

To achieve the goals of our field and fully harness the potential of neuroplasticity for cognitive benefit, we look forward to continued technological development, such as neuro-digital closed loops, and their integration with emerging design principles of cognitive training games. These technologies validated using randomized, controlled scientific evaluation methodologies will generate new understanding of how to translate cognitive neuroscience discoveries into new educational tools for healthy populations and mental

healthcare interventions for neuropsychiatric populations in need of cognitive remediation.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Health grants 5R01AG040333 (Adam Gazzaley), 5R24TW007988-05 subaward VUMC38412 (Jyoti Mishra). We would like to thank Joaquin Anguera for his feedback on this opinion article.

REFERENCES

- Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4357–4362. doi: 10.1073/pnas.1213555110
- Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., et al. (2013). Video game training enhances cognitive control in older adults. *Nature* 501, 97–101. doi: 10.1038/nature12486
- Arns, M., Heinrich, H., and Strehl, U. (2014). Evaluation of neurofeedback in ADHD: The long and winding road. *Biol. Psychol.* 95, 108–115. doi: 10.1016/j.neubiorev.2013.09.015
- Ball, K., Edwards, J. D., Ross, L. A., and McGwin, G. Jr. (2010). Cognitive training decreases motor vehicle collision involvement of older drivers. *J Am Geriatr Soc.* 58, 2107–2113. doi: 10.1111/j.1532-5415.2010.03138.x
- Bavelier, D., Green, C. S., Pouget, A., and Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annu. Rev. Neurosci.* 35, 391–416. doi: 10.1146/annurev-neuro-060909-152832
- Berry, A. S., Zanto, T. P., Clapp, W. C., Hardy, J. L., Delahunt, P. B., Mahncke, H. W., et al. (2010). The influence of perceptual training on working memory in older adults. *PLoS ONE* 5:e11537. doi: 10.1371/journal.pone.0011537
- Boot, W. R., Simons, D. J., Stothart, C., and Stutts, C. (2013). The pervasive problem with placebos in psychology: why active control groups are not sufficient to rule out placebo effects. *Perspect. Psychol. Sci.* 8, 445–454. doi: 10.1177/1745691613491271
- Borton, D., Micera, S., Millán, J. D. R., and Courtine, G. (2013). Personalized neuroprosthetics. *Sci. Transl. Med.* 5, 210rv2. doi: 10.1126/scitranslmed.3005968
- Carmena, J. M. (2013). Advances in neuroprosthetic learning and control. *PLoS Biol.* 11, e1001561. doi: 10.1371/journal.pbio.1001561
- Coffman, B. A., Clark, V. P., and Parasuraman, R. (2013). Battery powered thought: Enhancement of attention, learning, and memory in healthy adults using transcranial direct current stimulation. *Neuroimage* 85(Pt. 3), 895–908. doi: 10.1016/j.neuroimage.2013.07.083
- D’Angiulli, A., Lipina, S. J., and Olesinska, A. (2012). Explicit and implicit issues in the developmental cognitive neuroscience of social inequality. *Front. Hum. Neurosci.* 6:254. doi: 10.3389/fnhum.2012.00254
- Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., Vankov, A., et al. (2011).

- EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. *Comput. Intell. Neurosci.* 2011, 130714. doi: 10.1155/2011/130714
- Greenwood, P. M., and Parasuraman, R. (2010). Neuronal and cognitive plasticity: a neurocognitive framework for ameliorating cognitive aging. *Front. Aging Neurosci.* 2:150. doi: 10.3389/fnagi.2010.00150
- Gruzelier, J. H. (2013). EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants. *Neurosci. Biobehav. Rev.* doi: 10.1016/j.neubiorev.2013.09.015. [Epub ahead of print].
- Hayes, S. M., Hayes, J. P., Cadden, M., and Verfaellie, M. (2013). A review of cardiorespiratory fitness-related neuroplasticity in the aging brain. *Front. Aging Neurosci.* 5:31. doi: 10.3389/fnagi.2013.00031
- Karbach, J., and Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Dev. Sci.* 12, 978–990. doi: 10.1111/j.1467-7687.2009.00846.x
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends Cogn. Sci.* 14, 317–324. doi: 10.1016/j.tics.2010.05.002
- Kothe, C. A., and Makeig, S. (2013). BCILAB: a platform for brain-computer interface development. *J. Neural Eng.* 10, 056014. doi: 10.1088/1741-2560/10/5/056014
- Makeig, S., Kothe, C., Mullen, T., Bigdely-Shamlo, N., and Kreutz-Delgado, K. (2012). Evolving signal processing for brain-computer interfaces. *Proc. IEEE* 100, 1567–1584. doi: 10.1109/JPROC.2012.2185009
- Matta Mello Portugal, E., Cevada, T., Sobral Monteiro-Junior, R., Teixeira Guimarães, T., da Cruz Rubini, E., Lattari, E., et al. (2013). Neuroscience of exercise: from neurobiology mechanisms to mental health. *Neuropsychobiology* 68, 1–14. doi: 10.1159/000350946
- May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends Cogn. Sci.* 15, 475–482. doi: 10.1016/j.tics.2011.08.002
- McEwen, B. S., and Morrison, J. H. (2013). The brain on stress: vulnerability and plasticity of the prefrontal cortex over the life course. *Neuron* 79, 16–29. doi: 10.1016/j.neuron.2013.06.028
- Merzenich, M. M., and DeCharms, R. (1996). “Neural representations, experience, and change,” in *The Mind-Brain Continuum*, eds. R. Llinas and P. Churchland (Boston, MA: The MIT Press), 61–81.
- Merzenich, M. M., Recanzone, G. H., and Jenkins, W. (1991). “How the brain functionally rewires itself,” in *Natural and Artificial Parallel Computations*, eds. M. Arbib and J. A. Robinson (New York, NY: MIT Press).
- Mishra, J., Anguera, J. A., Ziegler, D. A., and Gazzaley, A. (2013). A cognitive framework for understanding and improving interference resolution in the brain. *Prog. Brain Res.* 207, 351–377. doi: 10.1016/B978-0-444-63327-9.00013-8
- Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2013). Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. *Conf. Proc. IEEE Eng. Med. Biol. Soc. Conf.* 2013, 2184–2187. doi: 10.1109/EMBC.2013.6609968
- Newberg, A. B., Serruya, M., Wintering, N., Moss, A. S., Reibel, D., and Monti, D. A. (2013). Meditation and neurodegenerative diseases. *Ann. N.Y. Acad. Sci.* 1307, 112–123. doi: 10.1111/nyas.12187
- Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., et al. (2010). Putting brain training to the test. *Nature* 465, 775–778. doi: 10.1038/nature09042
- Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., et al. (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. *J. Am. Geriatr. Soc.* 62, 16–24. doi: 10.1111/jgs.12607
- Rutledge, K. J., van den Bos, W., McClure, S. M., and Schweitzer, J. B. (2012). Training cognition in ADHD: current findings, borrowed concepts, and future directions. *Neurotherapeutics* 9, 542–558. doi: 10.1007/s13311-012-0134-9
- Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., et al. (2009). A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *J. Am. Geriatr. Soc.* 57, 594–603. doi: 10.1111/j.1532-5415.2008.02167
- Stevens, C., Fanning, J., Coch, D., Sanders, L., and Neville, H. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: electrophysiological evidence from language-impaired and typically developing children. *Brain Res.* 1205, 55–69. doi: 10.1016/j.brainres.2007.10.108
- Stopczynski, A., Stahlhut, C., Petersen, M. K., Larsen, J. E., Jensen, C. F., Ivanova, M. G., et al. (2013). Smartphones as pocketable labs: Visions for mobile brain imaging and neurofeedback. *Int. J. Psychophysiol.* 91, 54–66. doi: 10.1016/j.ijpsycho.2013.08.007
- Tallal, P., Miller, S. L., Bedi, G., Byrna, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271, 81–84. doi: 10.1126/science.271.5245.81
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2860–2865. doi: 10.1073/pnas.0030098100
- Vance, D. E., Roberson, A. J., McGuinness, T. M., and Fazeli, P. L. (2010). How neuroplasticity and cognitive reserve protect cognitive functioning. *J. Psychosoc. Nurs. Ment. Health Serv.* 48, 23–30. doi: 10.3928/02793695-20100302-01
- Vinogradov, S., Fisher, M., and de Villiers-Sidani, E. (2012). Cognitive training for impaired neural systems in neuropsychiatric illness. *Neuropsychopharmacology* 37, 43–76. doi: 10.1038/npp.2011.251
- Wang, J.-R., and Hsieh, S. (2013). Neurofeedback training improves attention and working memory performance. *Clin. Neurophysiol.* 124, 2406–2420. doi: 10.1016/j.clinph.2013.05.020
- Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., et al. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *J. Am. Med. Assoc.* 296, 2805–2814. doi: 10.1001/jama.296.23.2805
- Wolinsky, F. D., Vander Weg, M. W., Howren, M. B., Jones, M. P., and Dotson, M. M. (2013). A randomized controlled trial of cognitive training using a visual speed of processing intervention in middle aged and older adults. *PLoS ONE* 8:e61624. doi: 10.1371/journal.pone.0061624
- Zatorre, R. J. (2013). Predispositions and plasticity in music and speech learning: neural correlates and implications. *Science* 342, 585–589. doi: 10.1126/science.1238414

Conflict of Interest Statement: Jyoti Mishra is a part-time scientist at the Brain Plasticity Institute, PositScience, a company that develops cognitive training software. Adam Gazzaley is co-founder and chief science advisor of Akili Interactive Labs, a company that develops cognitive training software. Jyoti Mishra and Adam Gazzaley have a patent pending for “Methods of Suppressing Irrelevant Stimuli.” Adam Gazzaley has a patent pending for a game-based cognitive training intervention: “Enhancing cognition in the presence of distraction and/or interruption.”

Received: 30 December 2013; accepted: 27 March 2014; published online: 11 April 2014.

Citation: Mishra J and Gazzaley A (2014) Harnessing the neuroplastic potential of the human brain & the future of cognitive rehabilitation. *Front. Hum. Neurosci.* 8:218. doi: 10.3389/fnhum.2014.00218

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Mishra and Gazzaley. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Spaced cognitive training promotes training transfer

Zuowei Wang^{1*}, Renlai Zhou^{2,3} and Priti Shah¹

¹ Department of Psychology, Combined Program in Education and Psychology, University of Michigan, Ann Arbor, MI, USA

² Beijing Key Laboratory of Experimental Psychology, School of Psychology, Beijing Normal University, Beijing, China

³ State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

Edited by:

Guido P. H. Band, Leiden University, Netherlands

Reviewed by:

Dietsje Jolles, Stanford University, USA

Pier Prins, University of Amsterdam, Netherlands

*Correspondence:

Zuowei Wang, Combined Program in Education and Psychology, University of Michigan, 610 E. University Ave., Room 1400-J, Ann Arbor, MI 48109, USA
e-mail: zwwang@umich.edu

Cognitive training studies yield wildly inconsistent results. One dimension on which studies vary is the scheduling of training sessions (Morrison and Chein, 2011). In this study, we systematically address whether or not spacing of practice influences training and transfer. We randomly assigned 115 fifth grade children to an active control group or one of four training groups who received working memory training based on a “running span” task (Zhao et al., 2011). All groups received the same total amount of training: 20 sessions of training with 60 trials for an average of 20 min per session. The training was spread across 2, 5, 10, or 20 days. The active control group received 20-min sessions of math instruction for 20 sessions. Before and after training participants in all five groups performed a single transfer test that assessed fluid intelligence, the Raven’s Progressive Matrices Test. Overall, participants in all four training groups improved significantly on the training task (at least partially), as reflected by increased speed. More importantly, the only training group to show significant improvement on the Raven’s was the group who had the greatest amount of spacing (20 days group) during training and improvement in this group was significantly higher than that of the control group.

Keywords: cognitive training, working memory, transfer, schedule, spacing, primary school, children, fluid intelligence

INTRODUCTION

Working memory is the cognitive system that actively maintains and processes information for human problem solving (Miyake and Shah, 1999). Working memory training has been extensively studied in recent years (e.g., Klingberg, 2010; Morrison and Chein, 2011). Studies have shown that working memory training is beneficial to various subject populations, varying from young children to old adults, including healthy subjects as well as those with special needs. For example, working memory training has been found to reduce the symptoms of ADHD children (Klingberg et al., 2002, 2005), facilitate the recovery of cognitive functions in patients after stroke (Westerberg et al., 2007), enhance “old-old” adults’ memory performance (80 years old; Buschkuhl et al., 2008), and improve fluid intelligence in pupils (Jaeggi et al., 2011; Zhao et al., 2011) as well as college students (Jaeggi et al., 2008, 2014).

The significance of working memory training is largely dependent on the potential transfer effects to other untrained situations. Due to the various transfer effects identified by previous studies, some researchers view working memory training as promising for general cognition enhancement (see the review by Morrison and Chein, 2011). However, while researchers almost always find performance improvement on trained tasks, not all studies find transfer. In fact, some researchers argue that there is little to no solid evidence that true far transfer effects arise from working memory training (Shipstead et al., 2012; Melby-Lervåg and Hulme, 2013; Rapport et al., 2013; Redick et al., 2013).

Efforts have been taken to investigate factors that affect successful training and transfer. Three broad classes of factors are

likely relevant: individual characteristics of participants receiving training, the nature of the training task, and conditions of training. Individual characteristics that affect training and transfer may include initial ability of participants, the underlying source of any deficits in working memory performance, and motivational factors. Research in our laboratory has found, for example, that individuals who believe that intelligence is a malleable construct are more likely to benefit from training than those who believe intelligence is fixed (Jaeggi et al., 2014). Individuals who report that a training task is “too difficult” seem to disengage and show less improvement than individuals who enjoy the challenge of a task (Jaeggi et al., 2011). Other individual motivational factors may include the degree to which individuals are intrinsically motivated (Katz et al., in press; Shah et al., 2012).

The nature of the training task(s), of course, will also influence what types of transfer might be found. In a classic review of training and transfer, Schmidt and Bjork (1992) emphasize the importance of using a variety of tasks to increase the likelihood that training effects are not task-specific. Likewise, tasks that are adaptive and cognitively challenging may also limit the extent to which task-specific strategies that can be developed. In the original Jaeggi et al. (2008) study, the use of the dual n-back task that required maintaining and updating information in both visual and auditory modalities may have limited the ability of individuals to develop specific rehearsal strategies. More generally, Morrison and Chein (2011) found in their review of working memory training that studies that focus on training strategies were not as effective as those that focused on core working memory capacity. Some characteristics of the training task

may also interact with motivation or perceived difficulty of a task. We found, for example, that in a brief, 3-day cognitive training intervention that some game-like features may in fact serve as a distraction making cognitive training more difficult (Katz et al., *in press*, this issue).

A third class of factors that affect the effectiveness of cognitive training is the dosage or sheer amount of training (Jaeggi et al., 2008) as well as the training schedule (Schmidt and Bjork, 1992). It has been known for over 100 years that individuals remember information better under spaced practice conditions compared to massed practice (Ebbinghaus, 1885). This spacing effect has been extensively studied under numerous contexts since then. Most of these studies showed that spaced learning led to better learning outcomes (for a review, see Cepeda et al., 2006). For example, spaced learning is beneficial to foreign word learning (Bahrick et al., 1993), inductive learning (Kornell and Bjork, 2008; Vlach et al., 2008), and professional training (Hagman, 1980). The spacing effect is also found in cognitive skill acquisition tasks for both human and animal subjects (Shebilske et al., 1994, 1999; Sisti et al., 2007). For example, Sisti et al. (2007) trained rats using different schedules on a water maze task and they found that the rats that received more spaced training outperformed rats who received massed training. More interestingly, performance was correlated with the survival of new neurons in the hippocampus, suggesting that spaced training elicited more neural changes.

Shebilske et al. (1999) investigated the spacing effect in a complex skill acquisition task that is perhaps closest to the types of cognitive training tasks of interest in the current paper. The training task they used was Space Fortress, a task that required “short- and long-term memory loading, high workload demands, dynamic attention allocation, decision making, prioritization, resource management, continuous motor control, and discrete motor responses.” In their study, college students received training on this task for a total of 10 h with two different schedules, either within 10 days or within 2 days. Results showed that the more spaced training group had an advantage in the acquisition, retention and training transfer to a different device (from joystick to keyboard). In addition, the more spaced training group also showed less interference when asked to perform the Space Fortress task together with a secondary tapping task. The spacing effect in skill acquisition is also studied in the context of online gaming. Stafford and Dewar (2014) analyzed gameplay data of 854,064 players. They found that players who spread their practice tended to achieve higher score in the game.

There are many theoretical explanations for the spacing effect, most of which are not mutually exclusive. Spaced learning is consistent with rational models of memory that assume memory is adaptive (Anderson and Schooler, 1991). Exposure to information in a spaced manner is a clue to the memory system that the information may be needed again at a future date. By contrast, massed practice may support storing information for the short term, as the information is not again needed after a short period of practice. More specific theoretical explanations for spacing effects in memory have also been proposed, including the “deficient-processing hypothesis” (Greene, 1989), the “context variability” hypothesis (Greene, 1989), the incubation effect (for a review, see Sio and Ormerod, 2009), sleep-dependent

memory consolidation (Stickgold, 2005) and so on. These theoretical explanations are likely to hold not only in the context of general skill acquisition and memory, but also specifically in the context of cognitive training. The deficient-processing theory, for example, posits that when too much information is presented to participants in memory tasks, the information is processed with lower efficiency. The same rationale could also be applied to working memory training. Thus, massed training may not elicit the neural changes that are necessary for the training transfer.

In summary, the spacing effect in memory may shed light on the understanding of similar effects in cognitive training. However, cognitive training also has some unique features. The increase in the capacity and speed of cognitive processing cannot be treated similarly as the acquisition of new knowledge. For example, the spacing effect in cognitive training may also show different patterns than that in memory: the spacing effect in memory tasks may be a result of more covert rehearsal, whereas in skill acquisition (such as motor behavior), it is likely to be related to “effort, work, reactive inhibition, or fatigue” (Adams, 1987). In memory tasks the optimal spacing “gap” is greater when the delay between practice and final test is longer (Cepeda et al., 2008). However, the optimal spacing in skill acquisition situations still remains unknown (Stafford and Dewar, 2014).

Currently, we are not aware that any working memory training studies have systematically varied the schedule under which individuals are trained to investigate the effect of training on outcome. A potential spacing effect in working memory training has both theoretical significance and important practical implications. Theoretically, a systematic investigation of the spacing effect in working memory training may help clarify the mixed findings in the current working memory training literature. Studies have revealed different effect size in training gains and training transfers, which could be a result of uncontrolled training schedule. In practice, an optimized training schedule may produce stronger and broader training gains in a shorter time, which cuts the training cost and allows more people to benefit from it.

In the current study, we investigated the effect of different training schedules on the outcome of working memory training in 5th grade classes in Muling, China. We used the same intervention that was originally used in Zhao et al. (2011). In that study, the intervention was a running span task in which 4th grade children from China were presented with a sequence of either animal drawings or locations on a 3-by-3 grid. The task required them to recall the three most recent stimuli in the presented order when the sequence stopped. Using a pre-test, training, post-test paradigm, they found that 20 sessions of training significantly improved children’s performance on the Raven’s Standardized Progressive Matrices (SPM). In this study we used the same running span training task used by Zhao et al. (2011) for several reasons: (1) this updating task has already been shown to improve fluid reasoning in one study, (2) the task had already been used with children in China and was designed to be appealing and engaging to children of the same range with a similar cultural background. Our study differed from the Zhao et al. (2011) study in that we included an active control group that was educational in content (extra math instruction with their

teachers). All participants in the training groups received the identical amount of practice on the training task (20 total sessions, an amount of training that has been used in other training studies that found transfer of training such as Zhao et al., 2011). However, the groups differed in the spacing of the training sessions. One group of participants received all 20 sessions within 2 days (10 sessions per day), the second group received all 20 sessions within 5 days (four sessions per day), the third group received the training within 10 days (two sessions per day), and the fourth group, with the greatest spacing, received one session of training per day for 20 days.

Based on the body of research on spacing, memory, and skill acquisition, we predicted that training schedule would have a substantial impact on working memory training gain and transfer. Specifically, we predicted that the group(s) with the most spacing of training would improve most on the training task and furthermore show the most transfer. In addition to this primary goal, we wished to replicate the results of other studies that have trained memory updating and found transfer to fluid intelligence in children (Jaeggi et al., 2011; Zhao et al., 2011). The total number of studies in which updating is trained in typically developing children is rather small, and thus this study provides an additional data point with respect to the potential effects of updating training more generally.

MATERIALS AND METHODS

PARTICIPANTS

A total of 115 5th grade students (10–11 years old) from Muling Shiyuan Elementary School (Muling, China) were recruited to participate in the study. Before the training, they were told that upon finishing the training they would receive different gifts based on their performance in the training, including school bags, fountain pens and lockable notebooks. Twenty subjects were unable to strictly follow their assigned training schedule, or were absent during the pre-test or post-test thus dropped out from the study, resulting of 95 valid subjects in the data analysis (52 female). There was no group difference on the dropout rate, $\chi^2(4, n = 115) = 3.31, p = 0.51$. Due to computer error, two subjects' Grid task training data were lost, and one subject's Animal task training data were lost.

DESIGN

Participants were randomly assigned into one of the four training groups or an active control group. All the four training groups received the same total amount of training: 20 sessions of training with 60 trials for an average of 20 min per session. The training was spread across 2, 5, 10, or 20 days. The control group stayed with their teachers in their classrooms (after school) for 20 min each day for 20 days and received instruction focused primarily on math. The gender distribution in the five groups was: 20 Days—9 female 11 male; 10 Days—11 female 9 male; 5 Days—12 female 8 male; 2 Days—10 female 5 male; control group—10 female 10 male.

Before and after training, participants were all tested on a measure of fluid intelligence, the Raven's Standard Progressive matrices test. We compared pre-test to post-test improvements in the five groups (four training and one control) to assess transfer.

MATERIALS AND PROCEDURE

Training task

We used two forms of the “running span” task for the training (Zhao et al., 2011). In one task (Animal), subjects saw pictures of different animals presented sequentially on the computer screen. In the other task (Grid), subjects saw a cartoon figure moving sequentially to different locations in a 3×3 grid. In both of the tasks, the length of the sequence randomly varied from 5, 7, 9, or 11 items in a row and participants did not know ahead of time how many items would be presented prior to recall instructions. At the end of each sequence of trials, subjects were asked to recall the three most recent stimuli in the presented order. Because a subject could not predict when the sequence would end, they were required to continuously update their working memory with the most recent three items.

Each session consisted of participants performing one set of animal and one set of grid trials. Each set consisted of 30 trial sequences that were divided into six blocks of five trials each. Within each block, if subjects provided the correct response for three or more trials, the presentation time of each stimulus in the next block would decrease by 100 ms, thus making the task more difficult. If participants got fewer than three trials correct, the presentation time of each stimulus would increase by 100 ms, which made the next block easier. For both the Animal and the Grid tasks, the starting presentation time was 850 ms for the 20, 5, and 2 days groups. The starting presentation time for the 10 days group was about 1000 ms for the Animal task and 950 ms for the Grid task (this was due to a computer error: the recovery mode of the training computers wasn't turned off and the 10 days group started from where the 20 days group left after their first session).

We decided that subjects' performance on the training tasks could be reflected by the presentation time of the stimulus. To track their training performance, we calculated the averaged presentation time for each session (both Animal and Grid), and named this measure as “presentation time” for that set of the session hereafter. To encourage children to try their best on the training tasks, correct response on each trial would earn them one point, which was shown by adding one smiley face to a feedback chart which was located at the bottom of the screen. The total points (smiley faces) could be used to trade for different gifts (school bags, fountain pens and lockable notebooks) after the training. More detailed descriptions of the training tasks can be found in Zhao et al. (2011).

Math instruction for control group

When the 20 days training group received the daily training, the active control group remained in their classroom and worked with their math teachers. They received extra math exercises from a 5th grade mathematics workbook for 20 min every day. Students first worked on problems from the workbook, and then the teacher checked their answers and provided further instruction when necessary. Some example problems the students practiced include: solving equations with one variable, calculating the area of different shapes that required them to divide the shapes into regular shapes with known formulas for area calculation, word problems (e.g., calculating the distance of moving objects,

sometimes requiring the use of equations) etc. No rewards were provided for the control group.

Transfer task

The Raven's Standard Progressive Matrices (SPM) was used to evaluate the transfer effect of the training, following the design of a number of working memory training studies (e.g., Jaeggi et al., 2008; Zhao et al., 2011). Specifically, the 60 items in SPM were split into two subtests: odd numbered items and even numbered items, which were used in pre-test and post-test with counter-balance. We also used the TONI (Test of Nonverbal Intelligence; Brown et al., 2010) for some participants, but due to scheduling difficulties we were not able to collect TONI test scores for the 20 session training or active control group. Therefore, only Raven's scores are used in the analyses.

PROCEDURE

All children in the training groups were given one half of the SPM (even or odd items) as a pre-test before they started the training. Whether they received odd or even items at pre-test was counterbalanced. Children in all the groups were given the pre-test within the same 3 days prior to the training on the 20 days group. Thus, the distance between pre-test and post-test for all groups was approximately the same.

During the training, each training session consisted of one set of Animal task and one set of Grid task. Children in the 20 days group received one session every day after school, which took about 15–25 min. Children in the 10 days group received one session during the 2-h-long noon recess (for 15–25 min) and another session after school (another 15–25 min). Children in the 5 days group received two sessions during the noon recess (i.e., 30–50 min) and another two sessions after school (an additional 30–50 min). Children in the 2 days group received the training after the semester, and they finished the 20 total sessions within 2 days (approximately 10 sessions per day for a total of 150–250 min each day with rest and lunch breaks). For the 5 Days and 2 Days group, children were given a 5–10 min rest after approximately every 30–40 min of training. In all the groups, the very first training session was used as a practice session in which children were allowed to stop and ask questions. Thus, training data for the first session was not recorded and not included in the analysis.

After the training, children were given the alternate version of the SPM as a post-test. For the 5 and 2 days group, the SPM was administered the day following training completion to prevent decreased performance due to training fatigue. We were mindful about keeping the time interval between the pre-test and post-test the same for all the five groups. However, due to scheduling difficulties this interval for the 5 and 2 Days group was about 1 week longer.

Children in all the groups strictly adhered to their training schedule. Before weekends or holidays, we made arrangements with all the parents to make sure their children would come to school for the training. Three children in the 20 Days group and 2 in the 10 Days group who lived too far to get to the school received the training at home with the experimenter overseeing their training using remote desktop.

Children in none of the groups were given any information about how the Raven's pre- and post-tests may be related to training/math learning. Before working on the Raven's tests, they were just told that they were to work on some puzzles. Children in all the four training groups were motivated to earn more points by correctly recalling animals/grid locations during the training for better reward; children in the control group were motivated to learn math because they were to receive a math test after the 20 days.

The study was reviewed and approved by the Institutional Review Board at the University of Michigan. Informed consent was obtained from all the parents whose children participated in the study. Before each training/testing session, oral assent was also obtained from all the children who participated.

RESULTS

TRAINING GAIN

The five groups had similar scores in the Raven's pre-test, $F_{(4, 90)} = 0.20$, $p = 0.938$. Training performance is summarized in **Figures 1, 2**. Improvement on the two running span tasks is reflected by subjects' increasing capability to process faster presentation of the stimuli (i.e., increased speed). To evaluate the training gain, we subtracted the averaged presentation time of the first three sessions from that of the last three sessions, which represented the processing speed increase as a result of the training (here negative values indicate improvement). Based on this measure, all the four training groups showed significant improvement on the Grid task; both the 20 Days group and the 10 Days group made significant progress on the Animal task (one-sample t -test compared to the value "0"; **Table 1**). However, a comparison among the four groups did not show significant difference in training gain among the groups in either the Animal task, $F_{(3, 70)} = 1.288$, $p = 0.285$, nor the Grid task, $F_{(3, 69)} = 1.077$, $p = 0.365$.

Training gain can also be measured by the regression slope of presentation time (defined in Training task) on session number as

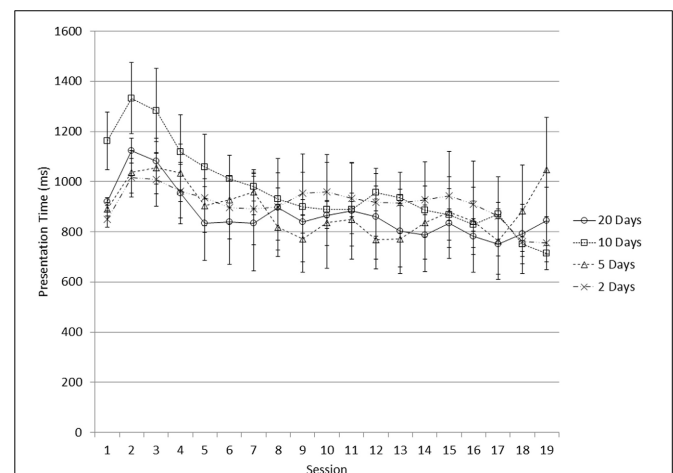


FIGURE 1 | Presentation time decreased with training on the Animal task; error bars represent standard error.

an indication of the session-by-session processing speed improvement. In both **Figures 1, 2**, the presentation time was the average presentation time of the six blocks within a session. From these figures, subjects' presentation time first showed a slight increase then a steady decrease, suggesting the starting presentation time was appropriate. **Table 2** shows a summary of these regression slopes for the four training groups on the two running span tasks. There was no significant difference in training gain (reflected by the regression slope) among the four training groups in either the Animal task, $F(3, 70) = 0.913$, $p = 0.439$, or the Grid task, $F(3, 69) = 0.534$, $p = 0.660$.

It should be noted that subjects' accuracy was also tracked. However, due to the nature of the task, subjects' accuracy only improved during the first few sessions and then remained stable. This is because if their accuracy in a given block (five trials), the presentation speed would become faster in the next block.

TRAINING TRANSFER

A Paired-sample t -test was performed for each of the four training groups together with the control group on SPM pre-test and post-test to evaluate the training transfer. Results show that only the 20

Days group showed significant improvement on the test (**Table 3**; for original score, see **Figure 3**). Because only the 20 Days training group showed evidence of improvement on the transfer task, we compared gain on the SPM (post-test minus pre-test scores) for the 20 Days group and the active control group. The SPM gain in the 20 Days group was significantly larger than the control group, $t_{(38)} = 1.832$, $p = 0.038$ (one-tail test), Cohen's $d = 0.59$.

Table 3 also shows that the effect size (Cohen's d) decreases as the students received a more massed training schedule. Thus, we conducted a regression analysis to see whether training schedule had an effect on the SPM post-test after controlling for pre-test scores. In the regression analysis, training schedule was entered as the number of days completing the training (i.e., 20, 10, 5, or 2), with the control group a value of "0." Results show that Training Schedule had an effect on the post-test score after controlling for pre-test score, $p = 0.052$ (2-tailed).

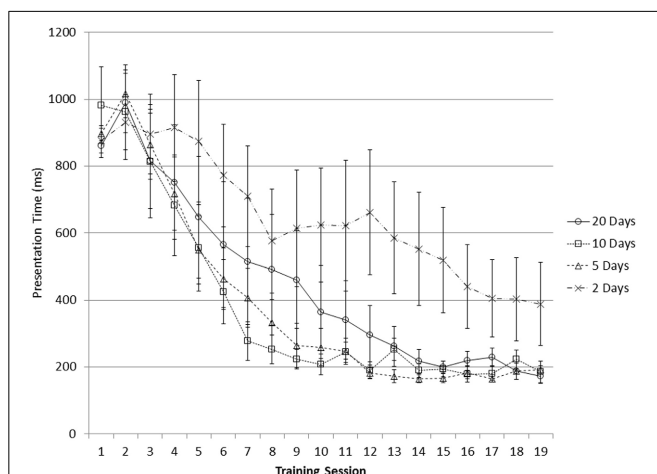


FIGURE 2 | Presentation time decreased with training on the Grid task; error bars represent standard error.

Table 2 | Regression slopes reflecting averaged session-wise stimulus presentation time decrease (in milliseconds; standard errors of the slopes provided in the parenthesis).

	Animal task	Grid task
20 Days Group	−74(32)	−261(47)
10 Days Group	−156(47)	−247(50)
5 Days Group	−43(75)	−265(32)
2 Days Group	−41 (43)	−187(54)

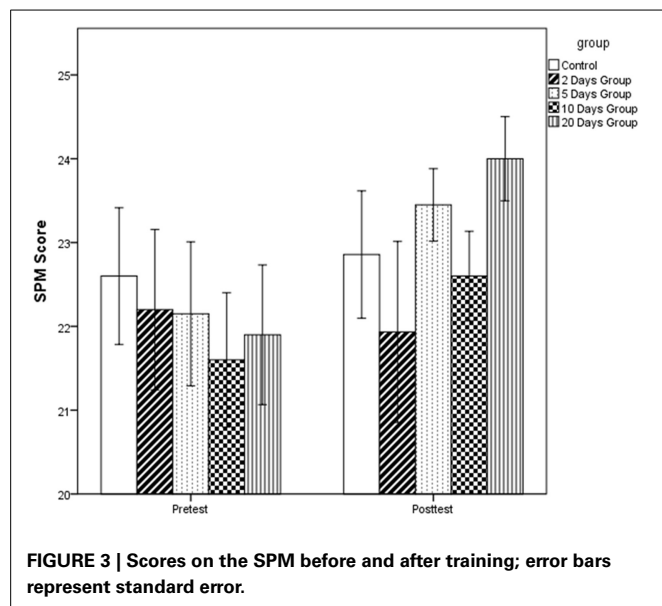
Table 3 | Improvement on SPM after the training as reflected by paired-sample t -test.

	t	df	p	Cohen's d
20 Days group	2.93	19	0.009	1.34
10 Days group	1.27	19	0.220	0.58
5 Days group	0.95	19	0.355	0.44
2 Days group	0.19	14	0.854	0.10
0 Days group(Control)	0.19	19	0.855	0.09

Table 1 | One-sample t -test showing training gain of the four training groups as measured by the presentation time decrease (the average of the last three sessions minus that of the first three sessions).

Animal task					Grid task			
	RT (ms) decrease	t	p	df	RT (ms) decrease	t	p	df
20 Days group	−246	−2.30	0.033	19	−691	−8.69	<0.001	19
10 Days group	−481	−3.57	0.002	18	−722	−5.12	<0.001	18
5 Days group	−124	−0.675	0.507	19	−743	−11.9	<0.001	19
2 Days group	−165	−1.54	0.147	14	−502	−4.62	<0.001	13

*Note: training data for one subject in the 10 Days group (both Animal and Grid) and one subject in the 2 Days group (Grid) were lost due to computer error; the starting stimulus presentation time for the 10 days group was longer than the other groups (~150 ms longer in the Animal task and ~100 ms longer in the Grid task) thus resulting in a relatively larger training gain in the Animal task, but not in the Grid task. See **Figures 1, 2** for details.



THE RELATIONSHIP BETWEEN TRAINING GAIN AND TRANSFER

To provide further support for the idea that practice on the running span task was directly related to improvement on the transfer measure of fluid intelligence, the SPM, we assessed the relationship between improvement on the training task and the magnitude of the transfer effect. Each subjects' training gain is measured by the averaged session-by-session presentation time decrease (i.e., processing speed increase), which is calculated as the regression slope of each subject's presentation time by the session number. The magnitude of the transfer effect is the score difference between SPM post-test and pre-test. **Table 5** summarizes the correlation coefficients of training gain and training transfer. It can be seen that only for the most spaced training schedule does the training gain significantly correlate with transfer, $r_{(20)} = 0.465$, $p < 0.05$.

DISCUSSION

This study assessed the effect of spacing of working memory training on the Raven's SPM. We predicted that spacing of training would affect both training gains as well as transfer, but found that there were no significant differences in training gain (as measured by reaction time) across the groups. Importantly, however, we did find a significant effect of training schedule on *transfer*. Only participants in the most spaced group, 20 sessions with one session per day, showed significant improvement on the Raven's SPM. Furthermore, improvement on SPM test performance for this group was significantly different from that of the active control group. More generally, there was a significant relationship between training schedule and transfer such that the more spaced the training, the greater the transfer (see **Table 4**).

We can draw two main conclusions from this study. First, training schedule has a significant impact on transfer of training. Second, the transfer effect of the 20-day group replicated results of a recent study that used the identical training and transfer tasks (Zhao et al., 2011). It is also consistent with studies that have

Table 4 | Regression analysis showing the effect of training schedule on SPM post-test.

	<i>B</i>	<i>SE B</i>	β	<i>p</i>
Model 1				
Pre-test	0.279	0.078	0.349	<0.001
Model 2				
Pre-test	0.289	0.077	0.361	<0.001
Training schedule	0.076	0.038	0.189	=0.052

Table 5 | Correlation between training gain (RT decrease) and magnitude of training transfer.

Score change in SPM	Averaged training gain
20 Days group ($n = 20$)	-0.465*
10 Days group ($n = 19^{**}$)	+0.046
5 Days group ($n = 20$)	-0.015
2 Days group ($n = 14^{**}$)	-0.135

* $p < 0.05$;

training gain data were lost from one subject in the 10 Days group and one subject from in 2 Days group so the n 's here do not match the df 's in **Table 3.

found improvements in typically developing children following working memory training (Thorell et al., 2009; Jaeggi et al., 2011; Zhao et al., 2011).

One question that arises is why this study and some others find far transfer effects whereas others do not. As discussed above, our study had several features associated with successful training studies: we tested children and not adults (Morrison and Chein, 2011), we used an adaptive training task that increased as performance increased (Jaeggi et al., 2008), we carefully monitored children's training schedules, children were given token extrinsic rewards not only for participation but also performance, multiple training tasks were used to ensure task variability and reduce the likelihood that children develop task-specific strategies (Schmidt and Bjork, 1992), and the working memory updating task taxed both active maintenance and interference resolution processes (Jaeggi et al., 2011). Further, the students who participated in this study appeared to be highly motivated; while the drop-out rate of many cognitive training studies is relatively high, 95 out of the initially recruited 115 completed the entire study.

Interestingly, our study found that training task improvements, at least in the group that received training across 20 days, was correlated with improvement in the transfer task. This result is consistent with the Jaeggi et al. (2011) study that found transfer only for participants who made significant improvements on the training task. Moreover, our study extends the picture. Training gain is not solely predictive of transfer outcome—participants who received the massed training schedules also improved on the training task the same extent as those who were in the most spaced group, but they did not show significant improvement on the transfer task. Thus, the relationship between training and transfer was significant only for the most spaced group. It is possible that individuals in the massed groups learned more task-specific strategies whereas those in the spaced group were able to focus

on the underlying cognitive skills. This explanation supports the deficient-processing hypothesis (Greene, 1989), which assumes that when too much information presented in a short period of time, it is processed with lower efficiency. In our working memory training context, this means that training that happens within a massed schedule does not induce deep processing that may contribute to training transfer. However, it is not entirely clear what exactly explains the dissociation between improvement and transfer for the massed groups. Future research should address the degree to which spacing might affect strategy use.

As the first study that we are aware of that explores the spacing effect in working memory training, the current study has some limitations that need to be addressed by future research. First, the four different training schedules we used did not fully represent the schedule variation of the current training studies. According to Morrison and Chein (2011), the schedule of current working memory training studies varied from 2 to 14 weeks. In the current study we only tested the lower end of this continuum. It is yet to be explored the magnitude of the transfer when the training is spaced over a longer period of time. Second, due to the small sample size, we used relatively less conservative statistical tests, including running planned contrast without adjustment of α level (we directly compared the 20 Days group with the control group using a one-tailed test with $\alpha = 0.05$). Based on the effect size of this study, we suggest future studies use a sample size twice as large as ours (i.e., 40 subjects per training schedule). Third, because we only used one training task and one transfer task, the reliability of this spacing effect in working memory training is limited to these specific tasks. In addition, the use of a single far transfer task did not allow us to evaluate the mechanism of the transfer effect. Thus, replicating these findings with a larger battery of transfer tasks including both near and far measures is an important next step.

In conclusion, this study demonstrated that training schedule has substantial impact on transfer of training. More research that investigates the moderators of training may help shed light on the debate of whether working memory training leads to broader cognitive improvement.

ACKNOWLEDGMENTS

The study was supported by the Rackham Graduate Student Research Grant awarded to Zuowei Wang. The authors thank Yanmei Ma, Yong Feng (school principals), Xiaoxia Xie, Changyan Liu, Xueqiang Liu, and Hongli Li and all other teachers at Muling Shiyan Elementary School for helping make arrangement with the data collection. The authors also thank all the students and parents for participating in the study.

REFERENCES

- Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychol. Bull.* 101, 41–74. doi: 10.1037/0033-2909.101.1.41
- Anderson, J. R., and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychol. Sci.* 2, 396–408. doi: 10.1111/j.1467-9280.1991.tb00174.x
- Bahrack, H. P., Bahrack, L. E., Bahrack, A. S., and Bahrack, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychol. Sci.* 4, 316–321. doi: 10.1111/j.1467-9280.1993.tb00571.x
- Brown, L., Sherbenou, R. J., and Johnsen, S. K. (2010). *TONI-4, Test of Nonverbal Intelligence*. Austin, TX: Pro-Ed.
- Buschkuhl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Däpp, C., Müller, M., et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychol. Aging* 23:743. doi: 10.1037/a0014342
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi: 10.1037/0033-2909.132.3.354
- Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., and Pashler, H. (2008). Spacing effects in learning a temporal ridgeline of optimal retention. *Psychol. Sci.* 19, 1095–1102. doi: 10.1111/j.1467-9280.2008.02209.x
- Ebbinghaus, H. (1885). *Über das Gedächtnis: Untersuchungen zur Experimentellen Psychologie*. Berlin: Duncker & Humblot.
- Greene, R. L. (1989). Spacing effects in memory: evidence for a two-process account. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 371–377. doi: 10.1037/0278-7393.15.3.371
- Hagman, J. D. (1980). *Effects of Presentation-and Test-Trial Training on Acquisition and Retention of Movement End-Location*. Fort Belvoir, VA: DTIC Document.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., and Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6829–6833. doi: 10.1073/pnas.0801268105
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., and Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108
- Jaeggi, S. M., Buschkuhl, M., Shah, P., and Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Mem. Cognit.* 42, 464–480. doi: 10.3758/s13421-013-0364-z
- Katz, B., Buschkuhl, M., Jaeggi, S., and Shah, P. (in press). Differential effect of motivational features on training improvements in school. *Front. Hum. Neurosci.* doi: 10.3389/fnhum.2014.00242
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends Cogn. Sci.* 14, 317–324. doi: 10.1016/j.tics.2010.05.002
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized training of working memory in children with ADHD-A randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 177–186. doi: 10.1097/00004583-200502000-00010
- Klingberg, T., Forssberg, H., and Westerberg, H. (2002). Training of working memory in children with ADHD. *J. Clin. Exp. Neuropsychol.* 24, 781–791. doi: 10.1076/j.jcen.24.6.781.8395
- Kornell, N., and Bjork, R. A. (2008). Learning concepts and categories is spacing the “enemy of induction”? *Psychol. Sci.* 19, 585–592. doi: 10.1111/j.1467-9280.2008.02127.x
- Melby-Lervåg, M., and Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Dev. Psychol.* 49, 270–291. doi: 10.1037/a0028228
- Miyake, A., and Shah, P. (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. New York, NY: Cambridge University Press.
- Morrison, A. B., and Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychon. Bull. Rev.* 18, 46–60. doi: 10.3758/s13423-010-0034-0
- Rapport, M. D., Orban, S. A., Kofler, M. J., and Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clin. Psychol. Rev.* 33, 1237–1252. doi: 10.1016/j.cpr.2013.08.005
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *J. Exp. Psychol. Gen.* 142, 359–379. doi: 10.1037/a0029082
- Schmidt, R. A., and Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychol. Sci.* 3, 207–217. doi: 10.1111/j.1467-9280.1992.tb00029.x
- Shah, P., Buschkuhl, M., Jaeggi, S., and Jonides, J. (2012). Cognitive training for ADHD: the importance of individual differences. *J. Appl. Res. Mem. Cogn.* 1, 204–205. doi: 10.1016/j.jarmac.2012.07.001
- Shebilske, W. L., Corrington, K., and Jordan, J. A. (1994). “Massed versus distributed practice in complex skill acquisition,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Nashville: SAGE Publications), 1215–1219. doi: 10.1177/154193129403801816

- Shebilske, W. L., Goettl, B. P., Corrington, K., and Day, E. A. (1999). Interlesson spacing and task-related processing during complex skill acquisition. *J. Exp. Psychol.* 5, 413–437. doi: 10.1037/1076-898X.5.4.413
- Shipstead, Z., Redick, T. S., and Engle, R. W. (2012). Is working memory training effective? *Psychol. Bull.* 138, 628–654. doi: 10.1037/a0027473
- Sio, U. N., and Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychol. Bull.* 135, 94–120. doi: 10.1037/a0014212
- Sisti, H. M., Glass, A. L., and Shors, T. J. (2007). Neurogenesis and the spacing effect: learning over time enhances memory and the survival of new neurons. *Learn. Mem.* 14, 368–375. doi: 10.1101/lm.488707
- Stafford, T., and Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychol. Sci.* 25, 511–518. doi: 10.1177/0956797613511466
- Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature* 437, 1272–1278. doi: 10.1038/nature04286
- Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x
- Vlach, H. A., Sandhofer, C. M., and Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition* 109, 163–167. doi: 10.1016/j.cognition.2008.07.013
- Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Östensson, M.-L., Bartfai, A., et al. (2007). Computerized working memory training after stroke-A pilot study. *Brain Inj.* 21, 21–29. doi: 10.1080/02699050601148726
- Zhao, X., Wang, Y., Liu, D., and Zhou, R. (2011). Effect of updating training on fluid intelligence in children. *Chin. Sci. Bull.* 56, 2202–2205. doi: 10.1007/s11434-011-4553-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 November 2013; accepted: 27 March 2014; published online: 10 April 2014.

Citation: Wang Z, Zhou R and Shah P (2014) Spaced cognitive training promotes training transfer. *Front. Hum. Neurosci.* 8:217. doi: 10.3389/fnhum.2014.00217

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Wang, Zhou and Shah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Differential effect of motivational features on training improvements in school-based cognitive training

Benjamin Katz^{1*}, Susanne Jaeggi², Martin Buschkuhl³, Alyse Stegman⁴ and Priti Shah⁴

¹ Combined Program in Education and Psychology, Department of Psychology, University of Michigan, Ann Arbor, MI, USA

² School of Education, University of California at Irvine, Irvine, CA, USA

³ MIND Research Institute, Irvine, CA, USA

⁴ Department of Psychology, University of Michigan, Ann Arbor, MI, USA

Edited by:

Michelle W. Voss, University of Iowa, USA

Reviewed by:

Martin Lövdén, Max Planck Institute

for Human Development, Germany

Pauline L. Baniqued, University of

Illinois at Urbana-Champaign, USA

Joshua Cosman, Vanderbilt University, USA

*Correspondence:

Benjamin Katz, Combined Program in Education and Psychology, Department of Psychology, University of Michigan, 2229 East Hall, 530 Church Street, Ann Arbor, MI 48104, USA

e-mail: benkatz@umich.edu

Cognitive training often utilizes game-like motivational features to keep participants engaged. It is unclear how these elements, such as feedback, reward, and theming impact player performance during training. Recent research suggests that motivation and engagement are closely related to improvements following cognitive training. We hypothesized that training paradigms featuring game-like motivational elements would be more effective than a version with no motivational elements. Five distinct motivational features were chosen for examination: a real-time scoring system, theme changes, prizes, end-of-session certificates, and scaffolding to explain the lives and leveling system included in the game. One version of the game was created with all these motivational elements included, and one was created with all of them removed. Other versions removed a single element at a time. Seven versions of a game-like n-back working memory task were then created and administered to 128 students in second through eighth grade at school-based summer camps in southeastern Michigan. The inclusion of real-time scoring during play, a popular motivational component in both entertainment games and cognitive training, was found to negatively impact training improvements over the three day period. Surprisingly, scaffolding to explain lives and levels also negatively impacted training gains. The other game adjustments did not significantly impact training improvement compared to the original version of the game with all features included. These findings are preliminary and are limited by both the small sample size and the brevity of the intervention. Nonetheless, these findings suggest that certain motivational elements may distract from the core cognitive training task, reducing task improvement, especially at the initial stage of learning.

Keywords: working memory, intervention, motivation, video games, n-back

INTRODUCTION

A key challenge in cognitive training research is how to keep participants engaged in training. Training programs are often challenging for participants to complete, and it is expected that they will remain focused on a task or set of tasks for 20–40 min at a time (Jaeggi et al., 2008; Thompson et al., 2013), for anywhere between a few days (Rueda et al., 2005) to 100 sessions (Schmiedek et al., 2010). Because transfer improvements generally require several hours of training (Jaeggi et al., 2008; Stepankova et al., 2013) it is important that participants in training paradigms remain compliant during training. Additionally, it may be necessary for participants to improve in the training program in order to experience transfer on untrained tests (Jaeggi et al., 2011).

Unfortunately the time commitment and effort required to complete a cognitive training study is often such that many participants do not complete the experiment. Studies often have high dropout rates, including some higher than 25% (Redick et al., 2013; Jaeggi et al., 2014). A variety of individual difference factors may contribute to a participant's ability to successfully engage in and complete the training, such as baseline ability in the training

task and one's intrinsic motivation to complete a training program (Jaeggi et al., 2014).

While individual difference factors are generally outside of the experimenters' control, the design of the training program may also contribute to a participant's engagement in the task, and these game design elements are often relatively simple to adjust. Cognitive training paradigms vary widely in the type of motivational elements they include, however, and while some studies have focused on recruiting unpaid, intrinsically motivated individuals that may be more likely to engage with and complete a training regimen (Jaeggi et al., 2014), others have utilized substantial financial compensation as a means of encouraging participants to complete the training (Redick et al., 2013; Thompson et al., 2013). Factors that may impact a participant's ability and willingness to engage, comply, and improve in training have been the subject of some interest in recent research. Studies with children often utilize prizes, certificates, and display of high scores to encourage individuals to excel at and complete the training (Holmes et al., 2009; Jaeggi et al., 2011; Wang et al., 2014).

One topic that has not gotten much attention is how game-based motivational elements may contribute to improvements in

training and transfer. This is somewhat surprising, considering that elements such as score, tutorials and scaffolding, theming, and feedback are often prominently featured in cognitive training programs. Cognitive psychologists and neuroscientists often find themselves in the role of game designer (Mané and Donchin, 1989; Anguera et al., 2013), and even some of the most basic training paradigms have at least included a motivational chart showing player improvements (Jaeggi et al., 2008). Other training programs, particularly those targeted at children (Klingberg et al., 2005; Jaeggi et al., 2011) look and feel more like traditional video games with appealing art and sound design. Cognitive training games are similar to certain types of entertainment games – specifically, those that Gee (2006) would describe as “problem games,” – that involve simple, repetitive mechanics, rather than large, open worlds for the player to explore. Almost all tasks used in cognitive training games, from n-back to useful field of view to conflict resolution tasks, can be translated into fairly simple gameplay mechanics (Klingberg et al., 2005; Rueda et al., 2005; Ball et al., 2010; Jaeggi et al., 2011; Alloway, 2012).

While game-based motivational elements have not been well-studied within cognitive training research, some of them have been examined by learning game researchers. For example, one popular game element, persistent scoring (the presentation of a number that represents player performance and changes as the player completes the task successfully) likely encourages engagement and motivation (Toups et al., 2009). However, the way this scoring is implemented, that is, whether points are earned specifically for completing tasks essential to the learning goal or are awarded for other non-core actions, can determine if scoring hinders or helps learning on the task (Habgood and Ainsworth, 2011). The inclusion of game features may either support or subvert participant motivation to engage depending on how well they tie in with the learning task and the participant’s pre-existing motivational framework. For example, imagine a cognitive training game that includes an extra bonus round where players perform some other task non-essential to the training component, such as answering a trivia question. If the number of points possible for the bonus round matches or is greater than that awarded during the core task, participants may be less motivated to perform well during the training portion of the game. This contrasts with situations where the reward is directly reinforcing of the performance task. In one related example, a review of reading incentive programs supported using literacy-related reward to motivate students (Fawson and Moore, 1999); one study found that students who received a book as a reward following a reading program were more motivated to participate than those who received a token prize (Marinak and Gambrell, 2008).

Psychologists who study motivation are also interested in game-based motivational features (Ryan et al., 2006; Przybylski et al., 2010), possibly because games are an ideal context for understanding the tension between intrinsic motivation and extrinsic reward. Elements such as scoring and feedback may impact a player’s intrinsic motivation and may also contribute to their success in learning the content included in the game. For example, in Malone’s examination of intrinsically and extrinsically

motivating game elements, different versions of the game *Breakout* were created that included elements of feedback such as persistent score and breaking bricks (Malone, 1981). Versions of the game with both of these elements were rated much more highly on a scale of enjoyment by players than versions where they were not present. Theming (referred to as “fantasy” by Malone) was also evaluated and found to significantly contribute to a child’s interest in the game, although gender differences were identified in the type of theme each child enjoyed the most.

More recently some focus has been applied to issues of motivational game elements in cognitive research, however, the research has thus far been inconclusive. Two recent studies compared game versions that included a variety of motivational elements such as those studied by Malone, to more basic versions of a task. While Prins et al. (2011) found that including game elements such as theming, game-like feedback, and animations increased motivation as well as performance for children completing a working memory training game, recent work from Hawkins et al. (2013), found that the addition of similar game features improved the player experience but not the quality of data collected during a cognitive task. One possible explanation of these mixed results is that the amount of time spent with the game experience also matters – while Prins et al. (2011) examined the effect of game features over 3 weekly sessions, the Hawkins et al. (2013) study included one single session of play for the games used.

It is also possible that the impact of scoring and other game-like features may differ from game to game, depending on factors such as the goals, difficulty, and demographics of the users. Conclusions drawn from one study cannot necessarily be applied more generally to other types of games or interactive experiences. Nevertheless, no study thus far has systematically examined the impact that *individual* game elements, rather than several features together, have on player performance; previous studies such as those from Hawkins et al. (2013) and Prins et al. (2011) compare versions of the game with a variety of features to versions of the game without any features present. Therefore findings from this present research will be of considerable interest to game designers beyond the cognitive training space. By separating out the most popular game-elements included in these training games, such as scoring, lives and leveling, prizes, and theming, we may better understand the extent that these elements contribute to participant engagement and improvements on the task.

To examine how these elements impacted performance on a visuospatial working memory training task, we designed several versions of a three-day working memory game based on a cognitive training task used in previous research (Jaeggi et al., 2011). In the original version of the task, many motivational elements were included, such as changing themes and art, display of score, lives and levels, and prizes and certificates awarded for player compliance and performance. We created new versions of this game, each with one of these elements removed, as well as one with several game-like motivational features absent from the training task. Even without persistent score, lives, prizes, and changing theme, the task was still *game-like*, with whimsical art and scoring presented between rounds.

This point brings up a significant additional note: why versions of the game with a single element removed were created rather than several versions with one single element added to a bare-bones version of the task. This would likely have been the approach taken if the experimenters had created a completely new game, however, each version of the task is a modification of a training game used in a previous cognitive training study (Jaeggi et al., 2011). Removing a single element generally did not impede gameplay but some elements are interdependent with each other. For example, the prizes students could pick at the end of each day in most conditions were offered based on the total score; students with a higher score could pick prizes of greater value. While other types of feedback (such as the display of leveling on screen) still gave sense of their performance and could be connected to earning better prizes, the addition of performance-based prizes without *any* additional context may not have made sense to the player. In this study the question of interest was whether removing any additional feature might have significant effect on motivation or training gain, and thus each version had one element removed. However, an alternative design, where a single feature is added to a completely bare-bones version of a task, offers an interesting possibility for future research.

We hypothesized that there would be a differential effect of motivational feature for learning on the training task. For example, given existing research on the potential negative effects of extrinsic reward, such as Marinak and Gambrell (2008), we expected that the removal of prizes might increase learning on the training task. However, in general, the findings from the Prins et al. (2011) study led us to expect that students in the group with all motivational elements included would outperform students in the no motivational element group. Additionally, students in a previous study using the same version of the game as in the “all motivational features” group who reported greater enjoyment of the task outperformed those who did not enjoy the task as much (Jaeggi et al., 2011). The results from Hawkins et al. (2013) and Prins et al. (2011) suggested that students in the group with the most motivational elements would rate more highly on self-report measures of intrinsic motivation or enjoyment; it is possible that the versions of the game that students enjoyed more would also be the versions where they experienced greater improvement on the training task. Thus we expected that removing other features commonly included in games, such as changing theme, scoring, and lives and levels would have a deleterious effect on learning the training task.

We included an outcome measure relatively similar (but not identical) to the training task, in which players were required to identify if a given object presented on screen matched an object presented on screen *n*-items earlier. Despite the similarity between the transfer task and the outcome measure, we did not expect to see significant transfer gain due to the limited three-day training duration. Rather, we primarily expected to find differences in player self-reported and observed motivation and performance on the task based on which elements were excluded. We hope that a better understanding of how the game-like elements included in this study impact motivation and performance will help researchers design better, more scientifically useful, cognitive training paradigms.

MATERIALS AND METHODS

PARTICIPANTS

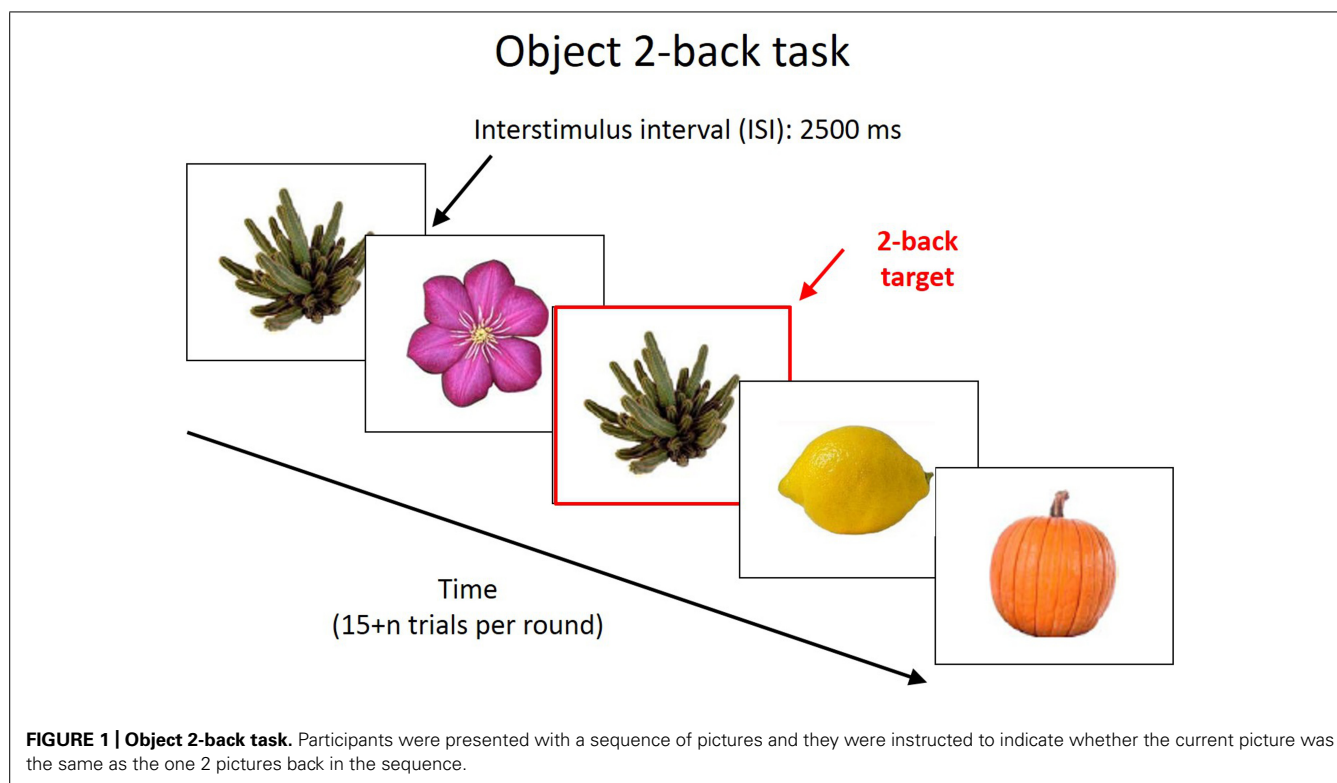
128 students were recruited from seven different school-based summer camps in the southern Michigan region (average age = 10.56 years, SD = 2.48, range 5–14, 37% girls). Students were invited to participate in a three day long experiment in which no compensation was provided outside of the possibility of prizes or certificates in some variants of the intervention; recruitment occurred at tables outside the entrance to the summer camps immediately prior to the start of each camp. Written informed consent was collected from both parents and students prior to participation. Students were also asked if they wished to continue the experiment prior to each training session and were informed that they could end their participation at any time. 21 students were not included in the analysis due to either not completing the entire three days of training and testing ($N = 13$), having taken part in previous cognitive training research ($N = 2$), or being too young to be included in the study (younger than 6 years, $N = 6$). Of the 13 students who dropped out and were not too young or participants in previous cognitive training research, no more than four dropped out of any individual condition. 107 students (average age = 10.65 years, SD = 2.36, range 6–14, 44% girls) were then included in the final analysis. Because students completed the tests, questionnaires, and training together as part of the camp, game versions were assigned randomly at the camp level to avoid children comparing the game and prizes amongst themselves and perhaps being disappointed when some received prizes or played more engaging games than others. Running the experiment within summer camps enabled us to evaluate motivational features in a real-world environment, however, one trade-off of this approach is that group sizes and ages differed somewhat depending on which camp students were recruited from. The demographic information for each condition is included in **Table 1**.

PROTOCOL

A pre-test was administered on the first day of the experiment prior to the training (**Figure 1**). The pre-test consisted of a computerized object 2-back assessment that presented participants with a sequence of images one at a time. Participants were required to determine whether each item matched the one presented two items previously and then press one of two keys to indicate their answer. An object was presented every 3 s, with a presentation

Table 1 | Demographic information.

Condition	<i>N</i>	Age (years)	Grade
All optional features included	25	11.28, SD 2.82	6.00, SD 2.38
No theme change	13	10.92, SD 1.71	5.77, SD 1.54
No points shown	19	12.21, SD 2.20	6.84, SD 1.64
No prizes	15	8.40, SD 2.03	3.33, SD 1.63
No explanation of lives/levels	11	9.82, SD 1.72	5.00, SD 1.61
No explanation of lives/levels or certificates	12	10.83, SD 0.835	5.42, SD 0.52
No optional features included	12	10.00, SD 1.86	4.83, SD 1.64



time of 500 ms and an inter-stimulus interval of 2,500 ms. The pre-test consisted of three blocks of 17 stimuli each and performance was measured as the proportion of correct answers minus the proportion of false responses. Each block included five target trials and 10 non-target trials after the presentation of the initial two stimuli. A few practice trials were included prior to the actual assessment to ensure that the children understood how to complete the task. Following the pre-test on the first day of the study, students began training with the n -back working memory game. After the training on each day, experimenters orally administered brief surveys with Likert-type questions asking how much students enjoyed the game, how exciting the game was, how difficult the game was, and how much effort each student had put into the game. These four questions were adapted for a previous cognitive training study from a factor analysis of the Intrinsic Motivation Inventory (McAuley et al., 1989; Jaeggi et al., 2011). Each of these variables was averaged over the course of three-days to create enjoyment, excitement, effort, and difficulty variables. Researchers also rated students on how engaged they seemed during each day of training using a Likert-type scale following each training session; this was also averaged over the course of the three days to create a final observer engagement score for each participant. Following the third day of training, participants completed the object 2-back assessment a second time.

Cognitive training game

Participants trained on a game-like computerized working memory task similar to that used in a previous study with children (Jaeggi et al., 2011). This spatial n -back task presented participants

with stimuli at one of six locations on the screen, at a rate of 3 s each, with 2,500 ms between stimuli and with each stimulus presented for 500 ms. Students were required to press the A key each time the current stimulus matched the location of the one presented n items previously, and the L key each time the current stimulus did not match. Participants completed 10 rounds of this task each day, each round consisting of 15 + n trials, and each round consisting of five targets and 10 + n non-targets. All versions of the game were adaptive in that the n level was adjusted depending on performance in each round. If a participant made four or more errors they would lose a single life; after losing three “lives” the participant’s n -level would be decreased by 1 in the following round. If a participant made three or less errors n increased by 1 in the following round.

Seven versions of the n -back training game was developed to examine the role of five motivational features: points, theming, explanation of lives and levels, prizes, and end-of-session certificates. One version of the game included all of these motivation features, while another included none of them. Four of the other versions excluded one of these features. Due to experimenter error, one additional version that was meant to exclude the certificates provided to players at the end of each session also excluded the display of lives and levels feature. However, because this group (with two interrelated elements) was of potential interest, it was included in the subsequent analysis.

Theming. Several different themes were developed to make the n -back task more appealing to students, that is, a frog jumping on lily pads, a cat appearing in windows of a haunted house, and a monkey jumping from sail to sail on a pirate ship (**Figure 2**). In

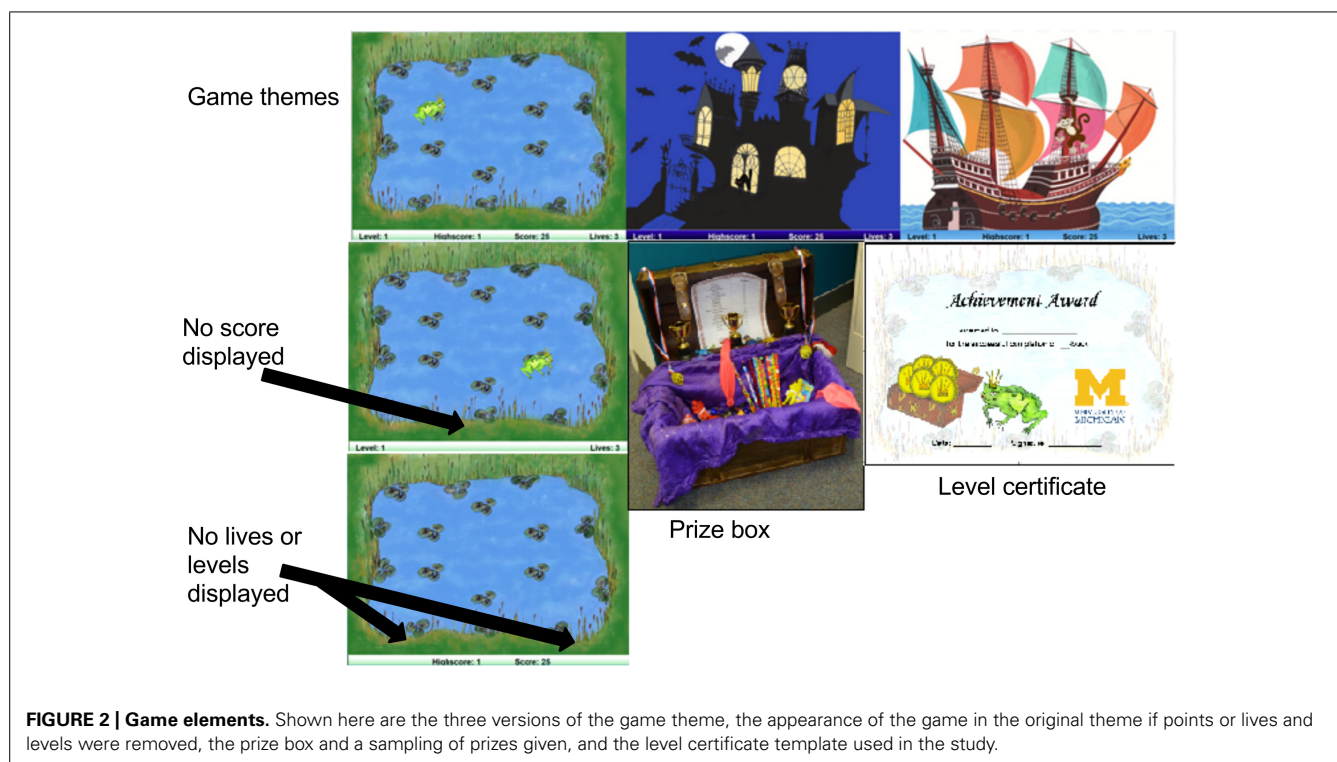


FIGURE 2 | Game elements. Shown here are the three versions of the game theme, the appearance of the game in the original theme if points or lives and levels were removed, the prize box and a sampling of prizes given, and the level certificate template used in the study.

all game versions except for the one excluding theming, the theme changed before the first round on the second and third day of training. In the “no theme” group as well as the “no motivational features” group, only the lily pad theme was included, and this theme remained persistent across the three days of training.

Score. A bar on the bottom of the screen displayed score as the player completed the n -back task. Points were earned for correctly identifying whether the location of the character on the screen matched the location presented n instances earlier. In versions of the game with prizes, players were instructed that they could trade in points earned for a prize at the end of each day. In the “no points” and “no motivational features” versions of the game, the persistent score was hidden during play (Figure 2). The score was still shown at the end of each round, however.

No display or explanation of lives or levels. Lives left and the current level were displayed on-screen during play. “Levels” indicated the n level the user was currently on, while “Lives” was used to indicate how many errors the participant could make before dropping an n -back level on the subsequent round. The “Lives” and “Levels” indicator were hidden on the bottom bar (Figure 2) for the “no lives or levels explanation” group, as well as the “no motivational features” group. Additionally, in these groups the experimenter did not mention lives and levels. The game remained adaptive as in the other conditions, however, and the participants still received a certificate after each day’s training with the n -back level he or she had reached.

Prizes. Prizes were offered each day after the completion of the game in exchange for “points” the participants had earned. In

the “no prizes” group and the “no motivational features” group, participants were given a prize at the very end of the study, but not each day during training. Additionally, participants in those groups were not told that prizes would be given prior to completing the post-test on day three. In the groups where prizes were present, students were allowed to see a treasure box (Figure 2) from which they would select items at the end of each day.

End of session certificates and no display or explanation of lives and levels. Players were awarded a certificate (Figure 2) at the end of each training day celebrating the level they reached. In the “no certificate” version of the game players were supposed to complete the standard version of the task but without a certificate given at the end of the round, however, the experimenters for this group incorrectly administered a version of the game without the display of lives or levels. Thus players in one of the seven groups were not aware of the role of lives or levels during the task, and additionally did not receive certificates at the end of each day

RESULTS

To identify differences in motivation, training performance over time, and pre/post-test performance on the object n -back measure, omnibus analysis of covariances (ANCOVAs) were conducted with all game conditions included; in the case of a significant effect of game-type on these variables, follow-up ANCOVAs were conducted comparing each game variant to the original version with all features included. Despite attempts to recruit summer camps with similar ages, there were significant differences in age across some of the training groups $F(6,100) = 5.46$, $p < 0.001$, partial $\eta^2 = 0.247$. Age predicted improvement in the training following

a regression analysis including the age of the pooled participants as predictor and the rate of improvement (operationalized as the slope of a linear model – see also below) of the task as outcome, $\beta = -0.202$, $t(105) = -2.108$, $p < 0.05$, $R^2 = 0.041$ (proportion of variance in slope explained $F(1,105) = 4.444$, $p = 0.037$). Thus, we included age as a covariate in our subsequent analysis.

TRAINING PERFORMANCE

To quantify each participant's training improvement over the three sessions of training, the slope of a linear regression model was calculated for each participant using the average n-back level per day of training (Figure 3). Due to the difference in ages across conditions (Table 1) and the variance in baseline performance across game versions (Figure 4), we included age and starting n-back level as covariates in our analyses. A univariate ANCOVA across conditions revealed a significant effect of game-version on training improvement as measured by linear slope $F(6,98) = 2.49$, $p = 0.028$, partial $\eta^2 = 0.132$).

To analyze the effect of each individual motivational feature on performance, we then computed a set of univariate ANCOVAs with training slope as the dependent variable, calculated from the average n-back on each day of the training task. We compared students playing the version of the game with the full set of motivational features to students playing each of the other versions with elements removed; see Table 2 and Figure 5. Students who played the version of the game without the persistent display of score performed significantly better at the training task over time versus students who played the version of the game with all motivational features, $F(1,40) = 7.22$, $p = 0.010$, partial $\eta^2 = 0.153$) as did students who completed the version of the game without the indication of lives or levels, $F(1,32) = 4.48$, $p = 0.042$, partial $\eta^2 = 0.123$. However, students in the group without theme changes did not perform significantly different from the group with all motivational features $F(1,34) = 0.07$, $p = 0.801$, partial $\eta^2 = 0.002$), nor did the group that did not receive prizes after each training session $F(1,36) = 0.01$,

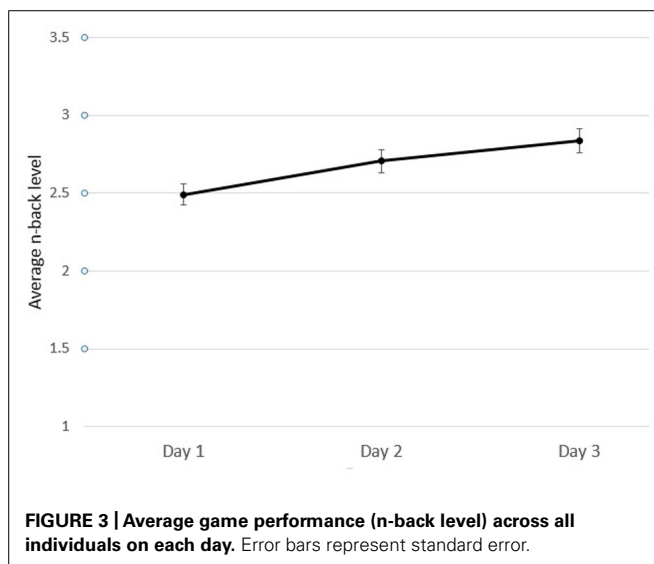


FIGURE 3 | Average game performance (n-back level) across all individuals on each day. Error bars represent standard error.

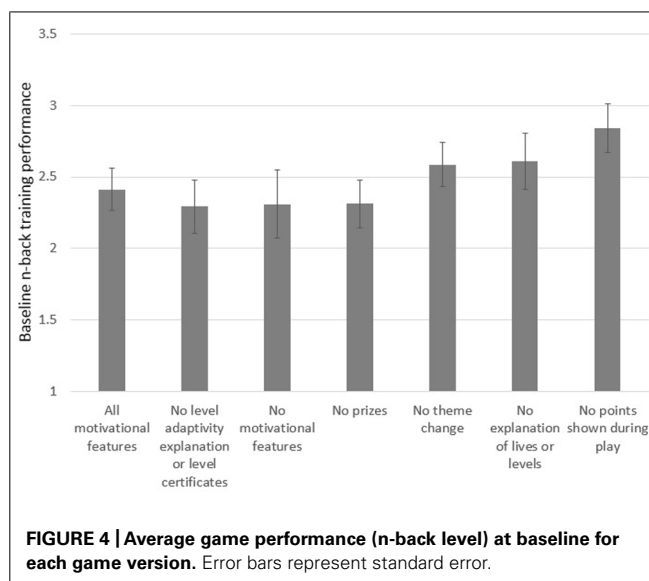


FIGURE 4 | Average game performance (n-back level) at baseline for each game version. Error bars represent standard error.

$p = 0.932$, partial $\eta^2 = 0.000$). The group that did not receive certificates after each day, and also did not see lives or level information during gameplay, trended worse than the all motivational features group, but not significantly so, $F(1,33) = 2.60$, $p = 0.116$, partial $\eta^2 = 0.073$. The group that completed the version with no motivational features trended higher but did not differ significantly in performance on the training task from the group with all features, $F(1,33) = 2.00$, $p = 0.167$, partial $\eta^2 = 0.057$).

An additional analysis was performed to further examine the most robust finding, that the display of points on screen may have had a deleterious effect on game performance, as well as to partially address the issue of small samples sizes in the study. A final univariate ANCOVA was thus conducted in a similar manner as above with both the group without any motivational features and the group with only the score removed ($N = 31$) compared to all other participants ($N = 75$, all of whom played a version of the game where points were displayed). This analysis further supported the original finding, as the combined task performance of all individuals who did not see points displayed was significantly better than the combined performance of all individuals who did have points displayed $F(1,103) = 7.937$, $p = 0.006$, partial $\eta^2 = 0.072$.

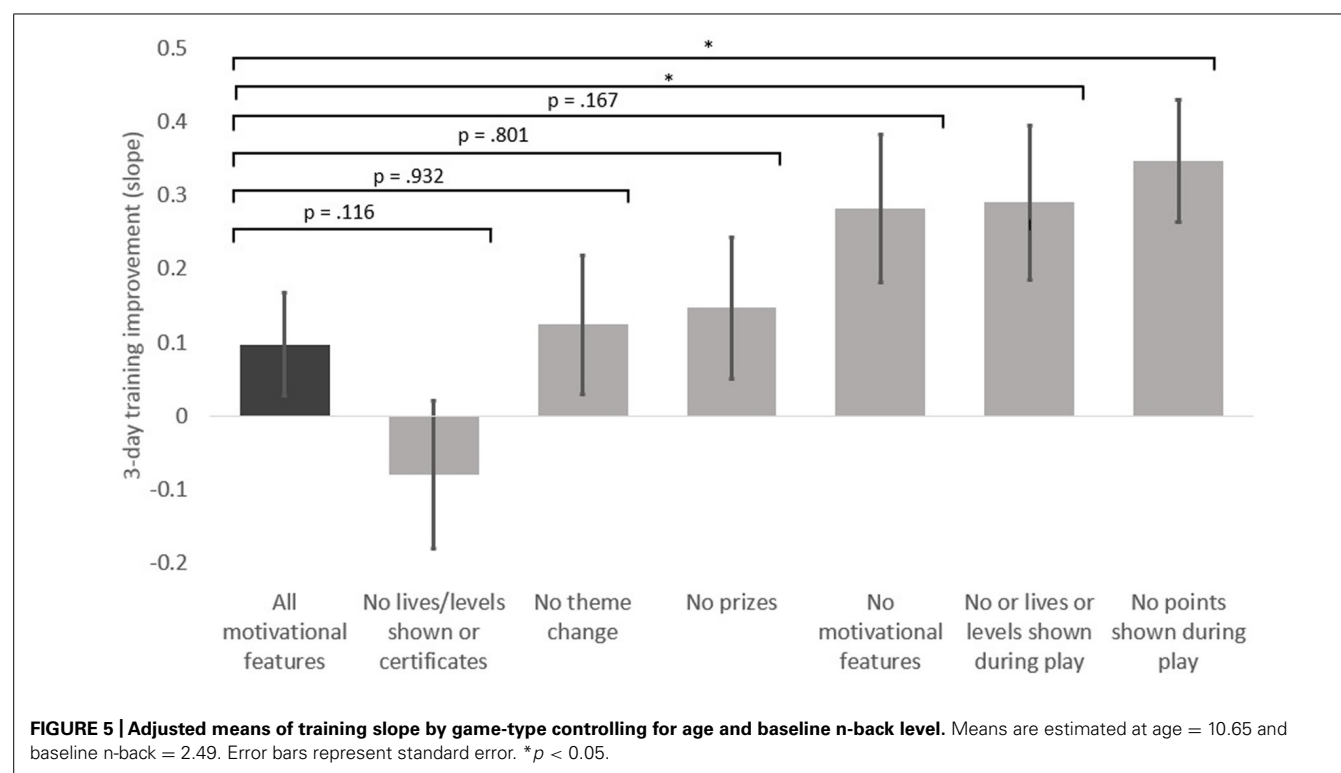
MOTIVATION

Participant self-ratings of task-related enjoyment, difficulty, effort, and excitement were averaged over the three days and examined as a function of game variant. ANCOVAs with game-type as the independent variable and age as a covariate did not find a significant effect of game-type on student self-report of enjoyment, $F(6,98) = 1.52$, $p = 0.180$, partial $\eta^2 = 0.084$, excitement, $F(6,98) = 1.43$, $p = 0.188$, partial $\eta^2 = 0.080$, or effort, $F(6,98) = 1.35$, $p = 0.241$, partial $\eta^2 = 0.076$, or student self-report of difficulty, $F(6,98) = 1.94$, $p = 0.082$, partial $\eta^2 = 0.105$. However, as in other studies of motivational factors in differently aged students (Lepper et al., 2005), a median

Table 2 | Training improvement by game variant.

Condition	<i>N</i>	Estimated mean slope	Standard error	<i>p</i>	partial η^2
All motivational features	25	0.10	0.07	–	–
No certificates given and no lives or levels displayed	12	−0.08	0.10	n.s.	0.07
No theme changes	13	0.12	0.10	n.s.	0.00
No prizes awarded	15	0.15	0.10	n.s.	0.00
No points displayed	19	0.35	0.08	*	0.15
No lives or levels displayed	11	0.29	0.10	*	0.12
No motivational features	12	0.28	0.10	n.s.	0.06

Slope here represents the linear model of training improvement calculated using the average *n*-back performance during each daily session. Slope is controlled for age and baseline *n*-back level and means are estimated at age = 10.65 and baseline *n*-back = 2.49. Significance and effect size are drawn from each follow-up ANCOVA that compares a single condition to the original, all motivational features group. Only the “No points displayed” and “No lives or levels displayed” differed significantly from the “all motivational features” comparison group. * $p < 0.05$.



split of students by age revealed that students 10 and under ($N = 47$, mean = 3.75, $SD = 0.79$) were significantly more excited to complete the task than students 11 and older ($N = 60$, mean = 3.29, $SD = 0.72$) to complete the task [$F(1,103) = 9.78$, $p = 0.002$, partial $\eta^2 = 0.085$]. On self-ratings of enjoyment, younger students ($N = 47$, mean = 3.89, $SD = 0.54$) were also more likely than older students ($N = 60$, mean = 3.57, $SD = 0.67$) to enjoy the task, $F(1,103) = 7.38$, $p = 0.008$, partial $\eta^2 = 0.066$, suggesting at least that the student questionnaires did accurately capture their personal feelings regarding engagement with the game. Additional analyses of motivational factors for the combined game versions without the display of points compared to the game versions with points on screen

did not identify a significant impact of this feature on any of the motivational factors, although students in the group that did not see a persistent score reported applying marginally less effort during gameplay ($N = 31$, $M = 3.76$, $SD = 0.55$) than those who did see a score ($N = 75$, $M = 3.48$, $SD = 0.67$), $F(1,103) = 3.901$, $p = 0.051$, partial $\eta^2 = 0.036$. Averaged observer ratings of player engagement over the three-days were also examined as a function of game variant. Again, an ANCOVA was conducted including researcher engagement ratings as the dependent variable, game-type as the independent variable, and age as a covariate. Game-type did not significantly predict experimenter ratings of engagement, $F(6,99) = 1.91$, $p = 0.086$, partial $\eta^2 = 0.104$.

OBJECT n-BACK TRANSFER TASK

Finally, performance on the object 2-back near-transfer task was examined through an ANCOVA with gain on the object n-back test as the dependent variable, game type as the independent variable, and age and pre-test performance on the object 2-back task as covariates. No differences in improvement were identified between any of the game variants, $F(6,98) = 1.54$, $p = 0.175$, partial $\eta^2 = 0.086$. There was a marginal effect of having score displayed when all individuals who played a version without persistent scoring ($N = 31$, mean object n-back gain = 0.06, $SD = 0.21$) were compared to the combined participants training with a version with persistent score ($N = 75$, mean object n-back gain = 0.02, $SD = 0.29$), $F(1,103) = 3.070$, $p = 0.083$, $\eta^2 = 0.029$. This is not surprising, however, as the training regimen were likely too short for sizable near-transfer effects to occur. The untrained object n-back task performance across all participants was not significantly higher after only three days of training ($M = 0.473$, $SD = 0.258$) than at the start ($M = 0.443$, $SD = 0.221$), as revealed through a paired-samples t -test, $t(106) = 1.14$, $p = 0.255$.

DISCUSSION

The results of this research should add nuance to our understanding of how popular “motivational” game features impact actual player performance. Over the three days of the study, students playing versions of the game without the persistent display of points and without the display of lives or levels improved significantly more on the game task than students using the original version of the game with all features present. Students playing game versions without changing theme, daily prizes, or end-of-session certificates and the display of lives and levels did not perform significantly differently than the comparison group. Game version did not significantly influence student motivation or performance on the object n-back task.

The effect of these game elements on training performance may seem counterintuitive at first. Why did only the “no score displayed” and “no lives or levels displayed” groups perform differently than the group with all features? And why was the removal of these motivational features associated with improved performance on the training task over the three sessions? It is worth noting that score and lives and levels were indicated on a persistent bar near the game space, a common feature in games. It is quite possible that any element that distracted the user from the challenging n-back task during the actual game would reduce performance. This is an interesting finding in light of the fact that cognitive training – and learning games in general – often include elements such as score or lives prominently in the game space. Given this possibility, one outstanding question is why the group without any motivational features did not perform significantly better than the group with all motivational features included. It is possible that although the no motivational features group did have fewer distracting elements, the exclusion of all other, non-distracting elements had a combined deleterious effect on performance. Determining whether there is a “happy medium” of motivational features that result in optimized performance is a worthwhile goal for future research. Additionally, the other motivational elements, such as awarding prizes and theme changes,

did not occur during core gameplay. Over the longer term these elements may impact performance differently, but this finding provides some evidence for removing motivational elements that may be distracting from the player in the early days of a cognitive training regimen.

Overall, the lack of an effect of game variant on student self-ratings of motivation and performance on the untrained object n-back task is not surprising. Each version of the training program still appeared game-like, and the removal of any individual feature may have a minimal effect on motivation. This suggests that cognitive game designers may be able to remove some of the game elements that were found to be distracting without any negative impact on a player’s own perceptions of enjoyment and excitement. Finally, it is not necessarily surprising that only the training improvements and not performance on the object n-back task was affected by condition within the limited three-day scope of the study. It is possible that, in an experiment utilizing a much longer training experience, differences in transfer might have been observed.

Several limitations inherent to the present study should be considered. Perhaps of greatest concern is the limited sample size and significant age differences across some of the conditions. While some of the groups are adequately powered, others, due to dropout or other extenuating factors, have as few as 11 participants. Age was included as a covariate in the analyses, but the small sample sizes mean that it is difficult to fully account for the influence of age on differences in training performance. Because these findings were not corrected for multiple comparisons, and the effect sizes found were fairly small, these findings must necessarily be seen as preliminary, and, while informative of future research, not conclusive.

Additionally, this is not a true randomized controlled study – while camps were assigned to conditions randomly, all participants within each camp trained on the same variant of the game. Both of these factors are tradeoffs resulting from the real-world nature of the study; students trained amongst their peers in an actual school environment. Finally, some features of the training regimen, such as the illustrative art style and display of score at the end of each round, exist in all versions of the game. These other features may impact student performance and engagement as well, and were not examined here. The fact that some of the more subtle motivational features, such as persistent score, had a significant impact on three-day performance improvements indicates that these other features should be a focus of future research. As mentioned in the introduction, one further consideration is the possibility that certain game elements may interact with each other and that this may influence participant engagement or performance on the training task. For example, it is possible that persistent scoring is more motivating when participants receive prizes based on their score at the end of each day. This is potentially a significant issue and one that is not examined in the present study.

Besides including additional game variants, future research could also focus on the impact of these motivational features over a longer-term training regimen. It is possible that some features that impede performance on the training task in this study have less of an effect in a longer training regimen. However, given evidence

that long-term improvement in the training task is necessary for transfer gains, any feature that impacts training performance is worth special consideration (Jaeggi et al., 2011). Given the fact that persistent scoring and the lives/levels feature did impact training performance, we recommend that developers of cognitive training exercise discretion when incorporating these features into their programs.

Our findings have broad implications not only for developers of cognitive training but game designers and cognitive psychologists more generally. Psychologists often make tasks game-like in an effort to drive user engagement. Likewise, within education there has recently been a movement toward game-like formative assessment to evaluate student performance (Wang, 2008). Our findings suggest that game-like elements should be added with caution. Adding game features to an already stressful testing situation may have a deleterious impact on student performance, particularly if the game features add irrelevant cognitive demands. Even seemingly innocuous features, such as displaying score or giving players a certain number of “lives,” may impact performance in a negative fashion. This does not mean that games cannot be effective teaching tools, instruments for cognitive training, or assessment mechanisms. On the contrary, this research provides further support for carefully matching game mechanics and features with the actual task. Researchers might take a look at venerable computerized training task, such as Space Fortress, and examine the impact that non-essential game-like elements included in those tasks have on performance.

While some research has supported the inclusion of game-like elements in cognitive training to improve motivation and training performance (Prins et al., 2011), our findings suggest that these features should be chosen judiciously. Combined with the results from Hawkins et al. (2013), our data suggest that game-like features may not improve the data one collects in research. Furthermore, distracting features may actually impair the participant's ability to improve quickly at the task. Certain “motivational” elements may at best be unnecessary for driving learning on the core task, and at worst have an effect counter to what is intended by the designer. Mae West may have said “the score never interested me, only the game,” but persistent display of score, like some other motivational features, might be distracting all the same.

REFERENCES

- Alloway, T. (2012). Can interactive working memory training improving learning? *J. Interact. Learn. Res.* 23, 197–207.
- Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., et al. (2013). Video game training enhances cognitive control in older adults. *Nature* 501, 97–101. doi: 10.1038/nature12486
- Ball, K., Edwards, J. D., Ross, L. A., and McGwin G. Jr. (2010). Cognitive training decreases motor vehicle collision involvement of older drivers. *J. Am. Geriatr. Soc.* 58, 2107–2113. doi: 10.1111/j.1532-5415.2010.03138.x
- Fawson, P. C., and Moore, S. A. (1999). Reading incentive programs: beliefs and practices. *Read. Psychol.* 20, 325–340. doi: 10.1080/027027199278385
- Gee, J. P. (2006). Are video games good for learning? *Nord. J. Digit. Lit.* 1, 172–182.
- Habgood, M. J., and Ainsworth, S. E. (2011). Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. *J. Learn. Sci.* 20, 169–206. doi: 10.1080/10580406.2010.508029
- Hawkins, G. E., Rae, B., Nesbitt, K. V., and Brown, S. D. (2013). Gamelike features might not improve data. *Behav. Res. Methods* 45, 301–318. doi: 10.3758/s13428-012-0264-3
- Holmes, J., Gathercole, S. E., and Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Dev. Sci.* 12, F9–F15. doi: 10.1111/j.1467-7687.2009.00848.x
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., and Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6829–6833. doi: 10.1073/pnas.0801268105
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., and Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108
- Jaeggi, S. M., Buschkuhl, M., Shah, P., and Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Mem. Cogn.* 42, 464–480. doi: 10.3758/s13421-013-0364-z
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized training of working memory in children with ADHD-A randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 177–186. doi: 10.1097/00004583-200502000-00010
- Lepper, M. R., Corpus, J. H., and Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: age differences and academic correlates. *J. Educ. Psychol.* 97, 184–196. doi: 10.1037/0022-0663.97.2.184
- Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cogn. Sci.* 5, 333–369. doi: 10.1207/s15516709cog0504_2
- Mané, A., and Donchin, E. (1989). The space fortress game. *Acta Psychol.* 71, 17–22. doi: 10.1016/0001-6918(89)90003-6
- Marinak, B. A., and Gambrell, L. B. (2008). Intrinsic motivation and rewards: what sustains young children's engagement with text? *Lit. Res. Instr.* 47, 9–26. doi: 10.1080/19388070701749546
- McAuley, E., Duncan, T., and Tammien, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. *Res. Q. Exerc. Sport* 60, 48–58. doi: 10.1080/02701367.1989.10607413
- Prins, P. J., Dovis, S., Ponsioen, A., Ten Brink, E., and Van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychol. Behav. Soc. Netw.* 14, 115–122. doi: 10.1089/cyber.2009.0206
- Przybylski, A. K., Rigby, C. S., and Ryan, R. M. (2010). A motivational model of video game engagement. *Rev. Gen. Psychol.* 14, 154–166. doi: 10.1037/a0019440
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *J. Exp. Psychol. Gen.* 142, 359–379. doi: 10.1037/a0029082
- Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102
- Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The motivational pull of video games: a self-determination theory approach. *Motiv. Emot.* 30, 344–360. doi: 10.1007/s11031-006-9051-8
- Schmiedek, F., Lövdén, M., and Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study. *Front. Aging Neurosci.* 2:27 doi: 10.3389/fnagi.2010.00027
- Stepankova, H., Lukavsky, J., Buschkuhl, M., Kopecek, M., Ripova, D., and Jaeggi, S. M. (2013). The malleability of working memory and visuospatial skills: a randomized controlled study in older adults. *Dev. Psychol.* doi: 10.1037/a0034913 [Epub ahead of print].
- Thompson, T. W., Waskom, M. L., Garel, K. L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., et al. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS ONE* 8:e63614. doi: 10.1371/journal.pone.0063614
- Toups, Z. O., Kerne, A., and Hamilton, W. (2009). “Motivating play through score,” in *ACM Computer Human Interaction 2009 Workshop on Engagement by Design*, Boston.

- Wang, T. H. (2008). Web-based quiz-game-like formative assessment: development and evaluation. *Comput. Educ.* 51, 1247–1263. doi: 10.1016/j.compedu.2007.11.011
- Wang, Z., Zhou, R., and Shah, P. (2014). Spacing and the transfer effect in game-based cognitive training. *Front. Hum. Neurosci.* 8:217. doi: 10.3389/fnhum.2014.00217

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 December 2013; accepted: 02 April 2014; published online: 24 April 2014.

Citation: Katz B, Jaeggi S, Buschkuhl M, Stegman A and Shah P (2014) Differential effect of motivational features on training improvements in school-based cognitive training. *Front. Hum. Neurosci.* 8:242. doi: 10.3389/fnhum.2014.00242

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Katz, Jaeggi, Buschkuhl, Stegman and Shah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Video game training and the reward system

Robert C. Lorenz^{1,2†}, Tobias Gleich^{1†}, Jürgen Gallinat^{1,3} and Simone Kühn^{4*}

¹ Department of Psychiatry and Psychotherapy, Charité-Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany

² Institute of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany

³ Department of Psychiatry and Psychotherapy, University Hospital Hamburg-Eppendorf, Hamburg, Germany

⁴ Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

Edited by:

Guido P. H. Band, Leiden University, Netherlands

Reviewed by:

Michelle W. Voss, University of Iowa, USA

Walter R. Boot, Florida State University, USA

*Correspondence:

Simone Kühn, Center for Lifespan Psychology, Max Planck Institute for Human Development, Lentzallee 94, 14195 Berlin, Germany
e-mail: kuehn@mpib-berlin.mpg.de

[†]These authors have contributed equally to this work.

Video games contain elaborate reinforcement and reward schedules that have the potential to maximize motivation. Neuroimaging studies suggest that video games might have an influence on the reward system. However, it is not clear whether reward-related properties represent a precondition, which biases an individual toward playing video games, or if these changes are the result of playing video games. Therefore, we conducted a longitudinal study to explore reward-related functional predictors in relation to video gaming experience as well as functional changes in the brain in response to video game training. Fifty healthy participants were randomly assigned to a video game training (TG) or control group (CG). Before and after training/control period, functional magnetic resonance imaging (fMRI) was conducted using a non-video game related reward task. At pretest, both groups showed strongest activation in ventral striatum (VS) during reward anticipation. At posttest, the TG showed very similar VS activity compared to pretest. In the CG, the VS activity was significantly attenuated. This longitudinal study revealed that video game training may preserve reward responsiveness in the VS in a retest situation over time. We suggest that video games are able to keep striatal responses to reward flexible, a mechanism which might be of critical value for applications such as therapeutic cognitive training.

Keywords: video gaming, training, reward anticipation, longitudinal, fMRI

INTRODUCTION

Over the last decades, the video gaming industry has grown into being one of the biggest multimedia industries in the world. Many people play video games on a day-to-day basis. For example in Germany 8 out of 10 people between 14 and 29 years of age reported to play video games, and 44% above age 29 still play video games. Taken together, based on survey data approximately more than 25 million people above 14 years of age (36%) play video games in Germany (Illek, 2013).

It seems as if human beings have a genuinely high motivation to play video games. Most frequently video games are played for the simple purpose of “fun” and a short-term increase in subjective well-being (Przybylski et al., 2010). Indeed, playing video games can satisfy different basic psychological needs, probably also dependent on the specific video game and its genre. Especially fulfillment of psychological needs like competence (sense of self-efficacy and acquisition of new skills), autonomy (personal goal-directed behavior in novel fictive environments), and relatedness (social interactions and comparisons) were associated with video gaming (Przybylski et al., 2010). Specifically, satisfaction of psychological needs might be mainly related to the various feedback mechanisms provided to the player by the game. This elaborate reinforcement and reward schedule has the potential to maximize motivation (Green and Bavelier, 2012).

Due to the high use, video games have come into the research focus of disciplines such as psychology and neuroscience. It has been shown that training with video games can lead to

improvement in cognitive performance (Green and Bavelier, 2003, 2012; Basak et al., 2008), and in health-related behavior (Baranowski et al., 2008; Primack et al., 2012). Further, it has been shown that video games can be used in the training of surgeons (Boyle et al., 2011), that they are associated with higher psychological quality of life in elderly participants (Allaire et al., 2013; Keogh et al., 2013), and that they can facilitate weight reduction (Staiano et al., 2013). Although it is known that video games are designed to be maximally rewarding by game developers, and video gamers achieve psychological benefits from the gaming, the underlying processes that account for psychological benefits are not fully understood. Green and Bavelier (2012) concluded from their research that beyond the improvements in cognitive performance, the “true effect of action video game playing may be to enhance the ability to learn new tasks.” In other words, the effects of video game training might not be limited to the trained game itself; it may foster learning across a variety of tasks or domains. In fact, video game players learned how to learn new tasks quickly and therefore outperform non-video game players at least in the domain of attentional control (Green and Bavelier, 2012).

The underlying neurobiological processes associated with video gaming have been investigated with different imaging techniques and experimental designs. A raclopride positron emission tomography (PET) study by Koepp et al. (1998) showed that video gaming (more specifically, a tank simulation) is associated with endogenous dopamine release in the ventral striatum (VS). Furthermore, the level of dopamine binding potential has been

related to performance in the game (Koepp et al., 1998). The VS is part of the dopaminergic pathways and is associated with reward processing and motivation (Knutson and Greer, 2008) as well as acquisition of learning in terms of prediction error signal (O'Doherty et al., 2004; Atallah et al., 2006; Erickson et al., 2010). Using magnetic resonance imaging (MRI) to measure gray matter volume, Erickson et al. (2010) showed that ventral and dorsal striatal volume could predict the early performance gains in a cognitively demanding video game (in particular, a two dimensional space shooter simulation). Additionally, Kühn et al. (2011) found that on the one hand frequent compared to infrequent video game playing was associated with higher structural gray matter volume and on the other hand was related to stronger functional activation during loss processing (Kühn et al., 2011). Further, striatal functional magnetic resonance imaging (fMRI) activity during actively playing or passively watching a video game (space shooter simulation, Erickson et al., 2010) or during completing a different non-video game related task (in particular an oddball task) predicted the subsequent training improvement (Vo et al., 2011). Taken together, these studies show that neural processes that are associated with video gaming are likely to be related to alterations of the neural processing in the VS, the core area of reward processing. Moreover, video gaming seems to be associated with structural and reward processing related functional changes in this area. However, it is not clear whether video game related structural and functional properties observed in earlier studies represent a *precondition*, which biases an individual toward playing video games or if these changes are the *result* of playing video games.

In summary, video games are quite popular and frequently used. One reason for that might be that video gaming may fulfill general human needs (Przybylski et al., 2010). Satisfied needs increase psychological well-being, which in turn is probably experienced as rewarding. Neuroimaging studies support this view by showing that video gaming is associated with alterations in the striatal reward system. Reward processing on the other hand is an essential mechanism for any human stimulus-response learning process. Green and Bavelier (2012) described video game training as a training for learning how to learn (learning of stimulus-response patterns is crucial to complete a video game successfully). We believe that video game training targets the striatal reward system (amongst other areas) and may lead to changes in reward processing. Therefore, in this study, we focus on striatal reward processing before and after video game training.

Here, we conducted a longitudinal study to be able to explore reward-related functional predictors in relation to performance and experience in the game as well as functional changes in the brain in response to video game training. We used a successful commercial video game, because commercial games are specifically designed to increase subjective well being (Ryan et al., 2006) and therefore game enjoyment and experienced reward during the game may be maximized. According to the prediction hypothesis, we expect that ventral striatal response in a reward task before video game training predicts performance as already shown in a previous study with a different task (Vo et al., 2011). Furthermore, we want to explore whether ventral striatal reward responsiveness

is related to experienced fun, desire, or frustration in the training group during the training episode. To investigate the effect of video game training, we conducted a second MRI scan after video game training had taken place. Based on the findings by Kühn et al. (2011) showing altered reward processing in frequent compared to infrequent video game players, we expected altered striatal reward signal during reward anticipation in participants that had received training compared to controls. If there are functional changes in the striatal reward system, these should be related to the effect of video game training. If not, the observed changes in the study by Kühn et al. (2011) may rather relate to a precondition of the frequent video game players.

MATERIALS AND METHODS

PARTICIPANTS

Fifty healthy young adults were recruited via newspaper and internet advertisements and randomly assigned to video game training group (TG) or control group (CG). Preferably, we recruited only participants that played little or no video games in the last 6 months. None of the participants reported to play video games more than 1 h per week in the last 6 months (on average 0.7 h per month, $SD = 1.97$) and never played the training game ["Super Mario 64 (DS)"] before. Furthermore, the participants were free of mental disorders (according to personal interview using Mini-International Neuropsychiatric Interview), right-handed, and suitable for the MRI scanning procedure. The study was approved by the local Ethics Committee of the Charité – Universitätsmedizin Berlin and written informed consent was obtained from all participants after participants were fully instructed on the procedures of the study. Data of anatomical gray matter maps of these participants have been previously published (Kühn et al., 2013).

TRAINING PROCEDURE

The TG ($n = 25$, mean age = 23.8 years, $SD = 3.9$ years, 18 females) was instructed to play "Super Mario 64 DS" on the "Nintendo Dual-Screen (DS) XXL" handheld console for at least 30 min per day over a period of 2 months. This extremely successful platformer game was chosen based on its high accessibility for video gaming naïve participants, as it offers a well-suited balance between reward delivery and difficulty and is popular among male and female participants. In the game, the player has to navigate through a complex 3D environment using buttons attached to the console used for movement, jumping, carrying, hitting, flying, stomping, reading, and character specific actions. Prior to the training, participants were instructed on general control and game mechanisms in a standardized way. During the training period, we offered different types of support (telephone, email, etc.) in case frustration or difficulties during game play arose.

The no-contact CG ($n = 25$, mean age = 23.4 years, $SD = 3.7$ years, 18 females) had no task in particular but underwent the same scanning procedure as the TG. All participants completed an fMRI scan at the beginning of the study (pretest) and 2 months after training or after a passive delay phase (posttest). The video game training for the TG began immediately after the pretest measurement and ended before posttest measurement.

QUESTIONNAIRES

During training, the participants of the TG were asked to record the amount of daily gaming time. Furthermore participants rated experienced fun, frustration and desire to play during video gaming on a 7-point Likert scale once a week in a word processing document (see, supplementary material for more details) and sent the electronic data files via email to the experimenters. The accomplished game-related reward (stars collected) was objectively assessed by checking the video gaming console after training period. The maximum absolute amount of stars was 150.

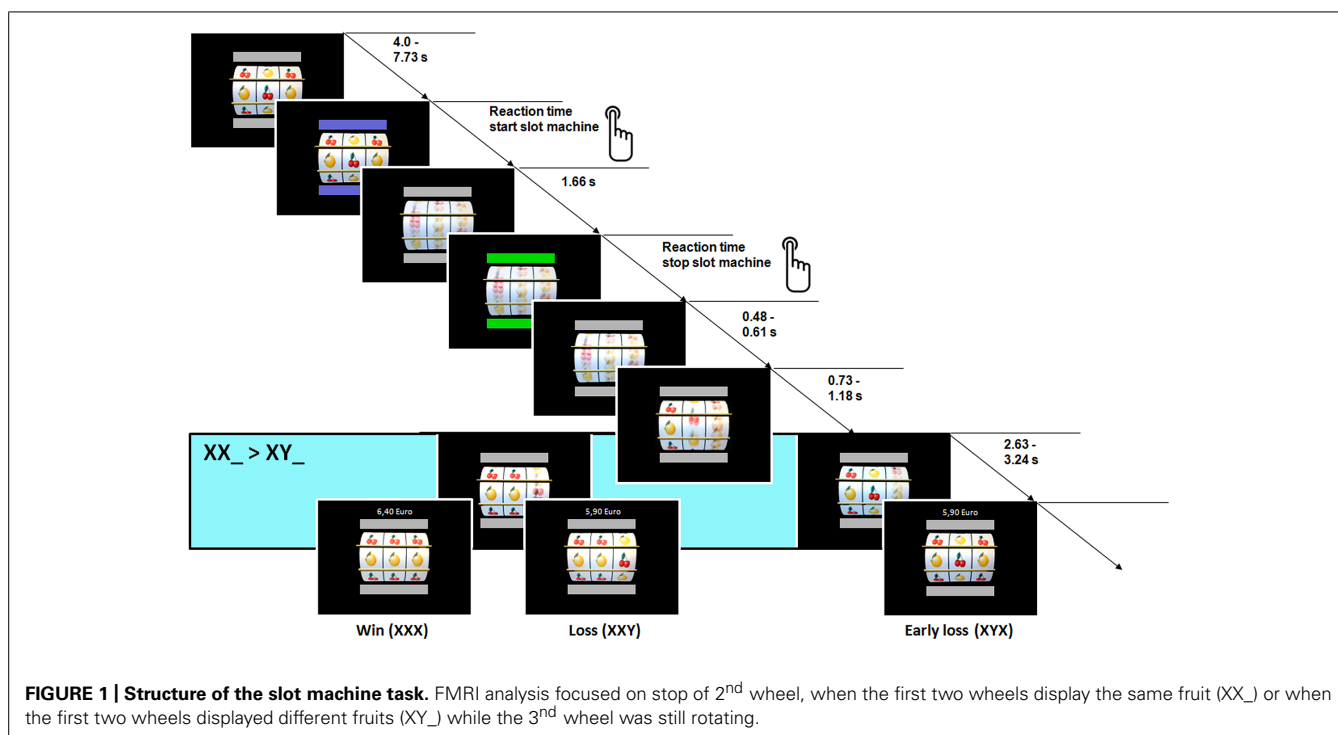
SLOT MACHINE PARADIGM

To investigate reward anticipation, a slightly modified slot machine paradigm was used that evoked strong striatal response (Lorenz et al., 2014). Participants had to go through the same slot machine paradigm before and after video game training procedure had taken place. The slot machine was programmed using Presentation software (Version 14.9, Neurobehavioral Systems Inc., Albany, CA, USA) and consisted of three wheels displaying two different sets of fruits (alternating fruit X and Y). At the two time points of measurement, a slot machine with cherries (X) and lemons (Y) or melons (X) and bananas (Y) were displayed in a counterbalanced fashion and equally distributed for the TG and CG. The color of two horizontal bars (above and below the slot machine) indicated the commands to start and stop the machine.

At the beginning of each trial, the wheels did not move and gray bars indicated the inactive state. When these bars turned blue (indicating the start of a trial), the participant was instructed to start the machine by pressing a button with the right hand. After a button press, the bars turned gray again (inactive state)

and the three wheels started to rotate vertically with different accelerations (exponential increasing from left to right wheel, respectively). When the maximum rotation velocity of the wheels was reached (1.66 s after button press) the color of the bars turned green. This color change indicated that the participant could stop the machine by pressing the button again. After another button press, the three wheels successively stopped rotating from the left to the right side. The left wheel stopped after a variable delay of 0.48 and 0.61 s after the button press, while the middle and right wheel were still rotating. The second wheel stopped after an additional variable delay of 0.73 and 1.18 s. The right wheel stopped rotating after the middle wheel with a variable delay of 2.63 and 3.24 s. The stop of the third wheel terminated the trial and a feedback about the current win and the total amount of reward was displayed on the screen. For the next trial, the button changed from gray to blue again and the next trial started after a variable delay that ranged between 4.0 and 7.73 s and was characterized by an exponential decreasing function (see Figure 1).

The experiment contained 60 trials in total. The slot machine was determined with a pseudo-randomized distribution of 20 win trials (XXX or YYY), 20 loss trials (XXY or YYX), and 20 early loss trials (XYX, YXY, XYY, or YXX). Participants started with an amount of 6.00 euro representing the wager of 0.10 euro per trial ($60 \text{ trials} \times 0.10 \text{ euro wager} = 6.00 \text{ euro wager}$) and gained 0.50 euro per trial, when all fruits in a row were of the same identity (XXX or YYY); if not, participants did not win (XXY, YYX, XYX, YXY, XYY, YYX) and the wager was subtracted from the total amount of money. Participants had no influence on winning or losing and the participants won the fixed amount of 10.00 euro ($0.50 \text{ euro gain} \times 20 \text{ win trials} = 10.00 \text{ euro gain}$) at the end of the



task. The participants were instructed to play the slot machine 60 times and that the aim in each trial is to get three fruit of the same kind in a row. Further, participants practiced the slot machine task before entering the scanner for 3–5 trials. No information was given that the task was a game of chance or any skill was involved.

SCANNING PROCEDURE

Magnetic resonance imaging scans were conducted on a three Tesla Siemens TIM Trio Scanner (Siemens Healthcare, Erlangen, Germany), equipped with a 12 channel phased array head coil. Via a video projector, the slot machine paradigm was visually presented via a mirror system mounted on top of the head coil. Functional images were recorded using axial aligned T2*-weighted gradient echo planar imaging (EPI) with the following parameters: 36 slices, interleaved ascending slice order, time to repeat (TR) = 2 s, time to echo (TE) = 30 ms, field of view (FoV) = 216×216 , flip angle = 80° , voxel size: $3 \text{ mm} \times 3 \text{ mm} \times 3.6 \text{ mm}$. For anatomical reference, 3D anatomical whole brain images were obtained by a three-dimensional T1-weighted magnetization prepared gradient-echo sequence (MPRAGE; TR = 2500 ms; TE = 4.77 ms; inversion time = 1100 ms, acquisition matrix = $256 \times 256 \times 176$, flip angle = 7° , voxel size: $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$).

DATA ANALYSIS

Image processing

Magnetic resonance imaging data was analyzed using Statistical Parametric Mapping software package (SPM8, Wellcome Department of Imaging Neuroscience, London, UK). EPIs were corrected for acquisition time delay and head motion and then transformed into the stereotactic normalized standard space of Montreal Neuroimaging Institute using the unified segmentation algorithm as implemented in SPM8. Finally, EPIs were resampled (voxel size = $3 \text{ mm} \times 3 \text{ mm} \times 3 \text{ mm}$) and spatially smoothed with a 3D Gaussian kernel of 7 mm full width at half maximum.

Statistical analysis

A two-stage mixed-effects general linear model (GLM) was conducted. On single subject level, the model contained the data of both fMRI measurements, which was realized by fitting the data in different sessions. This GLM included separate regressors per session for gain anticipation (XX_ and YY_) and no gain anticipation (XY_ and YX_) as well as the following regressors of no interest: gain (XXX and YYY), loss (XXY and YYX), early loss (XYX, XYY, YXY, and YXX), button presses (after bar changed to blue as well as green), visual flow (rotation of the wheels), and the six rigid body movement parameters. Differential contrast images for gain anticipation against no gain anticipation (XX_ vs. XY_) were calculated for pre- and posttest and taken to group level analysis. On the second level, these differential *T*-contrast images were entered into a flexible factorial analysis of variance (ANOVA) with the factors group (TG vs. CG) and time (pre- vs. posttest).

Whole brain effects were corrected for multiple comparisons using a Monte Carlo simulation based cluster size correction (AlphaSim, Song et al., 2011). One thousand Monte Carlo simulations revealed a corresponding alpha error probability of

$p < 0.05$, when using a minimum cluster size 16 adjacent voxels with a statistical threshold of $p < 0.001$. According to a meta-analysis by Knutson and Greer (2008), activation differences during reward anticipation were expected in the VS. Based on this a priori hypothesis, we further reported *post hoc* analysis within this brain area using a region of interest (ROI) analysis. To this end, we used a literature-based ROI for the VS (Schubert et al., 2008). These ROIs were created by combining previous functional findings regarding reward processing (predominantly monetary incentive delay task articles) with anatomical limits to gray matter brain tissue. Detailed information about the calculation of the VS ROI is described in supplementary material. Furthermore, we conducted a control analysis with the extracted mean parameters from the primary auditory cortex, because this region should be independent from the experimental manipulation in the reward task. Therefore we used an anatomical ROI of the Heschl's gyri as described in the Anatomical Labeling (AAL) brain atlas (Tzourio-Mazoyer et al., 2002).

RESULTS

PREDICTION-RELATED RESULTS (PRETEST)

Brain response during gain anticipation

At pretest, during the slot machine task in both groups, gain anticipation (against no gain anticipation) evoked activation in a fronto-striatal network including subcortical areas (bilateral VS, thalamus), prefrontal areas (supplementary motor area, precentral gyrus, and middle frontal gyrus, superior frontal gyrus), and insular cortex. Additionally, increased activation in the occipital, parietal and temporal lobes was observed. All brain regions showing significant differences are listed in supplementary Tables S1 (for TG) and S2 (for CG). Note that the strongest activation differences were observed in the VS in both groups (see **Table 1**; **Figure 2**). For the contrast TG > CG, a stronger activation in the right supplementary motor area [SMA, cluster size 20 voxel, $T(48) = 4.93$, MNI-coordinates [x y z] = 9, 23, 49] and for CG > TG a stronger activation in the right pallidum (cluster size 20 voxel, $T(48) = 5.66$, MNI-coordinates [x y z] = 27, 8, 7) were observed. Both regions are probably not associated to reward-related functions as shown in the meta-analysis by Liu et al. (2011) across 142 reward studies.

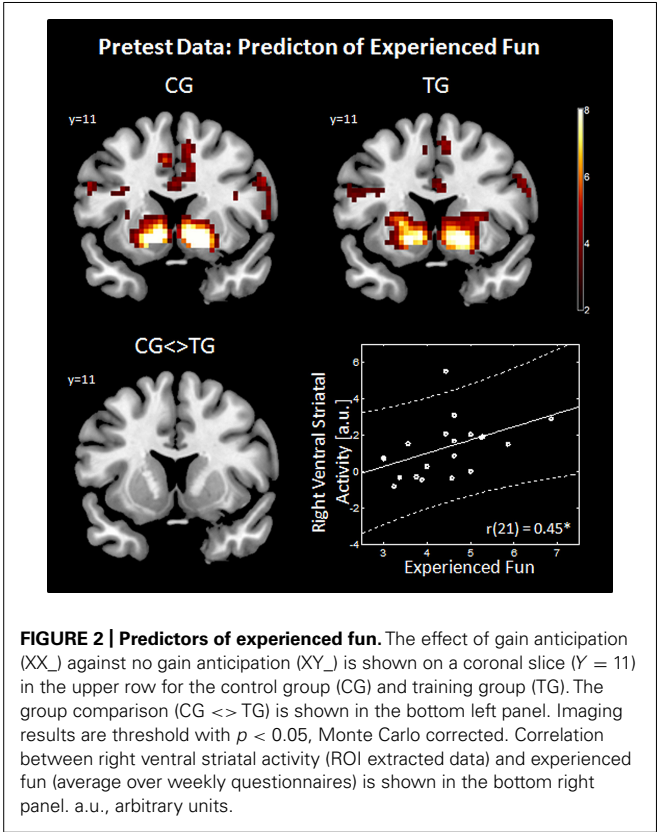
Association between ventral striatal activity and associated video gaming behavior

To test the hypothesis of the predictive properties of striatal reward signal toward video games, the ventral striatal signal was individually extracted using the literature-based ROI and correlated with questionnaire items as well as game success, which was assessed by checking the video gaming console. Due to a lack of compliance of participants, weekly questionnaire data of four participants was missing. Weekly questions about experienced fun ($M = 4.43$, $SD = 0.96$), frustration ($M = 3.8$, $SD = 1.03$) and video gaming desire ($M = 1.94$, $SD = 0.93$) were averaged across the 2 months. Participants collected 87 stars ($SD = 42.76$) on average during the training period.

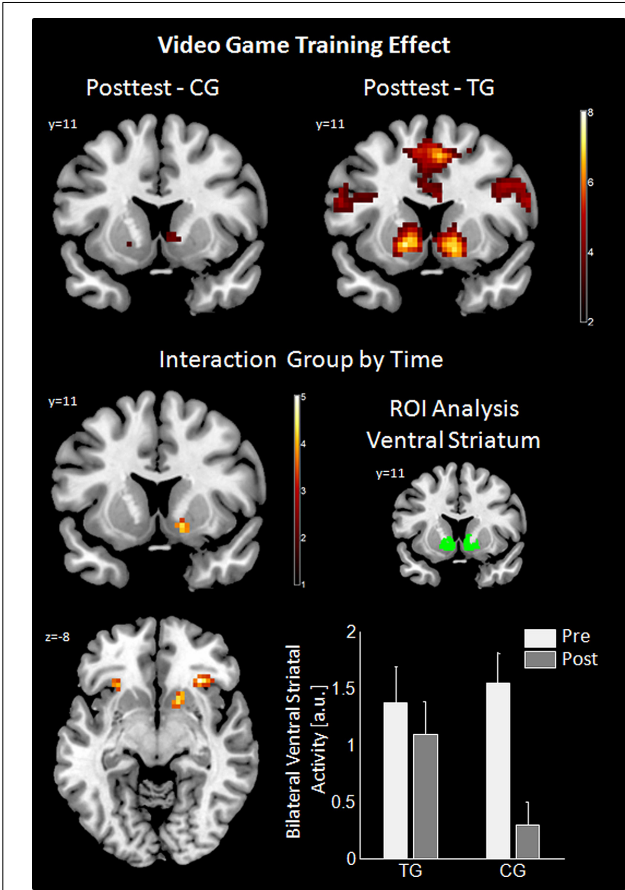
When applying Bonferroni correction to the calculated correlations (equal to a significance threshold of $p < 0.006$), none of the

Table 1 | Group by time interaction (TG: Post > Pre) > (CG: Post > Pre) of the effect of gain anticipation against no gain anticipation in the whole brain analysis using Monte Carlo corrected significance threshold of $p < 0.05$. TG, training group; CG, control group; H, hemisphere; MNI, Montreal Neurological Institute; L, left; R, right.

Brain structure	H	Cluster size (vox)	Z (peak)	MNI coord. (mm)		
				x	y	z
Supplementary motor area	R	36	5.32	9	20	49
Insula lobe/inferior frontal gyrus (p. orbitalis)	L	23	4.81	−24	23	−2
Precentral gyrus	R	22	4.58	39	5	31
Ventral striatum	R	22	4.27	15	11	−8
Insula lobe/inferior frontal gyrus (p. orbitalis)	R	20	5.10	30	26	−8



correlations were significant. Neither video gaming desire [left VS: $r(21) = 0.03$, $p = 0.886$; right VS: $r(21) = -0.12$, $p = 0.614$] nor frustration [left VS: $r(21) = -0.24$, $p = 0.293$; right VS: $r(21) = -0.325$, $p = 0.15$] nor accomplished game-related reward [left VS: $r(25) = -0.17$, $p = 0.423$; right VS: $r(25) = -0.09$, $p = 0.685$] were correlated with reward-related striatal activity. Interestingly, when using uncorrected significance threshold experienced fun during video gaming was correlated positively with the activity during gain anticipation in the right VS [$r(21) = 0.45$, $p = 0.039$] and a trend was observed in the left VS [$r(21) = 0.37$, $p = 0.103$] as shown in **Figure 2** (bottom right panel). However, when applying Bonferroni correction to this exploratory analysis, also the correlations between experienced fun and ventral striatal activity remained non-significant.



We further conducted a control analysis to investigate, whether this finding is specific for the VS. We correlated the same behavioral variables with the extracted parameter estimates of the Heschl's

gyri (primary auditory cortex). The analysis revealed no significant correlation (all p 's > 0.466).

EFFECT OF VIDEO GAME TRAINING (PRE- AND POSTTEST)

Analysis of gain anticipation against no gain anticipation during the slot machine task at posttest revealed activation differences in the TG in the same fronto-striatal network as observed at pretest (for details see Table S3). In the CG, this effect was similar, but attenuated (see Figure 3; Table S4). The interaction effect of group by time revealed a significant difference in reward-related areas (right VS and bilateral insula/inferior frontal gyrus, pars orbitalis) and motor-related areas (right SMA and right precentral gyrus) indicating a preserved VS activity in the TG between the time points, but not in the CG. *Post hoc* ROI analysis using the literature-based VS ROI confirmed the interaction result [Interaction group by time: $F(48,1) = 5.7$, $p = 0.021$]. ROI-analysis in the control region (Heschl's gyri) was non-significant. Additional t -tests revealed a significant difference between the time points within the CG group [$t(24) = 4.6$, $p < 0.001$] as well as a significant difference between the groups at posttest [$t(48) = 2.27$, $p = 0.028$]. Results for the interaction group by time are summarized in Table 1 and are illustrated in Figure 3.

DISCUSSION

The aim of the present study was twofold: We aimed at investigating how striatal reward responsiveness predicts video game related behavior and experience as well as the impact of video game training on functional aspects of the reward system. Regarding the prediction, we found a positive association between striatal reward signal at pretest and experienced fun during subsequent video game training. Regarding the effect of video gaming, a significant group by time interaction was observed driven by a decrease of the striatal reward signal in the CG.

STRIATAL REWARD RESPONSIVENESS AND ITS PREDICTIVE PROPERTIES FOR VIDEO GAMING EXPERIENCE

A relationship between striatal reward signal and game performance or experienced desire and frustration was not observed. However, we were able to demonstrate a positive association of the striatal reward signal with experienced fun during video game training. Thus, we believe that the magnitude of striatal activity during reward processing in a non-video gaming related reward task is predictive for experienced fun during game play. However, this finding has to be interpreted with caution, since the observed correlation did not remain significant after correction for multiple testing.

A possible explanation for the correlation between striatal reward signal and experienced fun during video gaming might be that the measured striatal reward signal during slot machine gambling reflects the individuals' reward responsiveness which may be associated with dopaminergic neurotransmission in the striatum. In accordance, previous studies showed that VS activity during reward anticipation is related to dopamine release in this region (Schott et al., 2008; Buckholz et al., 2010). It has further been shown that also video gaming was associated with dopamine release in the same area (Koepp et al., 1998). Thus, the VS seems to be crucially involved in neural reward processing

as well as video gaming, which involves many motivational and rewarding factors. Specifically, we are convinced that the observed relationship between VS activity and experienced fun might be related to a general responsiveness of the reward-related striatal dopamine system to hedonic stimuli. The VS has been associated with motivational and pleasure-elicited reactions in a recent review by Kringelbach and Berridge (2009). Thus, the observed association between ventral striatal activity and fun that refers to hedonic and pleasure-related experience during gaming seems well founded. Future studies should further investigate the relationship between striatal reward responsiveness and experienced fun during video gaming again to explore this relationship more deeply.

As mentioned above, striatal dopamine release (Koepp et al., 1998), volume (Erickson et al., 2010), and activity during gaming (Vo et al., 2011) were previously associated with video gaming performance. The results of the current study did not show an association between video gaming performance and VS activity. The achieved reward was operationalized by the number of accomplished missions/challenges in the game. Typical missions within the game are exemplified by defeating a boss, solving puzzles, finding secret places, racing an opponent, or gathering silver coins. These missions represent the progress in the game rather than the actual gaming performance. Thus, these variables may not be a sufficiently precise dependent variable of performance. We were, however, not able to collect more game-related variables, because "Super Mario 64 DS" is a commercial video game and a manipulation of this self-contained video game was impossible.

We further investigated the relationship between striatal reward signal and the experienced desire to play during video game training. Desire in this context is probably related to the need and expectations of video gaming's potential satisfaction and reward. Desire is not clearly separable from wanting, because it usually arises together with wanting. Neurobiologically, wanting involves not only striatal, but also prefrontal areas that are related to goal-directed behavior (Cardinal et al., 2002; Berridge et al., 2010). Therefore, a neural correlate of desire might not be limited to the striatal reward area. Indeed, Kühn et al. (2013) showed that structural gray matter volume changes in the dorsolateral prefrontal cortex induced by video game training are positively associated with the subjective feeling of desire during video game training. Thus, in the current study the striatal reward responsiveness might not be related to desire, because desire might rather be associated with prefrontal goal-directed neural correlates. Future studies may investigate this in detail.

We expected a negative correlation between striatal reward responsiveness and experienced frustration during video game training since the VS activity is decreased at the omission of reward relative to the receipt of reward (Abler et al., 2005). However, this relationship was not observed. Previous studies showed that the insula is selectively activated in the context of frustration (Abler et al., 2005; Yu et al., 2014). Thus, future studies might also investigate insular activity in the context of omitted reward.

EFFECT OF VIDEO GAME TRAINING ON THE REWARD SYSTEM

Kühn et al. (2011) showed in a cross-sectional study that frequent video game players (>9 h per week) demonstrated greater striatal

reward-related activity compared to infrequent video game players. However, the question remained, whether this finding was a predisposition toward or a result of video gaming. In our present longitudinal study, gain anticipation during slot machine task revealed VS activity, which was preserved in TG over the 2 months, but not in CG. We assume that the striatal reward signal might reflect the motivational engagement during the slot machine task, which was still high in the TG at the posttest. The participants of the TG might preserve the responsiveness in reward processing and motivational willingness to complete the slot machine task at the second time point in a similarly engaged state as during the first time. An explanation for that finding might be that the video game training has an influence on dopamine-related reward processing during gaming (Koepp et al., 1998). Our results support this view, as this effect might temporally not be limited to the gaming session, but rather might have an influence on general striatal reward responsiveness in rewarding situations not related to video games. Krügel and Berridge (2009) showed that activity in the VS might represent an amplifier function of reward, and thus, video games might preserve reward responsiveness during game play itself, and even in the context of other rewarding tasks through amplification of pleasure-related activity. Thus, the video game training might be considered as an intervention targeting the dopaminergic neurotransmitter system, which might be investigated in the future. There is evidence, that dopaminergic interventions in the context of pharmacological studies can have a therapeutic behavior changing character. A recent pharmacological study using a dopaminergic intervention on older healthy adults by Chowdhury et al. (2013) showed that age-related impaired striatal reward processing signal could be restored by dopamine targeted drugs. Future studies should investigate the potential therapeutic effects of video gaming training on cognitive demanding tasks involving dopaminergic striatal signal. It would be highly valuable to uncover the specific effect of video gaming in the fronto-striatal circuitry. Our findings suggested an effect on reward processing, which in turn is essential for shaping of goal-directed behavior and flexible adaption to volatile environments (Cools, 2008). Therefore, tasks involving reward-related decisions such as reversal learning should be investigated in future longitudinal studies in combination with video game training. Multiple pharmacological studies have shown that a dopaminergic manipulation may lead to an increase or decrease in reversal learning performance, which probably depends on task demand and individual baseline dopamine levels (Klanker et al., 2013).

The observed effect of video game training on the reward system was also driven by a decrease in striatal activity in the CG during posttest, which may in part be explained by a motivational decline in the willingness to complete the slot machine task at the re-test. A study by Shao et al. (2013) demonstrated that even a single training session with a slot machine task before the actual scanning session led to decreases in striatal reward activity during win processing compared to a group that did not undergo a training session. A further study by Fließbach et al. (2010) investigated the re-test reliability of three reward tasks and showed that the re-test reliability in VS during gain anticipation were rather poor, in contrast to motor-related reliabilities in primary motor

cortex that were characterized as good. A possible explanation of these findings might be the nature of such reward tasks. The identical reward at both time points may not lead to the same reward signal at the second time of task performance, because the subjective reward feeling may be attenuated by a lack of novelty.

Obviously, in the present study the re-test was completed by both groups, but the decrease of the striatal reward activity was only observed in the CG, not in the TG. This preservation result in the TG may in part be related to the video game training as discussed above. Nevertheless, the CG was a no-contact group and did not complete an active control condition and thus, the findings might also represent a purely placebo like effect in the TG. However, even if not the specific video game training itself was the main reason for the preserved striatal response, our study may be interpreted as evidence arguing that video games lead to a rather strong placebo-like effect in a therapeutic or training-based setting. If video games would represent a stronger placebo effect than placebo medication or other placebo-like tasks is an open question. Moreover, during the scanning session itself participants were in the same situation in the scanner and one can expect that both groups produce the same social desirability effects. Still, the preservation effect should be interpreted very carefully, because placebo effect might confound the result (Boot et al., 2011). Future studies focusing on the reward system should include an active control condition in the study design.

Another possible limitation of the study might be that we did not control the video gaming behavior of the CG. We instructed the participants of the CG not to change their video gaming behavior in the waiting period and not to play Super Mario 64 (DS). However, video gaming behavior in the CG might have changed and could have affected the results. Future studies should include active control groups and assess video gaming behavior during the study period in detail.

In this study we focused on the VS. Nonetheless, we observed a significant training-related effect also in the insular cortices, SMA, and precentral gyrus. A recent meta-analysis by Liu et al. (2011) including 142 reward studies showed that besides the “core area of reward” VS also insula, ventromedial prefrontal cortex, anterior cingulate cortex, dorsolateral prefrontal cortex, and inferior parietal lobule are part of the reward network during reward anticipation. The insula is involved in the subjective integration of affective information, for instance during error-based learning in the context of emotional arousal and awareness (Craig, 2009; Singer et al., 2009). The activation during reward anticipation in the slot machine task may reflect subjective arousal and motivational involvement in the task. We believe that this significant training effect in the insula might – similar to the effect in VS – represent a motivational engagement, which was preserved in the TG at the posttest. Future studies could test this e.g., by applying arousal rating scales and correlate these values with insular activity. According to the differences in SMA and precentral gyrus, we want to highlight that these areas might not be involved in reward anticipation as it is not part of the suggested network of the mentioned meta-analysis (Liu et al., 2011). Instead, the SMA is involved in learning of motor-related

stimulus-response associations among other functions (Nachev et al., 2008). With regard to the current study, SMA activity may reflect an updating process of the stimulus (slot machine with three rotating wheels) – response (button press to stop the slot machine) – consequence (here update of stop of the second wheel: XX_ and XY_) – chain. Speculatively, participants of the training group understand the slot machine after training as a video game, in which they could improve their performance by e.g., pressing the button at the right time point. In other words, the participants of the TG might have thought that they could impact the outcome of the slot machine by adapting their response pattern. Please note that the participants were not aware that the slot machine had a deterministic nature. As the precentral gyrus is also part of the motor system, the interpretation of the functional meaning of the SMA finding may be also valid for the precentral gyrus. Future studies might confirm these interpretations of SMA and precentral activation differences by systematically varying response-consequence-associations.

VIDEO GAMING, SUPER MARIO, MOTIVATION, SUBJECTIVE WELL-BEING, AND THE REWARD SYSTEM

From a psychological view, joyful video games provide highly effective reward schedules, perfectly adjusted difficulty levels and strong engagement (Green and Bavelier, 2012). These specific properties potentially contain the opportunity to satisfy basic psychological needs such as competence, autonomy and relatedness (Przybylski et al., 2010). A study by Ryan et al. (2006) showed that participants feeling volitionally motivated by a 20 min training session of Super Mario 64 had an increased well-being after playing. This increased well-being was further associated with increases in the feeling of competence (e.g., experienced self-efficacy) and autonomy (e.g., acting based on interest). Together with the current finding of the preservation of the reward signal in a non-trained task, we believe that video games harbor the potential of a powerful tool for specific (cognitive) training. Depending on the video gaming genre and individual properties of the game, video games demand very complex cognitive and motor interactions from players to be able to reach the goal of the game and thus a specific training effect. The rewarding nature of video games may lead to a constant high motivational level within the training session.

CONCLUSION

The current study showed that striatal reward responsiveness predicts the subsequent experienced video gaming fun suggesting that individual differences in reward responsiveness might affect motivational engagement of video gaming, but this interpretation needs confirmation in future studies. Furthermore, this longitudinal study revealed that video game training may preserve reward responsiveness in the VS in a re-test. We believe that video games are able to keep striatal responses to reward flexible, a mechanism which might be extremely important to keep motivation high, and thus might be of critical value for many different applications, including cognitive training and therapeutic possibilities. Future research should therefore investigate whether video game training might have an effect on reward-based decision-making, which is an important ability in everyday life.

ACKNOWLEDGMENTS

This study was supported by the German Ministry for Education and Research (BMBF 01GQ0914), the German Research Foundation (DFG GA707/6-1), and German National Academic Foundation grant to RCL. We are grateful for the assistance of Sonali Beckmann operating the scanner as well as David Steiniger and Kim-John Schlüter for testing the participants.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnhum.2015.00040/abstract>

REFERENCES

- Abler, B., Walter, H., and Erk, S. (2005). Neural correlates of frustration. *Neuroreport* 16, 669–672. doi: 10.1097/00001756-200505120-00003
- Allaire, J. C., McLaughlin, A. C., Trujillo, A., Whitlock, L. A., LaPorte, L., and Gandy, M. (2013). Successful aging through digital games: socioemotional differences between older adult gamers and non-gamers. *Comput. Hum. Behav.* 29, 1302–1306. doi: 10.1016/j.chb.2013.01.014
- Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., and O'Reilly, R. C. (2006). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.* 10, 126–131. doi: 10.1038/nn1817
- Baranowski, T., Buday, R., Thompson, D. I., and Baranowski, J. (2008). Playing for real: video games and stories for health-related behavior change. *Am. J. Prev. Med.* 34, 74–82.e10. doi: 10.1016/j.amepre.2007.09.027
- Basak, C., Boot, W. R., Voss, M. W., and Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychol. Aging* 23, 765–777. doi: 10.1037/a0013494
- Berridge, K. C., Ho, C.-Y., Richard, J. M., and Di Feliceantonio, A. G. (2010). The tempted brain eats: pleasure and desire circuits in obesity and eating disorders. *Brain Res.* 1350, 43–64. doi: 10.1016/j.brainres.2010.04.003
- Boot, W. R., Blakely, D. P., and Simons, D. J. (2011). Do action video games improve perception and cognition? *Front. Psychol.* 2:226. doi: 10.3389/fpsyg.2011.00226
- Boyle, E., Kennedy, A.-M., Traynor, O., and Hill, A. D. K. (2011). Training surgical skills using nonsurgical tasks—can Nintendo Wii™ improve surgical performance? *J. Surg. Educ.* 68, 148–154. doi: 10.1016/j.jsurg.2010.11.005
- Buckholz, J. W., Treadway, M. T., Cowan, R. L., Woodward, N. D., Benning, S. D., Li, R., et al. (2010). Mesolimbic dopamine reward system hypersensitivity in individuals with psychopathic traits. *Nat. Neurosci.* 13, 419–421. doi: 10.1038/nn.2510
- Cardinal, R. N., Parkinson, J. A., Hall, J., and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* 26, 321–352. doi: 10.1016/S0149-7634(02)00007-6
- Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Düzel, E., et al. (2013). Dopamine restores reward prediction errors in old age. *Nat. Neurosci.* 16, 648–653. doi: 10.1038/nn.3364
- Cools, R. (2008). Role of dopamine in the motivational and cognitive control of behavior. *Neuroscientist* 14, 381–395. doi: 10.1177/1073858408317009
- Craig, A. D. (2009). How do you feel — now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Erickson, K. I., Boot, W. R., Basak, C., Neider, M. B., Prakash, R. S., Voss, M. W., et al. (2010). Striatal volume predicts level of video game skill acquisition. *Cereb. Cortex* 20, 2522–2530. doi: 10.1093/cercor/bhp293
- Fliessbach, K., Rohe, T., Linder, N. S., Trautner, P., Elger, C. E., and Weber, B. (2010). Retest reliability of reward-related BOLD signals. *Neuroimage* 50, 1168–1176. doi: 10.1016/j.neuroimage.2010.01.036
- Green, C. S., and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature* 423, 534–537. doi: 10.1038/nature01647
- Green, C. S., and Bavelier, D. (2012). Learning, attentional control, and action video games. *Curr. Biol.* 22, R197–R206. doi: 10.1016/j.cub.2012.02.012
- Illek, C. (2013). Für die junge Generation gehören Computerspiele zum Alltag – BITKOM. Available at: http://www.bitkom.org/77030_77024.aspx [accessed August 21, 2013].

- Keogh, J. W. L., Power, N., Wooller, L., Lucas, P., and Whatman, C. (2013). Physical and psychosocial function in residential aged care elders: effect of Nintendo Wii sports games. *J. Aging Phys. Act.* 22, 235–44. doi: 10.1123/JAPA.2012-0272
- Klanker, M., Feenstra, M., and Denys, D. (2013). Dopaminergic control of cognitive flexibility in humans and animals. *Front. Neurosci.* 7:201. doi: 10.3389/fnins.2013.00201
- Knutson, B., and Greer, S. M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 3771–3786. doi: 10.1098/rstb.2008.0155
- Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., et al. (1998). Evidence for striatal dopamine release during a video game. *Nature* 393, 266–268. doi: 10.1038/30498
- Kringelbach, M. L., and Berridge, K. C. (2009). Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn. Sci.* 13, 479–487. doi: 10.1016/j.tics.2009.08.006
- Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., and Gallinat, J. (2013). Playing Super Mario induces structural brain plasticity: gray matter changes resulting from training with a commercial video game. *Mol. Psychiatry* 19, 265–271. doi: 10.1038/mp.2013.120
- Kühn, S., Romanowski, A., Schilling, C., Lorenz, R., Mörsen, C., Seiferth, N., et al. (2011). The neural basis of video gaming. *Transl. Psychiatry* 1:e53. doi: 10.1038/tp.2011.53
- Liu, X., Hairston, J., Schrier, M., and Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 35, 1219–1236. doi: 10.1016/j.neubiorev.2010.12.012
- Lorenz, R. C., Gleich, T., Beck, A., Pöhlend, L., Raufelder, D., Sommer, W., et al. (2014). Reward anticipation in the adolescent and aging brain. *Hum. Brain Mapp.* 35, 5153–5165. doi: 10.1002/hbm.22540
- Nachev, P., Kennard, C., and Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869. doi: 10.1038/nrn2478
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454. doi: 10.1126/science.1094285
- Primack, B. A., Carroll, M. V., McNamara, M., Klem, M. L., King, B., Rich, M., et al. (2012). Role of video games in improving health-related outcomes: a systematic review. *Am. J. Prev. Med.* 42, 630–638. doi: 10.1016/j.amepre.2012.02.023
- Przybylski, A. K., Scott, C., and Ryan, R. M. (2010). A motivational model of video game engagement. *Rev. Gen. Psychol.* 14, 154–166. doi: 10.1037/a0019440
- Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The motivational pull of video games: a self-determination theory approach. *Motiv. Emot.* 30, 344–360. doi: 10.1007/s11031-006-9051-8
- Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., et al. (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *J. Neurosci.* 28, 14311–14319. doi: 10.1523/JNEUROSCI.2058-08.2008
- Schubert, R., Ritter, P., Wüstenberg, T., Preuschhof, C., Curio, G., Sommer, W., et al. (2008). Spatial attention related SEP amplitude modulations covary with BOLD signal in S1—A simultaneous EEG—fMRI study. *Cereb. Cortex* 18, 2686–2700. doi: 10.1093/cercor/bhn029
- Shao, R., Read, J., Behrens, T. E. J., and Rogers, R. D. (2013). Shifts in reinforcement signalling while playing slot-machines as a function of prior experience and impulsivity. *Transl. Psychiatry* 3:e235. doi: 10.1038/tp.2013.10
- Singer, T., Critchley, H. D., and Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13, 334–340. doi: 10.1016/j.tics.2009.05.001
- Song, X.-W., Dong, Z.-Y., Long, X.-Y., Li, S.-F., Zuo, X.-N., Zhu, C.-Z., et al. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS ONE* 6:e25031. doi: 10.1371/journal.pone.0025031
- Staiano, A. E., Abraham, A. A., and Calvert, S. L. (2013). Adolescent exergame play for weight loss and psychosocial improvement: a controlled physical activity intervention. *Obesity (Silver Spring)* 21, 598–601. doi: 10.1002/oby.20282
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Vo, L. T. K., Walther, D. B., Kramer, A. F., Erickson, K. I., Boot, W. R., Voss, M. W., et al. (2011). Predicting Individuals' Learning Success from patterns of pre-learning MRI Activity. *PLoS ONE* 6:e16093. doi: 10.1371/journal.pone.0016093
- Yu, R., Mobbs, D., Seymour, B., Rowe, J. B., and Calder, A. J. (2014). The neural signature of escalating frustration in humans. *Cortex* 54, 165–178. doi: 10.1016/j.cortex.2014.02.013

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 November 2013; accepted: 15 January 2015; published online: 05 February 2015.

Citation: Lorenz RC, Gleich T, Gallinat J and Kühn S (2015) Video game training and the reward system. *Front. Hum. Neurosci.* 9:40. doi: 10.3389/fnhum.2015.00040

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2015 Lorenz, Gleich, Gallinat and Kühn. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The impact of binaural beats on creativity

Susan A. Reedijk, Anne Bolders and Bernhard Hommel*

Institute for Psychological Research and Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands

Edited by:

Michelle W. Voss, University of Iowa, USA

Reviewed by:

Kyle E. Mathewson, University of Illinois, USA

Hyunkyu Lee, Brain Plasticity Institute, USA

*Correspondence:

Bernhard Hommel, Cognitive Psychology Unit, Institute for Psychological Research and Leiden Institute for Brain and Cognition, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, Netherlands
e-mail: hommel@fsw.leidenuniv.nl

Human creativity relies on a multitude of cognitive processes, some of which are influenced by the neurotransmitter dopamine. This suggests that creativity could be enhanced by interventions that either modulate the production or transmission of dopamine directly, or affect dopamine-driven processes. In the current study we hypothesized that creativity can be influenced by means of binaural beats, an auditory illusion that is considered a form of cognitive entrainment that operates through stimulating neuronal phase locking. We aimed to investigate whether binaural beats affect creative performance at all, whether they affect divergent thinking, convergent thinking, or both, and whether possible effects may be mediated by the individual striatal dopamine level. Binaural beats were presented at alpha and gamma frequency. Participants completed a divergent and a convergent thinking task to assess two important functions of creativity, and filled out the Positive And Negative Affect Scale—mood State questionnaire (PANAS-S) and an affect grid to measure current mood. Dopamine levels in the striatum were estimated using spontaneous eye blink rates (EBRs). Results showed that binaural beats, regardless of the presented frequency, can affect divergent but not convergent thinking. Individuals with low EBRs mostly benefitted from alpha binaural beat stimulation, while individuals with high EBRs were unaffected or even impaired by both alpha and gamma binaural beats. This suggests that binaural beats, and possibly other forms of cognitive entrainment, are not suited for a one-size-fits-all approach, and that individual cognitive-control systems need to be taken into account when studying cognitive enhancement methods.

Keywords: creativity, binaural beats, gamma, alpha, cognitive enhancement

INTRODUCTION

Creativity is an important skill in the human cognitive repertoire, it is useful in art and science and essential in day-to-day life. Unfortunately, however, research into creativity is rather cluttered and mechanistic models about how creativity might work are not available (Dietrich and Kanso, 2010). It is thus not surprising that there is no single, widely accepted definition of creativity. What can be said, though, is that many cognitive processes seem to be involved, and that sub-functions underlying creativity depend on both state (Baas et al., 2008; Davis, 2009) and trait (Akbari Chermahini and Hommel, 2010) characteristics. Of all the processes involved in creativity, Guilford (1950, 1967) identifies divergent and convergent thinking as its two main ingredients. Together with insight (a possible sub-component of convergent thinking; see Bowden et al., 2005), these are nowadays still considered the most important processes in creativity (Dietrich and Kanso, 2010). Accordingly, it was these two processes that we considered in the present study.

Both divergent and convergent thinking have been assumed to be influenced by positive mood (e.g., Baas et al., 2008; Davis, 2009), but the mechanism underlying this impact remains unclear. Based on the observation that schizophrenic patients, who suffer from an overdose of the neurotransmitter dopamine (for a review, see Davis et al., 1991), sometimes exhibit

extraordinary creative performances (Keefe and Magaro, 1980; Nelson and Rawlings, 2008), some authors have assumed a strong link between creativity and dopamine (Eysenck, 1993). Indeed, positive-going mood is accompanied by phasic changes in the production and availability of dopamine in the mesolimbic and nigrostriatal systems of the brain (Ashby et al., 1999), which again is likely to facilitate cognitive search operations and related processes underlying creative behavior (Akbari Chermahini and Hommel, 2010; Hommel, 2012). If so, factors or techniques that are likely to modulate dopamine production or transmission could be suspected to have an impact on cognitive operations underlying creativity.

One phenomenon that has been suspected to propagate creativity is known under the name of “binaural beats”, an auditory illusion that can be considered a kind of cognitive or neural entrainment (Vernon, 2009; Turow and Lane, 2011). This phenomenon has encouraged sweeping claims about mind enhancement, and some websites even went as far as calling the illusion a “digital drug”. While binaural beats indeed seem to exert some effect on cognitive functioning and mood (Lane et al., 1998), and on neural firing patterns in the brain (Kuwada et al., 1979; Karino et al., 2006; Pratt et al., 2009; but see Vernon et al., 2012), it is as yet unclear how they do so. The binaural-beat illusion arises when two tones of a slightly different frequency are each presented to

different ears. For instance, when a tone of 335 Hz is presented to the right ear and a tone of 345 Hz to the left ear, this results in a subjectively perceived binaural beat of 10 Hz. Hence, instead of hearing two different tones, most individuals will hear just one tone that fluctuates in frequency or loudness: a beat (Oster, 1973).

How exactly the brain produces the perception of these beats is unclear, but the reticular activation system and the inferior colliculus seem to play a role (Kuwada et al., 1979; McAlpine et al., 1996; Turow and Lane, 2011). In animals, binaural-beat producing stimulus conditions have been shown to produce particular neural patterns of phase locking, or synchronization, beginning in the auditory system and propagating to the inferior colliculus (Kuwada et al., 1979; McAlpine et al., 1996). Even though the neural response to objectively presented beats is stronger, binaural beats seem to elicit similar neural responses in both humans and animals (Kuwada et al., 1979; McAlpine et al., 1996; Schwarz and Taylor, 2005; Karino et al., 2006), suggesting that the illusion arises through pathways normally associated with binaural sound detection (Kuwada et al., 1979; Pratt et al., 2010). As in humans binaural beats have been found to affect cognitive functioning and mood (Lane et al., 1998; Vernon, 2009), and responses to binaural beats are detectable in the human EEG (Schwarz and Taylor, 2005; Pratt et al., 2009), it can be assumed that neuronal phase locking spreads from the auditory system and the inferior colliculus over the cortex. A spreading pattern of neuronal activation and synchronization might affect short- and long-distance communication in the brain, processes which depend on neuronal synchronization and, presumably, on particular neurotransmitter systems (Schnitzler and Gross, 2005), thus affecting cognitive processing.

If binaural beats affect cognition through neural synchronization, it is possible that the frequency of the beat matters. For instance, short-range communication within brain areas is often associated with neural synchronization in the gamma frequency, while long-range communication is associated with neuronal phase locking in the slower frequency bands (von Stein and Sarnthein, 2000; Schnitzler and Gross, 2005). Moreover, a variety of frequency bands have been considered to represent the “messenger frequency” of cognitive-control signals. For instance, synchronization in the gamma frequency range seems to play a role in the top-down control of memory retrieval (Keizer et al., 2009), which should be relevant for many creativity tasks. Also of interest, phase locking in the alpha band has been associated with lower cortical arousal in general (Fink and Neubauer, 2006) and enhanced top-down control in creativity-related performance in particular (von Stein and Sarnthein, 2000; Fink et al., 2009). Especially divergent thinking seems to be associated with alpha wave synchronization (Fink et al., 2006, 2009). It could therefore be reasoned that inducing a state of lower cortical arousal by presenting people with alpha frequency binaural beats temporarily increases their performance on a divergent thinking task. Given that the available evidence highlights the alpha and gamma bands as possible messenger frequencies of control signals in creativity-related tasks, we investigated whether binaural beats presented at these two frequencies might affect performance in convergent- and divergent-thinking tasks—as compared to a control condition.

Performance in creativity tasks does not only depend on current states but is also affected by trait variables. As suggested by Eysenck (1993) and Ashby et al. (1999), creative performance seems to depend on an individual's basic supply of (striatal) dopamine. This suggestion fits with recent ideas about the interaction of frontal and striatal dopaminergic pathways in generating cognitive control. According to Cools and d'Esposito (2009), the frontal dopaminergic pathway (originating in the Ventral Tegmental Area) supports focusing on the current task while the striatal pathway (originating in the Substantia Nigra) facilitates the mental flexibility and switching between mental representations. Considering that this latter ability is particularly relevant for divergent thinking, it is not surprising that divergent thinking, but not convergent thinking, was found to be related to the spontaneous eye-blink rate (EBR; Akbari Chermahini and Hommel, 2010)—a clinical marker of striatal dopaminergic functioning (Karson, 1983; Shukla, 1985; Taylor et al., 1999).

Importantly for our study, markers of the individual striatal dopamine level (EBR) do not only predict individual performance in a divergent-thinking task, but also whether and how individuals are affected by state variables. Only recently, Akbari Chermahini and Hommel (2012) demonstrated that the creativity-enhancing effect of positive mood was restricted to individuals with low EBRs, i.e., low striatal dopamine levels. Indeed, tonic and phasic effects of neurotransmitters have often been assumed to interact in nonlinear fashions, in such a way that phasic changes can be more easily detected or are otherwise more effective if combined with a relatively low tonic baseline (e.g., Grace, 1991; Cohen et al., 2002). If so, and assuming EBRs reflect a fairly stable baseline level of tonic and phasic dopamine activity in the striatum, the hypothetical creativity-enhancing impact of binaural beats would be expected to be visible mainly in individuals with relatively low EBRs. We tested this hypothesis by analyzing performance in convergent- and divergent-thinking tasks, and beat-induced changes therein, as a function of low versus high EBR.

METHODS

Twenty-four first-year psychology or educational studies students (22 female, 2 male; 17–25 years) of Leiden University participated in exchange for course credit and/or pay. All participants had normal or corrected-to-normal sight and normal hearing, and no history of epileptic attacks or other neuropsychological illnesses. All participants were tested between 1 pm and 7 pm, and for each participant all sessions took place at the same time of the day. This was done to reduce variation due to normal daily fluctuations in EBR, mood, and related variables.

After the study procedure was explained to them by the experimenter, written informed consent was obtained from all participants. In the case of one underage participant, written informed consent was also obtained from the parents/caretakers. Participants were not made aware of the goal of the study beforehand, but all were debriefed after completing all sessions. The study was approved by the Leiden University Ethics Committee of the Institute of Psychology.

Participants came in for three sessions: one in which they were exposed to alpha frequency (10 Hz) binaural beat stimuli (the

Alpha condition), one in which they were exposed to gamma frequency (40 Hz) binaural beat stimuli (the Gamma condition), and one in which they listened to a constant tone of 340 Hz (the Control condition). The order of these three conditions and the two creativity tasks was counterbalanced across participants by means of a Latin square design. In every session a participant would complete the same tasks in the same order but with different items (see below) to avoid learning effects. The order of the items within each session was the same for all participants. Before starting the tasks, spontaneous EBRs were measured, and participants listened to a 3 min sound file (inducing the binaural beats, or the control sound) during which they did not complete any task. While the sound file continued to play, participants then carried out the creativity tasks using pencil and paper. At the beginning and end of the session participants' positive (PA) and negative affect (NA) was measured using the Positive And Negative Affect Scale—mood State questionnaire (PANAS-S), to assess possible mood changes over the session. To track possible changes in mood valence and arousal during the session, participants were also asked to rate their current mood on an affect grid immediately after completing each task. Both of the mood measures were completed on a computer. The task order for every session can be seen in **Figure 1**.

AUDITORY STIMULI

Auditory stimulation was presented through in-ear headphones (Etymotic Research ER-4B microPro), which provide 35 dB external noise attenuation. All sound files (44 kHz, 16 bit) were digitally generated in Audacity and played from the test computer using E-Prime 2. Sound levels at output were calculated from the voltages delivered at the earphone input as measured with an oscilloscope (Type Tektronix TDS2002) and the earphone efficiency as provided by the earphone manufacturer (180 dB sound pressure level (SPL) for 1 Vrms in a Zwislocki coupler, ER-4 datasheet, Etymotic Research, 1992). As beats are best perceived with a carrier frequency between 300 and 600 Hz (Licklider et al., 1950; Oster, 1973), both binaural beat conditions (10 Hz and 40 Hz) were based upon a 340 Hz carrier frequency. This 340 Hz carrier tone was presented to both ears in the control condition. The alpha frequency (10 Hz) beat was generated by presenting a tone of 335 Hz to the left ear and a tone of 345 Hz to the right ear, while the gamma frequency (40 Hz) beat was generated by presenting a tone of 320 Hz to the left ear and a tone of 360 Hz to the right ear. In all conditions, white noise (20

Hz–10 kHz band filtered) was added to the signal in both ears to enhance the clarity of the beats (Oster, 1973).

EYE BLINK RATE (EBR)

Participants' spontaneous EBRs were measured for 5 min at the start of each session using a BioSemi ActiveTwo system (BioSemi Inc., Amsterdam). During measurement of the blinks participants were not presented with any auditory stimuli. Spontaneous EBR was measured using six Ag/AgCL electrodes: two placed next to the outer canthus of each eye (measuring saccades), and two placed above and below the right eye (measuring the blink). Two electrodes placed on the mastoids served as a linked online reference. Participants were instructed to relax and look (but not stare) straight ahead at a paper with a fixation cross that was taped on the computer monitor. This monitor was turned off during EBR measurement. As EBR is stable during the day but goes up in the evening (after 8.30 pm; Barbato et al., 2000), participants were never tested after 7 pm. Blinks were identified automatically, and then manually checked for errors (such as noise segments wrongly identified as blink) in BrainVision Analyzer. Individual EBR was calculated by dividing the total amount of blinks during the 5 min measurement period by 5.

DIVERGENT THINKING: ALTERNATE USES TASK (AUT)

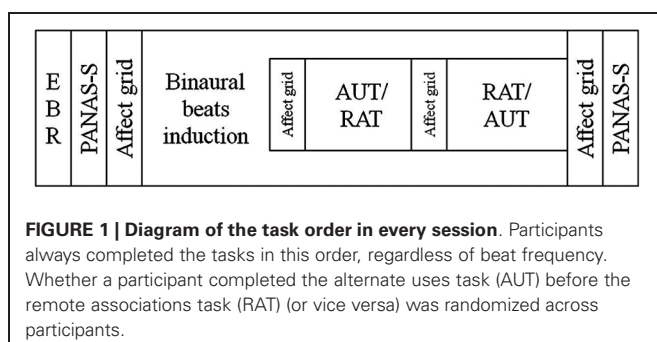
In this task, participants were to name as many uses for certain common household objects as possible. This task was scored on four components: originality, fluency, flexibility, and elaboration. However, as flexibility is most strongly and reliably connected to EBR scores (Akbari Chermahini and Hommel, 2010), we focused on this score, which reflects the number of different categories a participant uses in his or her answer for each item. For example, folding a hat of a paper or using it for origami counts as one category (folding), whereas writing a note on it counts as another (writing). We used a Dutch version of this task, which consisted of six items: *brick, shoe, paper, pen, bottle, and towel* (*baksteen, schoen, krant, pen, fles, and handdoek*, respectively). Per session, participants were given two items to solve in 10 min.

CONVERGENT THINKING: REMOTE ASSOCIATIONS TASK (RAT)

In this task, participants were presented with three seemingly unrelated words (e.g., “market”, “star” and “hero”) for which they had to find a single compound word that could be associated with all three of these words (in this case “super”; creating the words “supermarket”, “superstar” and “superhero”). We used the Dutch version of this task, which consists of a total of 30 items (Cronbach's alpha = .85; Akbari Chermahini et al., 2012). As our experiment consisted of three sessions per participant, we divided this task into three versions of 10 items each (Cronbach's alphas = .70, .67, and .70), matched by the items' discrimination value as reported in Akbari Chermahini et al. (2012). Participants were given 4 min to complete the 10 items.

POSITIVE AND NEGATIVE AFFECT SCHEDULE MOOD STATE QUESTIONNAIRE (PANAS-S)

This self-report mood scale consists of 20 items that provide a general measure of current mood in terms of PA and NA. Participants were given 10 positive (for instance, “interested” or



“alert”) and 10 negative (for instance, “upset” or “guilty”) words, and had to indicate how applicable a word was to their current mood on a Likert scale between 1 (very little or not at all) and 5 (very or extremely). The PANAS-S was completed on a computer, where participants used the mouse to select an option on the Likert scale.

AFFECT GRID

Participants indicated their current pleasure and arousal level by means of a computer mouse, which served to place a single cross in an arousal \times pleasure affect grid (Russell et al., 1989) presented on a computer monitor. The 9×9 grid was composed of a horizontal axis to code the current pleasure level (ranging from 1 [extremely unpleasant] on the left to 9 [extremely pleasant] on the right) and a vertical axis to code the current arousal level (ranging from 1 [low arousal] at the bottom to 9 [high arousal] at the top).

RESULTS

As a repeated measures analysis of variance (ANOVA) found no differences in EBR between the three sessions, $F(2, 46) = 1.77$, $p = 0.18$, we took the average across all three measures as an estimate of the individual EBR. To test whether binaural beats affected performance in the creativity tasks, repeated measures ANOVAs with auditory stimulation (Alpha, Gamma, Control) as within-participant factor were conducted.

The basic analysis of the AUT flexibility score (divergent thinking) showed no reliable effect, $F(2, 46) < 1$. However, adding the individual EBR as centered covariate (cf., van Breukelen and van Dijk, 2007) rendered the effect highly reliable, $F(2, 44) = 5.22$, $p = 0.009$, suggesting that the effect might be mediated by EBR. This was confirmed by regression analyses relating the individual Alpha benefit (performance in the Alpha condition minus performance in the Control condition) and the individual Gamma

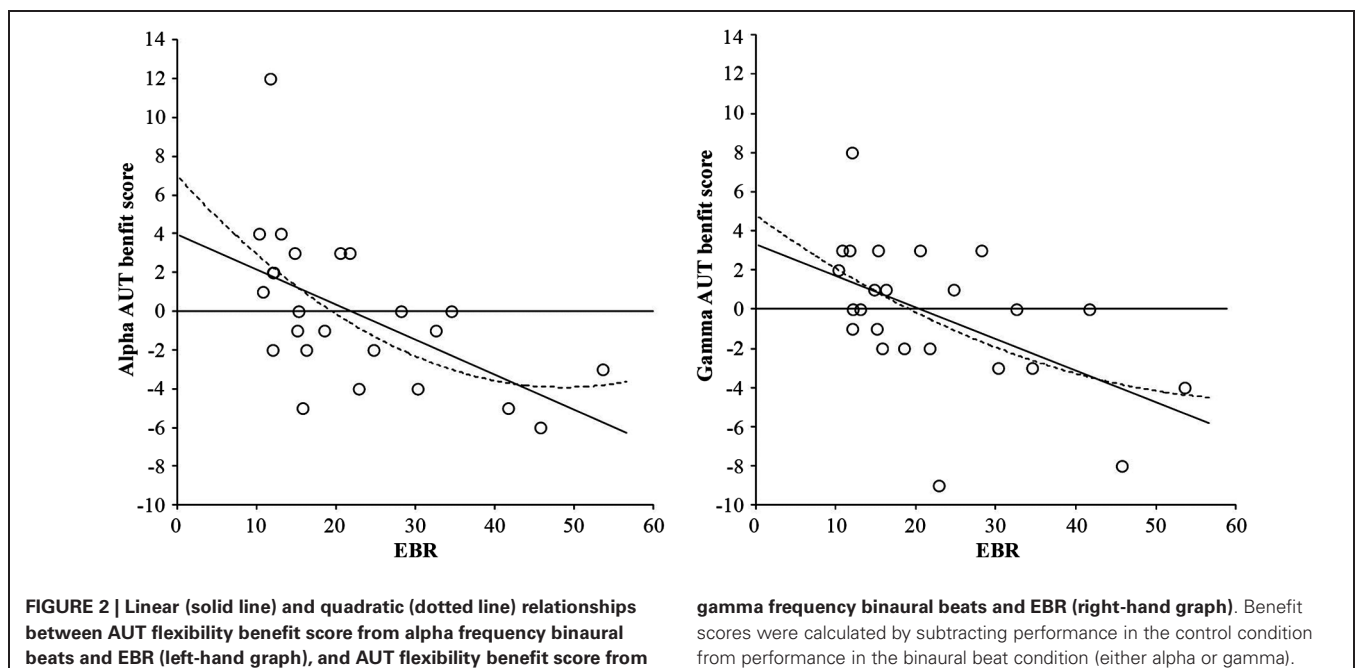
benefit (performance in the Gamma condition minus performance in the Control condition) to individual EBR. As shown in **Figure 2**, the relationships between both the Alpha benefit and EBR, $F(1, 22) = 9.71$, $p = 0.005$, and the Gamma benefit and EBR, $F(1, 22) = 8.51$, $p = 0.008$, followed highly reliable negative linear trends, while the quadratic trends explain lesser variance: $F(2, 21) = 5.31$, $p = 0.014$, and $F(2, 21) = 4.23$, $p = 0.029$, respectively. Interestingly, the distribution clearly crosses the zero line, suggesting that people with low EBRs (under 20 blinks per min) mostly benefit from both alpha, and benefit or are not impaired by gamma binaural beats, while people with higher EBRs do not benefit or are even impaired by binaural beat stimulation.

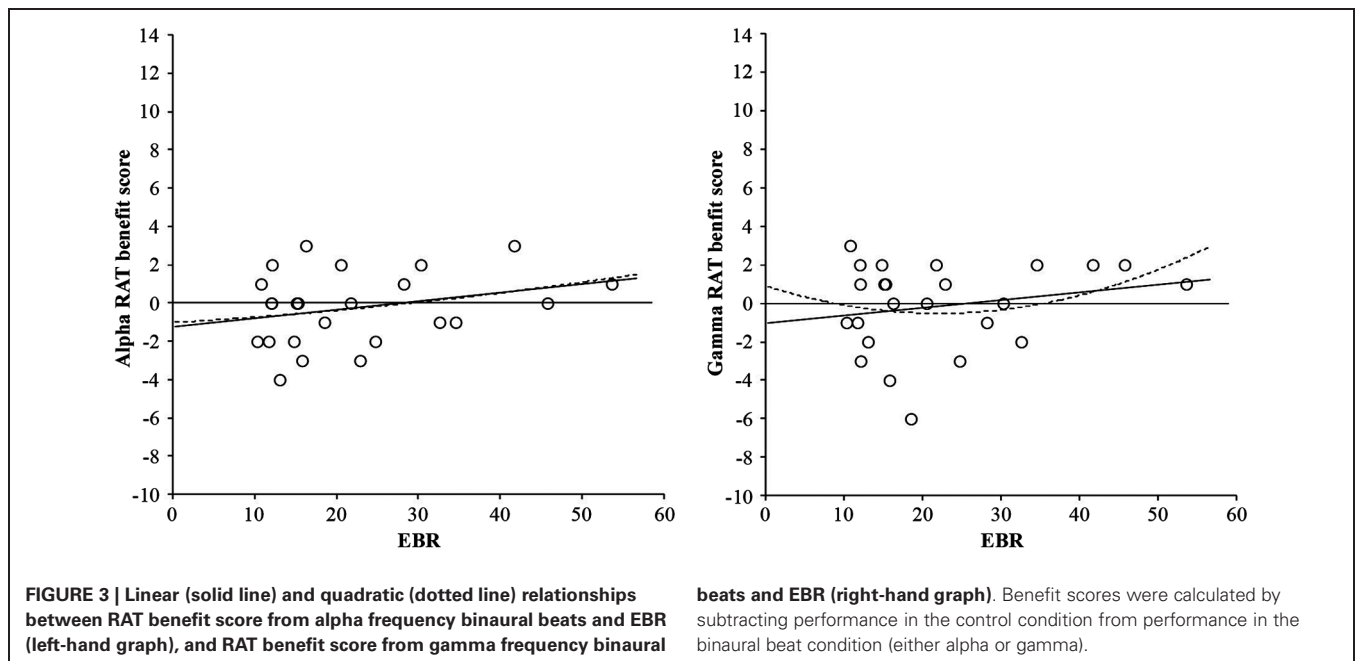
For the RAT score (convergent thinking) neither the basic ANOVA nor the ANCOVA with EBR as covariate yielded any reliable effect, $F(2, 46) < 1$. These non-significant relationships can be seen in **Figure 3**.

Given the assumption of a link between dopamine and mood or affect, we also explored whether the changes in performance were accompanied by changes in mood. This did not seem to be the case, however. For one, an (2×3) ANOVA of the NA score of the PANAS-S before and after the three sessions did not reveal any reliable effect, $ps > .3$. The same ANOVA of the PA PANAS-S scores produced a significant effect of time point, $F(1, 23) = 11.07$, $p = 0.003$ (indicating a slight reduction of positive mood from 29.9 to 28.3), but this effect did not interact with auditory condition, $p > .18$. The analysis of the affect grid data also found no indication for condition-specific effects on average pleasure, $F(2, 46) < 1$, or arousal, $F(2, 46) < 1$ scores.

DISCUSSION

The aim of this study was to investigate whether binaural beats, a form of cognitive entrainment, affect people's creative performance, and whether such impact might be mediated by the





individual striatal dopamine level, as assessed by means of EBR. The outcome provides a straightforward picture.

First, we found no evidence for any influence of binaural beats on convergent thinking, while divergent thinking was systematically affected depending on base-line EBR. This supports the assumption that convergent thinking, and other kinds of highly constrained top-down search processes, rely more on the frontal part of the frontal-striatal interaction constituting cognitive control (in the sense of Cools and d'Esposito, 2009), while divergent thinking, and other forms of mental flexibility, lean more towards the striatal part (Akbari Chermahini and Hommel, 2010; Hommel, 2012). Moreover, the observation of a differential effect on one of the two kinds of creative performance reinforces claims that human creativity is not a unitary function but consists of multiple components (Wallas, 1926; Guilford, 1967; Nijstad et al., 2010).

Second, we could not find any difference between the Alpha and the Gamma condition—both had the same kind and the same degree of impact on divergent thinking. This suggests that binaural beats do not so much trigger or facilitate a particular neural synchronization processes but rather support neuronal phase locking in general (Kuwada et al., 1979). For instance, they might impose some temporal structure on neural processes and thereby reduce cortical noise (Karino et al., 2006), which again may make task-specific processes that rely on neural communication and/or synchronization more reliable. In which frequency this temporal structure is operating might be less relevant.

Third, our findings clearly suggest that binaural beats do not represent a suitable all-round tool for cognitive enhancement. While participants with lower EBRs (20 blinks per min or lower) showed clear beat-induced benefits in divergent thinking, binaural beats impaired the performance of individuals with higher EBRs (20 blinks per min or higher; see Figure 2). As suspected,

this suggests that beat-induced cognitive enhancement depends on the individual striatal dopamine level—an observation that parallels Akbari Chermahini and Hommel's (2012) finding of equally selective mood effects on divergent thinking.

There are at least two possible, not mutually exclusive explanations for this observation. First, there is evidence that lower-than-average EBR levels are associated with less effective performance in divergent-thinking tasks, especially regarding flexibility (Akbari Chermahini and Hommel, 2010). Even though this difference just missed the significance criterion in our study (in the control condition, the flexibility scores of the low and high EBR groups were 10.58 and 12.83, respectively, $p = .08$), individuals with rather low striatal dopamine levels might have more room for improvement and are, thus, more sensitive to cognitive-enhancement procedures. For instance, it might be that binaural beats induce, or increase the size of phasic dopamine bursts, which might have a stronger impact in individuals with a relatively low tonic dopamine level. Individuals with a more suitable dopamine level may not need these extra or extra-sized bursts and may end up with more than optimal cortical noise. This would also suggest that EBRs mainly reflect tonic dopamine activity in the striatum, but this lies outside the scope of the current study and, thus, remains speculation for now.

Second, it might be that binaural beats do not operate directly on the individual dopamine level, be it tonic or phasic. Note that we did not find any systematic, beat-induced mood effects. To the degree that changes in dopamine levels are accompanied by changes in mood (Akbari Chermahini and Hommel, 2012), this might suggest that binaural beats facilitated or enabled processes that compensate for the individual lack of striatal dopamine. For instance, it might be that dopamine is functional in driving neural synchronization (Schnitzler and Gross, 2005). If so, a relatively

low level of striatal dopamine may thus make it more difficult to set up synchronized neural states, and this difficulty may somehow be overcome through other, compensatory processes that are induced or facilitated by binaural beats. As speculated earlier, binaural beats may increase the temporal structure of idling neural activities and thereby reduce cortical noise, which again might facilitate setting up synchronized states. Again, it is conceivable that individuals with more optimal dopamine levels do not need, or may even be impaired by this alternative way to create the necessary synchronized states.

Irrespective of which of these two scenarios will turn out to be more realistic, it is clear that binaural beats do not represent a one-size-fits-all enhancement technique. They can be effective in enhancing brainstorm-like creative thinking in individuals with low striatal dopamine levels, but they can at the same time impair performance in exactly the same kind of task in others. On the one hand, this calls for more care in the propagation of binaural beats as a cognitive-enhancement method and a better understanding of the underlying neural and cognitive mechanisms. On the other hand, however, it also implies that previous failures to find positive effects of binaural beats on cognitive performance need not be taken as evidence against the efficiency of the manipulation. In fact, careful selection of individuals involving a systematic evaluation of their cognitive control profiles is likely to yield evidence of cognitive enhancement, even under conditions that proved ineffective by previous research.

REFERENCES

- Akbari Chermahini, S., Hickendorff, M., and Hommel, B. (2012). Development and validity of a dutch version of the remote associates task: an item-response theory approach. *Think. Skills Creat.* 7, 177–186. doi: 10.1016/j.tsc.2012.02.003
- Akbari Chermahini, S., and Hommel, B. (2010). The (b)link between creativity and dopamine: spontaneous eye blink rates predict and dissociate divergent and convergent thinking. *Cognition* 115, 458–465. doi: 10.1016/j.cognition.2010.03.007
- Akbari Chermahini, S., and Hommel, B. (2012). More creative through positive mood? Not everyone! *Front. Hum. Neurosci.* 6:319. doi: 10.3389/fnhum.2012.00319
- Ashby, F. G., Isen, A. M., and Turken, A. U. (1999). A neuro-psychological theory of positive affect and its influence on cognition. *Psychol. Rev.* 106, 529–550. doi: 10.1037/0033-295x.106.3.529
- Baas, M., De Dreu, C. K. W., and Nijstad, B. A. (2008). A meta-analysis of 25 years of mood-creativity research: hedonic tone, activation or regulatory focus? *Psychol. Bull.* 134, 779–806. doi: 10.1037/a0012815.supp
- Barbato, G., Ficca, G., Muscettola, G., Fichelle, M., Beatrice, M., and Rinaldi, F. (2000). Diurnal variation in spontaneous eye-blink rate. *Psychiatry Res.* 93, 145–151. doi: 10.1016/s0165-1781(00)00108-6
- Bowden, E. M., Jung-Beeman, M., Fleck, J., and Kounios, J. (2005). New approaches to demystifying insight. *Trends Cogn. Sci.* 9, 322–328. doi: 10.1016/j.tics.2005.05.012
- Cohen, J. D., Braver, T. S., and Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Curr. Opin. Neurobiol.* 12, 223–229. doi: 10.1016/s0959-4388(02)00314-8
- Cools, R., and d'Esposito, M. (2009). "Dopaminergic modulation of flexible cognitive control in humans," in *Dopamine Handbook*, eds A. Björklund, S. Dunnet, L. Iversen and S. Iversen (Oxford, UK: Oxford University Press), 249–260.
- Davis, M. A. (2009). Understanding the relationship between mood and creativity: a meta-analysis. *Organ. Behav. Hum. Decis. Process.* 108, 25–38. doi: 10.1016/j.obhdp.2008.04.001
- Davis, K. L., Kahn, R. S., Ko, G., and Davidson, M. (1991). Dopamine in schizophrenia: a review and conceptualization. *Am. J. Psychiatry* 148, 1474–1486.
- Dietrich, A., and Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychol. Bull.* 136, 822–848. doi: 10.1037/a0019749
- Eysenck, H. J. (1993). Creativity and personality: suggestions for a theory. *Psychol. Inq.* 4, 147–178. doi: 10.1207/s15327965pli0403_1
- Fink, A., Grabner, R. H., Benedek, M., and Neubauer, A. C. (2006). Divergent thinking training is related to frontal electroencephalogram alpha synchronization. *Eur. J. Neurosci.* 23, 2241–2246. doi: 10.1111/j.1460-9568.2006.04751.x
- Fink, A., Grabner, R. H., Benedek, M., Reishofer, G., Hauswirth, V., Fally, M., et al. (2009). The creative brain: investigation of brain activity during creative problem solving by means of EEG and fMRI. *Hum. Brain Mapp.* 30, 734–748. doi: 10.1002/hbm.20538
- Fink, A., and Neubauer, A. C. (2006). EEG alpha oscillations during the performance of verbal creativity tasks: differential effects of sex and verbal intelligence. *Int. J. Psychophysiol.* 62, 46–53. doi: 10.1016/j.ijpsycho.2006.01.001
- Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41, 1–24. doi: 10.1016/0306-4522(91)90196-u
- Guilford, J. P. (1950). Creativity. *Am. Psychol.* 5, 444–454. doi: 10.1037/h0063487
- Guilford, J. P. (1967). *The Nature of Human Intelligence*. New York: McGraw-Hill.
- Hommel, B. (2012). "Convergent and divergent operations in cognitive search," in *Cognitive Search: Evolution, Algorithms, and the Brain*, eds P. M. Todd, T. T. Hills and T. W. Robbins (Cambridge, MA: MIT Press), 221–235.
- Karino, S., Yumoto, M., Itoh, K., Uno, A., Yamawaka, K., Sekimoto, S., et al. (2006). Neuromagnetic responses to binaural beat in human cerebral cortex. *J. Neurophysiol.* 96, 1927–1938. doi: 10.1152/jn.00859.2005
- Karson, C. N. (1983). Spontaneous eye-blink rates and dopaminergic systems. *Brain* 106, 643–653. doi: 10.1093/brain/106.3.643
- Keefe, J. A., and Magaro, P. A. (1980). Creativity and schizophrenia: an equivalence of cognitive processing. *J. Abnorm. Psychol.* 89, 390–398. doi: 10.1037//0021-843x.89.3.390
- Keizer, A. W., Verment, R. S., and Hommel, B. (2009). Enhancing cognitive control through neurofeedback: a role of gamma-band activity in managing episodic retrieval. *Neuroimage* 49, 3404–3413. doi: 10.1016/j.neuroimage.2009.11.023
- Kuwada, S., Yin, T. C. T., and Wickesberg, R. E. (1979). Response of cat inferior colliculus neurons to binaural beat stimuli: possible mechanisms for sound localization. *Science* 206, 586–588. doi: 10.1126/science.493964
- Lane, J. D., Kasian, S. J., Owens, J. E., and Marsh, G. R. (1998). Binaural auditory beats affect vigilance performance and mood. *Physiol. Behav.* 63, 249–252. doi: 10.1016/s0031-9384(97)00436-8
- Licklider, J. C. R., Webster, J. C., and Hedlun, J. M. (1950). On the frequency limits of binaural beats. *J. Acoust. Soc. Am.* 22, 468–473. doi: 10.1121/1.1906629
- McAlpine, D., Jiang, D., and Palmer, A. R. (1996). Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig. *Hear. Res.* 97, 136–152. doi: 10.1016/0378-5955(96)00068-8
- Nelson, B., and Rawlings, D. (2008). Relating schizotypy and personality to the phenomenology of creativity. *Schizophr. Bull.* 36, 388–399. doi: 10.1093/schbul/sbn098
- Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., and Baas, M. (2010). The dual pathway to creativity model: creative ideation as a function of flexibility and persistence. *Eur. Rev. Soc. Psychol.* 21, 34–77. doi: 10.1080/10463281003765323
- Oster, G. (1973). Auditory beats in the brain. *Sci. Am.* 229, 94–102. doi: 10.1038/scientificamerican1073-94
- Pratt, H., Starr, A., Mickalewski, H. J., Dimitrijevic, A., Bleich, N., and Mitterman, N. (2009). Cortical evoked potentials to an auditory illusion: binaural beats. *Clin. Neurophysiol.* 120, 1514–1524. doi: 10.1016/j.clinph.2009.06.014
- Pratt, H., Starr, A., Mickalewski, H. J., Dimitrijevic, A., Bleich, N., and Mitterman, N. (2010). A comparison of auditory evoked potentials to acoustic beats and to binaural beats. *Hear. Res.* 262, 34–44. doi: 10.1016/j.heares.2010.01.013
- Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. *J. Pers. Soc. Psychol.* 57, 493–502. doi: 10.1037//0022-3514.57.3.493
- Schnitzler, A., and Gross, J. (2005). Normal and pathological oscillatory communication in the brain. *Nat. Rev. Neurosci.* 6, 285–296. doi: 10.1038/nrn1650

- Schwarz, D. W. F., and Taylor, P. (2005). Human auditory steady state responses to binaural and monaural beats. *Clin. Neurophysiol.* 116, 658–668. doi: 10.1016/j.clinph.2004.09.014
- Shukla, D. (1985). Blink rate as clinical indicator. *Neurology* 35, 286. doi: 10.1212/wnl.35.2.286
- Taylor, J. R., Elsworth, J. D., Lawrence, M. S., Sladek, Jr. J. R., Roth, R. H., and Redmond, Jr. D. E. (1999). Spontaneous blink rates correlate with dopamine levels in the caudate nucleus of MPTP-treated monkeys. *Exp. Neurol.* 158, 214–220. doi: 10.1006/exnr.1999.7093
- Turow, G., and Lane, J. D. (2011). “Binaural beat stimulation: altering vigilance and mood states,” in *Music, Science, and the Rhythmic Brain: Cultural and Clinical Implications*, eds J. Berger and G. Turow (New York, NY: Routledge), 122–136.
- van Breukelen, G. J., and van Dijk, K. R. (2007). Use of covariates in randomized controlled trials. *J. Int. Neuropsychol. Soc.* 13, 903–904. doi: 10.1017/s1355617707071147
- Vernon, D. (2009). *Human Potential: Exploring Techniques Used to Enhance Human Performance*. London: Routledge.
- Vernon, D., Peryer, G., Louch, J., and Shaw, M. (2012). Tracking EEG changes in response to alpha and beta binaural beats. *Int. J. Psychophysiol.* doi: 10.1016/j.ijpsycho.2012.10.008. [Epub ahead of print].
- von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313. doi: 10.1016/s0167-8760(00)00172-0
- Wallas, G. (1926). *The Art of Thought*. New York: Harcourt Brace.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 September 2013; accepted: 30 October 2013; published online: 14 November 2013.

Citation: Reedijk SA, Bolders A and Hommel B (2013) The impact of binaural beats on creativity. *Front. Hum. Neurosci.* 7:786. doi: 10.3389/fnhum.2013.00786

This article was published in the Journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Reedijk, Bolders, and Hommel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer

Vessela Stamenova^{1*}, Janine M. Jennings², Shaun P. Cook³, Lisa A. S. Walker⁴, Andra M. Smith⁴ and Patrick S. R. Davidson⁴

¹ Rotman Research Institute, Baycrest – University of Toronto, Toronto, ON, Canada

² Department of Psychology, Wake Forest University, Winston-Salem, NC, USA

³ Department of Psychology, Millersville University, Millersville, PA, USA

⁴ School of Psychology, Faculty of Social Sciences, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada

Edited by:

Guido P. H. Band, Leiden University, Netherlands

Reviewed by:

Kristin Flegal, University of California, Davis, USA

Julie Bugg, Washington University in St. Louis, USA

*Correspondence:

Vessela Stamenova, Rotman Research Institute, Baycrest – University of Toronto, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada
e-mail: vstamenova@research.baycrest.org

Normal aging holds negative consequences for memory, in particular for the ability to recollect the precise details of an experience. With this in mind, Jennings and Jacoby (2003) developed a *recollection training* method using a single-probe recognition memory paradigm in which new items (i.e., foils) were repeated during the test phase at increasingly long intervals. In previous reports, this method has appeared to improve older adults' performance on several non-trained cognitive tasks. We aimed to further examine potential transfer effects of this training paradigm and to determine which cognitive functions might predict training gains. Fifty-one older adults were assigned to either recollection training ($n = 30$) or an active control condition ($n = 21$) for six sessions over 2 weeks. Afterward, the recollection training group showed a greatly enhanced ability to reject the repeated foils. Surprisingly, however, the training and the control groups improved to the same degree in recognition accuracy (d') on their respective training tasks. Further, despite the recollection group's significant improvement in rejecting the repeated foils, we observed little evidence of transfer to non-trained tasks (including a temporal source memory test). Younger age and higher baseline scores on a measure of global cognitive function (as measured by the Montreal Cognitive Assessment tool) and working memory (as measured by Digit Span Backward) predicted gains made by the recollection training group members.

Keywords: aging, memory, rehabilitation, recollection, familiarity

INTRODUCTION

Two broad approaches have been taken in cognitive rehabilitation of episodic memory. The first is to train people to use various strategies for memory encoding and/or retrieval. Although older adults can benefit from such training (for review, see Glisky and Glisky, 2008), many strategies are only appropriate for certain types of materials and/or relationships (e.g., in learning associations between faces and names), and once people are taught such strategies they are then left to apply them appropriately to daily life. The second broad approach is to attempt to repair or improve the cognitive processes that underlie episodic memory (for example see Buschkuhl et al., 2008; Brehmer et al., 2012). A major concern regarding this approach is generalization: Although it is relatively easy to produce improvements on the particular tasks being trained, it is more difficult to produce convincing evidence of improvement on tasks that were not trained but that depend on the cognitive process(es) that underwent training (e.g., Owen et al., 2010; Ranganath et al., 2011). In this paper we explored a potentially fruitful technique developed by Jennings and Jacoby (2003), which they argued can improve memory in healthy older people by focusing on recollection.

Dual-process theories propose that both recollection and familiarity can influence memory (Mandler, 1980, 2008; Jacoby, 1991;

Yonelinas, 2002). *Recollection* can be thought of as the essence of episodic memory: a vivid re-experiencing of the details of the encoding event (i.e., the “what,” “where,” and “when” of the event; Tulving, 2002). According to several theories, recollection often contributes significantly to performance on tests of recognition memory, and may contribute even more heavily to performance on tests of recall (Quamme et al., 2004; McCabe et al., 2011).

In contrast, *familiarity* entails a more general feeling that a stimulus or an event has been experienced in the past, and may contribute less strongly to recall than it does to recognition performance (McCabe et al., 2011). Aging is often associated with a decline in recollection, whereas familiarity remains relatively undisrupted (or, at least, can vary from study to study; for review, see Light, 2012).

One piece of evidence of recollection's decline in aging comes from Jennings and Jacoby's (1997) “opposition procedure.” On this task, which places recollection and familiarity in opposition to one another, the participant is given a study list, followed by a recognition test in which both old and new words appear. The participant's goal is to endorse only those items that had appeared on the study list. The critical words on the task are the new words (i.e., foils), because these are shown *twice* during the test. On its second presentation, a repeated foil feels more familiar but must

still be rejected. The only way to correctly do so is to recollect the context in which the word was seen previously (i.e., participants should remember that the word only appeared in the test list, and *not* in the study list.). If participants rely on familiarity alone during this decision, they will erroneously endorse the repeated foil and commit what Jennings and Jacoby (1997) call a *repetition error*. Older adults are much more likely to commit repetition errors than are young adults (Jacoby, 1999; Benjamin, 2001).

RECOLLECTION TRAINING WITH THE REPETITION-LAG PARADIGM: PROMISING PRELIMINARY RESULTS

To improve recollection in older adults, Jennings and Jacoby (2003) modified the opposition procedure by steadily lengthening the repetition lag (i.e., the interval over which foils were repeated during the test) across several training sessions. Before undergoing this “recollection training,” the older adults frequently made repetition errors, even when only a few words intervened between the first and second presentation of a given foil. After a week of training, however, the older adults could perform the task very well, avoiding repetition errors even when 28 items intervened between the first and second presentation of a foil. This improvement was observed only in the experimental group, in which the repetition lag was gradually lengthened for each participant once he or she had achieved a certain level of performance (i.e., the *adaptive* training group); little improvement was seen in a yoked active control group in which different length lags were presented randomly across training days (i.e., non-adaptive).

What makes this paradigm so interesting is that it is one of the few existing cognitive rehabilitation methods to show the potential to ameliorate recollection deficits in older adults and transfer those gains to non-trained tasks. Jennings et al. (2005) categorized their transfer tasks into *near-transfer* versus *far-transfer* on the basis of the underlying processes involved. For example, verbal n-back (Dobbs and Rule, 1989; Jonides et al., 1997), self-ordered pointing (Petrides and Milner, 1982) and source discrimination were designated as near-transfer because they all involved retrieval of contextual information (temporal, output, and source monitoring, respectively). Digit symbol substitution (Wechsler, 1997), reading span (Daneman and Carpenter, 1980), and verbal free recall [as assessed using the California Verbal Learning Test-Second Edition (CVLT-II); Delis et al., 2001] were categorized as far-transfer, due to the fact that they were considered to have different underlying processes. Their recollection training group improved mostly on the near transfer tests: one- and two-back working memory, self-ordered pointing, and source discrimination with far transfer gains seen only on digit symbol substitution. There were no significant changes on the three-back task or on the CVLT-II.

Other studies have replicated the benefits of this method to some of the same non-trained tasks used by Jennings et al. (2005): digit symbol substitution (Bailey et al., 2011) and self-ordered pointing (Bailey et al., 2011; but see Lustig and Flegal, 2008 for exception) in older adults, and 2-back working memory, and source discrimination in Alzheimer’s patients (Boller et al., 2012). Transfer effects have also been reported to Trail Making Part B (Lustig and Flegal, 2008), and the AX-CPT task (Quamme et al.,

2004) in older adults (Bailey et al., 2011) and delayed matching-to-sample in Alzheimer’s patients (Boller et al., 2012). Curiously, Boller et al. (2012) failed to observe transfer effects to reading span, free recall, or recognition memory in Alzheimer’s patients [recognition memory did improve in Boller et al. (2012), but did so in both the recollection training and the control training group].

THE PRESENT STUDY

Based on this previous work, we sought to answer two major questions:

Does recollection training produce clear improvements on non-trained memory tasks?

As reviewed above, past studies have shown evidence of transfer to non-trained tasks, although not universally (Jennings et al., 2005; Lustig and Flegal, 2008; Boller et al., 2012). One of the reasons for variability among the previous studies may be that they have used relatively small groups ($n_s = 12$ in Boller et al., 2012; $n_s = 12$ –17 in Jennings et al., 2005). In the present study we aimed to examine a number of measures that, based on past studies, appeared likely to show transfer, and to use larger sample sizes than previously, to increase statistical power. The strongest evidence of recollection training yielding a general benefit to memory would come from a comparison of two groups (recollection training versus a well-matched active control group with equal training gains expectations as recommended by Boot et al., 2013), assessed on non-trained tasks before and after training. Any improvements shown by the recollection training group above-and-beyond those seen in the active control group (in the form of a significant group by time interaction, Nieuwenhuis et al., 2011) could be attributed to the training, rather than to non-specific factors such as improvement in mood or comfort/reduction in stress, benefits of social/intellectual stimulation, and/or expectations of improvement (i.e., placebo).

We organized our non-trained measures by sorting them into those on which we should be most likely to see transfer effects, versus those that we should be less likely to see transfer. At the closest end of the spectrum, we designated two of our non-trained measures as “near transfer”: temporal order memory from an experimental source monitoring paradigm, and total across list intrusions on Long Delay Free Recall from the CVLT-II, both of which require participants to remember the conjunction between a specific item and the temporal context in which it appeared, which is essentially what the training in the Jennings and Jacoby (2003) paradigm involved.

We included several other “intermediate transfer” scores, all of which reflect episodic memory, and all of which we expected might show improvement based on theoretical and/or empirical grounds (Bailey et al., 2011; Boller et al., 2012). In addition to temporal order memory, our experimental source monitoring paradigm included measures of perceptual (i.e., voice) and spatial (i.e., location) source. Given these measures are measures of source memory, we considered them to be close to the contextually rich recollection memory that was being trained and thus we classified them as “intermediate transfer.” The CVLT-II

Long Delay Free Recall scores were also considered “intermediate transfer” as we were working under the assumption that recollection supports recall (McCabe et al., 2011). Jennings et al.’s (2005) recollection-training group showed numerically improved CVLT-II Long Delay Free Recall scores from baseline to post-training (improving from 69 to 75%). We hypothesized that a lack of power may have resulted in this numerical difference not attaining statistical significance and that with larger groups we would see significantly greater improvements on these measures in the recollection training group than in the active control group.

The Brief Visuospatial Memory Test-Revised (BVM-T-R; Benedict, 1997) was considered “far transfer,” given the different modality (visual as opposed to the verbal memory that was being training in the recollection-training task). Finally, given the above-listed evidence that recollection training might also improve performance on working memory tasks (verbal n-back, self-ordered pointing), we included a short-term/working memory measure: Digit Span (from the Wechsler Adult Intelligence Scale-III (WAIS-III), Wechsler, 1997), which has been shown in the past to correlate with n-back (Jaeggi et al., 2010). We considered this test to be a measure of “far-transfer,” given it does not require the support of long-term memory. The ranking of all transfer measures is displayed in **Figure 1**. Finally, we included the Multifactorial Memory Questionnaire (MMQ, Troyer and Rich, 2002), to examine whether participants’ self-perceptions and reported strategies might change from before to after training.

In addition to the group comparisons, we also asked whether, on an individual-by-individual basis, the gains made in training relate to the gains made on the transfer tasks. We were especially interested in whether those who show greater gains on the training task would also show greater gains on the transfer measures.

Do those who come into the training with better memory improve to a different degree from those with poorer memory?

In previous work with the recollection training paradigm, clearly some participants have improved more than others: For example, of Jennings and Jacoby’s (2003) $n = 12$ recollection training participants, three showed barely any advancement, three reached the maximum lag by the end of training, and the rest fell in between. In an attempt to explain these individual differences, Bissig and Lustig (2007) ranked older adults by how much they improved on a modified version of the recollection training paradigm. Those who were younger and had higher verbal intelligence (as measured by a vocabulary test) showed greater recollection training gains.

Note, however, that Bissig and Lustig (2007) modified the original method to allow for self-paced encoding. For this reason, we asked whether similar results would be obtained with the original training paradigm. Given previous reports that cognitive training tends to be more beneficial for individuals with higher cognitive function (Yesavage et al., 1990), our hypothesis was that people with better cognitive function may benefit more from the training. This is similar to the so called “Matthew Effect” proposed in the developmental literature by Stanovich (1986) who proposed that children who have better reading abilities early on tend to develop better reading and learning skills later in life, while children with lower reading abilities early on develop poorer reading skills. Higher cognitive function may also be related to higher cognitive reserve and participants with higher cognitive reserve may benefit more from cognitive training (Stern, 2002).

MATERIALS AND METHODS

PARTICIPANTS

Forty-three older adults were randomly assigned to either a *recollection training* ($n = 22$) or an active *recognition control* group ($n = 21$). At the end eight more participants were assigned directly to the recollection training condition, to increase its size to $n = 30$ (these participants did not differ from the rest of the sample on any of the variables the participant groups were matched on). The study was single-blind: participants were not aware of which condition they were in, although the examiner (V. S.) was. The sample size was based on Jennings et al.’s (2005) “repetition-lag” recollection training study which included 17 participants each in the recollection and in the recognition control group. Their treatment effect, as measured by the change in lag level from the beginning to the end, was large (Cohen’s $d = 1.79$). The participants were living independently in the community (Ottawa, ON, Canada) and recruited through local newspaper ads. There were no statistically significant group differences (see **Table 1**) in age, education, sex, handedness, cognitive status based on the Montreal Cognitive Assessment, (MoCA; Nasreddine et al., 2005) and mood as measured by the Center for Epidemiological Studies Depression scale (CES-D). The last scale was included after the beginning of the study, so the first nine participants did not complete it. MoCA scores ranged from 21 to 30. Although a few participants scored below the MoCA’s official recommended cut-off of 26, a more recent study indicates that a cut-off of 20 may be more appropriate, and for this reason we kept all participants (Waldron-Perrine and Axelrod, 2012). For the CES-D, we used a cut-off score of 27, which has been recommended for older adults

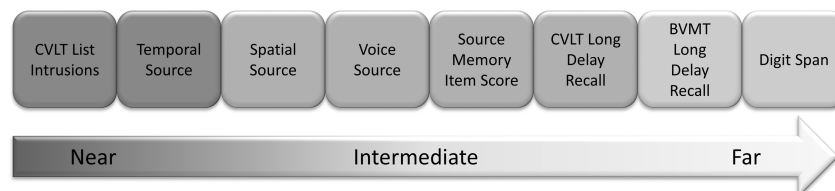


FIGURE 1 | Near-, intermediate- and far-transfer measures.

Table 1 | Participants' demographic information.

	Recognition control	Recollection training	<i>p</i> -value
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	
Age (years)	68.62 (6.39)	67.60 (5.56)	0.55
Education (years)	16.57 (2.99)	15.57 (3.09)	0.25
Sex (F/M)	19/2	23/7	0.28
Handedness (R/L/A)	19/1/1	29/0/1	0.46
MoCA (/30)	27.24 (2.48)	25.97 (2.52)	0.09
CES-D ^a	9.61 (7.72)	8.08 (5.69)	0.46

MoCA, Montreal Cognitive Assessment; CES-D, The Center for Epidemiological Studies Depression scale.

^aSample size (*n* = 21, *n* = 30).

(Schein and Koenig, 1997). Only one participant exceeded this level with a score of 31, but she reported that her 60th birthday had occurred that week and had made her feel worse than normal, and she asserted that she was not depressed. A number of participants (*n* = 12) indicated having ongoing mental health problems, mostly depression which was successfully treated. To ensure that these participants did not significantly affect our results, we re-ran all analyses without them and our pattern of results did not change in any way.

MATERIALS AND PROCEDURE

The project was approved by the University of Ottawa Research Ethics Board and all participants provided informed consent before taking part.

Training overview

Participants came in for three training sessions per week for 2 weeks. Each training day, they completed four sessions (each session involved studying one list and then performing the corresponding recognition memory test). Participants usually completed the four training sessions in 20–30 min. This schedule was similar to the one used in recollection training of older adults (Jennings et al., 2005) and identical to the one used with Alzheimer's patients (Boller et al., 2012).

Training tasks

Recollection training. The task was identical to that used by Jennings et al. (2005). Briefly, participants studied lists of 30 words presented one at a time for a 2-s period and were asked to read each word out loud and try to commit it to memory. Each study list was followed by a recognition test, which included the studied items plus 30 new items (each shown twice and repeated at a specific lag interval) and five “filler” words which were necessary to complete the interval lags. The words were nouns, verbs, and adjectives balanced across each list for frequency of occurrence in the language. There were 128 study word, 128 new word, and 128 filler word lists. On the test, participants were asked to respond to words they could remember having read on the study list with

“Yes” and to words that they could not remember from the original list with “No” by pressing one of two keys on the keyboard. The instructions also stated that when the participant was correct the word “CORRECT” would appear on the screen, but when the participant was not correct, nothing would appear. Finally, they were informed that some of the new words might reappear during the test but that they were still to say “No” to them, because they were not study list words; only new words would reappear during the test, whereas old (i.e., study list) words would appear only once during the test. In addition, they were asked to use the feedback they received to improve their performance if possible. The new words were split into two groups and each group of words was repeated at one of two different lag intervals. The lag interval pairs were 1 and 2; 1 and 3; 2 and 4; 2 and 8; 4 and 12; 4 and 16; 8 and 20; 8 and 24; 12 and 28; 12 and 32; 16 and 36; 16 and 40; 20 and 44; 20 and 48; and 24 and 52. All participants started their first set at lag 1 and 2, which means that half of the new words presented at the test were repeated at lag 1 and half were repeated at lag 2 (Figure 2). Participants had to reach a performance criterion at both lag levels in order to be moved up a level. The criterion was a maximum of one repetition error in identifying the repeated items for lags 1–4, and two repetition errors for lags 8–52. By having two different lag intervals in each run, one ensures that the participants always work at one lag interval level that they have mastered already. Participants had to meet criterion at *both* lag intervals in order to move to the next level. If participants did not achieve criterion at both lags, they remained at the same lag level for the next session and stayed at that level until they met criterion.

Active recognition control task. A single-probe verbal recognition task was designed as an active control condition using exactly the same words as on the recollection training task. Here participants studied 30 words, but on the recognition test the new words were not repeated. Instead, 65 new words were randomly intermixed with the 30 old words. Participants were instructed to press one of

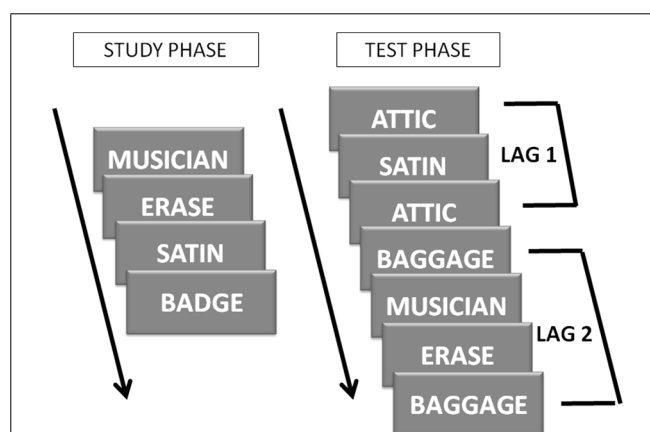


FIGURE 2 | Recollection training design. During the study phase, words were presented one at a time in the middle of the screen. During the test phase studied words and new words were presented. New words were repeated at two lag intervals. All participants started at lags 1 and 2, which is schematically represented above.

two keys to indicate their Yes/No response for each test item and they were also given feedback when they were correct in the same manner as the recollection training task. These participants came in on the same schedule as the other group (i.e., three times per week for 2 weeks, performing four study sessions per visit, which took the same amount of time as the recollection training did). This control condition was considered to be superior to the one used previously by Jennings et al. (2005), in that it used exactly the same study list length and identical word lists as in the recollection training.

Transfer tasks

California Verbal Learning Test-Second Edition (Delis et al., 2001). Participants studied 16 words (list A) that fell into one of four different categories. The list was read at a pace of one word per second and then participants were asked to recall as many words as possible. This was repeated four times. A second list (List B) of 16 words falling into four categories (two of these categories being the same as those on List A) was read. Participants were again asked to recall the list. This was followed by List A Short Delay Free and then Cued Recall. After a 20 min filled delay interval Long Delay Free Recall for List A was administered, followed by Long Delay Cued Recall for List A, and, finally, a yes-no recognition test for List A.

Brief Visuospatial Memory Test-Revised (Benedict, 1997). Participants were shown a page with six geometrical figures (2 per row) for 10 s and asked to commit them to memory (including the spatial location of the figures). They were then asked to draw them on a blank sheet of paper immediately afterward, drawing the figures as accurately as they could and placing them in the appropriate location on the page. This same procedure was performed two more times. After the third trial, participants were reminded not to forget the stimuli, because they might be asked to draw them later from memory. After a 25 min delay, the participants were asked to draw them one more time. After this delayed recall, participants were given a recognition test with six old and six new items and asked to respond by saying 'Yes' to the old items and 'No' to the new items.

Wechsler Adult Intelligence Scale-III (WAIS-III) Digit Span Forward/Backward (Wechsler, 1997). Digits were read at a one digit per second rate and participants were asked to repeat them either exactly as they were read to them (Digit Span Forward) or in backward order (Digit Span Backward).

Source memory. We administered a task developed by Cook (Cook, 2007; Davidson et al., 2013). One hundred and sixty sentences were recorded by three native English speakers, one female and two males. The sentences were emotionally neutral ("She put the rice on to boil and set the time to 20 minutes"). The task consisted of four different conditions: voice source, spatial source, temporal source, and item memory. The voice, spatial, and temporal source conditions each contained two practice items and 16 test items. The item memory test contained two practice items and 40 test items. In all conditions, as the participants were presented with the sentences they were asked to rate on a 5-point scale ('very likely,' 'likely,' 'no opinion,' 'unlikely,' and 'very unlikely')

how likely it is that the sentence would be heard on the radio. For the spatial source condition, all sentences were spoken by the same voice but were presented either on the left or on the right side through a loudspeaker and participants were asked to pay attention to the location because they were to be tested on this later. In the voice source condition, half of the sentences were spoken by a male and the other half by a female voice, and participants were asked to commit to memory which sentences were spoken by the man and which by the woman. In the temporal source condition, all sentences were spoken by a male and a single bell was rung halfway through the list. Participants were asked to indicate which sentences occurred before and which sentences occurred after the bell. For both the voice and temporal source conditions, the sentences were presented using the left and right speakers. Finally, for the item memory test, participants were asked to commit to memory all the sentences as best as they could. Each condition involved a forced choice recognition test. For the spatial, voice, and temporal source tests, participants were shown a sentence written on the screen and asked to indicate whether it had originally been presented in the Left/Right, Male/Female, or Before/After the bell context for each condition respectively by a key press. For the item memory test, participants were presented with pairs of sentences and asked to indicate by a key press which one of the two was a sentence that they had originally heard.

There were six different ways in which the administration of the source memory conditions were ordered (administration order). These orders ensured that each source memory task condition (spatial, temporal, or voice) appeared in each position (i.e., first, second, third) and that each source memory task condition was preceded and followed the same number of times by each of the other source memory task conditions. Item memory blocks were always last. Overall, there were eight sentence lists, which were rotated through the six administration orders to ensure that each sentence list was used at least once as targets for each source condition, as targets for the item condition, and as distractors for the item condition. This means there was some overlap in sentence lists between pre- and post-training for a given participant, but sentences that were present in both pre- and post-training versions never appeared in the same source memory condition. All sentences were presented using DMDX display software Version 4.0.6.0 (Forster, 2012).

The Multifactorial Memory Questionnaire (MMQ; Troyer and Rich, 2002). We included this questionnaire to assess self-reports on three dimensions of memory: contentment or satisfaction with one's own memory ability (MMQ-Contentment), perception of everyday memory ability (MMQ-Ability), and use of everyday memory strategies and aids (MMQ-Strategy). MMQ-Contentment includes 18 statements (e.g., 'I have confidence in my ability to remember things') each with one of five options for endorsement: Strongly Agree, Agree, Undecided, Disagree, and Strongly Disagree. MMQ-Ability includes 20 descriptions of abilities or problems (e.g., 'Not recall the name of someone you have known for some time.') with the following response options: All the time, Often, Sometimes, Rarely, and Never. Finally, MMQ-Strategy includes 19 descriptions of various strategies the

participant may be using (e.g., ‘Write things on a calendar, such as appointments or things you need to do.’) and response options were the same as in the MMQ-Contentment. Participants filled out the questionnaires themselves by checking one of the five options beside each statement.

Administration of transfer tasks. All transfer tasks were administered before and after the training. In most cases the pre-training session took place the week before the training started (i.e., Week 1), the training was conducted during Weeks 2 and 3, and the post-training session was completed the week after the training was completed (Week 4). All but three participants completed the study in a quiet, well-lit testing room at University of Ottawa (three participants in the recollection training group preferred to be tested at home). All participants were administered a general demographic and health questionnaire and then proceeded to the cognitive tests in the following order: (1) CVLT-II: five trials of List A Immediate Free Recall, (2) CVLT-II: List B, (3) CVLT-II: Short Delay Free Recall, (4) CVLT-II: Short Delay Cued Recall, (5) BVMT-R: three learning trials, (5) MMQ, (6) Digit Span Forward/Backward, (7) CVLT-II: Long Delay Free Recall, (8) CVLT-II: Long Delay Cued Recall, (9) CVLT-II: Recognition, (10) BVMT: Delayed Recall, (11) BVMT: Recognition, (12) MoCA, (13) source memory. The CVLT-II has two parallel versions, which were administered in counterbalanced order at each session. The BVMT-R has six versions; only versions 1 and 2 were administered and those were counterbalanced across the two assessments. The same version of the digit-span test was administered before and after training. Participants were administered different source memory administration order versions before and after (if on their first assessment they were administered version 1, after training they were administered version 2, if administered 2 at pre-training they were administered version 3 at post-training, and so on). All were counterbalanced across participants.

RESULTS

TRAINING

Recollection training progress

To determine progress through the training, we compared the longest lag interval reached by Session 3 on Day 1 (to give us a sense of baseline abilities) against the longest lag interval achieved by the end of the training (i.e., session 24, Day 6) following Jennings et al. (2005). These Day 1 scores did not appear to be obscured by a ceiling effect: Only one participant progressed as rapidly as possible through the initial (short) lags – all the others made repetition errors initially. A repeated measures *t*-test showed a significant improvement from the baseline lag interval reached by Session 3 on Day 1 ($M = 2.00$, $SD = 0.79$) to the maximum lag interval reached by the end of training ($M = 21.53$, $SD = 18.43$), $t(29) = 5.92$, $p < 0.001$, $d = 1.96$ (Figure 3).

Recognition memory on the training and control tasks

To assess recognition memory, we calculated the probability of responding “yes” to studied items (hits) and to new items on their first presentation (new false alarms). For both scores, we compared the average performance on Day 1 (collapsed across the four sessions that day) against the average performance on Day 6 (collapsed across the four sessions that day). Hit and false alarm rates and discrimination (d') and response bias (C ; Snodgrass and Corwin, 1988) indices are listed in Table 2. Please refer to Figure 4 for individual participant changes in discrimination index in each group. We conducted a 2×2 [Group (Recollection training, Recognition control) \times Time (Beginning, End)] mixed analysis of variance (ANOVA) for each of these four variables. Hits and false alarms were corrected by adding 0.5 to the number of hits (or false alarms respectively) and dividing by the number of items +1, to correct for situations in which scores were at ceiling or at floor (as per Snodgrass and Corwin, 1988).

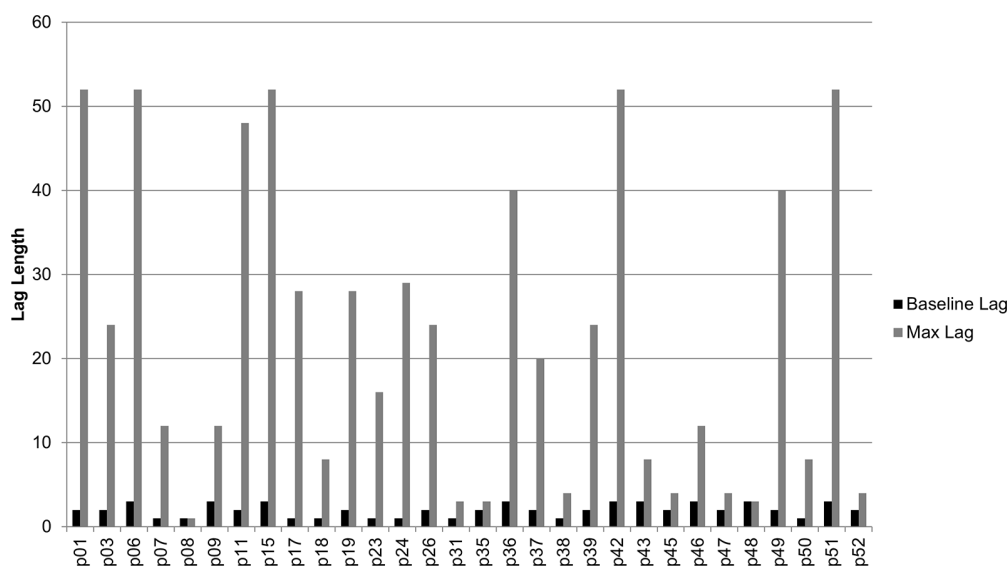


FIGURE 3 | Baseline vs. maximum lag reached by each participant in the recollection training group.

Table 2 | Hit and false alarms rates, and discrimination index (d') and response bias (C) at Day 1 and Day 6 of training.

	Recognition control		Recollection training	
	Day 1 <i>M (SD)</i>	Day 6 <i>M (SD)</i>	Day 1 <i>M (SD)</i>	Day 6 <i>M (SD)</i>
Hit rate	0.86 (0.08)	0.87 (0.07)	0.78 (0.1)	0.77 (0.12)
False alarm rate	0.09 (0.04)	0.06 (0.04)	0.11 (0.07)	0.07 (0.06)
Discrimination (d')	2.53 (0.49)	2.83 (0.68)	2.26 (0.53)	2.50 (0.74)
Bias (C)	0.15 (0.21)	0.20 (0.15)	0.27 (0.27)	0.45 (0.21)

The recognition control group obtained an overall higher *hit rate* than the recollection training group, $F(1,49) = 12.95$, $p = 0.001$, $\eta^2 = 0.21$. There was no significant time effect, $F(1,49) < 0.001$, $p = 0.99$, $\eta^2 = 0$ or interaction, $F(1,49) = 2.07$, $p = 0.16$, $\eta^2 = 0.04$.

Both groups decreased significantly in their *false alarm rates*, $F(1,49) = 14.81$, $p = 0.0003$, $\eta^2 = 0.23$ from the beginning to the end of training, but no group $F(1,49) = 0.73$, $p = 0.40$, $\eta^2 = 0.02$ or interaction effects, $F(1,49) = 1.21$, $p = 0.27$, $\eta^2 = 0.02$ were observed.

Both groups improved in their *discrimination (d')* over time $F(1,49) = 13.70$, $p = 0.001$, $\eta^2 = 0.22$. No group, $F(1,49) = 3.57$, $p = 0.07$, $\eta^2 = 0.07$, or interaction effect, $F(1,49) = 0.12$, $p = 0.73$, $\eta^2 = 0.002$, was observed.

Finally, both groups became more conservative in their responses over time, $F(1,49) = 9.46$, $p = 0.003$, $\eta^2 = 0.16$ and overall the recollection training group was more conservative than the recognition control group, $F(1,49) = 14.14$, $p = 0.0005$, $\eta^2 = 0.22$. The degree of change did not differ significantly between the two groups, however: interaction $F(1,49) = 2.77$, $p = 0.11$, $\eta^2 = 0.05$.

Predictors of training gains

As shown in **Figure 3** participants varied considerably in their progress through the training. Using the method of Bissig and Lustig (2007), we ranked participants based on their gains made during training. Participants were first ranked based on the longest lag level achieved (e.g., a lag of 8 was ranked lower than a lag of 12 with higher rank indicating larger gains in training). Ties were broken based on when the longest lag was achieved, with participants reaching their longest lag earlier being ranked higher than participants reaching it later (e.g., reaching lag 40 by session 23 was ranked higher than reaching it by session 24). Finally, any remaining ties were broken based on accuracy of performance on the repeated words, with participants with higher accuracy ranked higher.

Following Bissig and Lustig (2007), we conducted a hierarchical regression using rank-ordered training gains as the dependent variable. Given, Bissig and Lustig (2007) showed that two demographic characteristics, age and vocabulary, served as reliable predictors of rank, we entered age and years of education into the first model. Measures of cognitive status at the time of testing were entered in the second model and included the baseline MoCA score, the percentage accuracy baseline scores of the

CVLT-II Long Delay Free Recall, and the Digit Span Forward and Backward as the independent variables. The Source Memory score was excluded, because fewer participants completed the Source memory tasks (i.e., the sample size was smaller) due to technical issues with the task and inability to administer the task at a participant's home if s/he was assessed there (the number of participants in each source task are listed in **Table 3**). Using a hierarchical regression allowed us to examine the importance of demographic variables in predicting rank and estimate the amount of variability explained by the cognitive status of the participants over and above that explained by the demographic variables. The regression results are summarized in **Table 4**; Multiple R for the first block of regressors (age and YOE) was close to statistical significance, $F(2,29) = 2.96$, $p = 0.069$; multiple R for the next block of regressors was significant $F(6,29) = 3.12$, $p = 0.022$. The demographic variables (Age and YOE) explained 18% of the variance, while adding the cognitive status scores in block 2 of the analysis increased the amount of variability explained to 45%. This increase is significant by the F change test $F(4,23) = 2.81$, $p = 0.049$. Among the demographic variables, only age was significant, while among the cognitive status, Digit Span Backward ($p = 0.049$) was statistically significant, followed by the MoCA ($p = 0.098$), which was marginally significant.

Transfer effects

A series of repeated measures 2 (Group) \times 2 (Time) ANOVAs were run for each of the non-trained measures. Scores are shown in **Table 3**.

California Verbal Learning Test-Second Edition. To reduce the chance of a Type I error, the following subscores were selected for analysis: (1) five trials of List A immediate free recall; (2) Short Delay Free Recall (3) Long Delay Free Recall (both free recall tasks were chosen over cued to avoid any influence of familiarity on recall), and (4) total across list intrusions on Long Delay Free Recall. Recognition scores were at ceiling for many people and were therefore not used in this analysis. No significant main effects or interactions were observed $ps \geq 0.06$. (**Table 3**). Please note that one person from the recollection group was excluded from this analysis, because they had previously done one of the versions of the CVLT.

Brief Visuospatial Memory Test-Revised. No time, group or interaction effects were observed ($ps > 0.46$) in the performance of participants on Trial 1–3 or on delayed recall (**Table 3**).

Digit Span Forward and Backward. A significant time by group interaction on digits forward $F(1,48) = 4.49$, $p = 0.039$, $\eta^2 = 0.09$ indicated that recollection group participants improved after training whereas recognition group participants did worse than before. There were no group or time effects (**Table 3**; $ps \geq 0.87$).

In digits backward, there was a significant time effect $F(1,48) = 9.71$, $p = 0.003$, $\eta^2 = 0.17$, but no group effect or interaction ($ps \geq 0.31$), suggesting that both groups improved to a similar degree. (**Table 3**).

Source memory task. The only significant result among the four source memory measures was a significant group by time

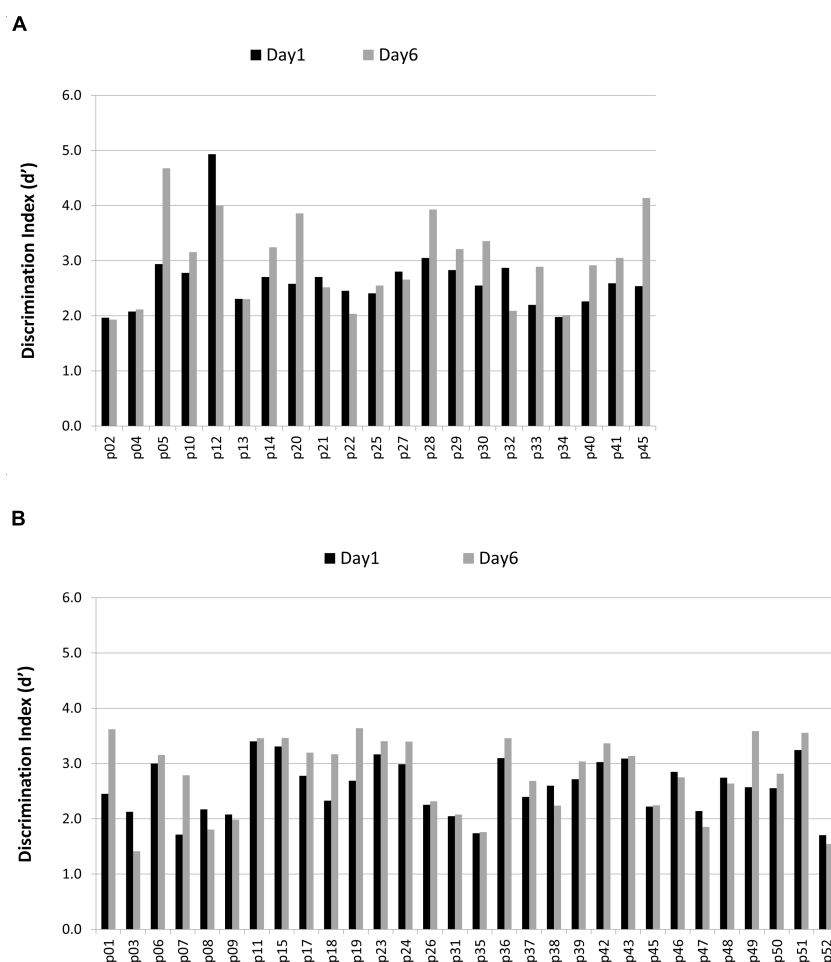


FIGURE 4 | Discrimination Index for each participant in (A) the recognition control group and (B) the recollection training group on Day 1 and Day 6 of training.

interaction on the spatial source memory task $F(1,38) = 4.21$, $p = 0.047$, $\eta^2 = 0.10$. This interaction effect stemmed from the recognition group's accuracy slightly decreasing from pre- to post-training assessment (3%), in the face of the recollection training group improving by almost 10% (Table 3).

Multifactorial Memory Questionnaire. No main effects or interactions were observed ($ps \geq 0.09$) in any of the three measures of the MMQ: Contentment, Ability or Strategy (Table 3; higher scores indicate higher levels of each measure).

Do those participants who show the greatest gains in training also show the greatest improvements on the transfer tasks?

We ran partial correlations to examine the relationship between rank and the change scores (between baseline and follow-up) on a selected set of transfer measures. We chose only one subscore from the CVLT-II (most subscores are highly intercorrelated)—the Accuracy Score of the Long Delay Free Recall as a classical measure of long-term memory, Digit Span Forward and Backward as measures of working memory, and the Spatial Source Memory

Score (given it was the only significant source memory test). Age, and baseline scores of Digit Span Backward were controlled for in the analysis, given these variables were the only ones that were significant predictors for rank. No correlations were seen between rank and the change scores on CVLT Long Delay Free Recall, or Digit Span Forward or Backward. There was, however, a significant correlation between the change in Spatial Source Memory and rank, $r = 0.37$, p (one-tailed) = 0.048, $n = 19$.

DISCUSSION

Our goal was to examine the efficacy of a “recollection training” paradigm (Jennings and Jacoby, 2003) in older adults, including possible transfer to non-trained measures of long-term and working memory. The recollection training group developed a greatly enhanced ability to reject the repeated foils in the training task, and both groups improved in recognition accuracy (d') on their respective training tasks. Despite this, performance on near-, intermediate- and far- transfer tests was affected little by the recollection training. Individual differences in cognitive ability appeared to play a role in the

Table 3 | Pre- and post- training scores on transfer measures.

	Recognition control		Recollection training		<i>p</i> -value		
	Pre <i>M</i> (SD)	Post <i>M</i> (SD)	Pre <i>M</i> (SD)	Post <i>M</i> (SD)	Group	Time	Group × time
CVLT-II list 1–5 proportion accuracy ^a	0.77 (0.10)	0.76 (0.13)	0.70 (0.15)	0.69 (0.18)	0.059	0.75	0.84
CVLT-II SD free proportion accuracy ^a	0.79 (0.17)	0.82 (0.18)	0.71 (0.22)	0.72 (0.25)	0.11	0.38	0.80
CVLT-II LD free proportion accuracy ^a	0.82 (0.18)	0.84 (0.19)	0.78 (0.21)	0.79 (0.23)	0.34	0.61	0.96
CVLT-II total across list intrusions	0.05 (0.22)	0.86 (1.31)	0.86 (1.90)	1.03 (2.47)	0.24	0.11	0.29
BVMTR T1–3 proportion accuracy	0.69 (0.21)	0.71 (0.21)	0.71 (0.11)	0.72 (0.13)	0.70	0.46	0.97
BVMTR DR proportion accuracy	0.83 (0.15)	0.84 (0.14)	0.83 (0.13)	0.83 (0.13)	0.94	0.73	0.95
Digit Span Forward ^b	10.60 (2.01)	10.20 (1.40)	10.23 (1.72)	10.70 (1.74)	0.89	0.87	0.04*
Digit Span Backward ^b	7.45 (2.42)	8.90 (2.36)	7.23 (2.33)	7.97 (2.63)	0.35	0.003**	0.31
Source memory spatial % ^c	62.50 (15.31)	59.19 (14.20)	58.15 (13.25)	67.39 (12.10)	0.54	0.34	0.047*
Source memory temporal % ^d	66.91 (15.11)	64.33 (11.43)	63.10 (13.67)	65.77 (17.52)	0.09	0.99	0.36
Source memory voice % ^d	72.79 (16.82)	77.57 (17.12)	75.60 (13.82)	73.51 (17.66)	0.89	0.61	0.20
Source memory item % ^d	94.26 (4.40)	93.24 (7.22)	92.07 (9.22)	90.87 (10.99)	0.37	0.35	0.94
MMQ contentment (/72)	39.33 (12.08)	42.38 (10.77)	40.73 (12.00)	41.73 (9.87)	0.90	0.10	0.40
MMQ ability (/80)	46.1 (9.20)	47.60 (9.28)	48.00 (10.78)	49.41 (10.72)	0.49	0.24	0.97
MMQ strategy (/76)	41.10 (8.74)	38.0 (7.21)	35.93 (10.92)	35.20 (8.31)	0.09	0.14	0.36

CVLT-II List 1–5, average proportion accuracy on Trials 1–5; CVLT-II SD, CVLT-II short delay; CVLT-II LD, CVLT-II long delay; BVMTR T1–3, average proportion accuracy on Trials 1–3; BVMTR DR, BVMTR delayed recall. MMQ: Multifactorial Memory Questionnaire; ^aSample size (*n* = 21, *n* = 29); ^bSample size (*n* = 20, *n* = 30); ^cSample size (*n* = 17, *n* = 23); ^dSample size (*n* = 17, *n* = 21). **p* < 0.05; ***p* < 0.01.

training: Those participants who achieved the longest lags on the recollection training were younger and had better working memory. While not significant, better global cognitive function (as represented by MoCA scores) also seemed to be a good predictor.

CLEAR IMPROVEMENTS ON THE TRAINING TASK

Initially, the older adults in the recollection training group made repetition errors in the training paradigm even when only a couple of items intervened between the first and second presentations of a foil. By the end of training, however, on average the group had reached a lag of 28 intervening items. These gains are commensurate with previous reports using this paradigm (Jennings and Jacoby, 2003; Jennings et al., 2005; Bailey et al., 2011). The recollection training group improved not only in the repetition error rate but also in the overall discrimination index (*d'*). Yet discrimination *also* improved in the recognition control group after training, which stands in contrast to previous reports (Jennings et al., 2005; Boller et al., 2012). Our control condition was more challenging than that used by Jennings et al. (2005), but yet similar to the one used by Boller et al. (2012) who also used a verbal recognition task with lists of equal length in both training and control conditions. Similar degree of improvement across treatment groups in the training task may be due merely to task practice effects (Verhaeghen et al., 1992; Engle et al., 1999; Owen et al., 2010).

PREDICTORS OF TRAINING GAINS

Similar to previous studies (Jennings and Jacoby, 2003; Boller et al., 2012), we found considerable variability among participants

in improvements on the recollection training paradigm (see **Figure 3**). Future work on this paradigm will benefit from our knowing which factors influence individuals' training gains. When we rank ordered all the recollection training participants by their progress through the lags, we found that those who were younger progressed further through the training. Bissig and Lustig (2007) reported that age was negatively associated with training benefits in the same training procedure, whereas verbal intelligence was positively associated with it. However, we did not find a relationship between years of education (i.e., a reasonable proxy for intelligence), and rank. Past meta-analyses of memory training have produced contradictory results when it comes to the effects of age on training gains, with some reporting a significant relationship (Yesavage et al., 1990) and others not (Gross et al., 2012). A separate model including several cognitive baseline measures (MoCA, Digit Span Forward and Backward, and CVLT-II Long Delay Free Recall) improved significantly the interpretation of rank scores over and above the influence of age and years of education. Among those cognitive measures, Digit Span Backward and MoCA were the best predictors. Although one would hope to see lower functioning people (who arguably need help the most) show greater benefits of recollection training, it appears that older adults with better cognitive status and working memory might be the ones who benefit more. This so-called "Matthew effect" (in which the cognitively rich get richer following training, and the cognitively poor do not benefit as much) is evident in other cognitive training studies in older adults (Yesavage et al., 1990; Verhaeghen et al., 1992) although it is far from universal (e.g., Belleville et al., 2006).

Table 4 | (A) Intercorrelations for Rank, MoCA, CVLT-II Long Delay Free Recall Proportion Accuracy (CVLT LD), and Digits Backward at baseline. **(B)** Hierarchical regression analysis summary for age, years of education (YOE), MoCA, CVLT-II Long Delay Free Recall (CVLT LD Free), Digits Forward, and Digits Backward at baseline predicting rank.

(A)							
	Rank	Age	YOE	MoCA	Digit Span Forward	Digit Span Backward	CVLT-II LD free
Rank	1.00	0.311*	−0.201	−0.472**	−0.049	−0.367*	−0.398*
Age		1.00	0.251	−0.155	−0.066	−0.281	−0.247
YOE			1.00	0.266	0.189	0.038	0.240
MoCA				1.00	0.263	0.227	0.336*
Digit Span Forward					1.00	0.435**	0.067
Digit Span Backward						1.00	−0.074
CVLT-II LD							1.00

* $p < 0.05$; ** $p < 0.01$.

(B)					
Model	Variable	B	SEB	Beta	p-value
1	Age	0.612	0.286	0.386	0.041
	YOE	−0.847	0.513	−0.297	0.110
2	Age	0.211	0.288	0.133	0.471
	YOE	−0.336	0.503	−0.118	0.511
	MOCA	−1.019	0.591	−0.303	0.098
	Digit Span Forward	1.271	0.909	0.248	0.176
	Digit Span Backward	−1.455	0.702	−0.385	0.049
	CVLT LD Free	−11.775	7.549	−0.280	0.132

Model 1 $R^2 = 0.18$; $R_{Adj} = 0.12$ ($N = 29$, $p = 0.069$).

Model 2 $R^2 = 0.45$; $R_{Adj} = 0.31$ ($N = 29$, $p = 0.022$).

LIMITED EVIDENCE OF TRANSFER

The more important question regarding the potential effectiveness of the training is whether we found transfer to non-trained tasks and materials. We used the recollection training paradigm in the present study because of the previous reports of transfer in older adults (Jennings and Jacoby, 2003; Jennings et al., 2005; Bailey et al., 2011). Yet, in the present study, although both groups improved significantly over time on their respective training tasks, we observed few convincing transfer effects. The only two cases in which the recollection training group improved to a greater degree than the recognition control group were in forward digit span and the spatial subtest of the source memory paradigm. In both cases, the recollection training group's scores increased slightly in the face of the recognition group's scores *decreasing* after training. We would be more confident in these effects if we had additionally found that those participants who improved to a greater degree on the recollection training task also improved to a greater degree on these two transfer tests, but this was not the case. In addition, it is puzzling why we might have found an effect of recollection training on source memory for spatial location, but not on the source memory subtests for voice or temporal context, especially given

that the recollection training required participants to remember *when* they saw a word and the task did not have a voice or spatial component. There is some evidence that temporal context memory may be more affected than spatial context memory by aging, which could cause them to be differentially affected by training (Parkin et al., 1995). Note also that *both* training groups improved significantly on Backward Digit Span. Whether this is a genuine training effect or merely a product of non-specific factors such as improvement in mood or comfort/reduction in stress, benefits of social/intellectual stimulation, and/or participants' expectations of improvement, cannot yet be determined. Digit span was considered a far-transfer measure and as such least likely to show benefits. At the same time, previous recollection training studies have shown transfer to working memory measures (Jennings et al., 2005; Bailey et al., 2011; Boller et al., 2012). This suggests that this training paradigm may tap more heavily into working memory (as opposed to episodic memory) than was originally intended. This is further supported by the fact that despite the larger sample size, we still failed to see transfer to free recall in the CVLT. In addition, Digit Span Backward was the only task that significantly predicted rank in the training group. These results, along with the

time by group interaction found with the Digit Span Forward task, may speak to the potential overlap between working memory and episodic memory processes, but at present it is too soon to tell. Further examination with experimental measures of recollection and working memory as transfer measures may be warranted to better delineate the two.

Why did we find only limited evidence of transfer in this study? Three possible explanations come to mind. First, any time one fails to find a predicted effect, the question of statistical power can be raised. Yet, we had an adequate sample size ($n = 30$) to detect large within-group recollection training effects, and large effects were reported by Jennings et al. (2005) and Bailey et al. (2011). Note also that in the current study we are not merely reporting null effects across the board; rather, we are reporting a clear dissociation between the recollection group's improved ability to reject repeated foils after training and a relative lack of change on the non-trained tests. In an attempt to balance the risks of Type I versus Type II errors in our statistical analyses of the transfer test scores, we adopted the strategy of running as few repeated-measures ANOVAs as possible (to reduce the likelihood of making a Type I error), and at the same time keeping our alpha at 0.05 (to reduce the likelihood of making a Type II error). Yet, even if we were to use a much more liberal statistical threshold (e.g., $p = 0.10$), none of the other Group \times Time interactions in **Table 3** would become significant.

Second, although we used a training schedule that was similar or identical to previous reports (Jennings and Jacoby, 2003; Jennings et al., 2005; Bailey et al., 2011; Boller et al., 2012), the participants might have shown more robust transfer effects if we had increased the intensity (i.e., the "dose") or the duration of training. Note, however, that a relative advantage of the current study was that all of our participants completed the training. Increasing the intensity, frequency, or length of training would increase the danger of a selective sample being recruited, and of at least some participants dropping out during training, which would complicate the interpretation of the data.

Third, the effects of our particular recollection training method might be restricted to specific cognitive processes. The repetition lag paradigm that we used for recollection training involved monitoring visually presented words during a yes-no recognition memory test to make a decision regarding the list membership (study versus test list) of each item. Although we employed two "near" transfer tests that we thought were very much akin to the repetition lag training task, these might not have been similar enough. It might be that only very precise test format-, modality-, and/or stimulus-specific gains should be particularly evident after training with the current protocol. For example, using this repetition lag protocol, Jennings et al. (2005) and Boller et al. (2012) *did* find a benefit of recollection training on source memory. In our study, we only found a significant transfer effect to Spatial Source Memory and these scores also served as good predictors of the gains made in the training as measured by rank. One key difference between studies may be that whereas we presented our source memory materials auditorily, the previous two studies presented theirs visually. We also used full sentences, while previous studies used single words/pictures. Similarly, in a very recent study, McDaniel et al. (2014) used Jennings

and Jacoby's (2003) recollection training procedure as one of three components of a cognitive training program. Older adults were trained with words as stimuli, but to examine transfer they completed a very similar task that incorporated repeated lures using sentences. Further, each sentence was presented in a specific context, which should have increased the contextual detail available for recollection of the stimulus. However, the authors failed to observe any transfer to this task despite its similarity to the actual training procedure. This further supports the notion that the benefits of the recollection-training procedure may have a limited ability to transfer to stimuli outside of those that were trained. If the benefits of recollection training do turn out to be relatively idiosyncratic, as found here, then the impact of this recollection training method on memory in everyday life will be limited.

Finally, on a positive note, we observed that after training, our participants did *not* alter their subjective ratings of memory Ability or Contentment on the MMQ (Troyer and Rich, 2002), which is in keeping with their lack of improvement on the objective tests they were given. The lack of change in subjective measures of memory is beneficial, since we do not want participants to *think* they have improved substantially when in fact they have not. Some cognitive training methods in the past have reported improvements in subjective memory ratings in both control and treatment groups (Berg et al., 1991). A necessity for any potential cognitive training method though is that it does not overinflate confidence.

CONCLUSION AND FUTURE DIRECTIONS

In summary, we found relatively weak transfer effects of recollection training, despite quite significant gains in performance on the training task itself. Two sets of questions must be addressed to allow progress with this (and other) cognitive training methods.

The first set of questions concerns the training itself. As outlined above, a more nuanced assessment of the processes affected by recollection training might yield clearer transfer effects. In addition, notwithstanding the problems inherent to increasing the dose or the duration of training, if one were able to use this training method over the long term, a different picture might emerge: That is, whereas the present study focused on potential *improvements* in memory performance over the relatively short term, if such training could be implemented over the long term it might help to stave off memory *decline* over time. For example, a more intensive training program with booster sessions, the ACTIVE study, has shown effects on reasoning and speed of processing that have persisted over a decade, and are associated with fewer self-reported difficulties in Instrumental Activities of Daily Living (Rebok et al., 2014), which are important for maintaining functional independence with age. If one wanted to address this question with the current training approach, drop-out could probably be attenuated by altering the paradigm (for example, by performing the training at home, rather than requiring frequent visits to the lab). Finally, combining recollection training with neurophysiological, pharmacological, or other behavioral therapies might yield clearer effects on everyday memory activities (Ranganath et al., 2011).

The second set of questions concerns who should receive (and who might benefit from) such training. In the present study,

higher-functioning older adults improved to a greater degree than did lower-functioning ones on the recollection training task itself (although this had little bearing on the transfer measure scores). Might there be a way to modify the paradigm so that it aids low-functioning older adults to a greater extent? For example, Boller et al. (2012) modified the recollection-training task to make it more suitable for Alzheimer's patients. They shortened the length of the word lists, increased the length of presentation of the words and lowered the maximum lag levels (albeit no patients reached the maximum lag level). Other modifications could include lengthening the encoding interval in a similar fashion as that done by Lustig and Flegal (2008). Modifications like these could prove to be more effective for lower functioning older adults, because they might allow the task to be more manageable and yet still challenging.

ACKNOWLEDGMENTS

This research was supported by funding from the Heart and Stroke Foundation Canadian Partnership for Stroke Recovery and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

- Bailey, H., Dagenbach, D., and Jennings, J. M. (2011). The locus of the benefits of repetition-lag memory training. *Aging Neuropsychol. Cogn.* 18, 577–593. doi: 10.1080/13825585.2011.591921
- Belleville, S., Gilbert, B., Fontaine, F., Gagnon, L., Menard, E., and Gauthier, S. (2006). Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: evidence from a cognitive intervention program. *Dement. Geriatr. Cogn. Disord.* 22, 486–499. doi: 10.1159/000096316
- Benedict, H. R. B. (1997). *Brief Visuospatial Memory Test-Revised Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 941–947. doi: 10.1037/10278-7393.27.4.941
- Berg, I. J., Koning-haanstraa, M., and Deelmana, B. G. (1991). Long-term effects of memory rehabilitation: a controlled study. *Neuropsychol. Rehabil.* 1, 97–111. doi: 10.1080/09602019108401384
- Bissig, D., and Lustig, C. (2007). Who benefits from memory training? *Psychol. Sci.* 18, 720–726. doi: 10.1111/j.1467-9280.2007.01966.x
- Boller, B., Jennings, J. M., Dieudonne, B., Verny, M., and Ergis, A. M. (2012). Recollection training and transfer effects in Alzheimer's disease: effectiveness of the repetition-lag procedure. *Brain Cogn.* 78, 169–177. doi: 10.1016/j.bandc.2011.10.011
- Boot, W. R., Simons, D. J., Stothart, C., and Stutts, C. (2013). The pervasive problem with placebos in psychology: why active control groups are not sufficient to rule out placebo effects. *Perspect. Psychol. Sci.* 8, 445–454. doi: 10.1177/1745691613491271
- Brehmer, Y., Westerberg, H., and Backman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Front. Hum. Neurosci.* 6:63. doi: 10.3389/fnhum.2012.00063
- Buschkuhl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Dapp, C., Muller, M., et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychol. Aging* 23, 743–753. doi: 10.1037/a0014342
- Cook, S. P. (2007). Are all sources equal? the roles of aging and the frontal lobes on multiple types of source memory using a repeated-measures design. *Diss. Abst. Int. B Sci. Eng.* 67, 5433.
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/s0022-5371(80)90312-6
- Davidson, P. S., Cook, S. P., McGhan, L., Bouchard, T., and Camicioli, R. (2013). Source memory in normal aging and Parkinson's disease. *J. Neuropsychol.* 7, 179–192. doi: 10.1111/jnp.12018
- Delis, D. C., Kramer, J. H., Kaplan, E., and Ober, B. A. (2001). *California Verbal Learning Test®*, 2nd Edn. New York: Psychological Corporation.
- Dobbs, A. R., and Rule, B. G. (1989). Adult age differences in working memory. *Psychol. Aging* 4, 500–503. doi: 10.1037//0882-7974.4.4.500
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331. doi: 10.1037/0096-3445.128.3.309
- Forster, J. (2012). *DMDX (Version 4.0.6.0)*. University of Arizona. Available at: <http://www.u.arizona.edu/~jforster/dmdx/>
- Glisky, E. L., and Glisky, M. L. (2008). "Memory rehabilitation in older adults," in *Cognitive Neurorehabilitation*, 2nd Edn, eds D. T. Stuss, G. Winocur, and I. H. Robertson (Cambridge: Cambridge University Press), 541–561.
- Gross, A. L., Parisi, J. M., Spira, A. P., Kueider, A. M., Ko, J. Y., Saczynski, J., et al. (2012). Memory training interventions for older adults: a meta-analysis. *Aging Ment. Health* 16, 722–734. doi: 10.1080/13607863.2012.667783
- Jacoby, L. L. (1991). A process dissociation framework – separating automatic from intentional uses of memory. *J. Mem. Lang.* 30, 513–541. doi: 10.1016/0749-596X(91)90025-F
- Jacoby, L. L. (1999). Ironic effects of repetition: measuring age-related differences in memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 3–22. doi: 10.1037/0278-7393.25.1.3
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., and Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory* 18, 394–412. doi: 10.1080/09658211003702171
- Jennings, J. M., and Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: telling effects of repetition. *Psychol. Aging* 12, 352–361. doi: 10.1037/0882-7974.12.2.352
- Jennings, J. M., and Jacoby, L. L. (2003). Improving memory in older adults: training recollection. *Neuropsychol. Rehabil.* 13, 417–440. doi: 10.1080/09602010244000390
- Jennings, J. M., Webster, L. M., Kleykamp, B. A., and Dagenbach, D. (2005). Recollection training and transfer effects in older adults: successful use of a repetition-lag procedure. *Aging Neuropsychol. Cogn.* 12, 278–298. doi: 10.1080/138255890968312
- Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *J. Cogn. Neurosci.* 9, 462–475. doi: 10.1162/jocn.1997.9.4.462
- Light, L. (2012). "Dual-process theories of memory in old age: an update," in *Memory and Aging*, eds M. Naveh-Benjamin and N. Ohta (New York: Taylor & Francis).
- Lustig, C., and Flegal, K. E. (2008). Targeting latent function: encouraging effective encoding for successful memory training and transfer. *Psychol. Aging* 23, 754–764. doi: 10.1037/a0014295
- Mandler, G. (1980). Recognizing: the judgment of previous occurrence. *Psychol. Rev.* 87, 252–271. doi: 10.1037/0033-295X.87.3.252
- Mandler, G. (2008). Familiarity breeds attempts a critical review of dual-process theories of recognition. *Perspect. Psychol. Sci.* 3, 390–399. doi: 10.1111/j.1745-6924.2008.00087.x
- McCabe, D. P., Roediger, H. L. III, and Karpicke, J. D. (2011). Automatic processing influences free recall: converging evidence from the process dissociation procedure and remember-know judgments. *Mem. Cogn.* 39, 389–402. doi: 10.3758/s13421-010-0040-5
- McDaniel, M. A., Binder, E. F., Bugg, J. M., and Waldum, E. R., Dufault, C., Meyer, A., et al. (2014). Effects of cognitive training with and without aerobic exercise on cognitively demanding everyday activities. *Psychol. Aging* 29, 717–730. doi: 10.1037/a0037363
- Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x
- Nieuwenhuis, S., Forstmann, B. U., and Wagenmakers, E. J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat. Neurosci.* 14, 1105–1107. doi: 10.1038/nn.2886
- Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A., et al. (2010). Putting brain training to the test. *Nature* 465, U775–U776. doi: 10.1038/nature09042
- Parkin, A. J., Walter, B. M., and Hunkin, N. M. (1995). Relationships between normal aging, frontal lobe function, and memory for temporal and spatial information. *Neuropsychology* 9, 304–312. doi: 10.1037/0894-4105.9.3.304
- Petrides, M., and Milner, B. (1982). Deficits on subject-ordered tasks after frontal-lobe and temporal-lobe lesions in man. *Neuropsychologia* 20, 249. doi: 10.1016/0028-3932(82)90100-2

- Quamme, J. R., Yonelinas, A. P., Widaman, K. F., Kroll, N. E., and Sauve, M. J. (2004). Recall and recognition in mild hypoxia: using covariance structural modeling to test competing theories of explicit memory. *Neuropsychologia* 42, 672–691. doi: 10.1016/j.neuropsychologia.2003.09.008
- Ranganath, C., Flegal, K. E., and Kelly, L. L. (2011). Can cognitive training improve episodic memory? *Neuron* 72, 688–691. doi: 10.1016/j.neuron.2011.10.022
- Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H. Y., King, J., et al. (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. *J. Am. Geriatr. Soc.* 62, 16–24. doi: 10.1111/jgs.12607
- Schein, R. L., and Koenig, H. G. (1997). The center for epidemiological studies depression (CES-D) scale: assessment of depression in the medically ill elderly. *Int. J. Geriatr. Psychiatry* 12, 436–446. doi: 10.1002/(SICI)1099-1166(199704)12:4<436::AID-GPS499>3.0.CO;2-M
- Snodgrass, J. G., and Corwin, J. (1988). Pragmatics of measuring recognition memory – applications to Dementia and Amnesia. *J. Exp. Psychol. Gen.* 117, 34–50. doi: 10.1037//0096-3445.117.1.34
- Stanovich, K. E. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21, 360–407. doi: 10.2307/747612
- Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *J. Int. Neuropsychol. Soc.* 8, 448–460. doi: 10.1017/S1355617702813248
- Troyer, A. K., and Rich, J. B. (2002). Psychometric properties of a new metamemory questionnaire for older adults. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 57, 19–27. doi: 10.1093/geronb/57.1.P19
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 10.1146/annurev.psych.53.100901.135114
- Verhaeghen, P., Marcoen, A., and Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: a meta-analytic study. *Psychol. Aging* 7, 242–251. doi: 10.1037//0882-7974.7.2.242
- Waldron-Perrine, B., and Axelrod, B. N. (2012). Determining an appropriate cutting score for indication of impairment on the Montreal Cognitive Assessment. *Int. J. Geriatr. Psychiatry* 27, 1189–1194. doi: 10.1002/gps.3768
- Wechsler, D. (1997). *Technical Manual for the Wechsler Adult Intelligence Test*, 3rd Edn. San Antonio: The Psychological Corporation.
- Yesavage, J. A., Sheikh, J. I., Friedman, L., and Tanke, E. (1990). Learning mnemonics-roles of aging and subtle cognitive impairment. *Psychol. Aging* 5, 133–137. doi: 10.1037//0882-7974.5.1.133
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* 46, 441–517. doi: 10.1006/jmla.2002.2864

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 June 2014; accepted: 21 October 2014; published online: 18 November 2014.

Citation: Stamenova V, Jennings JM, Cook SP, Walker LAS, Smith AM and Davidson PSR (2014) Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer. *Front. Hum. Neurosci.* 8:898. doi: 10.3389/fnhum.2014.00898

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2014 Stamenova, Jennings, Cook, Walker, Smith and Davidson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Corrigendum: Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer

Vessela Stamenova^{1*}, Janine M. Jennings², Shaun P. Cook³, Lisa A. S. Walker⁴, Andra M. Smith⁴ and Patrick S. R. Davidson⁴

¹ Rotman Research Institute, Baycrest – University of Toronto, Toronto, ON, Canada, ² Department of Psychology, Wake Forest University, Winston-Salem, NC, USA, ³ Department of Psychology, Millersville University, Millersville, PA, USA, ⁴ School of Psychology, Faculty of Social Sciences, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada

Keywords: aging, memory, rehabilitation, recollection, familiarity

A corrigendum on

Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer

by Stamenova, V., Jennings, J. M., Cook, S. P., Walker, L. A., Smith, A. M., and Davidson, P. S. R. (2014). *Front. Hum. Neurosci.* 8:898. doi: 10.3389/fnhum.2014.00898

Due to errors that were noticed recently in the California Verbal Learning Test-II (CVLT-II) scoring software version 1.0.0 with the software occasionally mis-counting items, we have recalculated all CVLT scores and re-run all relevant analyses. This update was completed with a software update (to version 1.0.2) (see <http://pearsonassessmentsupport.com/support/index.php?View=entry&EntryID=741>).

The majority of the statistical effects remain the same as in the original publication, except for two.

These are:

Section: *Predictors of training gains*

1. For the regression, the initial analysis showed that adding the additional variables led to a significant increase in prediction. Now this result is *marginally* significant, at $p = 0.056$. The initial increase was by 45% and now it is by 44%.
2. In the same regression, Model 2 Digit Span Backward was initially statistically significant at $p = 0.049$, but after corrections now $p = 0.058$. We state that this effect is *marginally* significant. Because of this initially significant effect, in a partial correlation conducted later in section entitled: *Do those participants who show the greatest gains in training also show the greatest improvements on the transfer tasks?*, we controlled for Digit Span backwards scores. We chose to still control for these scores, given that the significant test is so close to the cut-off.

No other effect changes were observed in any of our analyses involving the CVLT. The values, however, have changed slightly and we have updated those values in **Tables 3, 4** of the manuscript.

Overall, given the changes affect only one of our outcome measures and the fact that the results with this outcome measure have changed only for two statistical tests, we believe that the correction does not affect the scientific validity of the results.

OPEN ACCESS

Edited by:

Guido P. H. Band,
Leiden University, Netherlands

Reviewed by:

Kristin Flegel,
University of Glasgow, Scotland

*Correspondence:

Vessela Stamenova
vstamenova@research.baycrest.org

Received: 04 November 2015

Accepted: 19 November 2015

Published: 01 December 2015

Citation:

Stamenova V, Jennings JM, Cook SP, Walker LAS, Smith AM and Davidson PSR (2015) Corrigendum: Training recollection in healthy older adults: clear improvements on the training task, but little evidence of transfer. *Front. Hum. Neurosci.* 9:658. doi: 10.3389/fnhum.2015.00658

TABLE 3 | Pre- and post-training scores on transfer measures.

	Recognition control		Recollection training		<i>p</i> -value		
	Pre <i>M (SD)</i>	Post <i>M (SD)</i>	Pre <i>M (SD)</i>	Post <i>M (SD)</i>	Group	Time	Group × Time
CVLT-II List 1-5 proportion accuracy ^a	0.77 (0.10)	0.76 (0.13)	0.70 (0.15)	0.72 (0.12)	0.11	0.86	0.49
CVLT-II SD free proportion accuracy ^a	0.80 (0.16)	0.81 (0.18)	0.71 (0.22)	0.75 (0.21)	0.15	0.23	0.43
CVLT-II LD free proportion accuracy ^a	0.76 (0.24)	0.81 (0.26)	0.76 (0.21)	0.78 (0.23)	0.86	0.50	0.50
CVLT-II total across list intrusions	0.01 (0.02)	0.05 (0.09)	0.05 (0.12)	0.06 (0.15)	0.24	0.15	0.38
BVMT-R T1-3 proportion accuracy	0.69 (0.21)	0.71 (0.21)	0.71 (0.11)	0.72 (0.13)	0.70	0.46	0.97
BVMT-R DR proportion accuracy	0.83 (0.15)	0.84 (0.14)	0.83 (0.13)	0.83 (0.13)	0.94	0.73	0.95
Digit Span Forward ^b	10.60 (2.01)	10.20 (1.40)	10.23 (1.72)	10.70 (1.74)	0.89	0.87	0.04*
Digit Span Backward ^b	7.45 (2.42)	8.90 (2.36)	7.23 (2.33)	7.97 (2.63)	0.35	0.003**	0.31
Source memory spatial % ^c	62.50 (15.31)	59.19 (14.20)	58.15 (13.25)	67.39 (12.10)	0.54	0.34	0.047*
Source memory temporal % ^d	66.91 (15.11)	64.33 (11.43)	63.10 (13.67)	65.77 (17.52)	0.09	0.99	0.36
Source memory voice % ^d	72.79 (16.82)	77.57 (17.12)	75.60 (13.82)	73.51 (17.66)	0.89	0.61	0.20
Source memory item % ^d	94.26 (4.40)	93.24 (7.22)	92.07 (9.22)	90.87 (10.99)	0.37	0.35	0.94
MMQ contentment (/72)	39.33 (12.08)	42.38 (10.77)	40.73 (12.00)	41.73 (9.87)	0.90	0.10	0.40
MMQ ability (/80)	46.1 (9.20)	47.60 (9.28)	48.00 (10.78)	49.41 (10.72)	0.49	0.24	0.97
MMQ strategy (/76)	41.10 (8.74)	38.0 (7.21)	35.93 (10.92)	35.20 (8.31)	0.09	0.14	0.36

CVLT-II List 1-5, Average proportion accuracy on Trials 1-5; CVLT-II SD, CVLT-II Short Delay; CVLT-II LD, CVLT-II Long Delay; BVMT-R T1-3, Average proportion accuracy on Trials 1-3; BVMT-R DR, BVMT-R Delayed Recall. MMQ, Multifactorial Memory Questionnaire.

^aSample size (*n* = 21, *n* = 29).

^bSample size (*n* = 20, *n* = 30).

^cSample size (*n* = 17, *n* = 23).

^dSample size (*n* = 17, *n* = 21).

p* < 0.05; *p* < 0.01.

TABLE 4 | Intercorrelations for Rank, MoCA, CVLT-II Long Delay Free Recall Proportion Accuracy (CVLT LD), and Digits Backward at baseline.

	Rank	Age	YOE	MoCA	Digit Span Forward	Digit Span Backward	CVLT-II LD Free
Rank	1.00	0.311*	−0.201	−0.472**	−0.049	−0.367*	−0.392*
Age		1.00	0.251	−0.155	−0.066	−0.281	−0.225
YOE			1.00	0.266	0.189	0.038	0.259
MoCA				1.00	0.263	0.227	0.353*
Digit Span Forward					1.00	0.435**	0.058
Digit Span Backward						1.00	−0.059
CVLT-II LD							1.00

**p* < 0.05, ** *p* < 0.01.

Hierarchical Regression Analysis summary for Age, Years of Education (YOE), MoCA, CVLT-II Long Delay Free Recall (CVLT LD Free), Digits Forward, and Digits Backward at baseline predicting rank.

Model	Variable	B	SEB	Beta	<i>p</i> -value
1	Age	0.612	0.286	0.386	0.041
	YOE	−0.847	0.513	−0.297	0.110
2	Age	0.236	0.287	0.149	0.421
	YOE	−0.346	0.509	−0.122	0.503
	MOCA	−1.024	0.600	−0.305	0.101
	Digit span forward	1.231	0.917	0.240	0.192
	Digit span backward	−1.403	0.703	−0.371	0.058
	CVLT LD Free	−10.800	7.627	−0.256	0.170

Model 1. $R^2 = 0.18$, $R\text{ Adj} = 0.12$ (*N* = 29, *p* = 0.069).

Model 2. $R^2 = 0.44$, $R\text{ Adj} = 0.29$ (*N* = 29, *p* = 0.026).

Please see below the relevant changes made, with bold font for the edited or inserted text.

In section “Predictors of Training Gains,”

The regression results are summarized in **Table 4**; Multiple R for the first block of regressors (age and YOE) was close to statistical significance, $F_{(2, 29)} = 2.96$, $p = 0.069$; multiple R for the next block of regressors was significant $F_{(6, 29)} = 3.00$, **$p = 0.026$** . The demographic variables (Age and YOE) explained 18% of the variance, while adding the cognitive status scores in block 2 of the analysis increased the amount of variability explained to **44%**. This increase is **marginally** significant by the F change test $F_{(4, 23)} = 2.66$, **$p = 0.056$** . Among the demographic variables, only age was significant, while among the cognitive status, Digit Span Backwards (**$p = 0.058$**) was **marginally** statistically significant, followed by the MoCA ($p = 0.101$), which was marginally significant.

In section “Do those participants who show the greatest gains in training also show the greatest improvements on the transfer tasks?,” last sentence:

There was, however, a significant correlation between the change in Spatial Source Memory and rank, $r = -0.37$, **p (one-tailed) = 0.048, $df = 19$** .

Please see below relevant changes made to **Tables 3, 4**.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Stamenova, Jennings, Cook, Walker, Smith and Davidson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches

Erin C. Connors¹, Elizabeth R. Chrastil², Jaime Sánchez³ and Lotfi B. Merabet^{1*}

¹ The Laboratory for Visual Neuroplasticity, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA, USA

² Department of Psychology, Center for Memory and Brain, Boston University, Boston, MA, USA

³ Department of Computer Science, Center for Advanced Research in Education, University of Chile, Santiago, Chile

Edited by:

Heleen A. Slagter, University of Amsterdam, Netherlands

Reviewed by:

Michael Proulx, University of Bath, UK

Sara Haber, University of Texas at Dallas, USA

Achille Pasqualotto, Sabanci University, Turkey

*Correspondence:

Lotfi B. Merabet, Massachusetts Eye and Ear Infirmary, Harvard Medical School, 20 Staniford Street, Boston, MA 02114, USA
e-mail: lotfi_merabet@meei.harvard.edu

For profoundly blind individuals, navigating in an unfamiliar building can represent a significant challenge. We investigated the use of an audio-based, virtual environment called Audio-based Environment Simulator (AbES) that can be explored for the purposes of learning the layout of an unfamiliar, complex indoor environment. Furthermore, we compared two modes of interaction with AbES. In one group, blind participants implicitly learned the layout of a target environment while playing an exploratory, goal-directed video game. By comparison, a second group was explicitly taught the same layout following a standard route and instructions provided by a sighted facilitator. As a control, a third group interacted with AbES while playing an exploratory, goal-directed video game however, the explored environment did not correspond to the target layout. Following interaction with AbES, a series of route navigation tasks were carried out in the virtual and physical building represented in the training environment to assess the transfer of acquired spatial information. We found that participants from both modes of interaction were able to transfer the spatial knowledge gained as indexed by their successful route navigation performance. This transfer was not apparent in the control participants. Most notably, the game-based learning strategy was also associated with enhanced performance when participants were required to find alternate routes and short cuts within the target building suggesting that a ludic-based training approach may provide for a more flexible mental representation of the environment. Furthermore, outcome comparisons between early and late blind individuals suggested that greater prior visual experience did not have a significant effect on overall navigation performance following training. Finally, performance did not appear to be associated with other factors of interest such as age, gender, and verbal memory recall. We conclude that the highly interactive and immersive exploration of the virtual environment greatly engages a blind user to develop skills akin to positive near transfer of learning. Learning through a game play strategy appears to confer certain behavioral advantages with respect to how spatial information is acquired and ultimately manipulated for navigation.

Keywords: early blind, late blind, navigation, spatial cognition, games for learning, videogames, virtual environment, near transfer of learning

INTRODUCTION

Considerable interest has arisen regarding the use of virtual reality environments and video games for education, rehabilitation, as well as mental fitness training (Mayo, 2009; Bavelier et al., 2011, 2012; Lange et al., 2012). Moreover, simulation-based training combined with ludic-based approaches for learning have been associated with behavioral gains including the development and reinforcement of sensory, motor, and cognitive skills that might otherwise be more difficult, or even too dangerous, to learn under more typical training settings (e.g., Kuppersmith et al., 1996; Pataki et al., 2012; Rizzo et al., 2012). It has been proposed that realistic and immersive virtual environments allow individuals the opportunity to interact with objects and events in novel and meaningful ways, acquire relevant contextual

information, and “integrate knowledge by doing” (Shaffer et al., 2005). Furthermore, the open structure and self-directed discovery of information inherent in these virtual settings improves contextual learning and the transfer of situational knowledge (Shaffer et al., 2005; Dede, 2009). Thus, successfully leveraging these advantages in the education and rehabilitation arenas could have immense appeal by facilitating the learning of demanding tasks and the transfer of acquired skills.

While the exploration of virtual environments and video game play are typically ascribed to the visual modality, another potential application of these approaches could be for the training and rehabilitation of individuals with profound blindness. Blind individuals typically undergo formal instruction referred to as orientation and mobility (O&M) training as a means of learning

how to navigate independently through the environment. Unlike the sighted, blind individuals must rely on other sensory channels (such as hearing, touch, and proprioception Thinus-Blanc and Gaunet, 1997) to gather relevant spatial information for orientating, route planning, and path execution (Strelow, 1985; Ashmead et al., 1989; Loomis et al., 1993; Long and Giudice, 2010). The resultant mental representation of the surrounding space is referred to as a spatial cognitive map (Strelow, 1985), and generating an accurate and robust mental map is considered essential for efficient travel (Siegel and White, 1975; Blasch et al., 1997). Not surprisingly, situations where the environment is particularly complex or unfamiliar (or when familiar routes are no longer accessible) can represent a significant challenge when navigating without the benefit of sight. Certainly, many technical advancements and assistive devices have been developed to help blind individuals (including sensory substitution devices, digital maps, and GPS based systems) (e.g., Petrie et al., 1996; Loomis et al., 2005; Johnson and Higgins, 2006; Giudice et al., 2007; Kalia et al., 2010; Chebat et al., 2011; see also Giudice and Legge, 2008 for review). However, many of these approaches are difficult to learn, may require modifications to existing infrastructure, or are not readily adaptable to all situations. Moreover, from a learning and training standpoint, assistive devices are not typically designed for the purposes of training navigation skills.

Based on these observations, we developed an audio-based, virtual environment called Audio-based Environment Simulator (AbES) that can be explored to access contextually relevant spatial information for the purposes of surveying and learning the layout of an unfamiliar complex indoor environment. Key to this user-centered approach is the dynamic and interactive manner in which the spatial information is acquired, which engages the user to construct a spatial cognitive map of a designated space. The contextually relevant spatial information acquired can then be used for the purposes of navigation once the user arrives at the physical environment. This training strategy is comparable to the concept of near transfer of learning, which presupposes that there is a contextual overlap between the training and transfer settings, and that the training content is relevant to the task in question (Cormier and Hagman, 1987).

As a proof of concept (Merabet et al., 2012), we previously demonstrated that early and profoundly blind participants (i.e., documented prior to the age of three) who interacted with AbES were able to create an accurate spatial mental representation that corresponded to the spatial layout of an existing physical building. Furthermore, self-directed exploration carried out within a context of a video game metaphor allowed for the transfer of acquired spatial information for the purpose of navigating through an environment with which they were previously unfamiliar (Merabet et al., 2012).

Here, we present the results of a larger-scale study aimed at comparing the development and transfer of spatial information learned through self-directed game play with a structured, didactic approach. Specifically, we compared the exploration and learning of a virtual environment through either: (1) self-directed exploration and implicit learning under the pretext of a video game metaphor or (2) directed instruction and explicit learning with the aid of a sighted facilitator. Finally, as a control condition,

we also compared performance in a subset of participants who learned the spatial layout of a virtual environment following the same video game metaphor. However, in this latter condition, the virtual environment did not correspond to the target physical environment.

Employing this study design allowed for a direct comparison between the mode of interaction (i.e., self-directed, implicit learning through gaming vs. guided instruction, explicit learning through directed navigation) as well as controlling for the effect of contextual information (i.e., playing in a corresponding environment vs. non corresponding environment). By assessing the transfer of spatial information acquired from the exploration of the virtual environment, participants' spatial knowledge regarding the layout of the target building could be ascertained objectively. Given that the participants were completely unfamiliar with the layout of the target building and further, they were never explicitly trained to navigate the routes that were ultimately tested, we would interpret successful navigation performance as evidence of positive near transfer of learning. As a secondary goal, we also investigated the potential association of a number of factors of interest on navigation performance, including prior visual experience, age, gender, and verbal memory ability. The results from a subset of individuals (three) participating in our pilot study (Merabet et al., 2012) were incorporated into the analysis presented here.

Based on the possibility that cognitive behavioral gains may result from video game play, we hypothesized that participants who learned the environment through self-directed exploration would show evidence of transfer of spatial knowledge greater than or equal to those who were explicitly taught the environment through directed navigation. Further, these cognitive gains (i.e., near transfer of learning) would only arise if exploration occurred within an environment that corresponded to the target building where the acquired skills would ultimately be assessed. Secondly, given previous accounts suggesting that prior visual experience may have a beneficial effect on the ability to mentally represent surrounding space (Ashmead et al., 1989), we hypothesized that late blind participants (i.e., individuals having greater prior visual experience) would show a behavioral advantage compared to their early blind counterparts.

METHODS

PARTICIPANTS AND STUDY DESIGN

Thirty-eight profoundly blind individuals aged between 18 and 45 years (mean age 27.92 years \pm 8.51 SD; 20 males, 18 females) participated in the study. Blindness was defined as residual visual function no greater than perceived light perception, hand motion, color, or shadows. The etiology of blindness varied across participants, however all were of ocular related cause (e.g., retinitis pigmentosa, glaucoma, Leber's congenital amaurosis, retinopathy of prematurity). We defined early blind as documented profound blindness acquired prior to the age of three (i.e., typically prior to the development of high level language function and the retention of vivid visual memories). While the majority of the participants had diagnoses that could be considered of congenital cause, we relied on evidence of profound blindness based on documented visual functional assessment. In contrast, late blind was defined

as blindness acquired after the age of 14. In this latter group, profound vision loss occurred well after the development of high level language function and all participants had prior visual memories based on self-report. All participants had no other neurological or health concerns, had self-reported normal spatial hearing, and had received formal O&M training prior to participating in the study albeit with varying degrees of experience (mean $8.33 \text{ years} \pm 8.24 \text{ SD}$; see **Table 1**). The majority of the participants were right handed (based on self-report) but all used their right hand to operate the control keys of the software. Five additional participants were excluded from the study prior to any testing due to personal and/or medical reasons (unrelated to their participation in the study) and thus were unable to complete the behavioral assessments. All participants provided written informed consent in accordance with procedures approved by the investigative review board of the Massachusetts Eye and Ear Infirmary (Boston, MA, USA) and all training and performance assessments were carried out at the Carroll Center for the Blind (Newton, MA, USA). As participants had varying levels of residual visual function (e.g., hand motion, light perception), all wore a blindfold throughout the training and behavioral assessment sessions so as to eliminate the possibility of relying on any visual related cues. Participants were allowed to use their cane as a mobility aid during the behavioral assessments if they chose.

As potential factors of interest associated with navigation performance, we also collected age, gender, and verbal memory ability (**Table 1**). Verbal memory recall was assessed using the Wechsler Memory Scale; Third Edition (WAIS-III) Word List Test. For details regarding this assessment see WAIS-III (1997).

Using a stratified randomization strategy (i.e., based on early or late blind status), participants were relegated to one of three experimental groups; (1) gamers ($n = 15$) (2) directed navigators ($n = 14$), or (3) control ($n = 7$) (**Figure 1**). Training included three, 30-min sessions (for a total of 90 min) plus an initial familiarization period (roughly 10 min) to learn the key strokes and corresponding audio cues used in the software. For participants in the gamer and control groups, the rules and goals associated with game play were also presented. Prior to enrollment, we verified that all study participants were completely unfamiliar with the layout of target building (by formal questioning) as well as

to the overall purpose of the study. This was necessary to minimize any potential confounds related to expectation bias and prior experience on the assessments of performance.

Participants in the gamer group interacted with AbES within the context of a first-person video game designed to promote the full exploration of the virtual environment (**Figure 2A**). Following a goal-directed strategy, the game's premise is to explore the entire virtual building in order to collect as many jewels as possible (randomly hidden in various rooms) while avoiding roving monsters that are programmed to take away the jewels and hide them in other locations (**Figure 2B**). Once a jewel

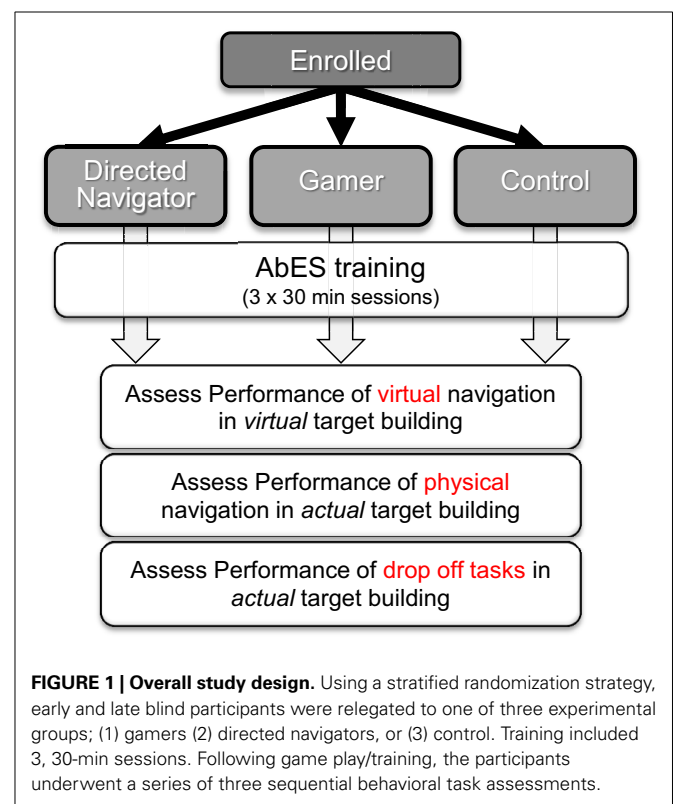


Table 1 | Participant characteristics separated by study subgroups.

Group	Sub group	Gender	Age (years \pm SD)	Verbal memory (score \pm SD)	O&M (years \pm SD)
Gamers	Early blind	4 m/4 f	24.00 \pm 6.76	81.13 \pm 15.29	14.25 \pm 7.78
	Late blind	5 m/2 f	29.86 \pm 7.34	72.43 \pm 18.16	1.0 \pm 0.00*
	Subgroup totals/means	9 m/6 f	26.73 \pm 7.42	77.07 \pm 16.69	8.07 \pm 8.78
Directed navigators	Early blind	4 m/3 f	28.29 \pm 9.05	87.86 \pm 6.82	10.43 \pm 9.91
	Late blind	4 m/3 f	31.71 \pm 10.45	83.43 \pm 11.49	1.0 \pm 0.00*
	Subgroup totals/means	8 m/6 f	30.00 \pm 9.56	85.64 \pm 9.36	5.71 \pm 8.32
Controls	Early blind	2 m/2 f	21.00 \pm 2.94	88.75 \pm 7.37	14.25 \pm 1.71
	Late blind	0 m/3 f	33.33 \pm 9.71	89.33 \pm 6.03	14.00 \pm 5.29
	Subgroup totals/means	2 m/5 f	26.29 \pm 8.90	89.00 \pm 6.27	14.14 \pm 3.29
Totals/means		20 m/18 f	27.92 \pm 8.51	82.72 \pm 13.25	8.33 \pm 8.24

*No more than 1 year of O&M experience.

is found, the player must remove it (i.e., bring it outside the building using one of three possible exits) before searching for the next jewel. The participants were encouraged to collect as many jewels as possible, but they were never instructed at any time to recall the spatial layout of the building while playing the game.

In comparison, participants relegated to the second “directed navigator” group were explicitly taught the spatial layout of the entire building using AbES through a series of pre-determined paths and with the assistance of a sighted facilitator (**Figure 2C**). Training involved a complete step-by-step instruction of the building layout such that all the room locations, exits, and landmarks were encountered in a serial and repeated fashion along the interior perimeter (following a clockwise direction) similar to a “shoreline” exploration strategy. The paths followed were representative of a virtual recreation of an O&M lesson and the instructions given by a professional O&M instructor for the purposes of learning the spatial layout of the target building.

Finally, participants randomized to the third control group interacted with the AbES software under the same self-directed exploratory strategy as the gamer group. However, in contrast to the gamer group, the virtual environment explored did not correspond to the target physical building (**Figure 2D**). As with the gamer group, the control participants were never instructed to recall the spatial layout of the building while playing the game.

SOFTWARE

The AbES software was developed using C++ programming language with Visual Studio.NET and framework 2.0 on a PC computer (Windows XP/7 operating system). The software runs using a 10 Mb HD, 1 Gb RAM Pentium processor using a standard laptop computer and sound card. Based on the original architectural floor plan of an existing building (located at the Carroll Center for the Blind; Newton, MA, USA), the rendered indoor virtual environment includes 23 rooms, a series of connecting corridors, three separate entrances and two stairwells (see **Figures 2B,C**). This building was selected by design given that it is physically removed from the main campus facilities and was not normally accessible to the clients of the Center. The design specifics of this user-centered audio-based interface have been described in detail elsewhere (see Connors et al., 2013). Briefly, using simple keystrokes, the user-centered software allows an individual to explore the virtual environment (space bar for moving forward, “L” for right, “J” for left, and “K” to open doors) and survey the spatial layout of the building. The “F” key could also be used to identify the individual’s location at any time. Scaling of the virtual environment is such that each virtual step approximates one step in the real physical building. While moving through the environment, contextual auditory and spatial information is acquired sequentially and is continuously updated, allowing the

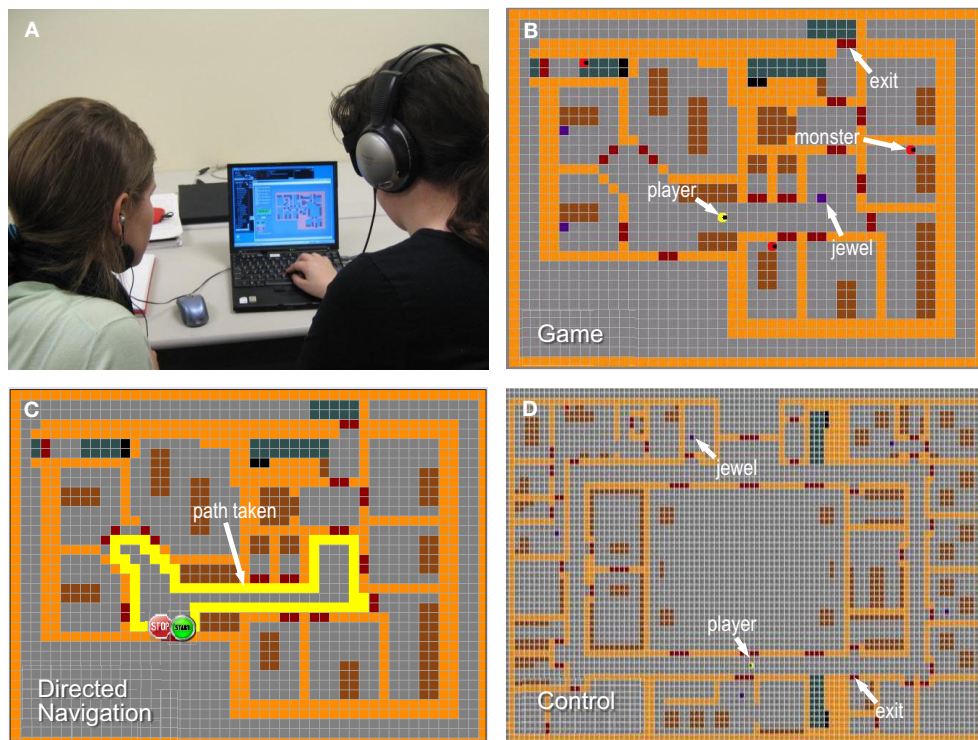


FIGURE 2 | Virtual rendering of an existing two story building (for simplicity, only the first floor is shown) represented in the AbES software used. (A) Blind participants (right) interacting with the AbES software while a facilitator (left) looks on. **(B)** In gamer mode, the user (yellow icon) navigates through the virtual environment using auditory cues to locate hidden jewels (blue squares) and avoid being caught by roving monsters (red

icons). In directed navigation mode **(C)**, the user learns the spatial layout of the building and the relative location of the rooms using a series of predetermined paths (shown in yellow) with the assistance of a facilitator (for simplicity, only one path is shown here). **(D)** For the control group, the user played in a virtual environment that did not correspond to the target building.

user to build a corresponding mental representation of the building's spatial layout. Spatial and situational information are based on iconic and spatialized sound cues provided after each step and updated to match the user's egocentric heading. For example, if a door is located on the user's right side, a door knocking sound is heard in the user's right ear. Conversely, if the user turns around 180° so that the same door is now located on their left side, the sound is heard in the left channel. Finally, if the user is facing the door, the same knocking sound is heard in both ears. Orientation is based on cardinal compass headings (e.g., "north" or "east") and text to speech (TTS) is used to provide further information regarding a user's current location, orientation and heading (e.g., "you are in the corridor, on the first floor, facing west") as well as the identity of objects and obstacles in their path (e.g., "this is a wall"). Distance cues are provided by modulating sound intensity (e.g., the sound of a nearby jewel increases as it is approached, pitch increases as the user walks up a flight of stairs). In this manner, the software plays an appropriate audio file as a function of the user's location and orientation and keeps track of their position as they move through the environment.

BEHAVIORAL TESTING

All participants interacted with AbES for the same amount of time (total of 90 min spread over three training sessions) regardless of the group they were relegated to. Following game play/training, all participants underwent a series of three sequential behavioral task assessments. These tasks were designed to evaluate their ability to transfer the spatial information while navigating within the virtual representation and corresponding physical environment modeled in AbES. The target paths used in the navigation assessments were never explicitly taught to any of the groups during the training/game play period.

A series of stop rules were implemented using criteria determined from pilot testing and performance assessments carried out prior to commencing the study. First, subjects were not allowed more than 6 min to carry out any given route task. If the participant was unable to complete the route task in the allotted time, a score of zero was given and the full 6 min was scored as the time taken. This upper time limit was defined as twice the standard deviation (SD) collected from the mean navigation times observed during pilot testing (thus by definition, a time greater than 6 min would be interpreted as an outlier response). Second, subjects were required to complete at least three out of the first five tested paths (either successfully or unsuccessfully, but within the designated time limit) for a given series of navigation tasks in order to proceed with the entire behavioral evaluation. These stop rules served two purposes. First, setting an upper limit on exploration time helped ensure that performance would be comparable across runs, tasks, and individuals, thus allowing for a more direct comparison and statistical analysis of performance. Second, from an ethical standpoint, enforcing stop rules would ensure participants would not be required to continue if their initial performance on a given behavioral assessment was too poor. As these participants were viewed as psychologically at-risk, it was deemed crucial to maintain their overall well-being and remain vigilant to any situation that may be perceived as exacerbating their sense of failure or personal frustration. Apart from

the individuals relegated to the control group (see results below), these criteria were met by all the participants randomized to the gamer and directed navigator arms of the study.

Virtual navigation

In the virtual navigation task, participants were instructed to complete a series of 10 predetermined paths in the virtual environment modeled in AbES. The paths used were a series of start and stop locations (i.e., rooms) and were all of comparable length and complexity in terms of path length and number of turns. The start-stop location pairs were loaded into the AbES software and presented automatically following a randomized order for each subject. Task success and navigation time were automatically scored and data was output to a text file for further analysis. Primary outcome measures included whether the participant was able to successfully complete the 10 navigation routes (i.e., number of correct paths expressed as percent correct) and the time taken (seconds) to reach the target.

Physical navigation

Following the first task assessment, participants were then taken to the physical building modeled in the AbES software and navigation performance was again assessed using a series of 10 predetermined routes of comparable length and complexity. Similar to the previous task assessment, participants were instructed to navigate a set of 10 predetermined targets presented in random order. Navigation performance was recorded by an experienced investigator following behind the study participant. Using a stopwatch, timing commenced once the subject took their first step and stopped when the subject verbally reported that they were in front of the door of the target destination. Navigation success (number of correct paths expressed as percent correct) and time to target (seconds) were collected.

Drop off

Finally, for the drop off task, participants were placed at five predetermined locations and instructed to exit the building using the shortest path possible relative to their starting point. To successfully carry out this task, subjects had to mentally choose one out of three possible exits relative to their starting position and navigate the route leading to that exit. Paths were scored such that the shortest possible path was given maximum points (i.e., three for the shortest path, two for the second, one for the longest, and no points for not being able to complete the task in the time allotted). Navigation time (seconds) was also collected.

While there was no direct measure of chance performance for these navigation tasks, it is important to note that there were 23 possible destinations (target rooms) from any given starting point. Further, scoring was based on the participant's first verbal response (in the case of the physical navigation tasks) once they reported arriving to the intended target. This latter rule was to ensure that subjects could not change their response once it was given, or give multiple responses within the allotted time in the hopes of arriving to the correct destination. Feedback on participants' navigation performance was not provided during the assessments.

DATA ANALYSIS

All data were analyzed using SPSS statistical software package (IBM, version 20). Three-Way ($2 \times 2 \times 2$) ANOVAs including condition (game/navigators) \times visual experience (early/late) \times gender (male/female) were performed for each outcome measure. *Post-hoc* tests were performed between groups following tests for interaction. We report mean and *SD* values with a significance level set at $p < 0.05$. Measures of association between the primary outcome of interest (i.e., percentage correct on physical navigation task) and additional factors (i.e., age, and verbal memory ability) were calculated using the Pearson product-moment correlation coefficient.

RESULTS

Overall, all participants were able to successfully interact with the AbES software. Following the navigation assessments, subjects in both the gamer and directed navigator groups demonstrated performance consistent with positive near transfer of learning. Specifically, these subjects applied the spatial information they acquired to successfully carry out navigation tasks within the virtual as well as physical environment modeled in AbES.

In contrast, it is important to note that this transfer of learning was not evidenced in the participants (both early and late blind) relegated to the control group of the study. That is, following game play in a non-contextual environment, these two subjects were unable to complete any of the navigation tasks successfully. Specifically, all subjects failed to find any target destinations on the first five paths tested (and timed out on each run) on all three of the navigation task assessments (i.e., virtual, physical, and drop off). As these control participants were unable to carry out any of the navigation task assessments, we interpreted their performance following training in the control arm as being functionally zero.

COMPARING PERFORMANCE IN GAMERS AND DIRECTED NAVIGATORS—ARRIVING TO TARGET

Evidence of transfer of learning on all the navigation task assessments was observed in both the gamer and directed navigator groups as well as in both early and late blind participants.

As a first level of analysis, Three-Way ANOVAs (2 condition \times 2 visual experience \times 2 gender) were used to confirm the effectiveness of the randomization procedure across groups. There were no differences between groups in terms of age [condition: $F_{(1, 27)} = 0.888$, $p = 0.357$, $\eta_p^2 = 0.041$; visual experience: $F_{(1, 27)} = 1.516$, $p = 0.232$, $\eta_p^2 = 0.067$; gender: $F_{(1, 27)} = 0.122$, $p = 0.730$, $\eta_p^2 = 0.006$]. No significant differences in verbal memory ability (Wechsler score) were apparent between gamers and directed navigators [$F_{(1, 27)} = 2.823$, $p = 0.108$, $\eta_p^2 = 0.119$], as well as the interaction of condition \times gender [$F_{(1, 27)} = 3.4230$, $p = 0.087$, $\eta_p^2 = 0.133$]. As expected, the early blind group had significantly more O&M experience than the late blind group [$F_{(1, 27)} = 18.245$, $p < 0.001$, $\eta_p^2 = 0.465$], but there was no statistical difference between gamers and directed navigators [$F_{(1, 29)} = 0.549$, $p = 0.467$, $\eta_p^2 = 0.025$].

Comparing performance on virtual navigation task, a Three-Way ANOVA (2 condition \times 2 visual experience \times 2 gender) revealed no significant main effects for condition [$F_{(1, 27)} =$

0.231 , $p = 0.636$, $\eta_p^2 = 0.011$], prior visual experience [$F_{(1, 27)} = 1.283$, $p = 0.27$, $\eta_p^2 = 0.058$], or gender [$F_{(1, 27)} = 3.857$, $p = 0.063$, $\eta_p^2 = 0.155$]. The interaction of condition and visual experience was not significant [$F_{(1, 25)} = 0.064$, $p = 0.803$, $\eta_p^2 = 0.003$], nor were any of the other interactions tested. Early blind gamers and directed navigators showed similar success on the navigation tasks following training with AbES (gamers: $85.00\% \pm 23.30$ correct, directed navigators: $82.86\% \pm 9.51$ correct; $p = 0.942$, n.s., Tukey test) (see **Figure 3A**). A similar profile of performance was observed in late blind gamers ($82.86\% \pm 28.70$) and directed navigators ($80.00\% \pm 38.30$) ($p = 0.997$, n.s., Tukey test). However, no significant difference was observed between gamers and directed navigators regardless of early and late blind status.

For the physical navigation task, a Three-Way ANOVA (2 condition \times 2 visual experience \times 2 gender) found no significant main effects for condition [$F_{(1, 27)} = 0.070$, $p = 0.793$, $\eta_p^2 = 0.003$], visual experience [$F_{(1, 27)} = 1.761$, $p = 0.199$, $\eta_p^2 = 0.077$], or gender [$F_{(1, 27)} = 1.127$, $p = 0.301$, $\eta_p^2 = 0.051$]. The interaction of condition and visual experience was not significant [$F_{(1, 25)} = 0.070$, $p = 0.793$, $\eta_p^2 = 0.003$], nor were any of the other interactions tested. Early blind gamers and directed navigators again showed comparable levels of performance in terms of their ability to transfer acquired spatial information to the real physical environment (gamers: $87.50\% \pm 10.35$; directed navigators: $88.57\% \pm 18.64$; $p = 0.998$, n.s., Tukey test). Similar performance levels were seen in late blind gamers and directed navigators as well (gamers: $92.86\% \pm 9.51$; directed navigators: $92.86\% \pm 9.51$; $p = 1.00$, n.s., Tukey test) (see **Figure 3B**). A repeated-measures ANOVA showed that the overall mean percentage correct performance on physical navigation (mean = $90.35\% \pm 12.10$) was not statistically different from virtual navigation (mean = $85.17\% \pm 25.86$) [$F_{(1, 28)} = 1.124$, $p = 0.298$, $\eta_p^2 = 0.039$]. Thus, as with the virtual navigation task, no significant difference was observed between gamers and directed navigators regardless of early and late blind status.

Finally, assessing performance on the drop off task (i.e., exiting the building using the shortest path possible) did reveal a significant difference in performance between gamers and directed navigators. A Three-Way ANOVA (2 condition \times 2 visual experience \times 2 gender) found a significant main effect of condition [$F_{(1, 27)} = 62.856$, $p < 0.001$, $\eta_p^2 = 0.750$], but no main effect of visual experience [$F_{(1, 27)} = 1.802$, $p = 0.194$, $\eta_p^2 = 0.079$] or gender [$F_{(1, 27)} = 0.003$, $p = 0.960$, $\eta_p^2 = 0.000$]. The interaction of condition and visual experience was not significant [$F_{(1, 25)} = 0.108$, $p = 0.745$, $\eta_p^2 = 0.005$], nor were any of the other interactions tested. By group, the performance of the gamers (mean = 2.71 points ± 0.41 out of a maximum of 3) was significantly better than the directed navigators (mean = 1.66 points ± 0.21). Comparing performance between early blind gamers and directed navigators revealed effects similar to the overall group. On average, early blind gamers scored more points (2.60 points ± 0.52) than their directed navigator counterparts (1.60 points ± 0.00) ($p < 0.001$, Tukey test) suggesting that on average, gamers were more likely to select the closest exit and navigate using the shortest path regardless of their initial starting point. In contrast, directed

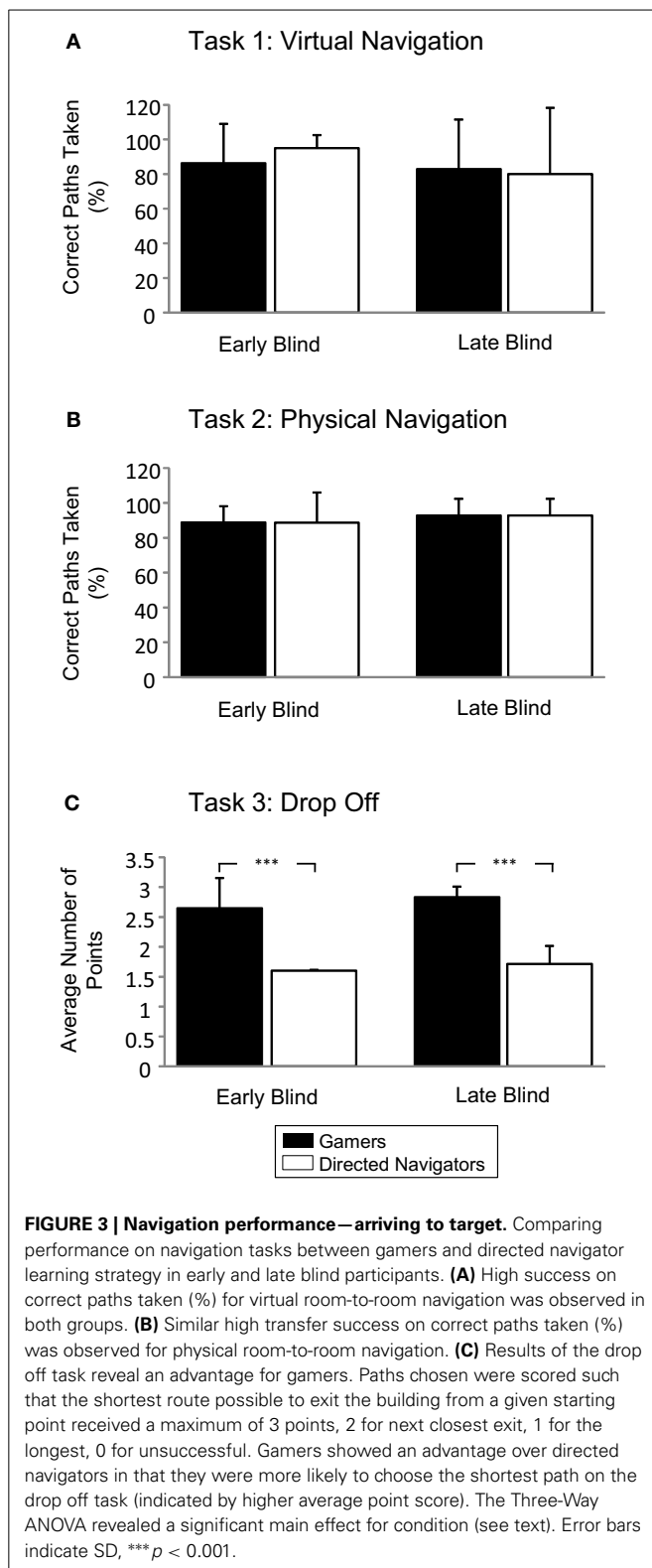


FIGURE 3 | Navigation performance—arriving to target. Comparing performance on navigation tasks between gamers and directed navigator learning strategy in early and late blind participants. **(A)** High success on correct paths taken (%) for virtual room-to-room navigation was observed in both groups. **(B)** Similar high transfer success on correct paths taken (%) was observed for physical room-to-room navigation. **(C)** Results of the drop off task reveal an advantage for gamers. Paths chosen were scored such that the shortest route possible to exit the building from a given starting point received a maximum of 3 points, 2 for next closest exit, 1 for the longest, 0 for unsuccessful. Gamers showed an advantage over directed navigators in that they were more likely to choose the shortest path on the drop off task (indicated by higher average point score). The Three-Way ANOVA revealed a significant main effect for condition (see text). Error bars indicate SD, *** $p < 0.001$.

navigators were more likely to use longer routes. A similar effect was seen in late blind gamers who scored on average 2.83 points ± 0.18 points, while directed navigators scored 1.71 points ± 0.30 ($p < 0.001$, Tukey test) (see **Figure 3C**).

COMPARING PERFORMANCE IN GAMERS AND DIRECTED NAVIGATORS—TIME TO TARGET

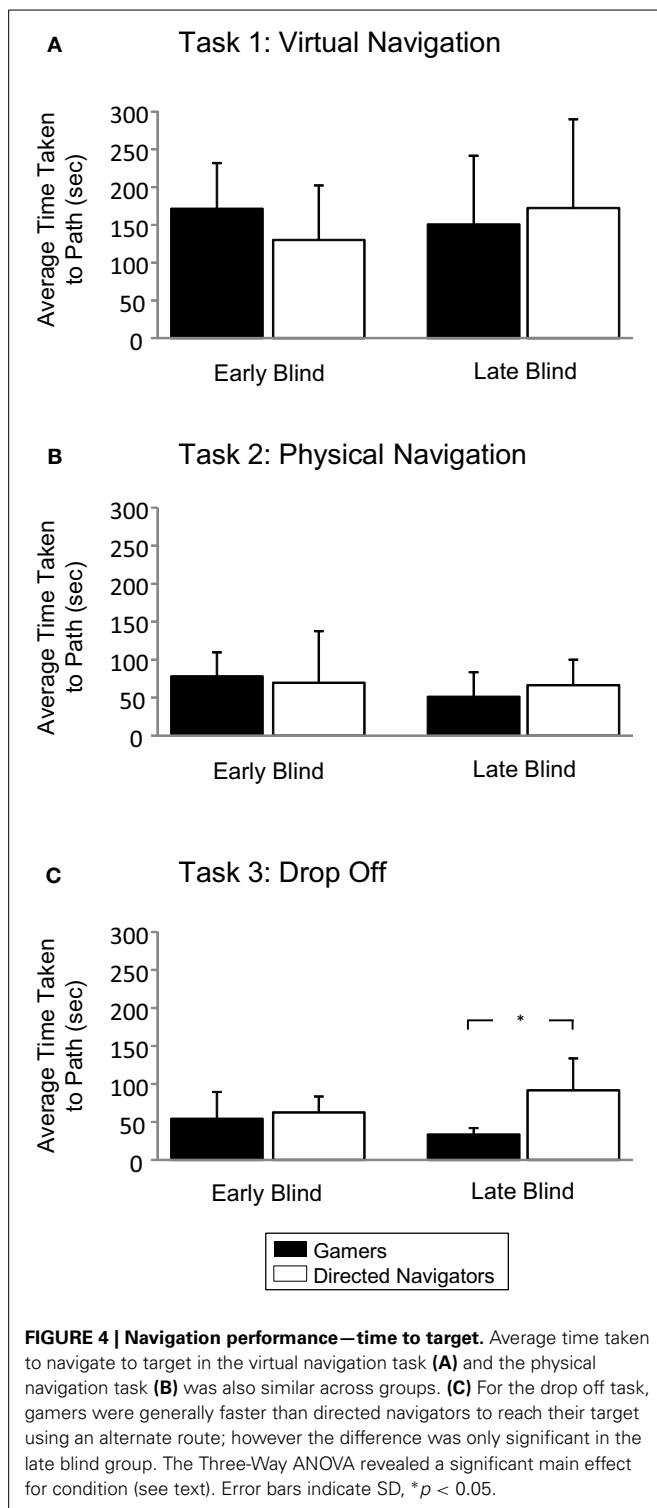
Assessing time to target for the virtual navigation task was performed using a Three-Way ANOVA (2 condition \times 2 visual experience \times 2 gender). This analysis found no main effect of condition [$F_{(1, 27)} = 0.164$, $p = 0.690$, $\eta_p^2 = 0.008$], visual experience [$F_{(1, 27)} = 0.251$, $p = 0.622$, $\eta_p^2 = 0.012$], or gender [$F_{(1, 27)} = 2.577$, $p = 0.125$, $\eta_p^2 = 0.109$]. The interaction of condition and visual experience was not significant [$F_{(1, 25)} = 0.206$, $p = 0.655$, $\eta_p^2 = 0.010$], nor were any of the other interactions tested. Performance was again comparable in both early blind groups (gamers: 170.56 s \pm 68.05; directed navigators: 136.00 s \pm 75.47; $p = 0.880$, n.s., Tukey test). Similar performance in terms of time taken to target was found in late blind participants (gamers: 150.70 s \pm 90.84; directed navigators: 172.34 s \pm 117.44; $p = 0.968$, n.s., Tukey test) (see **Figure 4A**).

For time taken to target on the physical navigation task, a Three-Way ANOVA (2 condition \times 2 visual experience \times 2 gender) found no main effect of condition [$F_{(1, 27)} = 0.377$, $p = 0.546$, $\eta_p^2 = 0.018$], visual experience [$F_{(1, 27)} = 1.112$, $p = 0.304$, $\eta_p^2 = 0.050$], or gender [$F_{(1, 27)} = 0.456$, $p = 0.507$, $\eta_p^2 = 0.021$]. The interaction of condition and visual experience was not significant [$F_{(1, 25)} = 0.420$, $p = 0.524$, $\eta_p^2 = 0.020$], nor were any of the other interactions tested. Performance in early blind gamers (75.26 s \pm 36.01) and directed navigators (71.34 s \pm 73.39) was not significantly different between groups ($p = 0.998$, n.s., Tukey test). For late blind gamers (51.34 s \pm 32.07) and directed navigators (66.47 s \pm 33.50), no significant difference in mean time was found ($p = 0.929$, n.s., Tukey test). A repeated-measures ANOVA found that, the navigation times were markedly shorter for the physical navigation task (mean = 66.42 s \pm 44.99) than for the virtual navigation task (mean = 157.94 s \pm 85.61) [$F_{(1, 28)} = 36.694$, $p < 0.001$, $\eta_p^2 = 0.567$] (see **Figure 4B**).

Finally, comparing time to target on the drop off task, a Three-Way ANOVA (2 condition \times 2 visual experience \times 2 gender) found a significant main effect of condition [$F_{(1, 27)} = 7.42$, $p = 0.013$, $\eta_p^2 = 0.261$], but no main effect of visual experience [$F_{(1, 27)} = 0.071$, $p = 0.792$, $\eta_p^2 = 0.003$] or gender [$F_{(1, 27)} = 0.000$, $p = 0.93$, $\eta_p^2 = 0.000$]. The interaction between condition and visual experience was not statistically significant [$F_{(1, 25)} = 3.429$, $p = 0.078$, $\eta_p^2 = 0.140$]. None of the other interactions showed significant differences. Further analysis revealed that early blind gamers' mean time (51.23 s \pm 42.36) compared to directed navigators (62.14 s \pm 22.72) was not statistically significant ($p = 0.916$, n.s., Tukey test). However, the mean navigation time for late blind gamers (33.26 s \pm 8.71) was significantly shorter than for directed navigators (91.66 s \pm 41.79) ($p = 0.013$ Tukey test). A repeated-measures ANOVA found that the mean overall time (59.28 s \pm 37.44) was not statistically different from the physical navigation task [$F_{(1, 28)} = 0.635$, $p = 0.432$, $\eta_p^2 = 0.022$] (see **Figure 4C**).

COMPARING PERFORMANCE WITH CONTROLS: ARRIVING AND TIME TO TARGET

All 7 of the control participants were unable to reach any of the target locations for the virtual, physical, and drop off tasks,



yielding a mean percent correct path score of 0. Likewise, all participants reached the maximum time limit (360 s) for time to target. Because all participants performed at floor levels, there was no variance in the data, precluding the use of ANOVAs or *t*-tests to compare the experimental groups with the control group. Instead, we used the performance levels of the control group as a measure

of chance, and compared the experimental groups against these values using 1-sample *t*-tests.

Examination of arrival at target yielded performance significantly greater than 0 in the gamer group for all three tasks [virtual: $t_{(14)} = 13.006$, $p < 0.001$; physical: $t_{(14)} = 34.857$, $p < 0.001$; drop-off: $t_{(14)} = 25.811$, $p < 0.001$; see **Figure 3**]. The directed navigator group also had performance significantly greater than 0 for the virtual [$t_{(13)} = 11.706$, $p < 0.001$], physical [$t_{(13)} = 23.583$, $p < 0.001$], and drop-off [$t_{(13)} = 29.000$, $p < 0.001$] navigation tasks (see **Figure 3**).

Time to target was compared against 360 s, the score for all participants in the control group. The gamer group was significantly faster than 360 s for the virtual [$t_{(14)} = -9.971$, $p < 0.001$], physical [$t_{(14)} = -32.522$, $p < 0.001$] and drop-off [$t_{(14)} = -38.545$, $p < 0.001$] tasks (see **Figure 4**). The directed navigator group was also significantly faster than 360 s for all three tasks [virtual: $t_{(13)} = -7.961$, $p < 0.001$; physical: $t_{(13)} = -19.852$, $p < 0.001$; drop-off: $t_{(13)} = -29.620$, $p < 0.001$; see **Figure 4**]. Thus, for both arrival and time to target, both the gamers and directed navigators were significantly above control performance for all three navigation tasks.

ASSOCIATIONS OF INTEREST

As a secondary analysis, we explored potential associations between navigation performance (assessed using percentage success on physical navigation as the primary outcome of interest) and the factors of age and verbal memory. Comparing individual navigation success with age in both conditions (collapsing early and late blind) revealed negative trends for both gamers [$r_{(13)} = -0.193$, $p = 0.492$] and directed navigators [$r_{(12)} = -0.503$, $p = 0.067$], although neither trend achieved statistical significance (**Figure 5A**). Similarly, no statistically significant association was evident comparing individual navigation performance with verbal memory recall (indexed by the Wechsler score) in either group [gamers: $r_{(13)} = -0.287$, $p = 0.300$; directed navigators: $r_{(12)} = 0.088$, $p = 0.766$] (**Figure 5B**). A second level analysis using a One-Way ANOVA with gender and group as factors confirmed a lack of association between navigation performance and gender for either the gamer [$F_{(1, 13)} = 0.263$, $p = 0.617$, $\eta_p^2 = 0.020$] or directed navigator group [$F_{(1, 12)} = 1.751$, $p = 0.210$, $\eta_p^2 = 0.127$]. Ancillary analyses exploring potential associations between the other assessments of navigation performance (i.e., virtual navigation and drop off task) with the factors of age and verbal memory revealed no further significant correlations.

DISCUSSION

In this study, we demonstrate that early and late blind individuals were able to effectively interact and explore an audio-based virtual environment for the purposes of acquiring relevant sensory information regarding a building's spatial layout. Furthermore, participants were able to generate a corresponding spatial cognitive map of the building they explored. The accuracy of this mental representation was confirmed by the fact that participants were able to successfully transfer acquired spatial information to a series of navigation tasks carried out in both the corresponding virtual environment explored and the physical building modeled

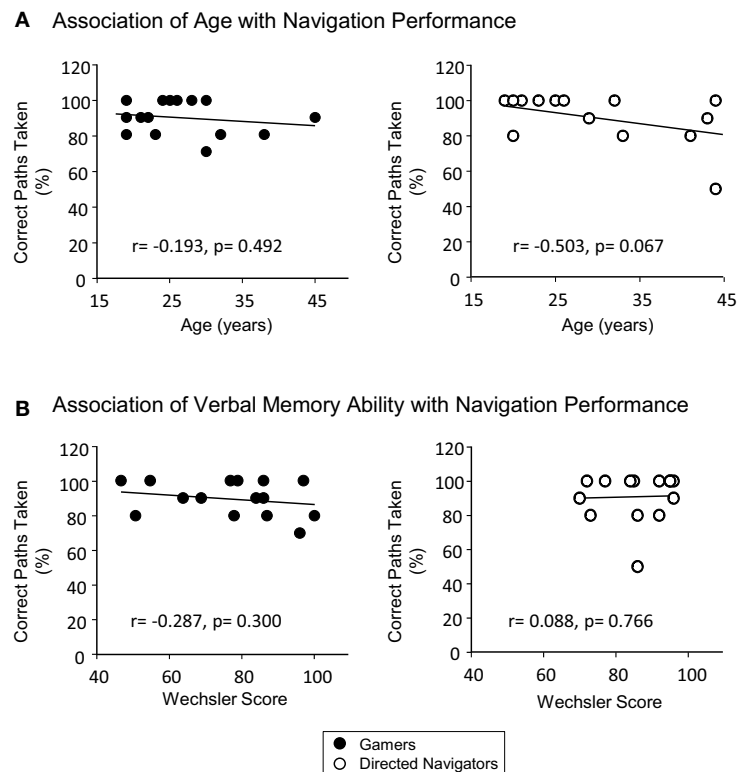


FIGURE 5 | Correlations between navigation task performance (% correct on physical navigation) and factors of interest (A) age and (B) verbal memory ability. No significant associations were observed (data from gamers and directed navigators collapsed).

in the AbES software. Furthermore, control subjects did not show any evidence of transfer. Specifically, the failure of the control subjects to carry out any of the navigation task assessments following game play suggests that a contextual overlap between the exploratory training and task environments is needed for the transfer of learning and further, game play alone cannot account for the behavioral performance observed.

In general, virtual and physical navigation performance was comparable whether participants learned the building's spatial layout implicitly through exploratory game play (gamers) or explicitly through structured and serial instructions (directed navigators). The similar performance between gamers and directed navigators suggests that both learning strategies (i.e., self-directed, implicit learning through gaming vs. guided instruction, explicit learning through directed navigation) allowed for the virtual exploration and subsequent generation of an accurate spatial representation that could be eventually transferred for the purposes of carrying out a large-scale and complex physical navigation task. As the target building contained two-stories, this also included successful navigation performance along the vertical dimension. This has been reported to be more difficult to achieve when acquiring information from virtual environments characterizing an indoor spatial layout (see Richardson et al., 1999, and later discussion). The most important difference between the two learning strategies compared was observed when the navigation task required the mental manipulation of this spatial information

for determining an alternate route or short cut. Specifically, participants from the gamer group showed a behavioral advantage on the drop off task compared to the directed navigator group. Gamers were on average more likely to use the shortest route possible relative to their starting point compared to their directed navigator counterparts. These results suggest that training experience through gaming provided for a more flexible use of the spatial information characterizing the layout of the target building.

Navigation times were also comparable across groups for the virtual and physical navigation tasks. However, overall navigation time was shorter when these tasks were carried out in the physical compared to the virtual environment. This latter finding is consistent with anecdotal reports provided by a number of participants describing their impressions that the physical building was perceived to be smaller than what they had mentally imagined following the virtual exploration of the same environment. This anecdotal finding is somewhat puzzling given that the scale of the AbES environment was designed such that one step corresponded to one step in the physical building. The discrepancy may be related to inter-individual differences in how the resultant spatial cognitive map is generated or associated physical differences of the participants (such as individual stride length). Moreover, as virtual navigation always preceded physical navigation, there remains the possibility of a carryover effect arising from the sequential assessment of performance. Regarding

the drop off task, navigation times for the gamers were faster than their directed navigator counterparts. This trend is consistent with the observation that participants on average chose the shortest possible route.

Finally, overall navigation performance did not appear to be strongly associated with other factors of interest including age, gender, and verbal memory skill.

The transfer of spatial knowledge from exploring virtual environments has been investigated previously under varied conditions in sighted subjects. In general, results from these studies have highlighted that immersion in a virtual environment can facilitate transfer of knowledge when the fidelity between the gaming environment and the real-world building is high (Waller et al., 1998; Farrell et al., 2003). The importance of contextual overlap between the learning and task settings is referred to as near transfer of learning (Cormier and Hagman, 1987). The results observed in this study can be viewed as consistent with this form of learning. This is also supported by the fact that the control group showed no evidence of transfer after game play in a virtual environment that did not correspond to the spatial layout of the target building. It is also worth highlighting that the gamer and directed navigator groups were never trained in the paths used in the behavioral assessments. This suggests that participants were not only able to acquire the spatial information needed to generate an accurate mental representation of the environment, but were also able to access and manipulate this information for the purposes of carrying out navigation tasks. It would be of considerable interest to investigate specifically the effect of prolonged training and game play (say, on the order of months) to see if enhanced performance (i.e., transfer of learning) would occur within other spatial related tasks or cognitive domains beyond the training context investigated here. In this direction, a specific longitudinal, long term, study incorporating a battery of pre and post-performance assessments would need to be carried out to answer this question conclusively.

In terms of rehabilitation and training for the blind, the use of virtual environments and game based learning strategies have been investigated in a number of studies (for applications including navigation and other forms of cognitive development such as short term memory and math skills; see Sánchez and Baloian, 2005; Sánchez and Maureira, 2007; Merabet and Sánchez, 2009; Afonso et al., 2010; Saenz and Sánchez, 2010; Lahav et al., 2011). Similar approaches have also been pursued for individuals with cognitive disabilities (Strickland, 1997; Salem et al., 2012) as well as for physical rehabilitation (Cho et al., 2012). Learning through a ludic-based approach could have specific benefits with regards to the transfer of spatial knowledge. Indeed, it has been proposed that game-based exploration of a virtual environment may have several potential advantages over those who are directed through a novel environment (Chrastil and Warren, 2012). First, they can make self-directed decisions about how and where they wish to navigate, which requires them to keep track of locations they have previously visited, and test theories about where they expect to be with every turn (Wilson et al., 1997; von Stulpnagel and Steffens, 2012). Second, exploring through game play may require additional mental manipulation of the spatial layout of the environment, which can be beneficial for learning (Bosco et al., 2004;

Coluccia et al., 2007). For example, a gamer would need to understand that rooms that were on the right side of the hallway when traveling in one direction are now on the left side when traveling in the opposite direction. Third, improved attentional processing may allow gamers to attend and recall navigation-relevant features of the environment, even when not given explicit instructions to do so (Magliano et al., 1995; Taylor et al., 1999). Finally, the immersive nature of video game play offers gamers the possibility to experience travel in the environment from multiple perspectives and reference frames, which might prove useful when trying to determine how to reach a particular goal and manipulate information to find an alternate route.

In this study, participants in the directed navigator group served as a reference group to characterize performance associated with the transfer of information based on explicit training. It is of interest that participants in the gaming and directed navigator groups showed comparable levels of performance despite the fact that gamers were never explicitly told to retain any information regarding the spatial layout of the building, nor were they aware that the overall purpose of the study. Moreover, the gamer group showed an advantage over the directed navigation group in the drop-off task in that gamers were more likely to find shorter routes (likely as a result of being able to better mentally manipulate their resultant spatial map). As outlined earlier, it is possible that the highly engaging, interactive, and immersive nature of game play supported the development of more flexible and robust spatial cognitive constructs, particularly in situations where spatial information had to be manipulated for the purposes of determining optimal and alternative routes. We propose that the differences in learning strategy and subsequent effects on behavioral performance observed are likely related to the method through which spatial information is acquired, organized, and how the resultant spatial cognitive map is developed. The results of this study confirm and validate our initial proof-of-concept findings that learning through game play may provide for superior contextual learning and transfer of situational knowledge resulting from greater understanding of the spatial inter-relations within a building environment (Merabet et al., 2012). While individuals who learned the building environment implicitly through gaming exhibited a transfer of learning, those who learned the same environment in an explicitly guided manner appeared to fail in capturing more global contextual and situation-relevant information, such as where to exit the building. This somewhat constrained response profile exhibited by the latter group is consistent with what is typically seen following more rote type learning strategies (Shaffer et al., 2005).

The fact that participants showed high levels of success in transferring spatial information to real world navigation tasks suggests that the mental representations generated were isomorphic with the real world spatial layout of the target environment (Richardson et al., 1999; Waller et al., 2001). Given that visual cues play a crucial role in capturing and organizing spatial formation, it has been assumed that blind individuals (and in particular, those who are born with profound blindness) would be impaired in creating an accurate mental spatial representation of their surroundings (von Senden, 1960; Ashmead et al., 1989; Thinus-Blanc and Gaunet, 1997; see also Blasch et al., 1997 for further

discussion). It is thus of interest that both early and late blind participants showed similar levels of performance in this study. Indeed, the effect of previous visual experience has long been an issue of debate not only in terms of teaching O&M skills, but also with regards to other spatial processing tasks (Cornoldi et al., 1991, 2009; Aleman et al., 2001; Vecchi et al., 2004; Tinti et al., 2006; Cattaneo et al., 2008). A review of this literature reveals contradictory results (particularly in relation to the role of prior visual experience), calling into question the validity of these earlier assumptions. In fact, some studies have reported that no differences exist in terms of how well blind individuals are able to mentally represent and interact with spatial environments (Landau et al., 1981; Passini and Proulx, 1988; Morrongiello et al., 1995; Noordzij et al., 2006) and in certain spatial navigation tasks, individuals with profound blindness have been shown to exhibit equal (Loomis et al., 2001) and in some cases, superior performance (Fortin et al., 2008) when compared to sighted control subjects.

In a recent and comprehensive review, Pasqualotto and Proulx (2012) discuss how visual development may be necessary for the normal capacities of spatial cognition. This argument is based on the ability of the visual modality to capture and convey parallel information, as well as to provide a basis for topographic representations necessary for high level (i.e., “allocentric”) spatial representations and multisensory integration (Pasqualotto and Proulx, 2012). In other words, spatial processing strategies and underlying representations may be different depending on the amount of prior visual experience available (see also Pasqualotto et al., 2013 for results demonstrating the preference of congenitally blind individuals in using egocentric based reference frames for the spatial representation of objects in space). In the present study, early and late blind participants exhibited comparable behavioral performance, suggesting that the immersive and exploratory nature of the virtual environment was sufficient to promote the accurate generation of a spatial cognitive map for the purposes of navigation. However, we are unable at this juncture to determine whether either group was preferentially employing strategies more aligned with egocentric or allocentric based representations. Furthermore, we acknowledge that while our early blind participants had documented evidence of profound blindness assessed prior to the age of three, disentangling the true effect of prior visual experience is less definitive based on the blindness history of the participants enrolled in this study. Finally, it is important to consider that overall behavioral performance may be related to not only the nature of the task demands, but also to the scale of the environment tested. At the same time, we recognize that early blind individuals typically will have spent more time learning compensatory strategies than late blind individuals, and thus the behavioral and associated neurophysiological adaptations at play may be very different between these two groups (Carroll, 1961). This remains a key question requiring careful study, as it has important implications not only in terms of understanding development and compensatory behavioral changes, but also for rehabilitation and educational strategies for the blind in general (Merabet and Pascual-Leone, 2010). A continuation of this study encompassing a larger scale and more complex mental map (e.g., incorporating indoor and outdoor environments) is

needed to confirm whether there are indeed discrepancies related to visual experience and overall navigation performance.

Finally, no strong associations were apparent between navigation performance (as indexed by the percentage correct on the physical navigation task) and other factors of interest including age and verbal memory ability. While such an absence of statistical evidence certainly does not rule out the possibility of a true underlying association, it does suggest that such a simulation-game based learning strategy can be effective across a wide demographic profile. Furthermore, no apparent differences in performance were observed as a function of gender. Indeed, differences in navigation and other spatial cognitive skills in men and women have long been reported and debated (Maguire et al., 1999; Bosco et al., 2004; Coluccia and Louse, 2004; Wolbers and Hegarty, 2010). Again, it is possible that the navigation task undertaken here did not allow for differences to be revealed between male and female participants or, more plausibly, that the nature of the training and assessments were such that any inherent gender-related differences in the acquisition and overall transfer of spatial skills were not made apparent (see Feng et al., 2007 reporting the use of an action based video game to reduce gender related differences of performance in spatial cognition tasks).

In conclusion, the findings from this study demonstrate that the highly interactive and immersive nature of the AbES system is effective for the learning of novel environments in both early and late blind individuals through a mechanism akin to near transfer of skill learning. Furthermore, active game play provided a more flexible cognitive representation of the environment, allowing gamers to mentally manipulate spatial information for the purposes of finding alternate routes. This learning approach may serve as a useful tool to learn the spatial layout of large-scale, three-dimensional spaces, and to transfer that knowledge into real-world navigation tasks.

AUTHOR CONTRIBUTIONS

Analyzed the data: Erin C. Connors, Elizabeth R. Chrastil, Lotfi B. Merabet. Designed the research: Lotfi B. Merabet. Collected data: Erin C. Connors, Lotfi B. Merabet. Contributed to writing the paper: Erin C. Connors, Elizabeth R. Chrastil, Jaime Sánchez, Lotfi B. Merabet.

ACKNOWLEDGMENTS

This work was supported by an NIH/NEI RO1 GRANT EY019924 (Lotfi B. Merabet) and also funded by the Chilean National Fund of Science and Technology, Fondecyt #1120330 and Project CIE-05 Program Center Education PBCT-Conicyt (Jaime Sánchez). The authors would like to thank the research participants, as well as Rabih Dow, Padma Rajagopal and the staff of the Carroll Center for the Blind (Newton MA, USA) for their support in carrying out this research.

REFERENCES

- Afonso, A., Blum, A., Katz, B. F., Tarroux, P., Borst, G., and Denis, M. (2010). Structural properties of spatial representations in blind people: Scanning images constructed from haptic exploration or from locomotion in a 3-D audio virtual environment. *Mem. Cogn.* 38, 591–604. doi: 10.3758/MC.38.5.591

- Aleman, A., van Lee, L., Mantione, M. H., Verkoijen, I. G., and de Haan, E. H. (2001). Visual imagery without visual experience: evidence from congenitally totally blind people. *Neuroreport* 12, 2601–2604. doi: 10.1097/00001756-200108080-00061
- Ashmead, D. H., Hill, E. W., and Talor, C. R. (1989). Obstacle perception by congenitally blind children. *Percept. Psychophys.* 46, 425–433. doi: 10.3758/BF03210857
- Bavelier, D., Green, C. S., Han, D. H., Renshaw, P. F., Merzenich, M. M., and Gentile, D. A. (2011). Brains on video games. *Nat. Rev. Neurosci.* 12, 763–768. doi: 10.1038/nrn3135
- Bavelier, D., Green, C. S., Pouget, A., and Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annu. Rev. Neurosci.* 35, 391–416. doi: 10.1146/annurev-neuro-060909-152832
- Blasch, B. B., Wiener, W. R., and Welsh, R. L. (eds.). (1997). *Foundations of Orientation and Mobility*, 2nd Edn. New York, NY: AFB Press.
- Bosco, A., Longoni, A. M., and Vecchi, T. (2004). Gender effects in spatial orientation: cognitive profiles and mental strategies. *Appl. Cogn. Psychol.* 18, 519–532. doi: 10.1002/acp.1000
- Carroll, T. J. (1961). *Blindness: What it is, What it Does, and How to Live with it*, 1st Edn. Boston, MA: Little.
- Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., et al. (2008). Imagery and spatial processes in blindness and visual impairment. *Neurosci. Biobehav. Rev.* 32, 1346–1360. doi: 10.1016/j.neubiorev.2008.05.002
- Chebat, D. R., Schneider, F. C., Kupers, R., and Ptito, M. (2011). Navigation with a sensory substitution device in congenitally blind individuals. *Neuroreport* 22, 342–347. doi: 10.1097/WNR.0b013e3283462def
- Cho, K. H., Lee, K. J., and Song, C. H. (2012). Virtual-reality balance training with a video-game system improves dynamic balance in chronic stroke patients. *Tohoku J. Exp. Med.* 228, 69–74. doi: 10.1620/tjem.228.69
- Chrastil, E. R., and Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychon. Bull. Rev.* 19, 1–23. doi: 10.3758/s13423-011-0182-x
- Coluccia, E., Bosco, A., and Brandimonte, M. A. (2007). The role of visuo-spatial working memory in map learning: new findings from a map drawing paradigm. *Psychol. Res.* 71, 359–372.
- Coluccia, E., and Louse, G. (2004). Gender differences in spatial orientation: a review. *J. Exp. Psychol.* 24, 329–340. doi: 10.1016/j.jenp.2004.08.006
- Connors, E. C., Yazzolino, L. A., Sanchez, J., and Merabet, L. B. (2013). Development of an audio-based virtual gaming environment to assist with navigation skills in the blind. *J. Vis. Exp.* 73. doi: 10.3791/50272
- Cormier, S. M., and Hagman, J. D. (1987). *Transfer of Learning: Contemporary Research Applications*. San Diego, CA: Academic Press Inc.
- Cornoldi, C., Cortesi, A., and Preti, D. (1991). Individual differences in the capacity limitations of visuospatial short-term memory: research on sighted and totally congenitally blind people. *Mem. Cogn.* 19, 459–468. doi: 10.3758/BF03199569
- Cornoldi, C., Tinti, C., Mammarella, I. C., Re, A. M., and Varotto, D. (2009). Memory for an imagined pathway and strategy effects in sighted and in totally congenitally blind individuals. *Acta Psychol.* 130, 11–16. doi: 10.1016/j.actpsy.2008.09.012
- Dede, C. (2009). Immersive interfaces for engagement and learning. *Science* 323, 66–69. doi: 10.1126/science.1167311
- Farrell, M. J., Arnold, P., Pettifer, S., Adams, J., Graham, T., and MacManamon, M. (2003). Transfer of route learning from virtual to real environments. *J. Exp. Psychol. Applied* 9, 219–227. doi: 10.1037/1076-898X.9.4.219
- Feng, J., Spence, I., and Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychol. Sci.* 18, 850–855. doi: 10.1111/j.1467-9280.2007.01990.x
- Fortin, M., Voss, P., Lord, C., Lassonde, M., Pruessner, J., Saint-Amour, D., et al. (2009). Memory for an imagined pathway and strategy effects in sighted and in totally congenitally blind individuals. *Acta Psychol.* 130, 11–16. doi: 10.1016/j.actpsy.2008.09.012
- Giudice, N. A., Bakdash, J. Z., and Legge, G. E. (2007). Wayfinding with words: spatial learning and navigation using dynamically updated verbal descriptions. *Psychol. Res.* 71, 347–358. doi: 10.1007/s00426-006-0089-8
- Giudice, N. A., and Legge, G. E. (2008). “Blind navigation and the role of technology,” in *Engineering Handbook of Smart Technology for Aging, Disability, and Independence*, eds A. Helal, M. Mokhtari, and B. Abdulrazak (John Wiley and Sons), 479–500.
- Johnson, L. A., and Higgins, C. M. (2006). A navigation aid for the blind using tactile-visual sensory substitution. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 1, 6289–6292. doi: 10.1109/IEMBS.2006.259473
- Kalia, A. A., Legge, G. E., Roy, R., and Ogale, A. (2010). Assessment of indoor route-finding technology for people with visual impairment. *J. Vis. Impair. Blind.* 104, 135–147.
- Kuppersmith, R. B., Johnston, R., Jones, S. B., and Jenkins, H. A. (1996). Virtual reality surgical simulation and otolaryngology. *Arch. Otolaryngol. Head Neck Surg.* 122, 1297–1298. doi: 10.1001/archotol.1996.01890240007002
- Lahav, O., Schloerb, D. W., and Srinivasan, M. A. (2011). Newly blind persons using virtual environment system in a traditional orientation and mobility rehabilitation program: a case study. *Disabil. Rehabil. Assist. Technol.* 7, 420–435. doi: 10.3109/17483107.2011.635327
- Landau, B., Gleitman, H., and Spelke, E. (1981). Spatial knowledge and geometric representation in a child blind from birth. *Science* 213, 1275–1278. doi: 10.1126/science.7268438
- Lange, B., Koenig, S., Chang, C. Y., McConnell, E., Suma, E., Bolas, M., et al. (2012). Designing informed game-based rehabilitation tasks leveraging advances in virtual reality. *Disabil. Rehabil.* 34, 1863–1870. doi: 10.3109/09638288.2012.670029
- Long, R. G., and Giudice, N. A. (2010). “Establishing and maintaining orientation for orientation and mobility,” in *Foundations of Orientation and Mobility*, 3rd ed. Vol. 1, *History and Theory*, eds B. B. Blasch, W. R. Wiener, and R. W. Welch (New York, NY: American Foundation for the Blind), 45–62.
- Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (2001). Navigating without vision: basic and applied research. *Optom. Vis. Sci.* 78, 282–289. doi: 10.1097/00006324-200105000-00011
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., and Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *J. Exp. Psychol. Gen.* 122, 73–91. doi: 10.1037/0096-3445.122.1.73
- Loomis, J. M., Marston, J. R., Golledge, R. G., and Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: a comparison of spatial displays for route guidance. *J. Vis. Impair. Blind.* 99, 219–232.
- Magliano, J. P., Cohen, R., Allen, G. L., and Rodrigue, J. R. (1995). The impact of a wayfinder's goal on learning a new environment: Different types of spatial knowledge as goals. *J. Environ. Psychol.* 15, 65–75. doi: 10.1016/0272-4944(95)90015-2
- Maguire, E. A., Burgess, N., and O'Keefe, J. (1999). Human spatial navigation: cognitive maps, sexual dimorphism, and neural substrates. *Curr. Opin. Neurobiol.* 9, 171–177. doi: 10.1016/S0959-4388(99)80023-3
- Mayo, M. J. (2009). Video games: a route to large-scale STEM education? *Science* 323, 79–82. doi: 10.1126/science.1166900
- Merabet, L. B., Connors, E. C., Halko, M. A., and Sanchez, J. (2012). Teaching the blind to find their way by playing video games. *PLoS ONE* 7:e44958. doi: 10.1371/journal.pone.0044958
- Merabet, L. B., and Pascual-Leone, A. (2010). Neural reorganization following sensory loss: the opportunity of change. *Nat. Rev. Neurosci.* 11, 44–52. doi: 10.1038/nrn2758
- Merabet, L., and Sánchez, J. (2009). Audio-based navigation using virtual environments: combining technology and neuroscience. *AER J.* 2, 128–137.
- Morronegiello, B. A., Timney, B., Humphrey, G. K., Anderson, S., and Skory, C. (1995). Spatial knowledge in blind and sighted children. *J. Exp. Child Psychol.* 59, 211–233. doi: 10.1006/jecp.1995.1010
- Noordzij, M. L., Zuidhoek, S., and Postma, A. (2006). The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. *Cognition* 100, 321–342. doi: 10.1016/j.cognition.2005.05.006
- Pasqualotto, A., and Proulx, M. J. (2012). The role of visual experience for the neural basis of spatial cognition. *Neurosci. Biobehav. Rev.* 36, 1179–1187. doi: 10.1016/j.neubiorev.2012.01.008
- Pasqualotto, A., Spiller, M. J., Jansari, A. S., and Proulx, M. J. (2013). Visual experience facilitates allocentric spatial representation. *Behav. Brain Res.* 236, 175–179. doi: 10.1016/j.bbr.2012.08.042
- Passini, R., and Proulx, G. (1988). Wayfinding without vision. *Environ. Behav.* 20, 227–252. doi: 10.1177/0013916588202006
- Pataki, C., Pato, M. T., Sugar, J., Rizzo, A. S., Parsons, T. D., St George, C., et al. (2012). Virtual patients as novel teaching tools in psychiatry. *Acad. Psychiatry* 36, 398–400. doi: 10.1176/appi.ap.10080118
- Petrie, H., Johnson, V., Strothotte, T., Raab, A., Fritz, S., and Michel, R. (1996). Mobic: designing a travel aid for blind and elderly people. *J. Navigation* 49, 45–52. doi: 10.1017/S0373463300013084

- Richardson, A. E., Montello, D. R., and Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Mem. Cogn.* 27, 741–750. doi: 10.3758/BF03211566
- Rizzo, A., Buckwalter, J. G., John, B., Newman, B., Parsons, T., Kenny, P., et al. (2012). STRIVE: stress resilience in virtual environments: a pre-deployment VR system for training emotional coping skills and assessing chronic and acute stress responses. *Stud. Health Technol. Inform.* 173, 379–385.
- Saenz, M., and Sánchez, J. (2010). Indoor orientation and mobility for learners who are blind. *Stud. Health Technol. Inform.* 154, 165–170.
- Salem, Y., Gropack, S. J., Coffin, D., and Godwin, E. M. (2012). Effectiveness of a low-cost virtual reality system for children with developmental delay: a preliminary randomised single-blind controlled trial. *Physiotherapy* 98, 189–195. doi: 10.1016/j.physio.2012.06.003
- Sánchez, J., and Baloian, N. (2005, June 27–July 2). “Modeling audio-based virtual environments for children with visual disabilities,” in *Paper Presented at the Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications ED-MEDIA*, (Montreal, QC).
- Sánchez, J., and Maureira, E. (2007). “Subway mobility assistance tools for blind users,” in *Lecture Notes in Computer Science, LNCS, 4397*, eds C. Stephanidis and M. Pieper (New York, NY: Springer), 386–404.
- Shaffer, D. W., Squire, K. R., Halverson, R., and Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan* 87, 104–111.
- Siegel, A. W., and White, S. H. (1975). The development of spatial representations of large-scale environments. *Adv. Child Dev. Behav.* 10, 9–55. doi: 10.1016/S0065-2407(08)60007-5
- Strelow, E. R. (1985). What is needed for a theory of mobility: direct perception and cognitive maps—lessons from the blind. *Psychol. Rev.* 92, 226–248. doi: 10.1037/0033-295X.92.2.226
- Strickland, D. (1997). Virtual reality for the treatment of autism. *Stud. Health Technol. Inform.* 44, 81–86.
- Taylor, H. A., Naylor, S. J., and Chechile, N. A. (1999). Goal-specific influences on the representation of spatial perspective. *Mem. Cognit.* 27, 309–319. doi: 10.3758/BF03211414
- Thinus-Blanc, C., and Gaunet, F. (1997). Representation of space in blind persons: vision as a spatial sense? *Psychol. Bull.* 121, 20–42. doi: 10.1037/0033-2909.121.1.20
- Tinti, C., Adenzato, M., Tamietto, M., and Cornoldi, C. (2006). Visual experience is not necessary for efficient survey spatial cognition: evidence from blindness. *Q. J. Exp. Psychol.* 59, 1306–1328. doi: 10.1080/17470210500214275
- Vecchi, T., Tinti, C., and Cornoldi, C. (2004). Spatial memory and integration processes in congenital blindness. *Neuroreport* 15, 2787–2790.
- von Senden, M. (1960). *Space and sight: the Perception of Space and Shape in the Congenitally Blind Before and After Operation*. Free Press.
- von Stulpnagel, R., and Steffens, M. C. (2012). Can active navigation be as good as driving? A comparison of spatial memory in drivers and backseat drivers. *J. Exp. Psychol. Appl.* 18, 162–177. doi: 10.1037/a0027133
- WAIS-III, W. M. S. (1997). *Wechsler Memory Scale, Third Edition (WAIS-III) World List Test*. Cleveland, OH: The Psychological Corporation.
- Waller, D., Hunt, E., and Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence* 7, 129–143.
- Waller, D., Knapp, D., and Hunt, E. (2001). Spatial representations of virtual mazes: the role of visual fidelity and individual differences. *Hum. Factors* 43, 147–158. doi: 10.1518/001872001775992561
- Wilson, P. N., Foreman, N., and Stanton, D. (1997). Virtual reality, disability and rehabilitation. *Disabil. Rehabil.* 19, 213–220. doi: 10.3109/09638289709166530
- Wolbers, T., and Hegarty, M. (2010). What determines our navigational abilities? *Trends Cogn. Sci.* 14, 138–146. doi: 10.1016/j.tics.2010.01.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 August 2013; accepted: 30 March 2014; published online: 01 May 2014.
 Citation: Connors EC, Chrastil ER, Sánchez J and Merabet LB (2014) Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches. *Front. Hum. Neurosci.* 8:223. doi: 10.3389/fnhum.2014.00223
 This article was submitted to the journal *Frontiers in Human Neuroscience*.
 Copyright © 2014 Connors, Chrastil, Sánchez and Merabet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task

Aki Nikolaidis^{1,2*}, Michelle W. Voss³, Hyunkyu Lee⁴, Loan T. K. Vo^{1,5} and Arthur F. Kramer^{1,6}

¹ Neuroscience Program, University of Illinois, Urbana-Champaign, Urbana, IL, USA

² Beckman Institute, University of Illinois Urbana-Champaign, Urbana, IL, USA

³ Department of Psychology, University of Iowa, Iowa City, IA, USA

⁴ Brain Plasticity Institute, San Francisco, CA, USA

⁵ Department of Electrical Engineering, Tan Tao University, Long An, Vietnam

⁶ Department of Psychology, University of Illinois Urbana-Champaign, Urbana, IL, USA

Edited by:

Heleen A. Slagter, University of Amsterdam, Netherlands

Reviewed by:

Sebastian J. Lipina, Unidad de

Neurobiología Aplicada (UNA,

CEMIC-CONICET), Argentina

Roberto Colom, Universidad

Autónoma de Madrid, Spain

*Correspondence:

Aki Nikolaidis, Beckman Institute,
University of Illinois

Urbana-Champaign, 405 North

Mathews Avenue, Urbana, IL 61801,

USA

e-mail: g.aki.nikolaidis@gmail.com

Researchers have devoted considerable attention and resources to cognitive training, yet there have been few examinations of the relationship between individual differences in patterns of brain activity during the training task and training benefits on untrained tasks (i.e., transfer). While a predominant hypothesis suggests that training will transfer if there is training-induced plasticity in brain regions important for the untrained task, this theory lacks sufficient empirical support. To address this issue we investigated the relationship between individual differences in training-induced changes in brain activity during a cognitive training videogame, and whether those changes explained individual differences in the resulting changes in performance in untrained tasks. Forty-five young adults trained with a videogame that challenges working memory, attention, and motor control for 15 2-h sessions. Before and after training, all subjects received neuropsychological assessments targeting working memory, attention, and procedural learning to assess transfer. Subjects also underwent pre- and post-functional magnetic resonance imaging (fMRI) scans while they played the training videogame to assess how these patterns of brain activity change in response to training. For regions implicated in working memory, such as the superior parietal lobe (SPL), individual differences in the post-minus-pre changes in activation predicted performance changes in an untrained working memory task. These findings suggest that training-induced plasticity in the functional representation of a training task may play a role in individual differences in transfer. Our data support and extend previous literature that has examined the association between training related cognitive changes and associated changes in underlying neural networks. We discuss the role of individual differences in brain function in training generalizability and make suggestions for future cognitive training research.

Keywords: cognitive training, neuroplasticity, transfer, working memory, video games

INTRODUCTION

Cognitive neuroscience has begun to explore the possibility of enhancing working memory through the use of videogame-based training products. It has been demonstrated that such videogame training can have a positive impact on the performance of untrained tasks (Green and Bavelier, 2003, 2007; Boot et al., 2008, 2010; Thorell et al., 2009; Van Muijden et al., 2012). A predominant hypothesis of how this occurs is that training affects untrained tasks when they share overlapping cognitive or neural processes with the training (Jonides, 2004; Dahlin et al., 2008b). This extends an older hypothesis in which transfer of training is based on behavioral overlap between trained and untrained tasks (Woodworth and Thorndike, 1901).

Working memory is a cognitive construct that represents the ability to encode, store, and manipulate information in memory

(Baddeley, 1992; D'Esposito et al., 1995). Several brain regions in the frontal and parietal cortices and striatum (caudate, putamen) are known to be involved in working memory, including the dorsal lateral pre-frontal cortex (Braver et al., 1997; D'Esposito et al., 2000; Funahashi, 2006), superior parietal lobe (SPL) and precuneus, (Cohen et al., 1997; Henson et al., 2000; Pessoa et al., 2002; Dahlin et al., 2008b; Koenigs et al., 2009), and caudate (Levy et al., 1997; Postle and D'Esposito, 1999, 2003; Bäckman et al., 2011). The involvement of these regions in working memory suggests that individual differences in the function of these regions may also be linked to individual differences in working memory performance (Kane and Engle, 2002). Similarly, previous research has demonstrated that individual differences in the volume of certain brain regions that are important for working memory and procedural learning, such as the striatum, predict learning

in complex videogame training (Erickson et al., 2010; Basak et al., 2011).

While previous research has demonstrated working memory training transfers selectively to untrained tasks that share cognitive and neural processes (measured by functional MRI activation) with the training task (Dahlin et al., 2008b), it is undefined how training-induced changes in the neural representation of the training task are related to performance changes in untrained tasks. The neural representation of a task can manifest in a variety of contexts and neuroimaging measurements, but in the current study we use this term to refer to patterns of activation during the training task as measured by an fMRI blood-oxygenation-level dependent (BOLD) contrast between task engagement and quiet rest. As a trainee learns a task or skill, the neural representation of the task changes considerably, both within and between training sessions, including increases and decreases in activation (Garavan et al., 2000; Kelly and Garavan, 2005; Kelly et al., 2006; Dayan and Cohen, 2011). How these changes in the neural representation relate to performance changes in untrained tasks has yet to be examined; however, extending the shared cognitive processing and neural overlap hypothesis, it is reasonable to predict that the plasticity in working memory associated brain regions following working memory videogame training, as measured by changes in brain activation patterns during game play, should relate to changes in the performance of an untrained working memory task.

While it is understood that training can induce changes in task-associated brain activity, it is unclear whether these changes will be an increase or decrease in activation (Buschkuhl et al., 2012); therefore in the current study we remain agnostic to the direction of change in training-associated brain activity. Instead we assert that greater performance changes in a working memory task (Sternberg Memory Search, SMS) should be mirrored by greater plasticity (measured by post-minus-pre brain activity in the training task) in the brain regions associated with working memory. Working memory describes the cognitive processes of storing, manipulating, and updating information in memory (Baddeley, 1992). Similarly, the SMS task taps working memory by asking participants to store and update sets of letters in memory (Sternberg, 1966), and accordingly many other studies have used this task or similar tasks as a measure of the storage and maintenance of information in working memory (Awh et al., 1996; Rypma and D'Esposito, 1999; Raghavachari et al., 2001; Jensen and Tesche, 2002). Previous research has demonstrated that individual differences in working memory performance, assessed independently of MRI scanning, are linked to working memory task-based activation in regions associated with working memory, such as the prefrontal cortex and regions of the parietal cortex (Kane and Engle, 2002; Todd and Marois, 2005). These findings offer further support for the prediction that individual differences in training-induced frontal-parietal plasticity during a working memory oriented training task would relate to individual differences in performance changes in an untrained working memory task.

In the current study, trainees performed two untrained tasks before and after training with Space Fortress for 15 2-h sessions. Before and after training, participants also were scanned

using fMRI while playing Space Fortress. To test whether performance changes in an untrained working memory task can be predicted by plasticity in regions associated with working memory, we first correlated pre-training brain activity during game play with changes in performance of two untrained tasks. We performed the same analyses on the post-training brain activity, and these “pre- and post-analyses” served as the basis of our “plasticity analysis,” in which we investigated the relationship between training-induced plasticity and performance changes in untrained tasks. To test whether the association between brain activity during Space Fortress would be related only to untrained tasks that shared cognitive overlap with Space Fortress, we used tasks that were cognitively similar or dissimilar to the processes occurring during the training task.

Space Fortress is an interactive, score based, complex videogame that has a long history of use as a multisensory training tool (Fabiani et al., 1989; Gopher et al., 1989; Donchin, 1995; Kramer et al., 1995, 1999; Vo et al., 2011); it makes high demands of working memory storage and updating, motor control, and attention. The structure of the SMS task has components that are directly related to activities in the Space Fortress training. Specifically, both tasks ask participants to store and update sets of letters for a subsequent response. Furthermore, the response pattern in the SMS task, whether a letter belonged to the most recent letter set, is mirrored directly by Space Fortress in asking participants whether the letter on screen refers to a “friend” or “foe,” which is determined by a letter set given before each Space Fortress trial. Given that Space Fortress engages working memory, and that the SMS task largely mirrors the working memory storage and updating components of Space Fortress, we hypothesize that individual differences in the neural representation of the Space Fortress game will relate to individual differences in performance changes in the SMS task. Furthermore, given that individual differences in the function of regions associated with working memory are likely related to individual differences in working memory performance (Kane and Engle, 2002), we further hypothesize that activity during Space Fortress in regions associated with working memory relates to individual differences in changes in SMS task performance. For our second untrained task, we used the Change Detection (CD) task. This task functions as a control task because while it taps into attention and working memory processes (Pashler, 1988; Rensink, 2002; Baddeley, 2003), the specific cognitive processes in the CD task are quite distinct from that of the Space Fortress game. For example, Space Fortress asks subjects to monitor changes in a symbol at the bottom of the screen, and if they respond when a dollar sign appears twice in a row, they can receive a bonus. This type of CD differs considerably from the CD task in which visual field changes are neither of a predictable type or location. Space Fortress is also a visually simple game, with easily discernable text symbols, and does not require identification of any masked changes; unlike the CD task, which involves both complex real street scenes, with subtly modified scenes, which are separated by a mask. Based on these differences in the both the dorsal and ventral visual attention components of these two tasks, we hypothesize that activity during Space Fortress will not be associated with individual differences in changes to performance in the CD task.

As hypothesized we show that individual differences in functional activation in pre and post fMRI sessions predict individual differences in performance changes to the SMS task; furthermore, we confirmed our hypothesis of no relationship between functional activation in either pre- or post-fMRI sessions and individual differences in performance changes to the CD task. The results of these two tasks taken together suggest that the neural representations of a training task relate more closely to learning in untrained tasks that share higher degrees of cognitive similarity with the training task, which supports previous research showing that training selectively affects untrained tasks with shared cognitive processes and neural overlap (Dahlin et al., 2008b). The results of these analyses gave us a set of regions to use in our subsequent plasticity analysis, in which we investigated the relationship between training-induced plasticity and performance changes in untrained tasks. As we hypothesized, our pre- and post-analyses only found significant results with the SMS task; therefore, we conducted the plasticity analysis on the SMS task and not the CD task.

Of these regions in which pre- and post-analyses identified a significant association with performance changes in the SMS task, we hypothesized that greater plasticity in working memory associated regions would occur in individuals with greater performance gains in the SMS task. Therefore, for our “plasticity analysis” we created spherical regions of interest (ROIs) surrounding the statistical peaks of the group-level maps from the pre- and post-analyses. To measure the plasticity in brain activity in these regions, we extracted the mean percent signal change from these regions, and took a post-minus-pre difference of the game play compared to fixation contrast. We then used a multiple regression model based on the activity differences in these ROIs to predict performance changes in our untrained tasks. Our results support the neural and overlap hypothesis because we show that the post-minus-pre activity differences in regions associated with working memory, such as the SPL, predicted a significant percentage of the variance in performance changes in the SMS task. These findings suggest that changes in a trainee’s neural representation of a training task may predict performance changes of untrained tasks that share cognitive or neural processes with training tasks. Furthermore, while some studies have found weakly significant or non-significant training-induced improvements in untrained tasks at the group level, our results demonstrate that analyzing the relationship between brain activity and untrained task performance at the individual level does reveal a significant association between training-induced plasticity and performance in untrained tasks.

METHODS

PARTICIPANTS

The University institutional review board (IRB) approved this study. We used flyers posted throughout campus as well as online advertisements to recruit participants. To determine eligibility, we asked potential participants to complete a survey about their video game habits, and experimenters determined the participants’ eligibility with individual in-person interviews. These in-person interviews addressed the participants’ health, as well as a more detailed assessment of their video game habits. All participants (1)

played videogames less than 4 h per week, (2) were right handed, (3) were free from psychiatric illness, neurological disease, and metallic implants, (4) had signed an informed consent form, (5) had normal color vision, (6) had a corrected visual acuity of 20/20 or better, and (7) were between the ages of 18 and 30. For our final sample of 45 trainees, there were $N = 27$ females, with a mean age of 21.74 years ($SD = 5.09$) and mean education of 15.71 years ($SD = 3.27$). We also had a minimal-contact control group of 25 participants; however, given that the current study focuses on individual differences in the effect of Space Fortress training, we did not include the control group in our analyses.

SPACE FORTRESS

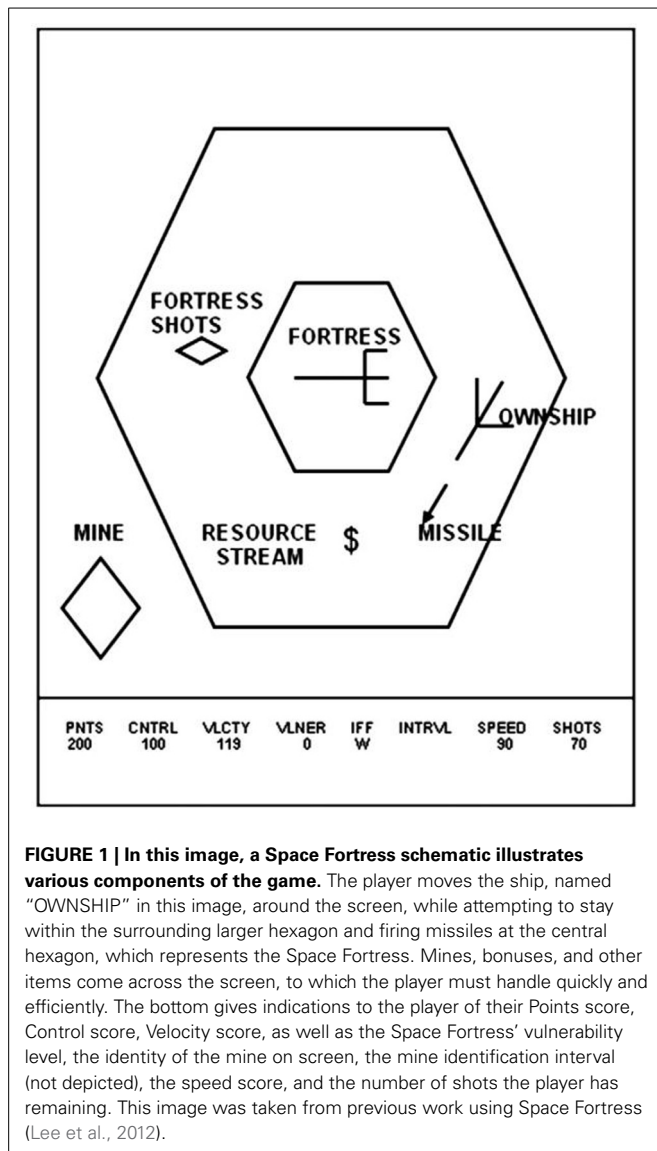
Space Fortress was developed to study the effect of different training strategies on learning, retention and transfer within the context of a rich and cognitively complex task (Figure 1). Playing Space Fortress requires complex motor skills, procedural learning, and working memory. The game score is compartmentalized into four subcategories measuring: (1) points: successfully destroying the space fortress with 10 successive missiles spaced 250 ms apart followed by rapid double shots spaced less than 250 ms; (2) velocity: keeping the ship’s movement within a predefined speed range; (3) control: moving the ship only within a predefined allowable area in a frictionless environment without braking; (4) speed: handling friendly and enemy mines quickly and precisely. In addition to these tasks, the participant must maintain three letters in working memory that identify mines as friend or foe. Furthermore, the participant must monitor a stream of symbols that will occasionally present two dollar signs (\$) in direct succession, which is an indication of a bonus for the player. For a more detailed explanation of Space Fortress, see Mané and Donchin (1989) or Lee et al. (2012).

Training procedure

All participants watched a 20-min instructional video that explained the details of the Space Fortress game, followed by a 5-min summary video and six 3-min games to practice before entering a 3T Siemens Trio scanner. Over the next several weeks, the trainees played the game for a total of 30 h, split into 15 2-h training sessions. Following training, all participants were scanned again with the pre-training protocol.

NEUROPSYCHOLOGICAL ASSESSMENT BATTERY

The participants performed pre and post training neuropsychological tasks (i.e., untrained tasks) in three general categories: visual-attention, memory, and multimodal task performance, to measure baseline cognitive abilities and changes of performance as a result of extended Space Fortress practice. These tests have been previously described in detail (Lee et al., 2012). In the current study, we focus on two tasks including the SMS task and a CD task. We focus on the SMS task because this task closely mirrors distinct cognitive components of the Space Fortress training. For example, the SMS task requires subjects to consolidate and maintain visual information within working memory in a fashion similar to Space Fortress. We focus on a CD task because similarly to the SMS task it is thought to tap into attention and the working memory (Pashler, 1988; Rensink, 2002; Baddeley,



2003). However, it is important to note that the stimulus-response processes involved and the contents to be remembered in this task (scenes) are quite distinct from those in the Space Fortress training and SMS task (letters and symbols).

Sternberg memory search task

In the SMS task, participants viewed a set of 3 or 5 letters (duration: 1200 ms) followed by a pause (1500 ms), and then a brief presentation of a letter (Sternberg, 1966). Participants needed to respond as quickly and accurately as possible whether this letter belonged to the previously viewed set of letters. Our participants received accuracy feedback for 32 practice trials before being tested on 96 trials without feedback. The SMS task uses reaction times and accuracy as outcome variables. We used accuracy alone to measure performance, because unlike the SMS task, during Space Fortress, there is a delay before the stimulus can be flagged as friend or foe, therefore subjects are encouraged to respond accurately rather than quickly. Conversely, each trial in

Space Fortress lasts, on average, longer than each trial of the SMS task, making larger demands on working memory maintenance and therefore accuracy of stimulus-recognition. Performance was measured by averaging accuracy scores in the 3 and 5 letter set conditions. The SMS task taps the storage and maintenance of information in working memory because participants are asked to store letter sets in memory over a delay period, and update this letter set in each trial (Sternberg, 1966).

Change detection

In a single trial the participants viewed a repeating cycle of four images: a street scene (240 ms), a gray interruption image (80 ms), a modified version of the original street scene (240 ms), and then another gray interruption image (80 ms), after which the cycle repeated. We asked participants to detect and report a difference between the original and modified image. If they did not detect the difference after 60 s of repeated cycling through the screens, they continued onto the next trial, for a total of 24 trials. We assessed CD accuracy by determining the percentage of correct trials out of all trials that contained a modified image (22 trials).

MRI DATA ACQUISITION

In the MRI sessions, we collected an MPAGE T1-weighted high-resolution structural volume with 144 contiguous axial slices, collected in an ascending fashion and parallel to the anterior posterior commissure line ($160 \times 192 \times 144$ voxels, voxel size $1.33 \times 1.33 \times 1.30$ mm, echo time (TE) = 3.87 ms, repetition time (TR) = 1800 ms, field of view (FOV) = 256 mm). Then, for the Space Fortress scans, we acquired three runs of T2* weighted EPI images for BOLD signal acquisition (TE = 25 ms, TR = 2 s, Flip angle 80° , voxel size $3.475 \times 3.475 \times 4$ mm, 28 slices, 64×64 voxels matrix, BOLD volumes in each functional scan = 115). While in the scanner, the participants alternated between 30-s blocks of fixating on a central cross (Fixation), passively viewing a recorded session of an experienced player's Space Fortress session, and playing the full Space Fortress game (Space Fortress). We began with a sample size of 50 trainees, and we excluded participants based on excessive motion artifacts. All images were collected on a 3T Siemens TRIO MRI scanner.

MRI pre-processing and analysis

All pre-processing and subsequent analyses of the MRI data were performed using FSL (FMRIB Software Library) (Smith et al., 2004; Woolrich et al., 2009; Jenkinson et al., 2012). We applied rigid body motion correction using MCFLIRT (Jenkinson et al., 2002), and then used BET to remove non-brain structures (Smith, 2002). We applied spatial smoothing using a Gaussian kernel with an 8.0 mm full width half maximum and applied a temporal high-pass filter of 220 s to remove low-frequency signal of non-interest. For the individual-level analyses of each participant, the hemodynamic response was modeled and convolved with a double gamma function in each of the three individual-level runs. Each of the three runs was registered linearly to the subject's MPAGE using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Then, individual-level statistical maps were forwarded to a fixed effects analysis, and these results were linearly registered to the standardized 2 mm ICBM-152 Montreal Neurological Institute(MNI) Template (Mazziotta et al., 2001).

See **Figure 2** for flow chart of the current study's analyses and results.

ASSOCIATING THE SPACE FORTRESS BOLD SIGNAL WITH CHANGES IN UNTRAINED TASK PERFORMANCE

It is likely that the BOLD signal in the Space Fortress > Fixation contrast is more informative of individual differences in working

memory processing in the context of a complex task, compared to a Passive > Fixation contrast. Similarly, we assert that the Space Fortress > Fixation contrast is more informative than a Space Fortress > Passive contrast for the reason that in the Passive condition the participants may still engage in working memory processes, which we are interested in investigating in our study. Therefore, contrasting Space Fortress with Passive viewing may remove such working memory-associated activity of interest. Because of these reasons, we chose to focus on the Space Fortress > Fixation contrast in our analyses, and therefore, all mentions of brain activity and refer to this contrast.

For our higher-level analysis, the individual level of fixed effect images of the Space Fortress > Fixation contrast were submitted to a mixed effects group analysis in FSL's FEAT (Worsley, 2001). We performed these analyses for two reasons, first to investigate whether individual differences in brain activity before or after training relate to individual differences in gains in untrained tasks, and second, to find a targeted set of areas to use in our subsequent plasticity analyses of whether individual differences in training-induced plasticity relate to individual differences in performance change in untrained tasks.

In order to examine the correlation between the BOLD signal and changes in performance in the untrained tasks, we used the performance change scores (i.e., the tasks performed before vs. after the completion of Space Fortress training) from the SMS and CD tasks as regression covariates in two separate group analysis design matrices for both the pre and post fMRI scan, yielding four total group analyses in total. We used a Z-statistic threshold of 1.96 and cluster *p*-value threshold of 0.01 for our mixed effects statistical maps. This Z threshold would be considered low for a standard GLM contrast; however, since in this analysis we correlated BOLD contrast with a behavioral variable, we believe that a Z threshold of 1.96 (two tailed Type I error rate of 0.05) and a cluster *p*-value threshold of 0.01 are reasonable. This analysis yields a statistical brain map of Z scores that reflect the strength of the association between individual differences in changes in untrained performance and the Space Fortress > Fixation contrasted BOLD signal, and we performed this mixed effects analysis in both pre- and post-fMRI scans separately (Table 1; Figure 3). The Harvard-Oxford Cortical Structural Atlas and Harvard-Oxford Subcortical Structural Atlas are the probabilistic atlases in FSL that defined our location labels of our ROIs.

MULTIPLE REGRESSION ANALYSIS OF BRAIN PLASTICITY

In order to investigate the effects of brain plasticity in Space Fortress on changes to SMS task performance, we created spherical ROIs, (10 mm in diameter and 264 mm³ in volume), surrounding the peak Z-statistic values from the pre- and post-mixed effects analyses. Since the pre- and post-mixed effects analyses looks for regions of the brain that have significant association with performance changes in the untrained tasks, we hypothesized that any generalized learning that occurs during Space Fortress training would manifest in brain plasticity in these same regions. Given that this analysis was based on these statistical peaks, the null mixed-effects results for the CD task prevented us from including the CD task in the multiple regression analyses.

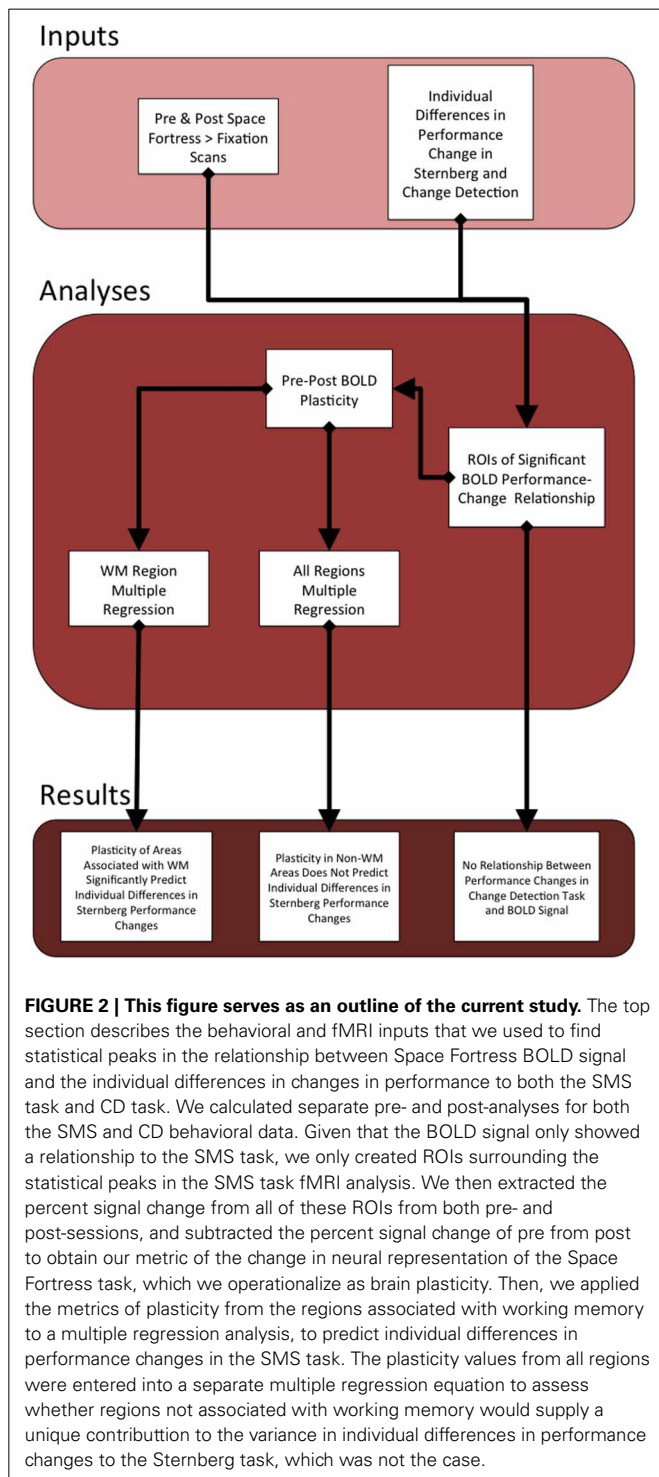


Table 1 | This table summarizes the locations of the fMRI activation predictors in the current study.

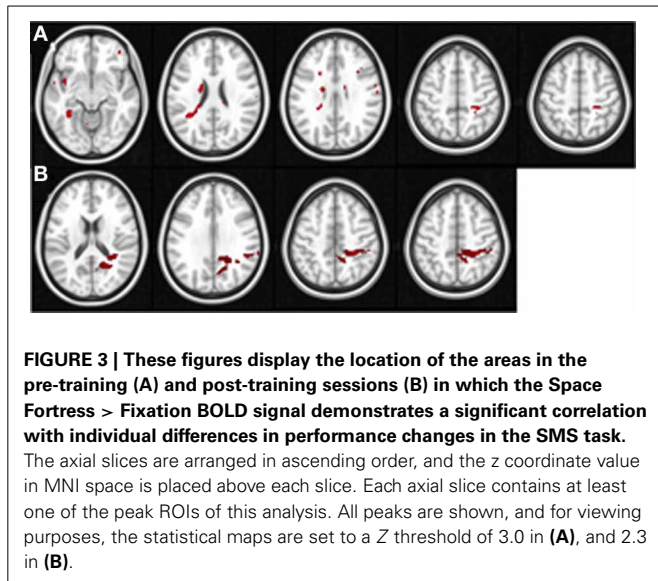
STERNBERG MEMORY SEARCH TASK			
	MNI Coordinates	Peak Z-Value	Space Fortress Pearson's r
PRE-TRAINING			
Caudate-L	−18, −8, 24	4.12	0.523
Fusiform cortex-L	−34, −46, −10	4.11	0.522
Insular cortex-L	−40, 4, −10	3.70	0.483
PCG-R	62, −14, 28	3.16	0.426
PCG SPL-R	28, −36, 52	3.98	0.510
SPL-R	26, −44, 50	3.24	0.435
Supramarginal gyrus-R	64, −20, 24	3.04	0.413
POST-TRAINING			
PCG-R	18, −36, 48	2.84	0.390
Precuneus-R	22, −56, 18	3.28	0.439
SPL-R	28, −38, 46	3.35	0.447
Supramarginal gyrus-R	50, −40, 30	2.81	0.386
CHANGE DETECTION TASK			
	MNI coordinates	Space Fortress Z-Value	Space Fortress Pearson's r
PRE-TRAINING			
No areas passed a threshold of $Z = 1.96$			
POST-TRAINING			
No areas passed a threshold of $Z = 1.96$			

The top portion of the table corresponds to ROIs that demonstrated a significant relationship between pre-training BOLD activity and performance changes to the SMS task, while the middle of this table corresponds to ROIs that demonstrated a significant relationship between post-training BOLD activity and performance changes to the SMS task. The pre-training BOLD activity peaks all belong to a single 15,725 voxel cluster, and the post-training BOLD activity peaks all belong to a single 1896 voxel cluster. The bottom portion of the table shows the null results found in the CD task. In the first column, the names of the approximate regions in which the ROIs are located, as according to the Harvard-Oxford Cortical and Subcortical Structural Atlases. Each ROI name is accompanied by an L to denote that it belongs to the left hemisphere, or an R to denote that it belongs to the right hemisphere. In the second column, MNI coordinates of the same ROIs are provided. The third column provides the significance level of the relationship between transfer and the BOLD signal for the Space Fortress > Fixation contrast at that given ROI. All ROIs in this column passed a $Z = 1.96$ significance threshold and a $p = 0.01$ cluster threshold. The fourth column offers a reverse of the Fisher's Z transformation to offer an averaged Pearson's r for all voxels contained in each 10 mm peak ROI. These r -values were calculated based on the following formula: $r = \frac{Z^2}{Z^2 + N}$, where Z is the significance value, and N corresponds to sample size.

We extracted percent signal change (%SC) from these ROIs from both sessions' fMRI scan and took the post-minus-pre difference. This step allowed us to restrict our search to areas of interest that were related to performance changes in the SMS task, while not causing problems of statistical resampling because we created new metrics of plasticity in these ROIs by creating a post-minus-pre training BOLD contrast, rather than using the brain activity from either session alone. By entering these values into a multiple regression we were able to assess the percentage of variance in changes in SMS task performance that are accounted for by changes in brain activation during performance of the Space Fortress game.

We used SPSS to calculate a backward multiple regression, in which the full model of all variables is considered, and then each variable is iteratively removed and the significance of the model is reassessed. In all iterations, the variable that contributes the least variance to the model is removed until each remaining variable contributes significant variance to the regression. This method results in a regression model in which each independent variable predicts a significant percentage of the variance in the dependent variable, but unlike the stepwise multiple regression model, the

backwards model is not biased by the order in which the variables are added to the model, since all are considered initially. Given that the working memory component of Space Fortress may contribute to changes in processing in the SMS task, our first multiple regression used plasticity values only from brain regions that have been shown to be involved in working memory: the SPL, caudate, precuneus, and postcentral gyrus (PCG) (Cohen et al., 1997; Levy et al., 1997; Henson et al., 2000; Pessoa et al., 2002; Dahlin et al., 2008a,b; Koenigs et al., 2009; Bäckman et al., 2011); we used backwards multiple regression to create a model from the plasticity in these regions. Since this analysis included only a restricted set of ROIs, a counter-hypothesis would be that any plasticity, regardless of brain region, could have significantly predicted performance changes to the Sternberg task. To confirm our hypothesis that the plasticity in working memory regions specifically accounted for the variance in performance changes to the SMS task, we compared the regression model of working memory regions alone, to a larger model, in which we added the remaining regions that were deemed significant in the mixed effects analysis: the supramarginal gyrus, temporal fusiform cortex, and the insular cortex.



RESULTS

PERFORMANCE CHANGES ON UNTRAINED TASKS AFTER SPACE FORTRESS TRAINING

On average, participants did not demonstrate significantly different pre-to-post scores in the SMS task [pre-training mean 94.4%; 95% CI = 93.2–95.6%; post-training mean 93.5%; 95% CI = 91.6–95.3%; paired t -test $t_{(44)} = 1.32$, $p > 0.05$]. We also did not find significant pre-to-post differences in the CD task [pre-training mean 85.3%; 95% CI = 82.3–88.3%; post-training mean 86.8%; 95% CI = 83.9–89.8%; paired t -test, $t_{(44)} = 0.935$, $p > 0.05$]. We did not find a reaction-time accuracy trade off from pre- to post-session for either of the tasks. The current study focuses on individual differences in performance changes of these tasks.

NEURAL REPRESENTATION OF SPACE FORTRESS VIDEOGAME PREDICTS INDIVIDUAL DIFFERENCES IN CHANGES TO PERFORMANCE IN AN UNTRAINED TASK

In our pre- and post-analyses we investigated the relationship between BOLD activity during Space Fortress performance and individual differences in performance changes in untrained tasks (SMS and CD tasks). We did so by using performance change scores as covariates of interest in a between-subjects analysis for both pre- and post-sessions separately (Table 1). We found that pre- and post-session BOLD signal in several frontal, parietal, and subcortical regions demonstrated a significant association with performance changes in the SMS task, but not in the CD task (Table 1). During the pre-training fMRI session, activity in the caudate, PCG, and SPL was positively associated to individual differences in performance changes in the SMS task (Figure 3A). In the post-training fMRI session, activity in the SPL, precuneus, and PCG were positively associated with performance changes in the SMS task (Figure 3B). These results corroborate our hypothesis that patterns of brain activation obtained during the performance of the Space Fortress task would be associated

with individual differences in performance change in the SMS task and that these relationships would be manifested in regions of the brain known to be involved in working memory, such as the SPL, caudate, precuneus. These results also support our hypothesis that this relationship would exist for performance changes in untrained tasks sharing cognitive processes with the training and not those using dissimilar cognitive processes. Since we did not find any brain regions with a significant association between signal to performance changes in the CD task, and the ROIs for our regression analyses were created by extracting the data surrounding the statistical peak activations in the mixed effects analysis, we could not include the CD task in the multiple regression analyses.

FRONTAL-PARIETAL BRAIN PLASTICITY PREDICTS INDIVIDUAL DIFFERENCES IN IMPROVEMENTS IN AN UNTRAINED WORKING MEMORY TASK

In our plasticity analysis we investigated whether changes in the neural representation of Space Fortress predicted a significant percentage of the variance in performance changes in the SMS task. First, we included the mixed effects derived ROIs in the SPL, caudate, precuneus, and PCG, which have all been associated with working memory (Cohen et al., 1997; Levy et al., 1997; Postle and D'Esposito, 1999, 2003; Henson et al., 2000; Pessoa et al., 2002; Dahlin et al., 2008b; Koenigs et al., 2009; Bäckman et al., 2011). We found that using backwards multiple regression, changes in %SC predicted 32% of the variance in the performance changes to the SMS task [Working Memory Model- R^2 : 0.37; adjusted R^2 : 0.32; $F_{(44)} = 8.040$, $p < 0.01$] (Table 2). Adjusted R^2 is an estimate of how well the same model would perform in an independent sample taken from the same population. These results support the notion that plasticity in regions important for working memory would have an impact on working memory processes of similar tasks. Greater decreases in activity in the SPL and PCG (Standardized Beta = -0.347 , and -0.264 , respectively), and greater increases in activity in the precuneus (Standardized Beta = 0.392) were associated with greater improvements to the SMS task. These standardized beta values indicate the importance of each variable in the model. Therefore, the increases in activity in the precuneus, and decreases in the SPL and PCG contribute to our model's significant prediction in declining order of importance.

To test our hypothesis that the plasticity in working memory regions alone would account for the variance in performance changes to the SMS task we used all the ROIs from the mixed effects analyses in a larger multiple regression analysis. We added post-minus-pre activity change scores from ROIs in the SPL, caudate, PCG, precuneus, supramarginal gyrus, temporal fusiform cortex, and the insular cortex to the analysis. We found a non-significant 3% R^2 improvement of the regression model from 37% (Figure 4) (working memory associated regions only) to 40% (all regions) ($F = 1.300$, $p > 0.05$). These results suggest that individual differences in activity-changes in working memory associated areas may be particularly important for predicting individual differences in performance changes in similar working memory tasks, and that plasticity in regions that have not been shown to be involved in working memory may not contribute

Table 2 | This table summarizes the progression of the backwards multiple regression model from including the plasticity of several regions associated with working memory to including only the Precuneus, PCG SPL, and PCG.

Model Summary				
Model	<i>R</i>	<i>R</i> ²	Adjusted <i>R</i> ²	Std. error of the estimate
1	0.614 ^a	0.377	0.259	0.084
2	0.614 ^b	0.377	0.279	0.082
3	0.614 ^c	0.377	0.297	0.081
4	0.613 ^d	0.376	0.314	0.080
5	0.609 ^e	0.370	0.324	0.080

In backwards multiple regression, the first model begins with all available information, and in each round removes the least important variable until it converges on an optimal solution. This iterative model attempts to maximize the adjusted *R*²-value, which is an estimate of the explanatory power of this model in an independent sample drawn from the same population.

^aPredictors: (Constant), SPL (32, -42, 62), SPL (26, -44, 50), Caudate (-18, -8, 24), Precuneus (22, -56, 18), PCG SPL (28, -36, 52), PCG (18, -36, 48), PCG (62, -14, 28).

^bPredictors: (Constant), SPL (32, -42, 62), SPL (26, -44, 50), Precuneus (22, -56, 18), PCG SPL (28, -36, 52), PCG (18, -36, 48), PCG (62, -14, 28).

^cPredictors: (Constant), SPL (32, -42, 62), SPL (26, -44, 50), Precuneus (22, -56, 18), PCG SPL (28, -36, 52), PCG (62, -14, 28).

^dPredictors: (Constant), SPL (32, -42, 62), Precuneus (22, -56, 18), PCG SPL (28, -36, 52), PCG (62, -14, 28).

^ePredictors: (Constant), Precuneus (22, -56, 18), PCG SPL (28, -36, 52), PCG (62, -14, 28).

to performance changes in such tasks. This result is important because it gives an insight into how individual differences in plasticity that occur during training can determine how the trainees change their performance on untrained tasks.

DISCUSSION

IMPLICATIONS OF THE CURRENT STUDY

The findings of our plasticity analysis demonstrate that changes in BOLD signal in the SPL, PCG, and precuneus, from pre- to post-training using a videogame with a working memory component, predict changes in performance in an untrained working memory task. Previous research has simultaneously found significant changes in activation in working memory associated regions, such as the SPL and caudate, in response to working memory training along with improvements in an untrained working memory task (Dahlin et al., 2008b). Our findings extend this research by demonstrating that the changes in functional activation that occur during working memory training predict individual differences in changes in untrained working memory task performance. These findings suggest that as the functional processing of Space Fortress changes following training, so does the functional processing of the SMS task, which supports previous notions that the training-induced plasticity in brain regions associated with training and untrained tasks is associated with transfer to untrained tasks (Jonides, 2004; Dahlin et al., 2008b). These findings also confirm hypotheses of others suggesting that the frontal-parietal network serves as a basis for transfer between working memory tasks (Klingberg, 2010).

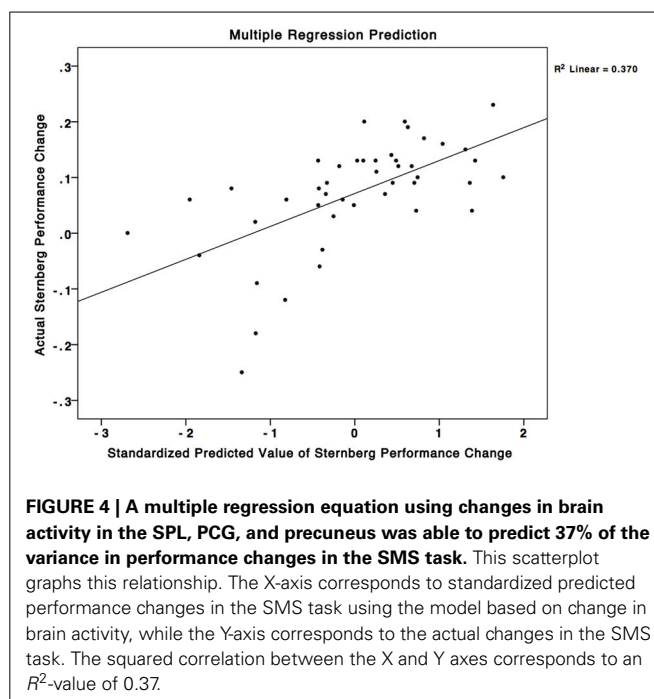


FIGURE 4 | A multiple regression equation using changes in brain activity in the SPL, PCG, and precuneus was able to predict 37% of the variance in performance changes in the SMS task. This scatterplot graphs this relationship. The X-axis corresponds to standardized predicted performance changes in the SMS task using the model based on change in brain activity, while the Y-axis corresponds to the actual changes in the SMS task. The squared correlation between the X and Y axes corresponds to an *R*²-value of 0.37.

In our pre and post analyses we also found a distributed set of brain regions in which the Space Fortress > Fixation BOLD signal at either pre or post fMRI scan correlated with performance changes in a working memory task. This analysis included regions that have been associated with working memory in previous research, such as the caudate (Levy et al., 1997; Postle and D'Esposito, 1999, 2003; Bäckman et al., 2011) and SPL (Cohen et al., 1997; Henson et al., 2000; Pessoa et al., 2002; Dahlin et al., 2008b; Koenigs et al., 2009). Given that previous literature has demonstrated that these regions play an important role in working memory, our findings suggest that these regions may also play a role in the relationship between training in a complex videogame, such as Space Fortress, and individual differences in performance changes in a working memory task.

Counter to our hypotheses, activation in brain regions aside from those associated with working memory and updating, such as the temporal-occipital fusiform cortex, were also associated with performance changes in an untrained working memory task. One interpretation of these findings is that the relationship between brain activity during Space Fortress and the untrained working memory task performance change is non-specific to regions associated with working memory. However, in follow-up analyses the multiple regression model that included these additional regions showed no improvement in model performance compared to the working memory model. This aids our interpretation of the results by indicating that the relationship between activity during Space Fortress and changes in performance in a working memory task are specifically explained by changes in activity in regions associated with working memory.

Space Fortress is a complex task, and it makes demands on working memory, motor control, and attention. While single components of the task are related to other cognitive processes, the training as a whole is different and more difficult than many

individual cognitive tasks. In light of the multimodal nature of the training, one may expect that the learning that occurs during training would only be represented by performance changes in tasks similar to the whole training task. However, given that we found relationships between Space Fortress brain activity and changes in performance to the cognitively similar SMS task but not the cognitively dissimilar CD task, our results suggest that training-component similarity may be sufficient for the training to affect untrained tasks. Furthermore, this similarity may be important for assessing how individual differences in brain plasticity predict changes in performance in untrained tasks, which agrees with previous notions that cognitive overlap between training and untrained tasks is critical in predicting transfer of training (Jonides, 2004; Dahlin et al., 2008b). We interpret these results as suggesting that training on a complex task induces a change in the representation of the training task across a variety of functional brain networks, and that these changes may affect a group of untrained tasks that are limited to those tasks which are functionally represented in a similar manner as a component of the complex training task that induces significant brain plasticity. This interpretation is supported by previous literature suggesting that training should preferentially affect those tasks that share elements of neural (Dahlin et al., 2008a,b) or behavioral similarity (Woodworth and Thorndike, 1901).

LIMITATIONS

While our findings offer several suggestions that are in line with previous cognitive training literature, they should be interpreted with some limitations in mind. First, it is well known that a properly controlled cognitive training experiment should include a group which trains with a control training task that is selected or created to minimize the difference in expectancy effects between the training and the control group (Boot et al., 2011). In other words, all participants should be blind to whether they belong to an experimental or control group, which is thought to minimize the effect of their own expectations on their training outcome. This is quite difficult to achieve in most laboratory settings. Nevertheless, to investigate questions of the efficacy of specific training components, some researchers have used modified high and low interference versions of a working memory task as experimental and active control conditions, respectively (Oelhafen et al., 2013). Our study does not include such an active control group with the removal of a single training component. Therefore, we cannot make strong conclusions on the specific effect of the working memory component of Space Fortress training on our untrained working memory task. Furthermore, we cannot make claims that Space Fortress uniquely had such an effect, as compared to a similarly complex multimodal training task. Thus, our hypotheses concern directly how variation in an individual's representation of the training, recorded by fMRI here but which could also be assessed with sophisticated behavioral metrics such as eye-tracking, predict the trainability of those individuals. We believe our findings provide a good beginning toward the understanding of the inter-relationships between transfer and individual differences in the representation of the training task, which may help future studies assess how to guide this dynamic relationship with the training task to increase the transfer of

training. While our findings agree with what previous findings and theory would suggest, future efforts should aim to include such targeted control groups to account for this effect.

IMPLICATIONS OF CURRENT FINDINGS ON QUESTIONS FOR FUTURE RESEARCH

The current study offers a variety of practical suggestions for future cognitive training studies. First, while a complex training task may have an effect on simpler untrained tasks that share little relationship with the training task in its entirety, the training task is more likely to have an effect on simpler untrained tasks that share cognitive processes with one or more components of the training task. Second, individual differences in the neural representation as well as changes in the neural representation of a training task can be predictive of how the training will affect untrained tasks. These findings suggest that such individual differences may play a critical role in the outcome of a cognitive training regimen. We suggest that future neuroimaging and cognitive training studies also perform assessments of participant motivation toward training, engagement in the task, and personality metrics, all of which may contribute to pre-training trait-like individual differences in how an individual will benefit from a cognitive training regimen. This information will help answer questions of how individual differences in cognitive training-based plasticity are related to individual differences in motivation, personality, and cognitive ability. Understanding these questions will not only allow for improvements in the development of cognitive training programs, but it will also help future researchers explore the topic of transfer with more clarity. Furthermore, in light of the importance of these individual differences, our findings support the suggestions of others that “one-size-fits-all” training regimens may be inappropriate, and training paradigms that cater to individual differences in trainability or other personal attributes may improve the effect of the cognitive training regimen (Jaeggi et al., 2011; Buschkuhl et al., 2012).

While the current study offers insight on how brain plasticity contributes to performance changes in untrained tasks, many questions remain. For example, one important future issue will be elucidating the difference in short-term vs. long-term brain plasticity that occurs within vs. between training sessions. By making these differences more clear, future cognitive training studies could investigate whether short-term and long-term learning uniquely contribute to performance changes in training tasks, or whether individual differences in these short-term and long-term learning lead to individual differences in generalizability of the training.

Finally, we suggest that future research focus on maximizing training-induced brain plasticity by combining cognitive training with other interventions that are thought to encourage states of plasticity, such as exercise, which increases many plasticity associated biomarkers, such as brain-derived neurotrophic factor (Cotman et al., 2007; Hillman et al., 2008; Van Praag, 2009), a protein that is thought to be important for the growth and differentiation of new neurons and synapses (Cotman et al., 2007). Trans-cranial current stimulation is another intervention technique that is thought to encourage neuroplasticity by raising levels of several biochemical markers of plasticity, including myoinositol (Hunter et al., 2013), which is associated with

the long-term potentiation second messenger system, which is important for the growth of synapses (Rango et al., 2008). Given the relationship we found between plasticity in Space Fortress training and performance changes in an untrained task, we suggest combining these other intervention techniques with cognitive training may increase training-induced plasticity thereby increasing the transfer of training.

REFERENCES

- Awh, E., Jonides, J., Smith, E., Schumacher, E., Koeppel, R., and Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. *Psychol. Sci.* 7, 25–31. doi: 10.1111/j.1467-9280.1996.tb00662.x
- Bäckman, L., Nyberg, L., Soveri, A., Johansson, J., Andersson, M., Dahlin, E., et al. (2011). Effects of working-memory training on striatal dopamine release. *Science* 333, 718. doi: 10.1126/science.1204978
- Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201
- Basak, C., Voss, M. W., Erickson, K. I., Boot, W. R., and Kramer, A. F. (2011). Regional differences in brain volume predict the acquisition of skill in a complex real-time strategy videogame. *Brain Cogn.* 76, 407–414. doi: 10.1016/j.bandc.2011.03.017
- Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiani, M., et al. (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychol.* 135, 349–357. doi: 10.1016/j.actpsy.2010.09.005
- Boot, W. R., Blakely, D. P., and Simons, D. J. (2011). Do action video games improve perception and cognition? *Front. Psychol.* 2:226. doi: 10.3389/fpsyg.2011.00226
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., and Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychol.* 129, 387–398. doi: 10.1016/j.actpsy.2008.09.005
- Braver, T., Cohen, J., Nystrom, L., and Jonides, J. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 49–62. doi: 10.1006/nimg.1996.0247
- Buschkuhl, M., Jaeggi, S. M., and Jonides, J. (2012). Neuronal effects following working memory training. *Dev. Cogn. Neurosci.* 2, S167–S179. doi: 10.1016/j.dcn.2011.10.001
- Cohen, J., Perlstein, W., Braver, T., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608. doi: 10.1038/386604a0
- Cotman, C. W., Berchtold, N. C., and Christie, L.-A. (2007). Exercise builds brain health: key roles of growth factor cascades and inflammation. *Trends Neurosci.* 30, 464–472. doi: 10.1016/j.tins.2007.06.011
- D'Esposito, M., Detre, J., Alsop, D., and Shin, R. (1995). The neural basis of the central executive system of working memory. *Nature* 378, 279–281. doi: 10.1038/378279a0
- D'Esposito, M., Postle, B., and Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Exp. Brain Res.* 133, 3–11. doi: 10.1007/s002210000395
- Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., and Nyberg, L. (2008a). Supporting online material for transfer of learning after updating training mediated by the striatum. *Science* 320, 1510. doi: 10.1126/science.1155466
- Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., and Nyberg, L. (2008b). Transfer of learning after updating training mediated by the striatum. *Science* 320, 1510–1512. doi: 10.1126/science.1155466
- Dayan, E., and Cohen, L. G. (2011). Review neuroplasticity subserving motor skill learning. *Neuron* 72, 443–454. doi: 10.1016/j.neuron.2011.10.008
- Donchin, E. (1995). Video games as research tools. *Behav. Res. Methods Instrum. Comput.* 27, 217–223. doi: 10.3758/BF03204735
- Erickson, K. I., Boot, W. R., Basak, C., Neider, M. B., Prakash, R. S., Voss, M. W., et al. (2010). Striatal volume predicts level of video game skill acquisition. *Cereb. Cortex* 20, 2522–2530. doi: 10.1093/cercor/bhp293
- Fabiani, M., Buckley, J., Gratton, G., Coles, M., and Donchin, E. (1989). The training of complex task performance. *Acta Psychol.* 71, 259–299. doi: 10.1016/0001-6918(89)90012-7
- Funahashi, S. (2006). Prefrontal cortex and working memory processes. *Neuroscience* 139, 251–261. doi: 10.1016/j.neuroscience.2005.07.003
- Garavan, H., Kelley, D., Rosen, A., Rao, S. M., and Stein, E. A. (2000). Practice-related functional activation changes in a working memory task. *Microsc. Res. Tech.* 51, 54–63. doi: 10.1002/1097-0029(20001001)51:1<54::AID-JEMT6>3.0.CO;2-J
- Gopher, D., Weil, M., and Siegel, D. (1989). Practice under changing priorities: an approach to the training of complex skills. *Acta Psychol.* 71, 147–177. doi: 10.1016/0001-6918(89)90007-3
- Green, C. S., and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature* 423, 534–537. doi: 10.1038/nature01647
- Green, C. S., and Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychol. Sci.* 18, 88–94. doi: 10.1111/j.1467-9280.2007.01853.x
- Henson, R. N. A., Burgess, N., and Frith, C. D. (2000). Recoding, storage, rehearsal and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia* 38, 426–440. doi: 10.1016/S0028-3932(99)00098-6
- Hillman, C. H., Erickson, K. I., and Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nat. Rev. Neurosci.* 9, 58–65. doi: 10.1038/nrn2298
- Hunter, M. A., Coffman, B. A., Trumbo, M. C., and Clark, V. P. (2013). Tracking the neuroplastic changes associated with transcranial direct current stimulation: a push for multimodal imaging. *Front. Hum. Neurosci.* 7:495. doi: 10.3389/fnhum.2013.00495
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., and Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 841, 825–841. doi: 10.1006/nimg.2002.1132
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *Neuroimage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015
- Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. doi: 10.1016/S1361-8415(01)00036-6
- Jensen, O., and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15, 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x
- Jonides, J. (2004). How does practice makes perfect? *Nat. Neurosci.* 7, 10–11. doi: 10.1038/nn0104-10
- Kane, M. J., and Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychon. Bull. Rev.* 9, 637–671. doi: 10.3758/BF03196323
- Kelly, A. M. C., and Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cereb. Cortex* 15, 1089–1102. doi: 10.1093/cercor/bhi005
- Kelly, C., Foxe, J. J., and Garavan, H. (2006). Patterns of normal human brain plasticity after practice and their implications for neurorehabilitation. *Arch. Phys. Med. Rehabil.* 87, 20–29. doi: 10.1016/j.apmr.2006.08.333
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends Cogn. Sci.* 14, 317–324. doi: 10.1016/j.tics.2010.05.002
- Koenigs, M., Barbey, A. K., Postle, B. R., and Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *J. Neurosci.* 29, 14980–14986. doi: 10.1523/JNEUROSCI.3706-09.2009
- Kramer, A., Larish, J., and Strayer, D. (1995). Training for attentional control in dual task settings: a comparison of young and old adults. *J. Exp. Psychol. Appl.* 1, 50–76. doi: 10.1037/1076-898X.1.1.50
- Kramer, A., Larish, J., Weber, T., and Bardell, L. (1999). *Training for Executive Control: Task Coordination Strategies and Aging. Attention and performance.* Cambridge, MA: The MIT Press.
- Lee, H., Boot, W. R., Basak, C., Voss, M. W., Prakash, R. S., Neider, M., et al. (2012). Performance gains from directed training do not transfer to untrained tasks. *Acta Psychol.* 139, 146–158. doi: 10.1016/j.actpsy.2011.11.003
- Levy, R., Friedman, H. R., Davachi, L., and Goldman-rakic, P. S. (1997). Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *J. Neurosci.* 17, 3870–3882.
- Mané, A., and Donchin, E. (1989). The space fortress game. *Acta Psychol. (Amst)* 71, 17–22. doi: 10.1016/0001-6918(89)90003-6

- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., et al. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 1293–1322. doi: 10.1098/rstb.2001.0915
- Oelhafen, S., Nikolaidis, A., Padovani, T., Blaser, D., Koenig, T., and Perrig, W. J. (2013). Increased parietal activity after training of interference control. *Neuropsychologia* 51, 2781–2790. doi: 10.1016/j.neuropsychologia.2013.08.012
- Pashler, H. (1988). Familiarity and visual change detection. *Percept. Psychophys.* 44, 369–378. doi: 10.3758/BF03210419
- Pessoa, L., Gutierrez, E., Bandettini, P. A., and Ungerleider, L. G. (2002). Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron* 35, 975–987. doi: 10.1016/S0896-6273(02)00817-6
- Postle, B. R., and D'Esposito, M. (1999). Dissociation of human caudate nucleus activity in spatial and nonspatial working memory: an event-related fMRI study. *Cogn. Brain Res.* 8, 107–115. doi: 10.1016/S0926-6410(99)00010-5
- Postle, B. R., and D'Esposito, M. (2003). Spatial working memory activity of the caudate nucleus is sensitive to frame of reference. *Cogn. Affect. Behav. Neurosci.* 3, 133–144. doi: 10.3758/CABN.3.2.133
- Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., et al. (2001). Gating of human theta oscillations by a working memory task. *J. Neurosci.* 21, 3175–3183. Available online at: <http://www.jneurosci.org/content/21/9/3175.short>
- Rango, M., Cogiamanian, F., Marceglia, S., Barberis, B., Arighi, A., Biondetti, P., et al. (2008). Myoinositol content in the human brain is modified by transcranial direct current stimulation in a matter of minutes: a 1H-MRS study. *Magn. Reson. Med.* 60, 782–789. doi: 10.1002/mrm.21709
- Rensink, R. A. (2002). Change detection. *Annu. Rev. Psychol.* 53, 245–277. doi: 10.1146/annurev.psych.53.100901.135125
- Rypma, B., and D'Esposito, M. (1999). The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6558–6563. doi: 10.1073/pnas.96.11.6558
- Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 15, 143–155. doi: 10.1002/hbm.10062
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, 208–219. doi: 10.1016/j.neuroimage.2004.07.051
- Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652
- Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x
- Todd, J. J., and Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cogn. Affect. Behav. Neurosci.* 5, 144–155. doi: 10.3758/CABN.5.2.144
- Van Muijden, J., Band, G. P. H., and Hommel, B. (2012). Online games training aging brains: limited transfer to cognitive control functions. *Front. Hum. Neurosci.* 6:221. doi: 10.3389/fnhum.2012.00221
- Van Praag, H. (2009). Exercise and the brain: something to chew on. *Trends Neurosci.* 32, 283–290. doi: 10.1016/j.tins.2008.12.007
- Vo, L. T. K., Walther, D. B., Kramer, A. F., Erickson, K. I., Boot, W. R., Voss, M. W., et al. (2011). Predicting individuals' learning success from patterns of pre-learning MRI activity. *PLoS ONE* 6:e16093. doi: 10.1371/journal.pone.0016093
- Woodworth, R., and Thorndike, E. (1901). The influence of improvement in one mental function upon the efficiency of other functions.(I). *Psychol. Rev.* 8, 247–261. doi: 10.1037/h0074898
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., et al. (2009). NeuroImage Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45, S173–S186. doi: 10.1016/j.neuroimage.2008.10.055
- Worsley, K. J. (2001). "Statistical analysis of activation images," in *Functional MRI: An Introduction to Methods*, eds P. Jefferard, P. M. Matthews, and S. M. Smith (Oxford University Press), 1–23.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2014; accepted: 07 March 2014; published online: 21 March 2014.
Citation: Nikolaidis A, Voss MW, Lee H, Vo LTK and Kramer AF (2014) Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task. *Front. Hum. Neurosci.* 8:169. doi: 10.3389/fnhum.2014.00169

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Nikolaidis, Voss, Lee, Vo and Kramer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read,
for greatest visibility



COLLABORATIVE PEER-REVIEW

Designed to be rigorous
– yet also collaborative,
fair and constructive



FAST PUBLICATION

Average 85 days from
submission to publication
(across all journals)



COPYRIGHT TO AUTHORS

No limit to article
distribution and re-use



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



SUPPORT

By our Swiss-based
editorial team



IMPACT METRICS

Advanced metrics
track your article's impact



GLOBAL SPREAD

5'100'000+ monthly
article views
and downloads



LOOP RESEARCH NETWORK

Our network
increases readership
for your article

Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on

