

Short- and long-term benefits of cognitive training

Susanne M. Jaeggi^{1,2}, Martin Buschkuhl^{1,2}, John Jonides, and Priti Shah

Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1043

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved May 17, 2011 (received for review March 1, 2011)

Does cognitive training work? There are numerous commercial training interventions claiming to improve general mental capacity; however, the scientific evidence for such claims is sparse. Nevertheless, there is accumulating evidence that certain cognitive interventions are effective. Here we provide evidence for the effectiveness of cognitive (often called “brain”) training. However, we demonstrate that there are important individual differences that determine training and transfer. We trained elementary and middle school children by means of a videogame-like working memory task. We found that only children who considerably improved on the training task showed a performance increase on untrained fluid intelligence tasks. This improvement was larger than the improvement of a control group who trained on a knowledge-based task that did not engage working memory; further, this differential pattern remained intact even after a 3-mo hiatus from training. We conclude that cognitive training can be effective and long-lasting, but that there are limiting factors that must be considered to evaluate the effects of this training, one of which is individual differences in training performance. We propose that future research should not investigate whether cognitive training works, but rather should determine what training regimens and what training conditions result in the best transfer effects, investigate the underlying neural and cognitive mechanisms, and finally, investigate for whom cognitive training is most useful.

n-back training | training efficacy | long-term effects | motivation

Physical training has an effect not only on skills that are trained, but also on skills that are not explicitly trained. For example, running regularly can improve biking performance (1). More generally, running will improve performance on activities that benefit from an efficient cardiovascular system and strong leg muscles, such as climbing stairs or swimming. This transfer from a trained to an untrained physical activity is, of course, advantageous; we do not have to perform a large variety of different physical activities to improve general fitness. Although the existence of transfer in the physical domain is not surprising to anyone, demonstrating transfer from cognitive training has been difficult (2, 3), but there is accumulating evidence that certain cognitive interventions yield transfer (4–6).

Fluid intelligence (Gf), defined as the ability to reason abstractly and solve novel problems (7), is frequently the target of cognitive training because Gf is highly predictive of educational and professional success (8, 9). In contrast to crystallized intelligence (Gc) (7), it is highly controversial whether Gf can be altered by experience, and if so, to what degree (10, 11). Nevertheless, it seems that Gf is malleable to a certain extent as indicated by the fact that there are accumulating data showing an increase in Gf-related processes after cognitive training (6). The common feature of most studies showing transfer to Gf is that the training regimen targets working memory (WM). WM is the cognitive system that allows one to store and manipulate a limited amount of information over a short period, and its functioning is essential for a wide range of complex cognitive tasks, such as reading, general reasoning, and problem solving (12, 13). Referring back to the analogy in the physical domain, we can characterize WM as taking the place of the cardiovascular system; WM seems to underlie performance in a multitude of tasks, and training WM results in benefits to those tasks.

Given the importance of WM capacity for scholastic achievement (14), even beyond its relationship to Gf (12, 15, 16), improving children’s WM is of particular relevance. Although there is some promising recent research demonstrating that transfer of cognitive training is an obtainable goal (4–6), there is minimal evidence for training and transfer in typically developing school-aged children. Furthermore, whether there are long-term transfer effects is largely unknown. Our goal in this study was to adapt WM training interventions that have been found effective for adults (17, 18) to train children’s WM skills with the aim of also improving their general cognitive abilities. We trained 62 children over a period of 1 mo (see Table 1 and *Materials and Methods* for demographic information). Participants in the experimental group trained on an adaptive spatial n-back task in which a series of stimuli was presented at different locations on the computer screen one at a time. The task of the participants was to decide whether a stimulus appeared at the same location as the one presented *n* items back in the sequence (Fig. 1) (17, 18). Participants in the active control group trained on a task that required answering general knowledge and vocabulary questions, thereby practicing skills related to Gc (Fig. S1) (7). Both training tasks were designed to be engaging by incorporating video game-like features and artistic graphics (19–22) (Fig. 2 and *Materials and Methods*). Before and after training, as well as 3 mo after completion of training, participants’ performance was assessed on two different matrix reasoning tasks (23, 24) as a proxy for Gf. Because the research on training and transfer sometimes yields inconsistent results (2), we also investigated the extent to which individual differences in training gain moderate transfer effects. Finally, we assessed whether transfer effects are maintained for a significant period after training completion—a critical issue if training regimens are to have any practical importance.

Results

Our analysis revealed a significant improvement on the trained task in the experimental group [performance gain calculated by subtracting the mean n-back level achieved in the first two training sessions from the mean n-back level achieved in the last two training sessions; $t(31) = 6.38$; $P < 0.001$] (Table 1 and Fig. S2). In contrast, there was no significant performance improvement in the active control group [rate of correct responses; performance gain two last minus two first sessions: $t(29) < 1$] (Table 1). However, despite the experimental group’s clear training effect, we observed no significant group \times test session interaction on transfer to the measures of Gf [group \times session (post vs. pre): $F(1, 59) < 1$; $P =$ not significant (ns); (follow-up vs. pre): $F(1, 53) < 1$; $P =$ ns; with test version at pretest (A or B) as a covariate] (Table 1). To

Author contributions: S.M.J., M.B., J.J., and P.S. designed research; M.B. programmed the training tasks; S.M.J., M.B., and P.S. performed research; S.M.J. and M.B. analyzed data; and S.M.J., M.B., J.J., and P.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹S.M.J. and M.B. contributed equally to this work.

²To whom correspondence may be addressed. E-mail: sjaeggi@umich.edu or mbu@umich.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103228108/-DCSupplemental.

Table 1. Demographic and descriptive data

Intervention group	Age	Grade	No. of training sessions	Training performance first two sessions	Training performance last two sessions	SPM pre	SPM post	SPM FU	TONI pre	TONI post	TONI FU	
Active control group	Mean (SD)	8.83 (1.44)	3.63 (1.47)	19.40 (2.09)	59.10 (12.80)	58.80 (17.70)	15.33 (4.32)	16.20 (5.10)	16.59 (3.95)	20.87 (6.04)	22.50 (5.06)	24.81 (7.30)
N-back group	Mean (SD)	9.12 (1.52)	3.72 (1.55)	19.09 (2.28)	2.17 (0.85)	2.93 (0.84)	15.44 (5.15)	16.94 (4.75)	16.52 (5.27)	20.41 (6.51)	22.03 (7.16)	24.07 (7.79)
Large training gain (n-back)	Mean (SD)	9.00 (1.79)	3.75 (1.65)	19.19 (2.26)	2.04 (0.75)	3.30 (0.75)	14.06 (4.91)	17.19 (4.74)	15.67 (5.07)	18.50 (4.91)	21.56 (6.93)	24.67 (8.15)
Small training gain (n-back)	Mean (SD)	9.25 (1.24)	3.69 (1.49)	19.00 (2.37)	2.30 (0.95)	2.55 (0.78)	16.81 (5.17)	16.69 (4.91)	17.43 (5.52)	22.31 (7.45)	22.50 (7.57)	23.43 (7.64)

Note: $n = 62$ (experimental group: $n = 32$, 14 girls; large training gain: $n = 16$, 7 girls; small training gain: $n = 16$; 7 girls; active control group: $n = 30$, 15 girls). There were no significant group differences on any of the pretest measures or the demographic variables. Training performance in the experimental group is given as mean n-back level; training performance in the active control group is given as percent correct responses. FU, follow-up test.

examine whether individual differences in training gain might moderate the effects of the training on Gf, we initially split the experimental group at the median into two subgroups differing in the amount of training gain. The mean training performance plotted for each training session as a function of performance group is shown in Fig. 3. A repeated-measures ANOVA with session (mean n-back level obtained in the first two training sessions vs. the last two training session) as a within-subjects factor, and group (low training gain vs. high training gain) as a between-subjects factor, revealed a highly significant interaction [$F(1, 30) = 43.37$; $P < 0.001$]. Post hoc tests revealed that the increase in performance was highly significant for the group with the high training gain [$t(15) = 14.03$; $P < 0.001$], whereas the group with the

small training gain showed no significant improvement [$t(15) = 1.99$; $P = ns$]. Inspection of n-back training performance revealed that there were no group differences in the first 3 wk of training; thus, it seems that group differences emerge more clearly over time [first 3 wk: $t(30) < 1$; $P = ns$; last week: $t(16) = 3.00$; $P < 0.01$] (Fig. 3). It is important to note that there were no significant differences between the training performance groups in terms of sex distribution, age, grade, number of training sessions, initial WM performance (performance in the first two training sessions), or pretest performance on the two reasoning tasks (Table 1).

Next, we compared transfer to Gf between these two training subgroups and the control group. Our results indicate that only those participants above the median in WM training improve-

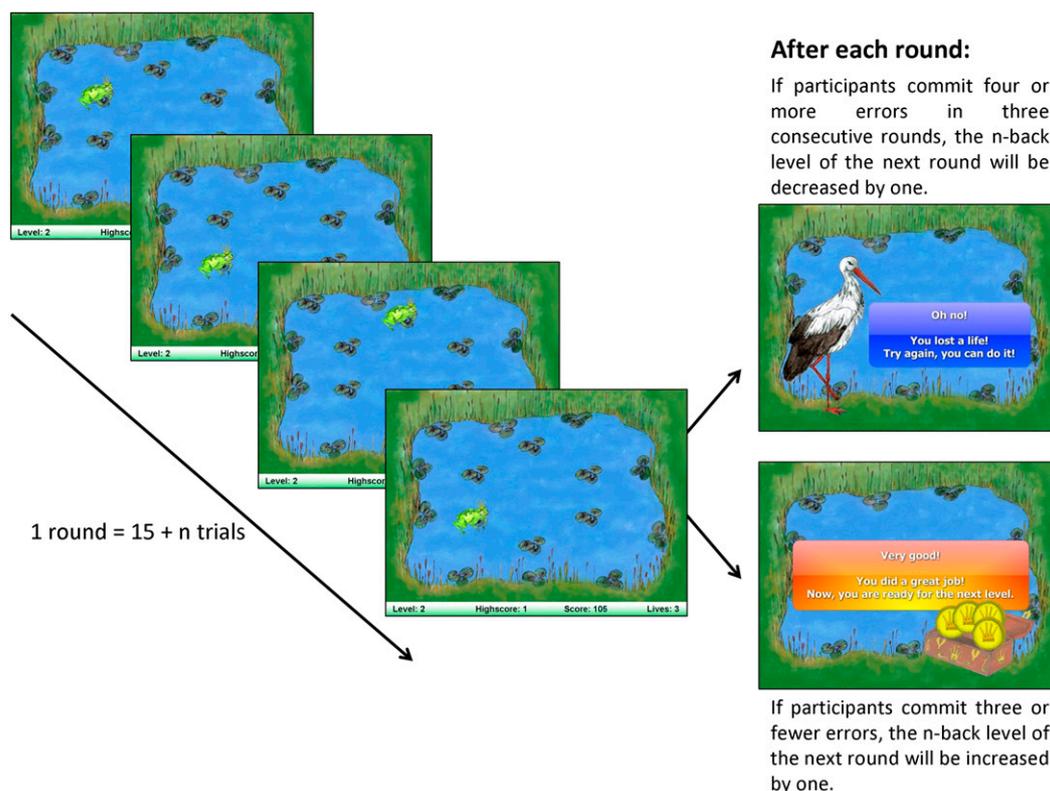


Fig. 1. N-back training task. Example of some two-back trials (i.e., level 2), along with feedback screens shown at the end of each round (*Materials and Methods*).



Fig. 2. Task themes. Outline of the four different training game themes. The treasure chest shown in the center is presented after each round and at the end of the training session when participants could trade in coins earned for token prizes (*Materials and Methods*).

ment showed transfer to measures of Gf [group \times session (post vs. pre); $F(2, 58) = 3.23$; $P < 0.05$ (Fig. 4A), with test version at pretest (A or B) as a covariate]. Planned contrasts revealed significant differences between the group with the large training gain and the other groups ($P < 0.05$; see Fig. 4A for effect sizes); there were no significant differences between the active control group and the group with the small training gain. Note that the pattern of transfer was the same for each of the individual matrix reasoning tasks. Furthermore, there was a significant positive correlation between improvement on the training task and improvement on Gf ($r = 0.42$, $P < 0.05$; Fig. S3), suggesting that the greater the training gain, the greater the transfer.* Unlike the training group, the active control group did not show differential effects on transfer to Gf as a function of training gain [$t(28) < 1$; $P = \text{ns}$; $r = -0.03$; $P = \text{ns}$; Table S1].

Interestingly, the group differences in Gf gain remained substantially in place even after a 3-mo hiatus (Fig. 4B). There was still a strong trend for differences between the group with the large training gain and the two other groups ($P < 0.055$; planned contrast); furthermore, the difference between the two n-back groups remained statistically significant ($P < 0.05$; see Fig. 4A for effect sizes), whereas there was no difference between the active control group and the group with the small training gain ($P = \text{ns}$). Whereas participants who trained well on the n-back task maintained their performance gain over the 3-mo period, the group with the small training gain and the active control group improved equally well over time from posttest to follow-up, probably a result of the natural course of development.

Because there are numerical differences at pretest between the experimental group's participants with the large and small training gain (Table 1), we calculated additional univariate

ANCOVAs with the mean standardized gain for both Gf measures, using standardized pretest scores (pretest score divided by the SD) as well as the test version at pretest (A or B) as covariates. The group effects were significant at posttest [pre- to postgain; $F(1, 28) = 3.06$; $P < 0.05$; one-tailed], and also at follow-up 3 mo later [pre- to follow-up gain; $F(1, 25) = 6.13$; $P = 0.01$; one-tailed].

Discussion

Our findings show that transfer to Gf is critically dependent on the amount of the participants' improvement on the WM task. Why did some children fail to improve on the training task and subsequently fail to show transfer to untrained reasoning tasks? Two

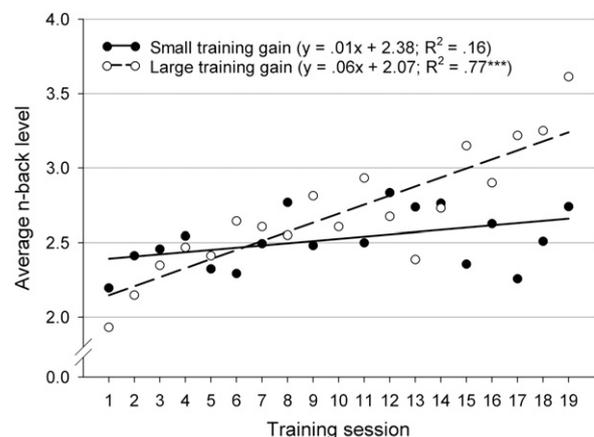


Fig. 3. Training performance. Training outcome plotted as a function of performance group. Each dot represents the average n-back level reached per training session. The lines represent the linear regression functions for each of the two groups. *** $P < 0.001$.

*The reported correlation is with one outlier participant removed; when this outlier is included, the correlation is $r = 0.25$ ($P = \text{ns}$). Note that all reported results are comparable regardless of inclusion or exclusion of this outlier.

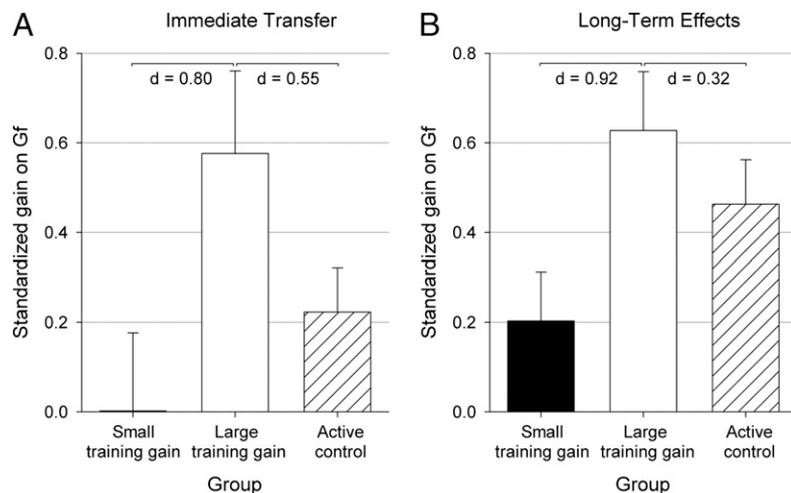


Fig. 4. Transfer effect on Gf. (A) Immediate transfer. The columns represent the standardized gain scores (posttest minus pretest, divided by the SD of the pretest) for the group with large training gain, the group with small training gain, and the active control group. (B) Long-term effects. Standardized gain scores for the three groups comparing performance at follow-up (3 mo after training completion) with the pretest. Error bars represent SEMs. Effect sizes for group differences are given as Cohen's *d*.

plausible explanations might be lack of interest during training, or difficulty coping with the frustrations of the task as it became more challenging. Data from a posttest questionnaire are in accordance with the latter: In general, children in both performance groups stated that they enjoyed training equally well [$t(28) = 1.64$; $P = 0.11$]. However, the children who improved the least during the n-back training rated the game as more difficult and effortful, whereas children who improved substantially rated the game as challenging but not overwhelming [$t(28) = 2.05$; $P = 0.05$]. This finding is consistent with the idea that to optimally engage participants, a task should be optimally challenging; that is, it should be neither too easy nor too difficult (25, 26). However, the fact that some of the participants rated the task as too difficult and effortful, and further, that this rating was inversely related to training gain, poses the question whether modifying the current training regimen might be beneficial for the group of children who did not show transfer to Gf. Although the adaptive training algorithm automatically adjusts the current difficulty level to the participants' capacities, the increments might have been too large for some of the children, and consequently, they may not have advanced as much as the other children. A more fine-grained scaffolding technique (e.g., by providing additional practice rounds with detailed instructions and feedback as new levels are introduced, or by providing more trials on a given level) might better support those students and ensure that they remain within their zone of proximal development (27).

One alternative explanation for these results could be that the children with a large training gain improved more in Gf because they started off with lower ability and had more room for improvement. A related explanation is that the children who did not show substantial improvement on the transfer tasks were already performing at their ceiling WM capacity at the beginning of training. Such factors might explain why there is more evidence for far transfer in groups with WM deficits (28–31). We note that in our sample, pretest as well as initial training performance in these two groups was not significantly different. Nevertheless, there was a small numerical difference between groups. Also, although participants in the large training gain group show greater improvement in Gf, they do not end up with significantly higher Gf scores at posttest (Table 1). Furthermore, children with an initially high level of Gf performance started with higher WM training levels than children with lower initial Gf performance (approximately one n-back level; $P < 0.01$) but showed less gain in

training [$t(30) = 3.19$; $P < 0.01$]. However, there were no significant group differences between the participants with high initial Gf performance and those with low initial Gf in terms of magnitude of transfer [$F(1,29) = 1.98$; $P = 0.17$; test version at pretest as covariate]. Finally, there was no correlation between Gf gain and initial n-back performance (Gf gain and the first two training sessions: $r = 0.00$). Thus, consistent with a prior meta-analysis (32), preexisting ability does not seem to be a primary explanation for transfer differences. Rather, our data reveal that what is critical is the degree of improvement in the trained task as well as the perceived difficulty of the trained task.

Finally, it might be that some of the transfer effects are driven by differential requirements in terms of speed between the two interventions: Whereas the active control group performed their task self-paced, the n-back task was externally paced (although speed was not explicitly emphasized to participants). Nevertheless, given previous work that shows some transfer after speed training (33, 34), we were interested in whether speed might account for some of the variance in transfer. However, our results show that the mean reaction times (RT) for correct responses to targets (i.e., hits) as well as for false alarms did not significantly change over time; in fact, RT slightly increased over the course of the training (Fig. S2) [hits: $t(31) = 1.45$; $P = 0.16$; false alarms: $t(31) = 0.93$; $P = 0.36$, calculated as gain subtracting the mean RT of the individual last two training sessions from the mean RT in the first two training sessions]. Of course, the numerical increase in RT is most likely driven by the increasing level of *n* on which participants trained (35). Thus, to control for difficulty, we calculated regression models for each participant as a function of n-back level using RT as the dependent variable, and session as the independent variable. Because only the minority of participants consistently trained at n-back levels beyond 4, we analyzed only levels 1–4. Inspection of the average slopes of the whole experimental sample revealed positive slopes for all four levels of n-back, both for hits and false alarms. Inspection of the slopes for the two performance groups revealed the same picture. That is, the slopes of the two training-gain groups did not significantly differ from each other at any level, neither for hits nor for false alarms (all $t < 0.24$). In sum, there is no indication that the differential transfer effects were driven by improvements in processing speed.

The current study has several strengths compared with previous training research and follows the recommendations of re-

cent critiques of this body of work (6, 36, 37): Specifically, in contrast to many other studies, we used only one well-specified training task; thus, the transfer effects are clearly attributable to training on this particular task. Second, unlike many previous training studies, we used an active control intervention that was as engaging to participants as the experimental intervention and designed to be, on the surface, a plausible cognitive training task. Finally, we report long-term effects of training, something rarely included in previous work. Although not robust, there was a strong trend for long-term effects, which, considering the complete absence of any continued cognitive training between posttest and follow-up, is remarkable. However, to achieve stronger long-term effects, it might be that as in physical exercise, behavior therapy, or learning processes in general, occasional practice or booster sessions are necessary to maximize retention (33, 38–41).

One potential downside of our current median split approach is that it does not provide a predictive value in the sense of a clearly defined training criterion a participant has to reach in order to show transfer effects. Of course, the definition of such a criterion will depend on the population, but also on the training and transfer tasks used. Though the strength of our approach is that it reveals the importance of training quality, future studies have to be designed that specify the degree of training gain required to achieve reliable transfer.

To conclude, the current findings add to the literature demonstrating that brain training works, and that transfer effects may even persist over time, but that there are likely boundary conditions on transfer. Specifically, in addition to training time (17, 42), individual differences in training performance play a major role. Our findings have general implications for the study of training and transfer and may help explain why some studies fail to find transfer to Gf. Future research should not investigate *whether* brain training works (2), but rather, it should continue to determine factors that moderate transfer and investigate how these factors can be manipulated to make training most effective. More generally, prospective studies should focus on (i) what training regimens are most likely to lead to general and long-lasting cognitive improvements (5); (ii) what underlying neural and cognitive mechanisms are responsible for improvements when they are found (43–46); (iii) under what training conditions might cognitive training interventions be effective (47), and, finally, (iv) for whom might training interventions be most useful (48).

Materials and Methods

Participants. Seventy-six elementary and middle school children from southeastern Michigan took part in the study. Because we included children from both the Detroit and Ann Arbor metropolitan areas, we had a broad range of socioeconomic status, race, and ethnicity. All participants were typically developing; that is, we excluded children who had been clinically diagnosed with attention-deficit hyperactivity disorder or other developmental or learning difficulties. Participants were pseudorandomly assigned to the experimental or control group (i.e., continuously matched based on age, sex, and pretest performance) and were requested to train for a month, five times a week, 15 min per session. Further, they were requested to return for a follow-up test session ~3 mo after the posttest session. For data analyses, we included only participants who completed at least 15 training sessions and who trained for at least 4 wk, but not longer than 6 wk, and who had no major training or posttest scheduling irregularities. The final sample used for the analyses consisted of 62 participants (see Table 1 for demographic information). Six participants (three from each intervention group) failed to complete the follow-up session, resulting in a total of 56 participants for the analyses involving follow-up sessions.

Training Tasks. Participants trained on computerized video game-like tasks. The experimental group trained on an adaptive WM task variant (spatial single n-back) (18). In this task, participants were presented with a sequence of stimuli appearing at one of six spatial locations, one at a time at a rate of 3 s (stimulus length = 500 ms; interstimulus interval = 2,500 ms). Participants were required to press a key whenever the currently presented stimulus was

at the same location as the one n items back in the series (targets), and another key if that was not the case (nontargets; Fig. 1). There were five targets per block of trials (which included 15 + n trials), and their positions were determined randomly.

The active control group trained on a knowledge- and vocabulary-based task for the same amount of time as the experimental group. In this self-paced task, questions were presented in the middle of the screen one at the time, and participants were required to select the appropriate answers of four alternatives presented below the question (Fig. S1). After each question, a feedback screen appeared informing the participant whether the answer was correct, along with additional factual information in some cases. There were six questions per round, and the questions that were answered incorrectly were presented again at the beginning of the next training session.

To maximize motivation and compliance with the training, we designed both tasks based on a body of research that identifies features of video games that make them engaging (20–22, 19). For example, the tasks were presented with appealing artistic graphics that incorporated four different themes (lily pond, outer space, haunted castle, and pirate ship; Fig. 2). The theme changed every five training sessions. There were background stories linked to the themes, providing context for the task (e.g., “You have to crack the secret code in order to get to the treasure before the pirate does”); in the case of the control training, vocabulary and knowledge questions were loosely related to each theme. Posttest questionnaires indicated that participants found the training tasks and active control tasks to be equally motivating [$t(56) = 1.22$; $P = ns$].

In each training session, participants in both intervention groups were required to complete 10 rounds. Each round lasted approximately 1 min, after which performance feedback was provided; thus, one training session lasted ~15 min. The performance feedback consisted of points earned during each round: For each correct response, participants earned points that they could cash in for token prizes such as pencils or stickers. In the experimental task, the difficulty level (level of n) (17, 18) was adjusted according to the participants' performance after each round (it increased if three or fewer errors were made, and decreased if four or more errors were made per round in three consecutive rounds). In the control task, the levels increased accordingly; that is, a new level was introduced if maximally one error was made, and decreased if three or more errors were made per round in three consecutive rounds. Further, once a question was answered correctly, it never reappeared. Thus, the control task adapted to participants' performance in that only new or previously incorrect questions were used. In both tasks, correct answers were rewarded with more points as the levels increased.

In addition to the points and levels, there were also bonus rounds (two per session, randomly occurring in rounds 4–10) in which the points that could be earned for correct answers were doubled. Finally, if participants reached a new level, they were awarded a high score bonus. The current high score (i.e., the highest level achieved) was visibly displayed on the screen.

Transfer Tasks. We assessed matrix reasoning with two different tasks, the Test of Nonverbal Intelligence (TONI) (23) and Raven's Standard Progressive Matrices (SPM) (24). Parallel versions were used for the pre-, post-, and follow-up test sessions in counterbalanced order. For the TONI, we used the standard procedure (45 items, five practice items; untimed), whereas for the SPM, we used a shortened version (split into odd and even items; 29 items per version; two practice items; timed to 10 min after completion of the practice items. Note that virtually all of the children completed this task within the given timeframe). The dependent variable was the number of correctly solved problems in each task. These scores were combined into a composite measure of matrix reasoning represented by the standardized gain scores for both tasks [i.e., the gain (posttest minus pretest, or follow-up test minus pretest, respectively) divided by the whole population's SD of the pretest].

Posttest Questionnaire. After the posttest, we assessed the children's engagement and motivation for training with a self-report questionnaire consisting of 10 questions in which they rated the training or control task on dimensions such as how much they liked it, how difficult it was, and whether they felt that they became better at it. Participants responded on five-point Likert scales that were represented with “smiley” faces ranging from “very positive” to “very negative.” A factor analysis (varimax rotation with Kaiser normalization) resulted in three factors explaining 67% of the total variance. These factors include interest/enjoyment, difficulty/effort, and perceived competence—dimensions that can be described as aspects of intrinsic motivation (49). Based on these factors, we created three variables, one for each factor, by averaging the products of a given answer (i.e., the individual score) and its factor loading.

ACKNOWLEDGMENTS. We thank the research assistants from the Shah laboratory for their help with collecting the data. This work was supported

by Institute of Education Sciences Grant R324A090164 (to P.S.) and by grants from the Office of Naval Research and the National Science Foundation (J.J.).

1. Suter E, Marti B, Gutzwiller F (1994) Jogging or walking—comparison of health effects. *Ann Epidemiol* 4:375–381.
2. Owen AM, et al. (2010) Putting brain training to the test. *Nature* 465:775–778.
3. Salomon G, Perkins DN (1989) Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educ Psychol* 24:113–142.
4. Klingberg T (2010) Training and plasticity of working memory. *Trends Cogn Sci* 14: 317–324.
5. Lustig C, Shah P, Seidler R, Reuter-Lorenz PA (2009) Aging, training, and the brain: A review and future directions. *Neuropsychol Rev* 19:504–522.
6. Buschkuhl M, Jaeggi SM (2010) Improving intelligence: A literature review. *Swiss Med Wkly* 140:266–272.
7. Cattell RB (1963) Theory of fluid and crystallized intelligence: A critical experiment. *J Educ Psychol* 54:1–22.
8. Deary IJ, Strand S, Smith P, Fernandes C (2007) Intelligence and educational achievement. *Intelligence* 35:13–21.
9. Rohde TE, Thompson LA (2007) Predicting academic achievement with cognitive ability. *Intelligence* 35:83–92.
10. Herrnstein RJ, Murray C (1996) *Bell Curve: Intelligence and Class Structure in American Life* (Free Press, New York).
11. Jensen AR (1981) Raising the IQ: The Ramey and Haskins study. *Intelligence* 5:29–40.
12. Engle RW, Tuholski SW, Laughlin JE, Conway ARA (1999) Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *J Exp Psychol Gen* 128:309–331.
13. Shah P, Miyake A (1999) Models of working memory: An introduction. *Models of Working Memory: Mechanism of Active Maintenance and Executive Control*, eds Miyake A, Shah P (Cambridge Univ Press, New York), pp 1–26.
14. Pickering S, ed (2006) *Working Memory and Education* (Elsevier, Oxford).
15. Gray JR, Chabris CF, Braver TS (2003) Neural mechanisms of general fluid intelligence. *Nat Neurosci* 6:316–322.
16. Kane MJ, et al. (2004) The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *J Exp Psychol Gen* 133:189–217.
17. Jaeggi SM, Buschkuhl M, Jonides J, Perrig WJ (2008) Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci USA* 105:6829–6833.
18. Jaeggi SM, et al. (2010) The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence* 38:625–635.
19. Gee JP (2003) *What Video Games Have to Teach Us About Learning and Literacy* (Palgrave Macmillan, New York).
20. Prensky M (2001) *Digital Game-Based Learning* (McGraw-Hill, New York).
21. Squire K (2003) Video games in education. *Int J Intell Simulations Gaming* 2:1–16.
22. Malone TW, Lepper MR (1987) Making learning fun: A taxonomy of intrinsic motivations for learning. *Aptitude, Learning and Instruction: III. Cognitive and Affective Process Analyses*, eds Snow RE, Farr MJ (Erlbaum, Hillsdale, NJ), pp 223–253.
23. Brown L, Sherbenou RJ, Johnsen SK (1997) *TONI-3: Test of Nonverbal Intelligence* (Pro-Ed, Austin, TX), 3rd Ed.
24. Raven JC, Court JH, Raven J (1998) *Raven's Progressive Matrices* (Oxford Psychologist Press, Oxford).
25. Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 55:68–78.
26. Bjork EL, Bjork RA (2011) Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, eds Gernsbacher MA, Pew RW, Hough LM, Pomerantz JR (Worth, New York), pp 56–64.
27. Vygotsky L (1978) Interaction between learning and development. *Mind in Society: The Development of Higher Psychological Processes*, eds Cole M, John-Steiner V, Scribner S, Souberman E (Harvard Univ Press, Cambridge, MA), pp 79–91.
28. Kerns AK, Eso K, Thomson J (1999) Investigation of a direct intervention for improving attention in young children with ADHD. *Dev Neuropsychol* 16:273–295.
29. Klingberg T, 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.
30. Klingberg T, Forssberg H, Westerberg H (2002) Training of working memory in children with ADHD. *J Clin Exp Neuropsychol* 24:781–791.
31. Holmes J, Gathercole SE, Dunning DL (2009) Adaptive training leads to sustained enhancement of poor working memory in children. *Dev Sci* 12:F9–F15.
32. Colquitt JA, LePine JA, Noe RA (2000) Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *J Appl Psychol* 85: 678–707.
33. Ball K, et al.; Advanced Cognitive Training for Independent and Vital Elderly Study Group (2002) Effects of cognitive training interventions with older adults: A randomized controlled trial. *JAMA* 288:2271–2281.
34. Healy AF, Wohldmann EL, Sutton EM, Bourne LE, Jr. (2006) Specificity effects in training and transfer of speeded responses. *J Exp Psychol Learn Mem Cogn* 32:534–546.
35. Jaeggi SM, Schmid C, Buschkuhl M, Perrig WJ (2009) Differential age effects in load-dependent memory processing. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* 16:80–102.
36. Shipstead Z, Redick TS, Engle RW (2010) Does working memory training generalize? *Psychol Belg* 50:245–276.
37. Chein JM, Morrison AB (2010) Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychon Bull Rev* 17:193–199.
38. Whisman MA (1990) The efficacy of booster maintenance sessions in behavior therapy: Review and methodological critique. *Clin Psychol Rev* 10:155–170.
39. Bell DS, et al. (2008) Knowledge retention after an online tutorial: A randomized educational experiment among resident physicians. *J Gen Intern Med* 23:1164–1171.
40. Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D (2006) Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol Bull* 132:354–380.
41. Haskell WL, et al.; American College of Sports Medicine; American Heart Association (2007) Physical activity and public health: Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* 116:1081–1093.
42. Basak C, Boot WR, Voss MW, Kramer AF (2008) Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychol Aging* 23:765–777.
43. Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L (2008) Transfer of learning after updating training mediated by the striatum. *Science* 320:1510–1512.
44. Jonides J (2004) How does practice makes perfect? *Nat Neurosci* 7:10–11.
45. McNab F, et al. (2009) Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* 323:800–802.
46. Kelly AM, Garavan H (2005) Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex* 15:1089–1102.
47. Schmidt RA, Bjork RA (1992) New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychol Sci* 3:207–217.
48. Minear M, Shah P (2006) Sources of working memory deficits in children and possibilities for remediation. *Working Memory and Education*, ed Pickering S (Elsevier, Oxford), pp 274–307.
49. McAuley E, Duncan T, Tammen VV (1989) Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Res Q Exerc Sport* 60:48–58.