ARTICLE IN PRESS

Intelligence xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

Multi-modal fitness and cognitive training to enhance fluid intelligence

Ana M. Daugherty[a,g,*], Christopher Zwilling[a,g], Erick J. Paul[a,g], Nikolai Sherepa[a,g], Courtney Allen[a,g], Arthur F. Kramer[b,c,d], Charles H. Hillman[c], Neal J. Cohen[a,e], Aron K. Barbey[a,f,g,h,i,*]

[a] *Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 N Matthews Ave, Urbana, IL 61801, United States*
[b] *Office of the Provost, Northeastern University; 100 CH, 360 Huntington Ave, Boston, MA 02115, United States*
[c] *Department of Psychology, Northeastern University; 125 Nightingale Hall, 360 Huntington Ave, Boston, MA 02115, United States*
[d] *Department of Mechanical & Industrial Engineering, Northeastern University, 360 Huntington Ave, Boston, MA 02115, United States*
[e] *Center for Nutrition, Learning, and Memory, University of Illinois at Urbana-Champaign, IL 61801, United States*
[f] *Department of Psychology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States*
[g] *Decision Neuroscience Laboratory, University of Illinois at Urbana-Champaign, IL, United States*
[h] *Department of Bioengineering, University of Illinois at Urbana-Champaign, IL, United States*
[i] *Neuroscience Program, University of Illinois at Urbana-Champaign, IL, United States*

ARTICLE INFO

Keywords:
Reasoning
Mindfulness
Fitness
Latent modeling

ABSTRACT

Improving fluid intelligence is an enduring research aim in the psychological and brain sciences that has motivated public interest and scientific scrutiny. At issue is the efficacy of prominent interventions—including fitness training, computer-based cognitive training, and mindfulness meditation—to improve performance on untrained tests of intellectual ability. To investigate this issue, we conducted a comprehensive 4-month randomized controlled trial in which 424 healthy adults (age 18–43 years) were enrolled in one of four conditions: (1) Fitness training; (2) Fitness training and computer-based cognitive training; (3) Fitness, cognitive training, and mindfulness meditation; or (4) Active control. Intervention effects were evaluated within a structural equation modeling framework that included repeated-testing gains, as well as novel tests of fluid intelligence that were administered only at post-intervention. The combination of fitness and cognitive training produced gains in visuospatial reasoning that were greater than in the Active Control, but not in performance on novel tests administered only at post-intervention. Individuals more variably responded to multi-modal training that additionally incorporated mindfulness meditation (and less time spent on cognitive training), and those who demonstrated repeated-testing gains in visuospatial reasoning also performed better on novel tests of fluid intelligence at post-intervention. In contrast to the multi-modal interventions, fitness only training did not produce Active Control-adjusted gains in task performance. Because fluid intelligence test scores predict real-world outcomes across the lifespan, boosting intelligence ability via multi-modal intervention that is effective even in young, healthy adults is a promising avenue to improve reasoning and decision making in daily life.

1. Introduction

An enduring research aim in the psychological and brain sciences is to enhance brain health and to deliver sustainable cognitive gains that benefit daily living. A central question in this effort is whether experimental interventions can enhance general intelligence. General intelligence captures the statistical regularities in performance across a wide range of cognitive domains, including reasoning, problem solving, and decision making (Barbey, 2017; Spearman, 1927). Within this framework, fluid intelligence (G_f) that encompasses pattern detection and problem solving is distinguishable from static knowledge and skills in crystallized intelligence (Carroll, 1993; Cattell, 1963). Higher intelligence scores predict real-world outcomes across the lifespan: better scholastic achievement (Gottfredson, 1997; Kuncel & Hezlett, 2007), job performance (Hunter, 1986; Salgado et al., 2003), and career success (Hagmann-von Arx, Gygi, Weidmann, & Grob, 2016). Although it can be conceived as a stable trait (Carroll, 1993; Jensen, 1998), the prospect of enhancing intelligence—thereby improving reasoning and decision making in daily life—remains intriguing. This pursuit has renewed vigor following recent reports of training gains in G_f (e.g.,

---

* Corresponding authors at: Decision Neuroscience Laboratory, Beckman Institute for Advanced Science and Technology, 405 N Matthews Ave, Urbana, IL 61801, United States. URL: https://www.DecisionNeuroscienceLab.org

E-mail addresses: adaugher@illinois.edu (A.M. Daugherty), barbey@illinois.edu (A.K. Barbey).

https://doi.org/10.1016/j.intell.2017.11.001
Received 11 April 2017; Received in revised form 22 October 2017; Accepted 6 November 2017

Please cite this article as: Daugherty, A.M., Intelligence (2017), https://doi.org/10.1016/j.intell.2017.11.001

Jaeggi, Buschkuehl, Jonides, & Perrig, 2008), yet it has been met with mixed results and inconsistent replication (e.g., Chooi & Thompson, 2012; Harrison et al., 2013; Jaeggi et al., 2010; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Redick et al., 2013; Thompson et al., 2013). Discrepant evidence may in part be due to differences in intervention methods and an incomplete theoretical model of the relevant mechanisms (see Greenwood & Parasuraman, 2015 for a review).

Aerobic exercise that delivers global effects to brain health has long been a focus for interventions aimed to promote cognitive function. The brain carries hefty metabolic demands that are serviced by its large vascular endothelial network (Attwell & Laughlin, 2001). The link between vascular health, brain integrity and cognitive function is well documented (Hillman, Erickson, & Kramer, 2008; Raz & Rodrigue, 2006; Warsch & Wright, 2010). Physical activity that promotes endothelial function is associated with better cognitive outcomes (Colcombe & Kramer, 2003; Smith et al., 2010), including G_f (Talukdar et al., 2017; Elsayed, Ismail, & Young, 1980; Reed, Einstein, Hahn, Gross, & Kravitz, 2010), working memory (Pontifex et al., 2014; Pontifex, Hillman, Fernhall, Thompson, & Valentini, 2009), and executive functions (Scott, Souza, Koehler, Petkus, & Murray-Kolb, 2016). Poor cardiovascular health and chronic neuroinflammation are associated with worse G_f (Spryidaki et al., 2014), and frequent exercise reduces these risk factors and promotes G_f ability across the lifespan (Karr, Areshenkoff, Rast, & Garcia-Barrera, 2014; Reed et al., 2010; Singh-Manoux, Hillsdon, Brunner, & Marmot, 2005). The magnitude of gains is dependent upon the level of activity and duration of intervention (Colcombe & Kramer, 2003; Karr et al., 2014; Smith et al., 2010), but even moderate-level aerobic activity over several weeks has demonstrated benefits. These cognitive effects are plausibly conferred by microstructural changes throughout the brain, including synaptogenesis, neurogenesis, increased production of nerve growth factors and other important cellular and molecular changes (Schwarb et al., 2017; Erickson, Hillman, & Kramer, 2015; Gomez-Pinilla & Hillman, 2013; Voss, Vivar, Kramer, & van Praag, 2013), and changes in functional activation (Kleemeyer et al., 2017). Exercise-related microstructural changes in the brain are considered to be dynamic and to persist beyond prescribed intervention duration to produce potentially long-term effects (Colcombe & Kramer, 2003; Gomez-Pinilla & Hillman, 2013); although frequent, habitual activity is expected to produce more sustainable change (Erickson et al., 2015). In this manner, global benefits of cardiorespiratory fitness to brain function, even following a relatively short intervention period, may encourage better response to other cognitive-based interventions aimed to bolster intelligence.

Cognitive interventions to promote higher G_f commonly target working memory abilities that appear closely related to performance on intelligence tests (Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004; Barbey, Colom, Paul, & Grafman, 2014; Barbey et al., 2012; also see Martinez et al., 2011). Working memory capacity is central to other cognitive abilities (Barbey, Koenigs, & Grafman, 2013; Engle et al., 1999; Oberauer, Schulze, Wilhelm, & Süss, 2005), especially when performing complex mental operations with demands on attention and inhibition (Harrison et al., 2013). Thus, interventions aimed to improve working memory capacity may buttress general cognitive ability. Several studies have reported the transfer of working memory task gains to general cognition (e.g., Jaeggi et al., 2008; Jaeggi et al., 2010; Klingberg et al., 2005; Klingberg, Fonsberg, & Westerberg, 2002; Dahlin, Nyberg, Bäckman, & Neely, 2008; Baniqued et al., 2013; but also see Harrison et al., 2013). However, effect sizes are highly variable (see Melby-Lervåg & Hulme, 2013; Danielsson, Zottarel, Palmqvist, & Lanfranchi, 2015; Melby-Lervåg, Redick, & Hulme, 2016 for meta-analyses), producing contradictory views of cognitive training regimens that motivate further research and debate (Shipstead, Redick, & Engle, 2012, Buschkuehl & Jaeggi, 2010; Morrison & Chein, 2011; Schwaighofer, Fischer, & Buhner, 2015; Dougherty, Hamovitz, & Tidwell, 2016; van Heugten, Ponds, & Kessels, 2016 for reviews). Working memory relies on several neural correlates, including the striatum and prefrontal cortex (Barbey et al., 2013), that are sensitive to changes in cardiorespiratory fitness (Diamond, 2013). Therefore, aerobic activity that bolsters function of relevant neural substrates may facilitate cognitive training and its transfer to fluid intelligence.

An alternative to directly training working memory ability is to indirectly promote it and its contribution to G_f with interventions that target other aspects of cognitive performance (Ward et al., 2017). For example, training in mindfulness—the ability to monitor one's thoughts and limit mind wandering—may minimize the impact of distraction during test taking to indirectly improve indices of cognitive ability. Mind wandering is negatively correlated with scores on tests of working memory, G_f, and scholastic aptitude (Mrazek et al., 2012), and mindfulness training appears to prevent this to improve test scores (Banks, Welhaf, & Srour, 2015; Brown et al., 2011; Mrazek, Franklin, Phillips, Baird, & Schooler, 2013; Noone, Bunting, & Hogan, 2016). Training in mindfulness technique improves self-referential thought that fosters better executive functioning, including attentional control (Tang, Holzel, & Posner, 2015), which is also a putative mechanism of cognitive training effects on G_f (Greenwood & Parasuraman, 2015). Thus, better task attention via mindfulness may improve cognitive training and boost performance on tests of G_f. Moreover, the combination of aerobic exercise, mindfulness meditation and cognitive training, that each promotes executive functions, may produce additive gains that surpass exercise alone. Each of these intervention strategies has been considered before, and here we test multi-modal interventions that may optimally engage the neural and cognitive constituents of fluid intelligence.

Fundamental to determining the relevant mechanism to promote G_f function is the assessment of intervention efficacy via testing gains. Foremost, the study of "gain" requires a longitudinal, pre-post test design and appropriate statistical tests of change (McArdle, 2009). Repeated-testing gains, or "practice effects", confound the interpretation of interventions aimed at improving cognition, and thus comparison to a randomized control group is a second consideration. However, repeated-testing gains theoretically reflect the function of intact cognitive systems for which the tests are designed to measure (Thorvaldsson, Hofer, Hassing, & Johansson, 2005) and characterizing individual differences in the magnitude (and direction) of change is a means to evaluate these functions that scaffold intelligence (e.g., Baltes, Dittmann-Kohli, & Kliegel, 1986; Hertzog & Schaie, 1986, 1988; Hertzog, von Oertzen, Ghisletta, & Lindenbrger, 2008; McArdle, 2009). Additional measures of fluid intelligence ability were administered only at post-intervention to avoid the contribution of practice effects and therefore to provide further insight into individual differences in response to interventions. When accounting for pre-intervention cognitive ability, higher scores in intervention groups as compared to control on the novel tasks of G_f taken at post-intervention may indicate a boost to intelligence. In the absence of such effects, greater repeated-testing gains that are associated with higher post-intervention G_f scores on novel tests may indicate the transfer of components relevant to task performance other than general intelligence—e.g., attention, motivation, and strategy (Hayes, Petrov, & Sederberg, 2015).

We investigate the efficacy of multi-modal interventions to enhance G_f within a four-month randomized control trial of young adults assigned to either an active control, or an experimental condition— fitness (Fit); fitness and cognitive training with a suite of adaptive computer games, Mind Frontiers (Fit-MF); fitness, cognitive training, and mindfulness training (Fit-MF-Mind). G_f was assessed in two ways: repeated-testing pre- and post-intervention with parallel forms of a canonical fluid intelligence test (Figure Series) and an achievement test of analogical reasoning (Law School Admission Test; LSAT), as well as a collection of fluid intelligence indices that were assessed only at post-intervention. Within a latent modeling framework (McArdle, 2009) we test three hypotheses (1) As compared to Active Control, interventions will account for better post-intervention G_f assessed by novel tests (defined by letter series, number series, matrix reasoning, and Shipley abstraction tests) and differentially higher scores in conditions engaged in multi-modal training. (2) Based upon individual differences, intervention groups will show greater post-intervention gains in scores on LSAT and Figure Series as compared to individuals in the control group. (3) Finally, individuals who show greater repeated-testing gains are expected to perform better on novel tests of fluid intelligence. Due to putative test-taking benefits of the training activities, we expect intervention group membership will moderate the magnitude of repeated-testing gains transferred to post-intervention G_f scores on novel tests.

2. Methods

2.1. Participants

Participants were recruited from the Champaign-Urbana, IL metro region and following pre-screening procedures, 424 adults (age $M = 23.35$, $SD = 4.84$; 46% female; 50% Caucasian) were enrolled. To be eligible for the study, participants were age 18–44 years; had at least a high school education; spoke English fluently; had normal or corrected-to-normal vision and hearing; no current or recent medications affecting the central nervous system or presenting a risk during aerobic exercise; no history of psychological, neurological, or endocrine disease, concussion within the past two years, or learning disorders; did not smoke $>$ 10 cigarettes per day; body mass index $<$ 35; and did not respond negatively to all items on the physical activity readiness questionnaire revised (Scott, Reading, & Shephard, 1992).

Participants were randomly assigned to four groups (see Fig. 1) that were similar in distributions of age ($F(3420) = 0.55$, $p = 0.65$), sex (χ^2 (3) ≥ 1.37, $p \geq 0.61$) and education level (χ^2 (3) ≥ 0.60, $p \geq 0.90$). Of the eligible sample, 234 individuals returned for testing after completing the intervention. The proportion of attrition was similar across intervention groups (χ^2 (3) $= 2.21$, $p = 0.53$) and was unrelated to demographic variables (all $p \geq 0.07$); individuals who returned performed similar to those who did not on Figure Series (t (422) $= 0.09$, $p = 0.93$) and LSAT at pre-intervention (t (422) $= -0.50$, $p = 0.62$). Intent-to-treat analyses were completed with the total sample and missing data were handled under the assumption of missing at random via full information maximum likelihood (FIML; Muthén, Kaplan, & Hollis, 1987).

2.2. Intervention protocol

Intervention regimens were 48 sessions over the course of 4 months, each 70 min long; all participants completed at least 85% of training sessions. The intervention conditions were designed to combine several activities and thus training duration of a particular activity was determined by intervention condition but all groups completed a total of 56 intervention hours. The Fit condition only participated in fitness training. The Fit-MF condition completed 28 sessions of fitness training and 20 sessions of cognitive training, begining in month 1 with 12 sessions of fitness training, in month 2 with 8 sessions of fitness training and 4 sessions of cognitive training, and in months 3 and 4 with 4 sessions of fitness training and 8 sessions of cognitive training. The Fit-MF-Mind condition completed 10 sessions of mindfulness mediation, 28 sessions of fitness training, and 10 sessions of cognitive training, begining in month 1 with 2 sessions of fitness training and 10 sessions of mindfulness meditation, in month 2 with 12 sessions of fitness training, in month 3 with 10 sessions of fitness training and 2 sessions of cognitive training, and in month 4 with 4 sessions of fitness training and 8 sessions of cognitive training.

2.2.1. Fitness training

The physical fitness intervention was conducted in a group setting of no more than five participants per trainer. Each session consisted of dynamic stretch and light activity warm-up (10 min), intervals of speed and recovery of static walking or running (10 min), high-intensity cardio-resistance training (30 min), whole-body drill and skill training (e.g., battle ropes or sand bags; 10 min), and yoga-inspired flexibility training and static stretching (10 min).

2.2.2. Cognitive training

Participants in the Fit-MF and Fit-MF-Mind groups completed training with Mind Frontiers, a suite of games designed to engage visuospatial working memory, executive functions, and analogical reasoning (adapted from Baniqued et al., 2013). Briefly, the cognitive training included seven Western-themed mini-games that each engaged different cognitive function. All mini-games were administered on tablets, each for 8.5 min per session, presented in a random order. A modification to the original version (Baniqued et al., 2013) included adaptive difficulty during game play that was defined uniquely for each game and participant (see supplementary material for details; Fig. S1). *Mini-game tasks.* In "Ante Up", a Tower of London style card game

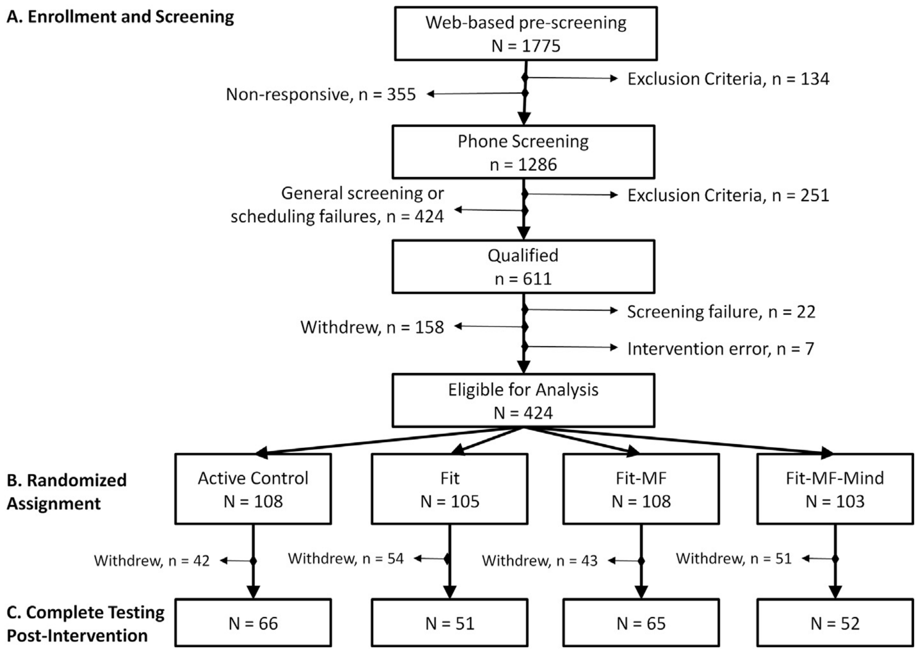

Fig. 1. Enrollment and attrition of the eligible sample in analyses. Participants were recruited and screened based on enrolment criteria. Based upon subject eligibility, responsiveness, and scheduling availability, 424 persons were enrolled and randomly assigned to one of four experimental conditions. Intent-to-treat analyses included all eligible participants.

designed to engage executive function (Shallice, 1982; Zook, Davalos, DeLosh, & Davis, 2004), participants were required to arrange cards to match an array within a minimum number of moves. The "Pen 'Em Up" mini-game was a dual task-switching executive function test (Karbach & Kray, 2009) in which participants sorted items based on one of two criteria and were required to memorize two sorting rules and the order of their application prior to starting. Another executive function task was "Supply Run" in which participants filled orders for items within themed categories and were required to remember the last item requested from each category (Dahlin et al., 2008). "The Irrigator" mini-game employed visuospatial reasoning (Mackey, Hill, Stone, & Bunge, 2011) as the participant built lines of irrigator pipes from a start to a series of targets. "Riding Shotgun" was a visuospatial working memory task (Klingberg et al., 2005; Olesen et al., 2004) that required participants to memorize target locations within a complex presentation of five 4 × 4 grids, arranged as a cross-hair. In a dual n-back working memory task (Jaeggi et al., 2008; Redick et al., 2013), "Sentry Duty", stimuli were presented in a grid and paired with a spoken word. Participants were required to compare the current combination of spoken word and location to that *n* times previously. Finally, "Trader Jack's" was an analogical reasoning task (Wechsler, 2008): participants determined the in-game value equivalency of an item compared to a target, based upon a depicted rule. As a reward in each mini-game, participants received game currency, which could be spent to build a civilization in a meta-game component intended to motivate game play.

2.2.3. Mindfulness training

In 10 sessions, participants completed questionnaires (5 min), breathing exercises (5 min), a period of guided meditation (5 min), gentle movement/stretching (15 min), two periods of silent meditation (30 min total), and a question-answer period (10 min). During the meditation period, a soft chime rang every 5 min to reorient participants to their practice.

2.2.4. Active control

The Active Control condition completed change detection and visual search tasks (Gaspar, Neider, Simons, McCarley, & Kramer, 2013; Redick et al., 2013) that were adaptive in difficulty. In the change detection task, the participant viewed two arrays of objects with an intermediate mask, and indicated which one object in the second array had changed from the first. Difficulty was manipulated by decreasing the amount of time available to view the initial array of items and by increasing the number of stimuli in the array. Stimuli variants included cars, toys, and street signs. In the visual search task, the participant searched for a target amidst distractors and tapped the screen to indicate which direction the target faced (left or right). Varying the number and heterogeneity of distractors served as difficulty manipulations and stimuli variants include F, hand, and colored P targets.

2.3. Assessment of intervention-targeted outcomes

To determine the efficacy of each training activity in the targeted function, relevant outcome measures were tested for change from pre- to post-intervention. All intervention groups participated in fitness classes, for which maximal aerobic capacity was tested for change from pre- to post-intervention. The Fit-MF and Fit-MF-Mind conditions completed the suite of adaptive cognitive training games, for which increasing adaptive difficulty level within a mini-game was used as a proxy indicator of improvement in the task. Finally, mindfulness training was evaluated at the first and last training session by the Mindful Attention Awareness Scale (MAAS)—a theory-based self-report measure that evaluates self-awareness (Carlson & Brown, 2005).

2.3.1. Aerobic capacity assessment

Maximal aerobic capacity was evaluated at pre- and post-intervention in all participants with a graded exercise test designed to measure maximal oxygen consumption (VO_2 max) during physical activity. VO_2 max scores were assessed using an indirect computerized calorimetry system (Parvo Medics TrueOne® 2400) during a modified Balke treadmill test (American College of Sports Medicine, 2014). During the test, the participant's heart rate was constantly monitored using a Polar heart rate monitor (Polar WearLink + 31, Polar Electro, Finland) and individual subjective rate of perceived exertion was assessed every 2 min using the Borg scales of perceived exertion (American College of Sports Medicine, 2014). The VO_2 max test included an initial warm-up period in which the treadmill gradually increased in speed; after which, the treadmill speed remained constant and the graded incline increased 2–3% every 2 min. Average oxygen consumption and respiratory exchange ratio were sampled every 30 s during the test using a mouthpiece. Maximum oxygen consumption (VO_2 max) is reported relative to fat-free mass (ml/kg/min) that was determined by a dual energy x-ray absorpitometry scan. Maximum effort was defined when two or more of the following criteria were met: (1) age-defined maximum heart rate norms (i.e., heart rate $>$ 85% of predicted maximum heart rate); (2) respiratory exchange ratio (CO2/O2) $>$ 1.1; (3) subjective rate of perceived exertion $>$ 17 of 20; (4) leveling of VO_2 consumption despite increasing aerobic demand.

2.4. Tests of fluid intelligence

Fluid intelligence was assessed via standardized tests that were adapted to be administered on a computer with an online interface for data recording. Based on an independent sample, all tests had high internal consistency (Chronbach's $\alpha > 0.80$) and were administered with a time limit determined from response times in the 75th percentile, thus response speed was not considered as a confound to the assessment of G_f.

2.4.1. Repeated measures: LSAT and figure series

Two tests of G_f were administered at pre- and post-intervention. The LSAT is a standardized achievement test used to determine admittance into law school and by design evaluates logical reasoning (Roussos & Norton, 1998). Here, two parallel forms of the analogical reasoning section, each composed of the 25 items from the standardized test. Total number of correct responses within 35 min was used as an indicator of the G_f latent variable. Figure Series is a canonical measure of G_f in which participants must choose the correct item missing in a series, presumably by deducing the rule governing the series (Hogrefe, 2008). Two parallel forms were created each with 30 items (response time limit of 60 s per item) and the total number correct was the outcome measure in the analysis reported here. For both tests, items not completed within the time limits were considered incorrect and administration of the parallel forms was counterbalanced at pre- and post-intervention.

2.4.2. Post-intervention G_f

Four additional tests of G_f were administered only at post-intervention. Letter Series presents a string of letters and participants were required to choose the missing letter in the sequence (Schaie, 1985). This task included 30 items (60 s response time limit) and total number correct responses within 30 min were recorded. Number Series was in a similar format but presented digits (McGrew, Schrank, & Woodcock, 2007), which included 35 test item (90 s response time limit) and total number correct within 30 min was recorded. Matrix reasoning also tested deductive reasoning of a pattern but within a matrix (Hogrefe, 2008) and included 30 items (120 s response time limit), total number correct within 32 min was recorded. Shipley abstraction required sequence completion (Shipley, Gruber, Martin, & Klein, 2009), included 25 items and the total number correct within 12 min was recorded.

2.5. Statistical analysis

We test three complementary hypotheses (1) Group differences in post-intervention tests and differentially greater gains in conditions engaged in multi-modal training. (2) Group differences in the magnitude of post-intervention gains in scores on Figure Series and LSAT. (3) Group differences in the degree to which repeated-testing gains predict performance on novel tests. Due to putative test-taking benefits of the training activities, we expect intervention group membership will moderate the magnitude of repeated-testing gains transferred to post-intervention G_f scores on novel tests. These hypotheses were assessed within an SEM framework with a grouped modeling procedure in MPlus software (Muthén and Muthén).

Repeated measures were modeled as latent change scores that provide tests of change and individual differences therein that are free of measurement error (McArdle, 2009; McArdle & Nesselroade, 1994). Post-intervention G_f was represented as a latent variable of the four tests, in line with psychometric definitions. Latent models were assessed in the entire sample ($N = 424$), handling missing data via full information maximum likelihood (FIML)—a non-imputation method that leverages all available data during latent model estimation (Larsen, 2011; Muthén et al., 1987) and is recommended for handling attrition in longitudinal analyses (Little & Card, 2013). Finally, all three hypotheses were assessed within a single model and group comparisons were made simultaneous to hypothesis testing, eliminating the need to correct for multiple comparisons; thus significance testing was set at $p < 0.05$.

Prior to hypothesis testing, measurement invariance was evaluated longitudinally in repeated measures of LSAT and Figure Series in the total sample, as well as between-group invariance in all indicators and latent variable specification. The assumption of measurement invariance that is necessary for the interpretation of longitudinal change and group differences in latent scores requires some constraints to the model parameters (Wicherts, 2016). The degree of measurement invariance (i.e., configural to strict factorial invariance) was evaluated by successively imposing the respective model constraints and evaluating chi-square fit index, change in the comparative fit index (Δ CFI < 0.01) and change in root mean square error of approximation and comparison of its 90% confidence intervals (Cheun & Rensvold, 2002). To facilitate model specification, all data were normed to total sample means and standard deviations, and post-intervention LSAT and Figure Series were normed to pre-intervention scores; thus group-specific effects can be interpreted relative to the sample average. The latent test scores of Figure Series and LSAT at each time point were identified each by two indicators that were the sum correct responses on even and odd items of the tests. The loadings of the even and odd item indicators were each fixed at 1 with estimated measurement residuals to create latent scores, thus the latent estimates are conceptually similar to a total sum score, free of measurement error. The indicators letter series, number series, matrix reasoning and Shipley abstraction tests identified a single latent variable, each with fixed loadings at 1 and freely estimated measurement residuals. Hypothesized group differences in performance were evaluated with a grouped modeling procedure that constrained all aspects of the model to be equal between groups, except for the means, variances and paths being tested.

The model construction evaluated mean change in LSAT and Figure Series and individual variability in change as an estimate of post-intervention gains in scores on these tests (e.g., Thorvaldsson et al., 2005). In addition, we tested group differences in post-intervention G_f, in the magnitude of post-intervention gains, and possibly in the degree to which repeated-testing gains predict performance on novel measures of G_f at post-intervention. Model fit was assessed by several accepted indices (Hu & Bentler, 1999): non-significant normal theory weighted chi-square (χ^2), comparative fit index (CFI > 0.90), root mean square error of approximation (RMSEA < 0.05), and standardized root mean residual (SRMR < 0.08). Model fit was determined for the total sample and with grouped modeling procedures. Unstandardized effects of latent means and group differences are reported, unless otherwise stated. To avoid spurious effects due to smaller group sizes, models were bootstrapped (5000 draws) with bias-correction to produce 95% confidence intervals of unstandardized coefficients (BS 95% CI; Hayes & Scharkow, 2013), which if not overlapping zero is evidence supporting an effect at $p < 0.05$. Group moderation of path effects (Hypothesis 3) was conservatively tested in a random effects framework as an omnibus effect (similar to an interaction term) and only significant effects were decomposed for specific comparisons to the Active Control and Fit groups.

3. Results

3.1. Model specification

Before testing hypotheses in a latent modeling framework, latent model construction and critical assumptions of the analyses were assessed. First, a G_f latent variable at pre- and post-intervention identified by LSAT and Figure Series was tested and fit poorly: $\chi^2 = 224.36$ (19), $p < 0.001$; CFI = 0.79; RMSEA = 0.16; SRMR = 0.12. Inspecting the factor loadings, performance on Figure Series even and odd items were similar (standardized loadings at pre: 0.80 and 0.76, at post: 0.90 and 0.88, all $p < 0.001$) and LSAT scores weakly identified a common latent variable with Figure Series (pre: 0.38 and 0.42, post: 0.41 and 0.59, all p's < 0.001). Thus, all further model construction treated the LSAT and Figure Series tests separately.

Additional tests of measurement invariance informed model specification. Raw scores on Figure Series were moderately correlated from pre- to post-intervention ($r = 0.57$) and groups were confirmed to have similar pre-intervention raw scores (F(3420) = 0.62, $p = 0.61$; see Table 1). The Figure Series latent variable demonstrated strong factorial invariance between groups (i.e., equal factor loadings and indicator means between groups at each assessment) but indicator means and

Table 1
Performance on tests of fluid intelligence across intervention conditions.

Time	Test	Condition			
		Fit	Fit-MF	Fit-MF-Mind	Active control
Pre-intervention	LSAT	12.43 ± 4.45	12.65 ± 4.38	12.02 ± 3.74	12.51 ± 3.70
	Figure series	18.50 ± 5.17	19.35 ± 4.88	18.98 ± 4.22	18.91 ± 5.17
Post-intervention	LSAT	12.96 ± 4.13	12.08 ± 4.15	12.06 ± 4.29	12.65 ± 4.81
	Figure series	21.28 ± 5.68	21.94 ± 5.34	19.54 ± 6.75	20.11 ± 6.41
	Letter series	23.47 ± 6.81	24.55 ± 5.98	22.44 ± 7.35	22.88 ± 7.22
	Number series	9.18 ± 6.33	9.69 ± 7.48	7.67 ± 6.88	10.22 ± 14.02
	Matrix reasoning	15.84 ± 6.95	17.00 ± 6.42	16.19 ± 7.62	16.25 ± 7.49
	Shipley abstraction	110.80 ± 11.94	112.05 ± 11.60	106.85 ± 11.68	108.12 ± 18.21

Note: Average tests scores and standard deviations (total number correct responses) are reported.

Table 2
Summary of measurement invariance testing between groups and between repeated measures at pre- and post-intervention.

	Invariance class	Figure series change score			LSAT change score			Post-intervention Gf		
		CFI	Δ CFI	RMSEA (90% CI)	CFI	Δ CFI	RMSEA (90% CI)	CFI	Δ CFI	RMSEA (90% CI)
Between-Group	Configural	0.998		0.02 (0.00/0.09)	1.000		0.00 (0.00/0.08)	1.000		0.00 (0.00/0.06)
	Weak factorial	0.998	0.000	0.01 (0.00/0.08)	1.000	0.000	0.00 (0.00/0.07)	1.000	0.000	0.00 (0.00/0.05)
	Strong factorial	1.000	− 0.002	0.00 (0.00/0.07)	1.000	0.000	0.00 (0.00/0.07)	1.000	0.000	0.00 (0.00/0.02)
	Strict factorial	0.990	0.010	0.04 (0.00/0.08)	1.000	0.000	0.00 (0.00/0.07)	1.000	0.000	0.00 (0.00/0.00)
Repeated Measure	Configural	1.000		0.00 (0.00/0.00)	0.995		0.04 (0.00/0.12)			
	Weak factorial	1.000	0.000	0.00 (0.00/0.05)	0.997	0.000	0.03 (0.00/0.09)			
	Strong factorial	0.978	0.022	0.09 (0.04/0.13)	0.999	− 0.002	0.01 (0.00/0.08)			
	Strict factorial	0.973	0.005	0.08 (0.04/0.11)	0.999	0.000	0.01 (0.00/0.07)			

Note: Constraints for invariance testing were added step-wise and change in CFI < 0.01 and similar RMSEA with overlapping 90% confidence intervals support progressively more stringent measurement invariance. CFI—comparative fit index; Δ CFI—change in CFI; RMSEA—root mean square error of approximation; CI—confidence intervals.

variances were not equal between pre- and post-intervention in the total sample (i.e., weak factorial invariance; see Table 2). Therefore, constraints on indicator means were specified between groups but not longitudinally. Raw scores on LSAT were moderately correlated from pre- to post-intervention ($r = 0.52$) and groups were similar in performance at pre-intervention (F(3420) = 0.47, $p = 0.71$; Table 1), and the latent variable demonstrated strict factorial invariance longitudinally and between groups (i.e., equal factor loadings, indicator means and variances; Table 2). Finally, the four novel tests of Gf that were administered only at post-intervention were moderately correlated ($r = 0.51$–0.60). In contrast to the repeated measures, a single latent variable identified performance on these tests well: $\chi^2 = 3.83$ (2), $p = 0.15$; CFI = 0.99; RMSEA = 0.06; SRMR = 0.02; and the tests had comparable standardized loadings, the lowest loading by matrix reasoning (0.70) and the highest, Shipley abstraction (0.76). Fixing all factor loadings to 1 and freely estimating all measurement residuals produced similar model fit ($\chi^2 = 4.68$ (9), $p = 0.87$; CFI = 1.00; RMSEA = 0.00; SRMR = 0.02) and strict factorial invariance was supported (Table 2), therefore all hypotheses were tested with this model specification.

3.2. Intervention manipulation evaluation

To confirm the various training activities affected the targeted functions, we tested pre- to post-intervention differences in relevant indicators within each group using uncorrected paired *t*-tests in the sample of individuals who returned for assessment at post-intervention. All groups participating in fitness training showed improvement in cardiorespiratory fitness (maximum volume of oxygen consumption, ml/kg/min): Fit (t (51) = 5.10, $p < 0.001$), Fit-MF (t (65) = 4.22, $p < 0.001$), Fit-MF-Mind (t (55) = 3.40, $p = 0.001$), and Active Control (t (65) = − 1.99, $p = 0.05$). The two intervention groups that completed Mind Frontiers cognitive training both showed increases in mini-game difficulty from pre- to post-intervention (all games $t \geq 9.21$, $p \leq 0.01$) in line with improvements within the tasks. Finally, MAAS questionnaire scores were higher at the last intervention session than at the first in those participating in the mindfulness training (t (55) = 4.45, $p < 0.001$). Therefore, the training activities appear to have improved the directly targeted functions and we went on to evaluate possible intervention group differences in G_f test scores.

3.3. Latent modeling

In the total sample (not accounting for group membership), Figure Series demonstrated significant improvement over time (mean change = 0.42, $p < 0.001$; BS 95% CI: 0.32/0.52; d = 0.51) whereas LSAT was stable (mean change = − 0.06, $p = 0.15$; BS 95% CI: -0.13/0.01; d = − 0.09), and individuals significantly varied in the magnitude of change in both tests (0.66, $p < 0.001$ and 0.14, $p = 0.01$, respectively). The model of correlated change in Figure Series and LSAT had excellent fit: χ^2 (24) = 24.80, $p = 0.42$; CFI = 1.00; RMSEA = 0.01; SRMR = 0.03.

Individual differences in post-intervention gains is in line with the notion that other influences besides repeated test exposure may be relevant to the assessment of change in G_f scores. Higher scores at pre-intervention were associated with lesser gains in LSAT (− 0.06, $p = 0.02$; BS 95% CI: − 0.12/− 0.01) and Figure Series (− 0.13, $p = 0.01$; BS 95% CI: 0.08/0.21), likely due to a ceiling effect in the magnitude of change within this sample. Greater repeated-testing gains in Figure Series (0.28, $p = 0.001$; BS 95% CI: 0.13/0.53), but not LSAT (− 0.07, $p = 0.51$; BS 95% CI: -0.67/0.26), explained better post-intervention G_f scores, even when accounting for its relationship with pre-intervention measures (Fig. 2). Thus, repeated-testing gains in visuospatial reasoning (Figure Series) over the course of the intervention were associated with individual differences in performance on novel tests of G_f. Yet, pre-intervention scores and longitudinal change in these tests only accounted for a small portion of the individual differences in post-intervention G_f ($R^2 = 0.09$) and intervention condition effects may explain additional variability. The hypothesis model had excellent fit in the total sample: $\chi^2 = 69.18$ (60), $p = 0.20$; CFI = 0.99; RMSEA = 0.02; SRMR = 0.04; as did the grouped model: $\chi^2 = 297.03$ (269), $p = 0.12$ (group χ^2: Fit = 84.56, Fit-MF = 66.49, Fit-MF-Mind = 68.92, Active Control = 77.07); CFI = 0.98; RMSEA = 0.03; SRMR = 0.09.

Hypothesis 1. *Group differences in post-intervention G_f.* Intervention groups showed no significant differences in post-intervention G_f latent scores as compared to the Active Control group (all $p > 0.23$; see Fig. 3 and Table 3). Scores in the Fit-MF condition were significantly better than that in Fit-MF-Mind (difference = 0.30, $p = 0.04$; BS 95% CI: 0.05/0.54) and the Fit group was similar to both (vs. Fit-MF, difference = − 0.11, $p = 0.44$; BS 95% CI: -0.35/0.13; vs. Fit-MF-Mind, difference = 0.19, $p = 0.24$; BS 95% CI: -0.08/0.45). Thus, the intervention conditions did not produce control-adjusted gains in G_f at post-intervention, nor any evidence of additive gains from the multimodal interventions. To confirm that the treatment of missing data via FIML estimation did not introduce bias to the intent-to-treat analyses, the identical model was evaluated again in the sample with complete longitudinal data. In this restricted sample, intervention groups were similar to Active Control (vs. Fit = 0.06, $p = 0.68$; vs. Fit-MF = 0.17, $p = 0.22$; vs. Fit-MF-Mind = − 0.13, $p = 0.41$) and the hypothesis model fit comparably well: $\chi^2 = 298.36$ (269), $p = 0.11$ (group χ^2: Fit = 86.00, Fit-MF = 62.87, Fit-MF-Mind = 68.14, Active Control = 81.36); CFI = 0.98; RMSEA = 0.04; SRMR = 0.09.

Hypothesis 2. *Group differences in post-intervention gains.* The mean improvement in Figure Series aligns with expected practice effects and we further evaluated possible benefits of intervention based on the variability in post-intervention gains. As illustrated by the moderate-large effect sizes displayed in Table 4, intervention groups experienced greater gains in Figure Series, but only group Fit-MF (difference vs.

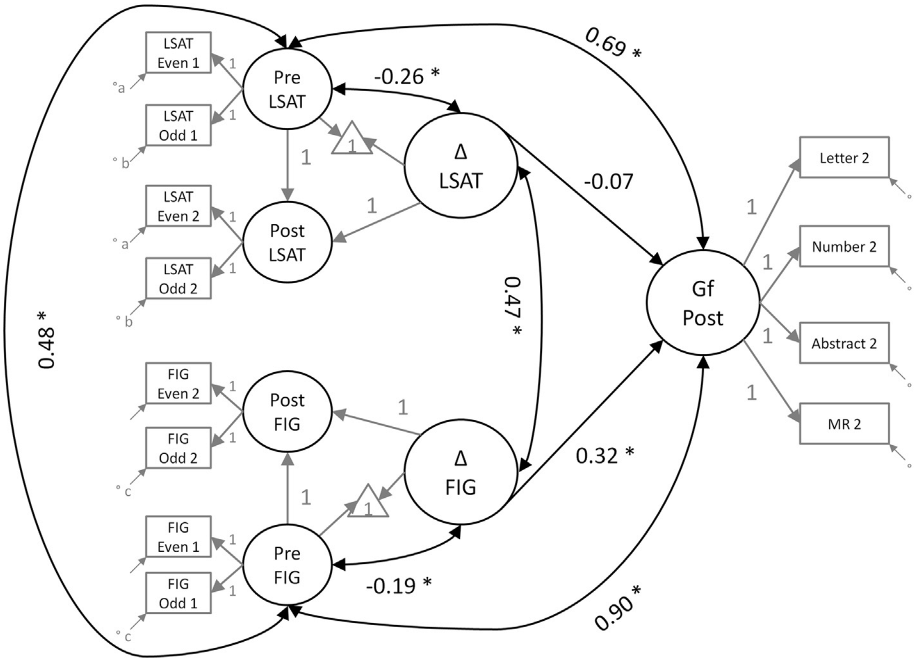

Fig. 2. Hypothesis 3 model evaluating repeated-testing gains in relation to higher test scores on novel tests of G_f that were administered only at post-intervention. Results for the total sample are shown, unstandardized coefficients are reported with standardized coefficients in parentheses, *$p < 0.05$. The measurement model is illustrated in gray; °a-b indicate estimated residual terms that were constrained to be equal across time; paths and factor loadings constrained to 1 were used to identify the model. Straight arrows indicate regression paths whereas curved, double-headed arrows are correlations. When accounting for the relationship of pre-intervention test scores with post-intervention G_f, greater repeated-testing gains in Figure Series (but not LSAT) explained better performance on novel tests of G_f.

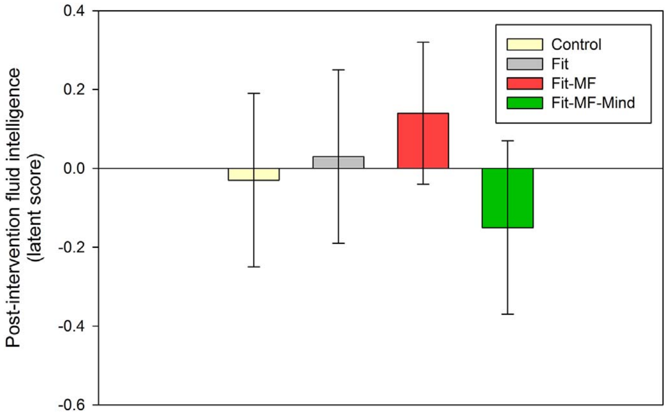

Fig. 3. Group differences in mean post-intervention G_f latent factor scores. Error bars represent two standard errors of the mean. Values are unstandardized, normed to the total sample average. Post-intervention G_f means can be interpreted as deviation from the total sample average. Control—active control ($n = 108$); Fit—fitness only intervention ($n = 105$); Fit-MF—fitness and cognitive training intervention ($n = 108$); Fit-MF-Mind—fitness, cognitive training and mindfulness training ($n = 103$).

Table 3
Group differences in post-intervention G_f as estimated in the grouped model.

Group	Mean post-intervention G_f [BS 95% CI]	Variance in post-intervention G_f	Difference from control	
			Mean [BS 95% CI]	d
Control (n = 108)	− 0.03 [− 0.21/0.16]	0.61*		
Fit (n = 105)	0.03 [− 0.17/0.22]	0.48*	0.06 [− 0.20/0.31]	0.08
Fit MF (n = 108)	0.14 [− 0.04/0.14]	0.47*	0.17 [− 0.06/0.39]	0.22
Fit MF Mind (n = 103)	− 0.16 [− 0.35/0.05]	0.55*	− 0.13 [− 0.40/0.13]	− 0.17

Note: BS 95% CI—bias-corrected bootstrapped 95% confidence intervals. d— standardized effect size. Reported coefficients are unstandardized for data normed to sample average. Control—active control; Fit—aerobic fitness only intervention; Fit-MF—fitness and cognitive training intervention; Fit-MF-Mind—fitness, cognitive training and mindfulness training intervention.

* $p < 0.05$.

Active Control = 0.34, $p = 0.02$; Fig. 4) showed significantly greater gains than Active Control, suggesting benefits of that intervention to repeated-testing gains on visuospatial tests of G_f. However, performance in the Fit-MF (difference vs. Fit = 0.18, $p = 0.30$) and Fit-MF-Mind (difference vs. Fit = − 0.10, $p = 0.60$) conditions were similar to those with exercise alone. Comparable results were found in analysis restricted to the sample with complete data: all groups demonstrated gains in Figure Series (mean change = 0.28–0.62, all $p < 0.02$), and group Fit-MF showed greater gains relative to Active Control (difference = 0.34, $p = 0.02$) whereas groups Fit (difference = 0.15, $p = 0.39$) and Fit-MF-Mind (difference = 0.07, $p = 0.69$) did not. Yet, repeated-testing gains in Fit-MF (difference = 0.20, $p = 0.26$) and Fit-MF-Mind (difference = − 0.08, $p = 0.70$) were similar to that in Fit.

Hypothesis 3. *Intervention group membership moderates the relationship between repeated-testing gains and post-intervention G_f scores.* Finally, we assessed possible group differences in repeated-testing gains predicting performance on novel tests of G_f taken at post-intervention. The random effects model supported significant group moderation of change in Figure Series predicting post-intervention G_f (0.05, p = 0.02) but not of change in LSAT (− 0.13, p = 0.08). The group differences in this effect were driven by an association between repeated-testing gains and post-intervention G_f in the Fit-MF-Mind group (0.36, p = 0.003; BS 95% CI: 0.18/0.73;) and there was no evidence of this effect in the other groups. The Fit group did not demonstrate the same association (0.06, p = 0.57; BS 95% CI: -0.09/0.22) and this relationship was significantly less than in the Fit-MF-Mind condition (difference = 0.30, p = 0.04). The same effect was also not significant in the Active Control group (0.28, p = 0.10; BS 95% CI: − 0.05/0.71), but the effect magnitude was statistically similar as in the Fit-MF-Mind condition (difference = 0.08, p = 0.81; BS 95% CI: − 0.43/0.52). Finally, the Fit-MF intervention did not show this relationship (0.002, p = 0.99; BS 95% CI: − 0.21/0.28), and both the Fit-MF and Fit groups did not differ from Active Control (p ≥ 0.38). Accounting for repeated-testing gains in Figure Series explained approximately 20% of the variance in post-intervention G_f latent scores in the Fit-MF-Mind condition ($R^2 = 0.21$), and notably lesser

Table 4
Group differences in longitudinal latent change in figure series and LSAT.

Test	Group	Pre-intervention variance	Change			Difference from control	
			Mean [BS 95% CI]	Variance	d	Mean [BS 95% CI]	d
Figure Series	Control (n = 108)	0.79*	0.27* [0.08/0.44]	0.38*	0.30		
	Fit (n = 105)	0.79*	0.44* [0.22/0.67]	1.07*	0.50	0.17 [− 0.11/0.46]	0.28
	Fit MF (n = 108)	0.72*	0.61* [0.43/0.79]	0.55*	0.72	0.34* [0.11/0.60]	0.55
	Fit MF Mind (n = 103)	0.43*	0.33* [0.09/0.57]	0.72*	0.50	0.07 [− 0.23/0.60]	0.11
LSAT	Control (n = 108)	0.32*	− 0.06 [− 0.20/0.08]	0.12	− 0.11		
	Fit (n = 105)	0.53*	− 0.004 [− 0.17/0.16]	0.23	− 0.01	0.05 [− 0.16/0.27]	0.14
	Fit MF (n = 108)	0.50*	− 0.10 [− 0.22/0.01]	0.07	− 0.14	− 0.05 [− 0.22/0.14]	− 0.14
	Fit MF Mind (n = 103)	0.35*	− 0.05 [− 0.20/0.11]	0.14	− 0.08	0.01 [− 0.20/0.22]	0.03

Note: BS 95% CI—bias-corrected bootstrapped 95% confidence intervals; d—standardized effect size. Coefficients are unstandardized for data normed to the sample average at pre-intervention. Control—active control; Fit—fitness only intervention; Fit-MF—fitness and cognitive training intervention; Fit-MF-Mind—fitness, cognitive training and mindfulness training intervention.

* $p < 0.05$.

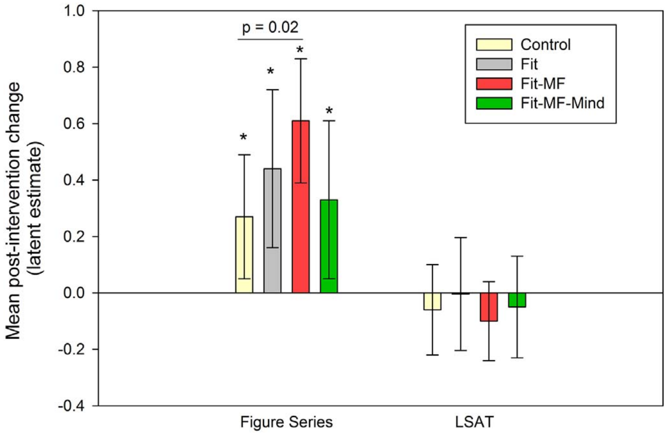

Fig. 4. Group differences in mean latent longitudinal change in Figure Series and LSAT. *$p < 0.05$; error bars represent two standard errors of the mean. Coefficients are unstandardized and normed to the pre-intervention average of the total sample. All groups demonstrated significant improvement in Figure Series, and group Fit-MF showed significantly greater gains as compared to the control group. Control—active control (*n* = 108); Fit—fitness only intervention (*n* = 105); Fit-MF—fitness and cognitive training intervention (n = 108); Fit-MF-Mind—fitness, cognitive training and mindfulness training (*n* = 103).

variance in the Active Control ($R^2 = 0.06$), Fit ($R^2 = 0.01$) and Fit-MF conditions ($R^2 < 0.01$), and additional covariates are expected to explain individual differences in test performance. Similar results were identified in the analysis of the sample with complete data. Within the Fit-MF-Mind condition, greater repeated-testing gains in Figure Series was associated with better performance on post-intervention G_f tests (0.33, $p < 0.01$) and the same effect was not significant in Active Control (0.28, $p = 0.09$), Fit-MF (0.002, $p = 0.99$) or Fit conditions (0.03, $p = 0.74$). Thus, post-intervention G_f scores in Fit-MF-Mind were low on average, but individuals within this condition who showed greater repeated-testing gains on Figure Series demonstrated relatively higher post-intervention G_f scores.

4. Discussion

Fluid intelligence is central to higher-order cognition, but the prospect of boosting its function has been met with mixed results and controversy (e.g., Hayes et al., 2015; Melby-Lervåg & Hulme, 2013) and the mechanism underlying intervention benefits to G_f is unclear (Greenwood & Parasuraman, 2015). Here we examined fitness, cognitive and mindfulness intervention activities that may optimally engage the neural and cognitive constituents of G_f ability when combined, for which we find no definitive evidence of differential gains from multi-modal intervention. However, the group engaged in equal parts fitness and cognitive training (Fit-MF) showed greater post-intervention gains than the Active Control in a canonical test of G_f (Figure Series), but not in verbal analogical reasoning (LSAT) or in performance on novel G_f tests that were administered only at post-intervention. The fitness-only condition produced no control-adjusted gains, therefore there appear to be limited benefits of a multi-modal intervention that combines fitness and adaptive cognitive training. A different pattern of evidence emerged for the intervention group that also engaged in mindfulness training: no apparent control-adjusted difference in group mean G_f, but the magnitude of repeated-testing gains in Figure Series predicted better test scores on novel tests of G_f within that group and this effect was greater than with exercise alone. Taken together, individuals appear to differ in their response to mindfulness intervention when combined with cognitive and fitness training, and on average, this type of multi-modal intervention may interfere with G_f task performance.

These two patterns of results provide insight into the discrepant findings within the cognitive intervention literature. We find greater post-intervention gains in visuospatial reasoning as compared to Active Control in the Fit-MF condition, in line with other intervention studies with pre-post comparisons (see Melby-Lervåg et al., 2016; Melby-Lervåg & Hulme, 2013 for reviews). Differentially greater gains in cognitive function from the combination of exercise and cognitive training have been reported in middle-aged and older adults (see Lauenroth, Loannidis, & Teichmann, 2016 for a review), yet older age and poorer health are expected to increase responsiveness to such intervention (Colcombe & Kramer, 2003). Here, we find evidence of gains in one task of G_f from a multi-modal intervention even in a sample consisting largely of healthy, young adults (Ward et al., 2017). Repeated-testing is a standard for assessing intervention benefits (e.g., Jaeggi et al., 2011; Redick et al., 2013; Thompson et al., 2013) and as applied here provides evidence of G_f gains as compared to a randomized Active Control condition, but there was no significant difference in performance on novel tests.

The mechanism of intervention-related benefits to a repeated test of G_f does not appear to mutually affect performance on novel tests to the same degree. This cannot be attributed to repeat exposure alone but it is possible that the sensitivity of psychometric tests of G_f is altered once an individual is familiar with the task design. Indeed, test-retest stability of fluid intelligence scores over time tend to be lower than that of crystallized intelligence (e.g., Heaton et al., 2014; Kaufman & Kaufman, 1993). In line with the evidence here, high-performing individuals commonly show lesser repeated-testing gains in fluid intelligence (Vernon, 1954) and this may contribute to variability in scores over time. Additional factors, including motivation (Duckworth, Quinn, Lynam, Loeber, & Stouthamer-Loeber, 2011) and attention (Greenwood & Parasuraman, 2015; Unsworth, Fukuda, Awh, & Vogel, 2014) are expected to impact performance on these tests and individuals may differ in the effects of prior exposure and intervention activities may interact with this. Indeed, individuals in the Fit-MF-Mind intervention who showed relatively greater gains in Figure Series also performed better on novel tests of G_f at post-intervention, which accounted for approximately 20% of individual differences in post-intervention G_f test scores within that group. In comparison, the Fit-MF group showed greater repeated-testing gains in G_f test scores relative to control, but no relationship between the magnitude of these gains and novel test performance.

Mindfulness meditation is known to promote attentional control (Tang et al., 2015) and minimize anxiety-related rumination that impairs executive functions (Beauchemin, Hutchins, & Patterson, 2008; Brown et al., 2011; Hofmann, Sawyer, Witt, & Oh, 2010; Mrazek et al., 2013; Yang, Cao, Shields, Teng, & Liu, 2016), which together are expected to bolster cognitive training effects on G_f (Greenwood & Parasuraman, 2015), but we find mixed effects within this condition. On average, the Fit-MF-Mind group did not differ from Active Control on tests of G_f, and produced lesser gains in Figure Series as compared to Fit-MF. However, within the Fit-MF-Mind group, individuals varied in their responsiveness to the intervention, and those who showed relatively greater gains in Figure Series also performed better on novel tests of G_f at post-intervention. Mindfulness is expected to require cognitive and attentional resources that are also demanded during task performance, and it appears that individuals differ in their ability to adapt to these demands and benefit from the exercise. Moreover, this effect was evident in performance on Figure Series—a canonical research tool of visuospatial reasoning—and not in an adapted achievement test of verbal analogical reasoning (LSAT). This is incongruent with prior reports of unimodal mindfulness interventions that improve performance on standardized aptitude tests, including the graduate record examinations (GRE) that is constructed similar to the LSAT (Mrazek et al., 2013). It is noteworthy that LSAT did not demonstrate significant repeated-testing gains in this sample, and thus appears to be less vulnerable to repeated-testing effects as compared to Figure Series. One difference between this study and prior reports with GRE is the introduction of mindfulness training as part of a multi-modal intervention including adaptive cognitive and fitness training. The cause of differential test sensitivity to this hypothesized attentional control component in the context of multi-modal intervention is unknown; but the evidence is consistent with extant reports of larger cognitive-training intervention effects in tests of visuospatial ability (Melby-Lervåg & Hulme, 2013). An intentional study of individual differences in these various cognitive and affective components, including measures of executive function, would be necessary to disentangle the contributions of each to G_f task performance.

The intervention-related effects are strikingly different between the Fit-MF and Fit-MF-Mind conditions. The intervention design differed between these groups in two regards: the Fit-MF-Mind condition completed fewer cognitive training sessions and engaged in additional mindfulness meditation. Thus, it is unclear if the addition of mindfulness training was detrimental to the intervention-related benefits of combined fitness and cognitive training, or if the reduced time spent on cognitive training in that group was insufficient to produce gains. Cognitive training duration explains some variability between studies (Melby-Lervåg & Hulme, 2013) and longer intervention duration (e.g., > 10 h) may minimize the transfer of working memory training to G_f (Melby-Lervåg et al., 2016). The total duration of intervention for each group reported here was the same (56 h total), but the Fit-MF-Mind group spent only 10 h on cognitive training whereas the Fit-MF condition completed 20 h. The pattern of effects here does not conform with a dose-response curve identified in a recent meta-analysis (Melby-Lervåg et al., 2016) and the possible trade-off between mindfulness training and cognitive training duration remains to be addressed in future studies that include a fully-crossed design of the intervention activities.

The study of individual differences proved to be critical in the present analyses and brought to light variability in the response to mindfulness training on G_f test performance. However, the analyses were limited to tests of G_f and repeated-testing gains that were assessed with LSAT and Figure Series tests separately, limiting the interpretation to test-specific performance. The examined effects explained only a small amount of variability in test scores within each group. Thus, it is likely that other covariates are relevant, including possible intervention group differences in processing speed and working memory capacity (Ackerman, Beier, & Boyle, 2005; Fry & Hale, 2000; Harrison et al., 2013), as well as brain structure and function (Demirakca, Cardinale, Dehn, Ruf, & Ende, 2016; Thomas, Dennis, Bandettini, & Johansen-Berg, 2012; Voss et al., 2010). Unlike cognitive benefits of fitness interventions among older adults (Colcombe & Kramer, 2003), the younger adults who engaged only in fitness training (Fit) showed no effects in G_f scores. Each intervention included fitness training, designed similar to scaled doses, and it is unlikely that differences in fitness training duration explained the effects observed in the Fit-MF and Fit-MF-Mind groups. However, G_f test scores may be too distal an outcome of fitness-related benefits to detect with group-level differences among young adults. Measures of brain structural integrity and functional connectivity (Demirakca et al., 2016; Thomas et al., 2012; Voss et al., 2010) may be more sensitive to individual differences in aerobic capacity that in turn explain additional variability in cognitive performance (Talukdar et al., 2017). We aim to examine these possible effects in future reports.

The results presented here should be interpreted in the context of several limitations. First, the sample was intentionally recruited to be young, healthy adults. Individual differences in G_f, although statistically significant, were likely smaller as compared to samples including older adults with compromised health. Meta-analyses demonstrate that physical fitness interventions produce larger effects in older and less-healthy samples (Colcombe & Kramer, 2003; Smith et al., 2010), and working memory training effects are more pronounced in older adults and patient populations (Danielsson et al., 2015; Karr et al., 2014; Kelly, Loughrey, & Lawlor, 2014; Spencer-Smith & Klingberg, 2015; Wicherts, 2016). Here we report small-moderate effect sizes of change in repeated measures of G_f and group differences therein that likely reflect the sample selection. The tests of these effects may be underpowered in the smaller group sizes, although this was a relatively large sample. Moreover, sample size was compromised by approximately 45% attrition. The total sample was included in the latent models to improve the validity of the hypothesis testing and missing data were handled via FIML—a non-imputation approach that is the recommended practice for longitudinal studies (Little & Card, 2013)—yet, we cannot dismiss the possibility of bias introduced by attrition. Nonetheless, comparable evidence was identified in the analyses restricted to the sample with complete data, and we consider this source of bias negligible. A further consideration when interpreting change in Figure Series performance is the weak factorial measurement invariance longitudinally in the entire sample. While this qualifies interpretation of the change in scores overall, the assumption of strong measurement invariance between groups was satisfied and thus the comparison between groups is still meaningful. Examining performance on the individual items of the Figure Series test and individual differences in item-level responses is necessary to further understand the sources of weak measurement invariance as a repeated-measure. Finally, our aim was to assess cognitive training, mindfulness training, and physical exercise as components of the putative mechanism of intervention gains to G_f, but the models do not directly assess causality. We can only speculate about the underlying source of the shared variance between repeated-testing gains on one test of G_f and performance on novel tests of the same cognitive construct. Moreover, the source of group differences between Fit-MF and Fit-MF-Mind may also in part be due to the order of the intervention, although both groups engaged in cognitive training immediately before the final assessment. The sustainability of intervention-related gains in repeated tests or in novel G_f scores is unclear as we made the assessment immediately after the conclusion of the intervention. Future studies may consider counterbalancing the order of each intervention modality within a multi-modal protocol and including another assessment after a long delay, which may clarify the possible trade-off in G_f ability when engaged in multiple intervention activities.

5. Conclusion

In a four-month, randomized control trial we evaluated a multi-modal intervention that engaged working memory, attentional control, and physiological contributors to G_f. The combination of fitness and adaptive cognitive training (Fit-MF) produced greater repeated-testing gains in a canonical G_f test of visuospatial reasoning (Figure Series) as compared to Active Control but the same effect was not seen in repeated testing of verbal analogical reasoning (LSAT) or novel tests that were only administered at post-intervention. In contrast, Fit-MF-Mind condition produced mixed benefits—on average, the group was similar to Active Control, but individuals within that condition who showed gains in Figure Series also performed better on novel tests of G_f. The mechanism of G_f function and the action of interventions aimed to promote it are complex—supported by multiple neural systems and cognitive functions (Barbey, 2017). Here, we demonstrate that a multi-modal intervention pairing equal time spent on fitness and adaptive cognitive training produced greater repeated-testing gains in a test of visuospatial reasoning. However, as evident in the Fit-MF-Mind condition, multiple activities may promote task attention to improve performance scores on tests of G_f without improving G_f ability per se. This underscores the complexity of the mechanism and the many avenues through which interventions may furnish effects. As intervention studies proliferate, the field requires a better understanding of the relevant mechanism to bolster G_f and the role of attentional control as one of its constituents.

Acknowledgments

The research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via Contract 2014-13121700004 to the University of Illinois at Urbana-Champaign (PI: Barbey). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes not- withstanding any copyright annotation thereon. Ana M. Daugherty was supported by a Beckman Institute Postdoctoral Fellowship at the University of Illinois at Urbana-Champaign, with funding provided by the Arnold and Mabel Beckman Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.intell.2017.11.001.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs. *Psychological Bulletin, 131*, 30–60.

American College of Sports Medicine (2014). *ACSM's guidelines for exercise testing and prescription* (9th ed). Lippincott, Williams, & Wilkins.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism, 21*, 1133–1145.

Baltes, P. B., Dittmann-Kohli, F., & Kliegel, R. (1986). Reserve capacity of the elderly in aging-sensitive tests of fluid intelligence: Replication and extension. *Psychology and Aging, 1*(2), 172–177.

Baniqued, P. L., Lee, H., Voss, M. W., Basak, C., Cosman, J. D., DeSouza, S., ... Kramer, A. F. (2013). Selling points: What cognitive abilities are tapped by casual video games? *Acta Psychologica, 142*(1), 74–86.

Banks, J. B., Welhaf, M. S., & Srour, A. (2015). The protective effects of brief mindfulness meditation training. *Consciousness and Cognition, 33*, 277–285.

Barbey, A. K. (2017). Network neuroscience theory of human intelligence. *Trends in Cognitive Sciences* [Epub ahead of print].

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex, 49*, 1195–1205.

Barbey, A. K., Colom, R., Paul, E. J., & Grafman, J. (2014). Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Structure & Function, 219*(2), 485–494.

Barbey, A. K., Colom, R., Solomon, J., Kreuger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain, 135*, 1154–1164.

Beauchemin, J., Hutchins, T. L., & Patterson, F. (2008). Mindfulness meditation may lessen anxiety, promote social skills, and improve academic performance among adolescents with learning disabilities. *Journal of Evidence-Based Complementary & Alternative Medicine, 13*, 34–35.

Brown, L. A., Forman, E. M., Herbert, J. D., Hoffman, K. L., Yuen, E. K., & Goetter, E. M. (2011). A randomized controlled trial of acceptance-based behavior therapy and cognitive therapy for test anxiety: A pilot study. *Behavior Modification, 35*, 31–53.

Buschkuehl, M., & Jaeggi, S. M. (2010). Improving intelligence: A literature review. *Swiss Medical Weekly, 140*(19–20), 266–272.

Carlson, L. E., & Brown, K. W. (2005). Validation of the mindful attention awareness scale in a cancer population. *Journal of Psychomsomatic Reseach, 58*(1), 29–33.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1–22.

Cheun, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233–255.

Chooi, W. T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Int, 40*, 531–542.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14*(2), 125–130.

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

Danielsson, H., Zottarel, V., Palmqvist, L., & Lanfranchi, S. (2015). The effectiveness of working memory training with individuals with intellectual disabilities—A meta-analytic review. *Frontiers in Psychology, 6*, 1230.

Demirakca, T., Cardinale, B., Dehn, S., Ruf, M., & Ende, G. (2016). The exercising brain: Change in functional connectivity induced by an integrated multimodal cognitive and whole-body coordination training. *Neural Plasticity, 2016*, 8240894.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168.

Dougherty, M. R., Hamovitz, T., & Tidwell, W. (2016). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin & Review, 23*, 303–316.

Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences of the United States of America, 108*(19), 7716–7720.

Elsayed, M., Ismail, A. H., & Young, R. J. (1980). Intellectual differences of adult men related to age and physical fitness before and after an exercise program. *Journal of Gerontology, 35*(3), 383–387.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology. General, 128*(3), 309–331.

Erickson, K. I., Hillman, C. H., & Kramer, A. F. (2015). Physical activity, brain, and cognition. *Current Opinion in Behavioral Science, 4*, 27–32.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychiatry, 54*, 1–34.

Gaspar, J. G., Neider, M. B., Simons, D. J., McCarley, J. S., & Kramer, A. F. (2013). Change detection: Training and transfer. *PLoS One, 8*(6), e67781.

Gomez-Pinilla, F., & Hillman, C. (2013). The influence of exercise on cognitive abilities. *Comprehensive Physiology, 3*(1), 403–428.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence, 24*, 79–132.

Greenwood, P. M., & Parasuraman, R. (2015). The mechanisms of far transfer from cognitive training: review and hypothesis. *Neuropsychology, 30*(6), 742–755.

Hagmann-von Arx, P., Gygi, J. T., Weidmann, R., & Grob, A. (2016). Testing relations of crystallized and fluid intelligence and the incremental predictive validity of conscientiousness and its facets on career success in a small sample of German and Swiss workers. *Frontiers in Psychiatry, 7*, 500.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science, 24*(12), 2409–2419.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter?

Association for Psychological Science, 24(10), 1918–1927.

Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2015). Do we really become smarter when our fluid-intelligence test scores improve? *Intelligence, 48*, 1–14.

Heaton, R. K., Akschoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., ... Gershon, R. (2014). Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *Journal of the International Neuropsychological Society, 20*(6), 588–598.

Hertzog, C., & Schaie, K. W. (1986). Stability and change in adult intelligence: 1. Analysis of longitudinal covariance structures. *Psychology and Aging, 1*(2), 159–171.

Hertzog, C., & Schaie, K. W. (1988). Stability and change in adult intelligence: 2. Simultaneous analysis of longitudinal means and covariance structures. *Psychology and Aging, 3*(2), 122–130.

Hertzog, C., von Oertzen, T., Ghisletta, P., & Lindenbrger, U. (2008). Evaluating the power of latent growth curve models to detect individual differences in change. *Structural Equation Modeling: A Multidisciplinary Journal, 15*, 541–563.

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews. Neuroscience, 9*(1), 58–65.

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 169–183.

Hogrefe (2008). *Measuring intelligence with culture fair tests.* Oxford: Hogrefe Ltd.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alterantives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*(3), 340–362.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Science, 105*(19), 6829–6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive taining and transfer. *Memory and Cognition, 42*, 464–480.

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., YF, S., Jonides, J., & Perig, W. J. (2010). The relationship between n-back performance and matrix reasoning—Implications for training and transfer. *Intelligence, 38*, 625–635.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Praeger.

Kane, M. J., Hambrick, D. Z., Tuhoski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12*(6), 978–990.

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

Kaufman, A. S., & Kaufman, N. L. (1993). *Kaufman adolescent and adult intelligence test.* Pearson.

Kelly, M. E., Loughrey, D., & Lawlor, B. (2014). The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: A systematic review and meta-analysis. *Ageing Research Reviews, 15*, 28–43.

Kleemeyer, M. M., Polk, T. A., Schaefer, S., Bodammer, N. C., Brechtel, L., & Lindenberger, U. (2017). Exercise-induced fitness changes correlate with changes in neural specificity in older adults. *Frontiers in Human Neuroscience, 11*, 123.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., ... Westerberg, H. (2005). Computerized training of working memory in children with ADHD—A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 177–186.

Klingberg, T., Fonsberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*, 781–791.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Forum, 315*, 1080–1081.

Larsen, R. (2011). Missing data imputation versus full information maximum likelihood with second-level dependencies. *Structural Equation Modeling: A Multidisciplinary Journal, 18*(4), 649–662.

Lauenroth, A., Loannidis, A. E., & Teichmann, B. (2016). Influence of combined physical and cognitive training on cognition: A systematic review. *BMC Geriatrics, 16*, 141.

Little, T. D., & Card, N. A. (2013). *Longitudinal structural equation modeling.* Guilford Press.

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science, 14*(3), 582–590.

Martinez, K., Burgaleta, M., Román, F. J., Escorial, S., Shih, P. C., Quiroga, M. A., & Colom, R. (2011). Can fluid intelligence be reduced to 'simple' short-term storage? *Intelligence, 39*, 473–480.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605.

McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen, & H. W. Reese (Eds.). *Life span developmental psychology: Methodological contributions* (pp. 223). Erlbaum.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Woodcock-Johnson III normative update: Technical manual.* Riverside Publishing.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychobiology, 49*(2), 270–291.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*(4), 512–534.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*(1), 46–60.

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurement of general aptitude. *Journal Experimental Psychology: General, 141*(4), 788–798.

Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science, 24*(5), 776–781.

Muthén, B. O., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that not missing completely at random. *Psychometrika, 52*, 431–462.

Noone, C., Bunting, B., & Hogan, M. J. (2016). Does mindfulness enhance critical thinking? Evidence for the meditating effects of executive functioning in the relationship between mindfulness and critical thinking. *Frontiers in Psychiatry, 6*, 2043.

Oberauer, K., Schulze, R., Wilhelm, O., & Süss, H. M. (2005). Working memory and intelligence—Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*(1), 61–65.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7*(1), 75–79.

Pontifex, M. B., Hillman, C. H., Fernhall, B., Thompson, K. M., & Valentini, T. A. (2009). The effect of acute aerobic and resistance exercise on working memory. *Medicine and Science in Sports and Exercise, 41*(4), 927–934.

Pontifex, M. B., Parks, A. C., O'Neil, P. C., Egner, A. R., Warning, J. T., Pfeiffer, K. A., & Fenn, K. M. (2014). Poorer aerobic fitness relates to reduced integrity of multiple memory systems. *Cognitive, Affective, & Behavioral Neuroscience, 14*, 1132–1141.

Raz, N., & Rodrigue, K. M. (2006). Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience and Biobehavioral Reviews, 30*(6), 730–748.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., ... Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology. General, 142*(2), 359.

Reed, J. A., Einstein, G., Hahn, S. P., Gross, V. P., & Kravitz, J. (2010). Examining the impact of integrating physical activity on fluid intelligence and academic performance in an elementary school setting: A preliminary investigation. *Journal of Physical Activity & Health, 7*(3), 343–351.

Roussos, L. A., & Norton, L. L. (1998). *LSAT item-type validity study: Law school admission council technical report 98-01.* Law School Admission Council, Inc.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *The Journal of Applied Psychology, 88*, 1068–1081.

Schaie, K. W. (1985). *Schaie-Thurstone adult mental abilities test manual.* Consulting Psychologists Press.

Schwaighofer, M., Fischer, F., & Buhner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*, 138–166.

Schwarb, H., Johnson, C. L., Daugherty, A. M., Hillman, C. H., Kramer, A. F., Cohen, N. J., Barbey, A. K., et al. (2017). Aerobic fitness, hippocampal viscoelasticity, and relational memory performance. *Neuroimage, 153*, 179–188.

Scott, S. P., De Souza, M. J., Koehler, K., Petkus, D. L., & Murray-Kolb, L. E. (2016). Cardiorespiratory fitness is associated with better executive function in young women. *Medicine & Science in Sports & Exercise, 48*(1), 1994–2002.

Scott, T., Reading, J., & Shephard, R. J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sport Sciences, 17*(4), 338–345.

Shallice, T. (1982). Specific impairments of planning. *Philosophical transactions of the Royal Society of London, 298*(1089), 199–209.

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shipley-2.* Western Psychological Services.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138*(4), 628–654.

Singh-Manoux, A., Hillsdon, M., Brunner, E., & Marmot, M. (2005). Effects of physical activity on cogntiive functioning in middle age: Evidence from the Whitehall II prospective cohort study. *American Journal of Public Health, 95*(12), 2252–2258.

Smith, P. J., Blumenthal, J. A., Hoffman, B. M., Cooper, H., Strauman, T. A., Welsh-Bohmer, K., ... Sherwood, A. (2010). Aerobic exercise and neurocognitive performance: A meta-analytic review of randomized controlled trials. *Psychosomatic Medicine, 72*(3), 239–252.

Spearman, C. (1927). *The abilities of man.* Macmillan.

Spencer-Smith, M., & Klingberg, T. (2015). Benefits of a working memory training program for inattention in daily life: A systematic review and meta-analysis. *PLoS One, 10*(3), e0119522.

Spryidaki, E. C., Simos, P., Avgoustinaki, P. D., Dermitzaki, E., Venihaki, M., Bardos, A. N., & Margioris, A. N. (2014). The association between obesity and fluid intelligence impairment is mediated by chronic low-grade inflammation. *The British Journal of Nutrition, 112*(10), 1724–1734.

Talukdar, T., Nikolaidis, A., Zwilling, C. E., Paul, E. J., Hillman, C. H., Cohen, N. J., ... Barbey, A. K. (2017). Aerobic fitness explains individual differences in the functional brain connectome of healthy young adults. *Cerebral Cortex,* 1–10 [Epub ahead of print].

Tang, Y. Y., Holzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews. Neuroscience, 16*, 213–225.

Thomas, A. G., Dennis, A., Bandettini, P. A., & Johansen-Berg, H. (2012). The effects of aerobic activity on brain structure. *Frontiers in Psychology, 3*, 86.

Thompson, T. W., Waskom, M. L., Garel, K. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., & Gabriali, J. D. E. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS One, 8*(5), 1–15.

Thorvaldsson, B., Hofer, S. M., Hassing, L. B., & Johansson, B. (2005). Cognitive change as conditional on age heterogeneity in onset of mortality-related processes and repeated testing effects. In S. M. Hoffer, & D. F. Alwin (Eds.). *Handbook of cognitive Aging: Interdisciplinary Perspectives* (pp. 248). Sage Publications.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71*, 1–26.

van Heugten, C. M., Ponds, R. W. H. M., & Kessels, R. P. C. (2016). Brain training: Hype or hope? *Neuropsychological Rehabilitation, 26*(5–6), 639–644.

Vernon, P. E. (1954). Practice and coaching effects in intelligence tests. *Ed Forum, 18*(3), 269–280.

Voss, M., Vivar, C., Kramer, A. F., & van Praag, H. (2013). Bridging animal and human models of exercise-induced brain plasticity. *Trends in Cognitive Sciences, 17*(10), 525–544.

Voss, M. W., Erickson, K. I., Prakash, R. S., Chaddock, L., Malkowski, E., Alves, H., ... Kramer, A. F. (2010). Functional connectivity: A source of variance in the association between cardiorespiratory fitness and cognition? *Neuropsychologia, 48*(5), 1394–1406.

Ward, N., Paul, E., Watson, P., Cooke, G. E., Hillman, C. H., Cohen, N. J., ... Barbey, A. K. (2017). Enhanced learning through multimodal training: Evidence from a comprehensive cognitive, physical fitness, and neuroscience intervention. *Sciences Report, 7*, 5808.

Warsch, J. R., & Wright, C. B. (2010). The aging mind: Vascular health in normal cognitive aging. *Journal of the American Geriatrics Society, 58*(Suppl. 2), S3119–24.

Wechsler, D. (2008). *Wechsler adult intelligence scale* (Fourth ed.). Pearson.

Wicherts, J. (2016). The importance of measurement invariance in neurocognitive ability testing. *The Clinical Neuropsychologist, 30*(7), 1006–1016.

Yang, Y., Cao, S., Shields, G. S., Teng, Z., & Liu, Y. (2016). The relationship between rumination and core executive functions: A meta-analysis. *Depression and Anxiety, 00*, 1–14.

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition, 56*(3), 286–292.